Probability in the Engineering and Informational Sciences, **16**, 2002, 1–17. Printed in the U.S.A.

COMPARISONS BETWEEN CERTAIN QUEUING SYSTEMS

TORGNY LINDVALL
Department of Mathematics
University of Göteborg
41296 Göteborg, Sweden
E-mail: lindvall@math.chalmers.se

We use a Poisson imbedding technique to investigate the possibilities of generalizing some pathwise and stochastic monotonicity results from the $M/M/k$ queues to systems with monotone failure rate service time distribution. A dichotomy between the decreasing failure rate (DFR) and the increasing failure rate (IFR) cases is revealed: In the DFR case, we achieve that the number of customers in the systems is stochastically increasing if it is idle at time 0, whereas in the IFR case, there is an alternating character that renders a bound between two identical systems that have different initial conditions. We also explore how our methods work in a comparison between systems with different numbers of stations but the same maximal capacity.

1. INTRODUCTION

Consider two standard $M/M/k$ queuing systems with the same arrival and service intensities, but different initial conditions. Let Y_t and Y_t' denote the number of customers in the systems (to be referred to as the first and second one) at time $t \geq 0$, and let λ and λ' denote their respective initial distributions. If λ is stochastically dominated by λ', $\lambda \overset{\mathcal{D}}{\leqq} \lambda'$ for short, then it is easy to produce a coupling of $Y = (Y_t)_0^\infty$ and $Y' = (Y_t')_0^\infty$ such that

$$Y_t \leq Y_t' \quad \text{for all } t \geq 0. \tag{1}$$

In fact, this holds for a wide class of systems, at least when $Y_0 = 0$; we shall return to that aspect later, but, for now, we remain in the $M/M/k$ setting.

If $Y_0 = 0$, then

$$Y_t \text{ is stochastically increasing in } t. \tag{2}$$

© 2002 Cambridge University Press 0269-9648/02 $12.50

https://doi.org/10.1017/S0269964802161018 Published online by Cambridge University Press

To prove that well-known monotonicity result, inequality (1) comes into use. It is probably less known that if $Y_0 = 0$, then

$$\mathbb{E}[Y_t] \text{ is concave in } t. \tag{3}$$

For that result, a careful coupling and further theory are necessary. The reader is referred to Lindvall [2, Sect. V.4] for all the details concerning (2) and (3). Also, a few hints are given in Section 2, where we recall the theory for monotonicity of Markov processes that sustains (2) and (3) and a generalization in this article.

Our main topic is the following: If we rather consider two $M/G/k$ queuing systems, where G is a distribution with monotone failure rate, is it then possible to achieve (2) and possibly (3)? In the theory we rely on, we need an inequality that is sharper than (1) for our analysis.

It turns out that there is a dichotomy. If G has a decreasing failure rate (DFR) function, then (1) and (2) can always be achieved. In contrast, our methods give us no hope that (2) holds in general if G has an increasing failure rate (IFR) function. In the latter case it is the possibility to obtain bounds for $Y_t' - Y_t$ that is interesting: If $k = 1, Y_0' = i_0'$, and $Y_0 = 0$, then

$$0 \le Y_t' - Y_t \le i_0' \quad \text{for all } t \ge 0. \tag{4}$$

There is a similar dichotomy in renewal theory when the life-length distribution involved is of the DFR or IFR type (cf. Lindvall [2, Sects. V.22 and V.23]).

We finish by noting a number of observations on another topic, namely comparisons between systems with different numbers of service stations; assume now that the second system has k' stations where $k < k'$. For a comparison between $M/M/k$ and $M/M/k'$ queuing systems, the natural condition is the following:

$$k\delta = k'\delta', \tag{5}$$

where δ and δ' are the service intensities in the first and second system, respectively. Equality (5) means that the intensity for finishing a service is the same in the two systems if all stations are busy. It turns out that if the initial conditions are the same, then (1) can be achieved; that speaks in favor of the principle to use fewer but quicker stations.

Is that principle valid in general? For our analysis, assume that the service time distributions in the two systems have failure rate functions r and r', respectively (beware of the notation r'!), such that

$$kr(x) = k'r'(x) \quad \text{for } x > 0, \tag{6}$$

which is a natural generalization of (5). It turns out that our methods fail in the attempt to obtain interesting results under condition (6). The concluding section contains a number of reflections on that.

Considerable attention has been paid to the comparisons of queues (cf. Stoyan [4, Chaps. 5 and 6] and Wolff [6, Chap. 11]). Pathwise methods and stochastic monotonicity of imbedded Markov chains often come into use. For a result in the

https://doi.org/10.1017/S0269964802161018 Published online by Cambridge University Press

spirit of this article, see, for example, Stoyan [4, Thm. 6.2.3], which proves that if we compare two k-station systems that are idle at time 0 and where the services are slower in the second system, then we have stochastic domination for the queue lengths at any time. The strand of theory leading to this result is related to ours, in certain respects.

Imbedding in Poisson processes is a versatile method, often efficient and simple to use. Since we limit ourselves to service times with densities, we are able to use a Poisson process in the first quadrant to govern the completions of service times; that is crucial in our study.

The plan is as follows. After the Markov process domination results in Section 2, the Poisson imbedding to be used is presented in Section 3. Our analyses of the DFR and IFR service time cases are in Sections 4 and 5, respectively. The comparisons of systems with different numbers of stations are carried out in Sections 6 and 7. The brief Section 6 presents a suitable coupling of $M/M/k$ and $M/M/k'$ processes, whereas the DFR and IFR cases are discussed in Section 7.

2. STOCHASTIC MONOTONICITY OF MARKOV PROCESSES

We refer to Lindvall [2, Sect. IV.3] or Lindvall [3, Sect. 2] for detailed accounts of stochastic monotonicity of Markov processes; here, we make a brief recollection of the aspects relevant for us.

Let $(E, \mathcal{E})$ be a measurable space. We endow E with a partial ordering $\preccurlyeq$. A real-valued function f on E is increasing if

$$x \preccurlyeq y \Rightarrow f(x) \le f(y).$$

For probability measures P and P' on $(E, \mathcal{E})$, P is stochastically dominated by P' if

$$\int f\,dP \le \int f\,dP' \quad \text{for all increasing and bounded measurable functions } f \text{ on } E.$$

We write that $P \overset{\mathcal{D}}{\preccurlyeq} P'$. For random elements Y and Y' in E, the meaning of $Y \overset{\mathcal{D}}{\preccurlyeq} Y'$ is the obvious one. We say that a process $(Y_t)_0^\infty$ with E as state space is stochastically increasing if $Y_s \overset{\mathcal{D}}{\preccurlyeq} Y_t$ whenever $s \le t$.

Now, suppose that there is a minorizing element a in E:

$$a \preccurlyeq x \quad \text{for all } x \in E. \tag{7}$$

Let $(P_t)_0^\infty$ be a semigroup of transition probabilities with the following property:

> For every distribution μ on $(E, \mathcal{E})$, there exists a coupling (X, X') of Markov processes $X = (X_t)_0^\infty$ and $X' = (X'_t)_0^\infty$ governed by $(P_t)_0^\infty$, with $X_0 = a$ and $X'_0 \overset{\mathcal{D}}{=} \mu$, such that $X_t \preccurlyeq X'_t$ for all $t \ge 0$. **(8)**

https://doi.org/10.1017/S0269964802161018 Published online by Cambridge University Press

If (8) is satisfied, then a simple application of the Markov property renders

$$(X_t)_0^\infty \text{ is stochastically increasing if } X_0 = a. \tag{9}$$

It follows immediately that if g is an increasing measurable function on E, then $(g(X_t))_0^\infty$ is stochastically increasing.

We now have sufficient background for statement (2) on $M/M/k$ queues. Indeed, (8) is satisfied for any birth and death process. To prove that, use a Doeblin coupling: Let $X_0 = 0$ and $X_0' \stackrel{\mathcal{D}}{=} \mu$, and let X_t and X_t' evolve independently until they meet. After that, let X_t and X_t' coincide. Obviously, $X_t \le X_t'$ for all $t \ge 0$ since birth and death processes have skip-free paths.

Let us return to the general theory. We say that

$$\begin{aligned}&\text{the real-valued process } Y = (Y_t)_0^\infty \text{ has stochastically}\\ &\text{decreasing increments if } Y_{t+\epsilon} - Y_t \text{ is stochastically}\\ &\text{decreasing in } t \text{ for every } \epsilon > 0.\end{aligned} \tag{10}$$

In short, $(Y_t)_0^\infty$ has the stochastically decreasing increments (SDI) property. If so, it is immediate that $\mathbb{E}[Y_t]$ is concave in t.

Now, assume

$$\begin{aligned}&\text{The pathwise domination in (8) holds for any initial}\\ &\text{values } x \text{ and } x' \text{ of } (X_t)_0^\infty \text{ and } (X_t')_0^\infty \text{ such that } x \preccurlyeq x'.\end{aligned} \tag{11}$$

Let $g : E \to \mathbb{R}$ and assume further that

$$\begin{aligned}&\text{the coupling of } (X_t)_0^\infty \text{ and } (X_t')_0^\infty \text{ is such that}\\ &g(x') - g(x) \ge g(X_\epsilon') - g(X_\epsilon) \text{ for all } \epsilon > 0,\\ &\text{whenever } x \preccurlyeq x'.\end{aligned} \tag{12}$$

Under (11) and (12), we have

$$g(X_t),\ t \ge 0, \text{ has the SDI property if } X_0 = a. \tag{13}$$

We return to the $M/M/k$ queue. For $g(x) = x$, one may, indeed, produce a coupling satisfying (11) and (12). Hence, (13) holds if $X_0 = 0$, and we have explained the background of (3).

For our analysis of the $M/G/k$ queue, with G of DFR type, the choice of E and $\preccurlyeq$ becomes crucial.

3. POISSON IMBEDDING

Recall that a Poisson process ξ, say, in $\mathbb{R}_+^2$ with expectation measure ℓ is a point process such that

https://doi.org/10.1017/S0269964802161018 Published online by Cambridge University Press

(i) $\xi(B) \stackrel{\mathcal{D}}{=} \text{Poi}(\ell(B))$ for $B \in \mathcal{R}_+^2$
(ii) $\xi(B_1), \xi(B_2), \ldots, \xi(B_n)$ are independent for any class of disjoint sets $B_1, B_2, \ldots, B_n \in \mathcal{R}_+^2$. **(14)**

Here, ℓ is the Lebesgue measure on $\mathbb{R}_+^2$ and $\mathcal{R}_+^2$ is the class of Borel sets in $\mathbb{R}_+^2$. Let H be a distribution on $(0,\infty)$ with failure rate function ρ. Set

$$B = \{(x,y); x > 0, y \geq 0, y \leq \rho(x)\}$$

and

$$B_t = \{(x,y) \in B; x \leq t\} \quad \text{for } t > 0.$$

Define the random variable Z by

$$Z = \inf\{t > 0; \xi(B_t) \geq 1\}$$

(i.e., the x coordinate for the first Poisson point under the curve $(x, \rho(x)), x > 0$). We have $Z \stackrel{\mathcal{D}}{=} H$ because for any $x > 0$, it holds that

$$\begin{aligned}\mathbb{P}(Z > x) &= \mathbb{P}(\xi(B_x) = 0) \\ &= \exp(-\ell(B_x)) = \exp\left(-\int_0^x \rho(s)\, ds\right),\end{aligned}$$

which equals $\bar{H}(x)$, as is well known.

Now, let $H_1, \ldots, H_n$ be life-length distributions with failure rate functions $\rho_1(x), \ldots, \rho_n(x)$, $x > 0$. With

$$B_j = \left\{(x,y); x > 0, y \geq 0, \sum_{i=1}^{j-1} \rho_i(x) < y \leq \sum_{i=1}^{j} \rho_i(x)\right\} \qquad \textbf{(15)}$$

and

$$B_{jt} = \{(x,y) \in B_j; x \leq t\}, \qquad t > 0,$$

for $1 \leq j \leq n$, we can use the above method to produce independent variables $Z_1, \ldots, Z_n$ by

$$Z_j = \inf\{t > 0; \xi(B_{jt}) \geq 1\},$$

due to (14) (ii). This will be crucial to us.

Note that once $\min(Z_1, \ldots, Z_n)$ has been observed, we may delete the strip (one of $B_1, \ldots, B_n$) where the first Poisson point in $\cup_{i=1}^n B_i$ appeared. Also, note that the restriction of ξ to $(\min(Z_1, \ldots, Z_n), \infty) \times \mathbb{R}_+$, with a shift of the zero point to $\min(Z_1, \ldots, Z_n)$, is a Poisson process with expectation measure 1 on $(0,\infty) \times \mathbb{R}_+$. This new Poisson process is independent of $\min(Z_1, \ldots, Z_n)$.

For accounts of several uses of Poisson imbedding, see Lindvall [2, Sects. V.4–V.6].

https://doi.org/10.1017/S0269964802161018 Published online by Cambridge University Press

4. THE DFR CASE

Let the service time distribution on $(0,\infty)$ for this and the following section be denoted by G and its failure rate function by r. To avoid some trite technicalities, we shall restrict our attention to service time distributions with unbounded support; that is automatic in the DFR case.

Suppose a service has gone on for a time units. The term "attained service time" for a is accepted, but we shall use "age" for convenience and say that a customer "is of age a"; this makes it possible to use, for example, the quick term "age configuration." Note that for a customer of age a, the failure rate of her remaining service time is

$$r(a+x), \qquad x>0, \tag{16}$$

which we denote by $r_a(x)$.

A state space for an $M/G/k$ system has to grasp two things: the number of customers and the age configuration of those at service. We will denote a configuration of ages by

$$\mathbf{a} = [a_1, a_2, \ldots, a_n]$$

when there are n customers, $n \le k$, at service. For this section, we use the convention that

$$\begin{aligned}&\text{the enumeration } a_1, a_2, \ldots, a_n \text{ of the configuration } \mathbf{a} \text{ of}\\ &\text{ages in the } \textit{first} \text{ system is always done such that}\\ &a_1 \le a_2 < \cdots \le a_n;\end{aligned} \tag{17}$$

that is, we put the label 1 on the youngest customer, the label 2 on the second youngest, …, and the label n on the oldest of the n customers at service in the first system.

We avoid the set notation $\{\cdot\}$ for these configurations in order to allow multiple entries of equal ages, but we shall not hesitate to use standard operations for sets; for example,

$$[a_1,\ldots,a_n]\backslash[a_1] = [a_2,\ldots,a_n]$$

and

$$[a_1,\ldots a_n] \cup [b] = [a_1,\ldots,a_n,b].$$

Let $\mathcal{A}$ be the class of possible age configurations; we denote that of an idle system by ϕ. As state space for each of the two systems to be compared, we choose

$$E = \{(i,\mathbf{a});\, i \in \mathbb{Z}_+,\, \mathbf{a} \in \mathcal{A},\, \#\mathbf{a} = i \wedge k\}.$$

https://doi.org/10.1017/S0269964802161018 Published online by Cambridge University Press

The partial ordering $\preccurlyeq$ on E is defined as follows: for $(i, \mathbf{a})$ and $(i', \mathbf{a}')$, where $\mathbf{a}, \mathbf{a}' \in \mathcal{A}$ and $\mathbf{a} = (a_1, \ldots, a_n)$, let

$$(i, \mathbf{a}) \preccurlyeq (i', \mathbf{a}')$$

exactly when

$$i \leq i' \tag{18}$$

and

$$\begin{aligned} &a_m \leq a'_m,\ 1 \leq m \leq n, \\ &\text{for some enumeration } a'_1, \ldots a'_{n'} \\ &\text{of the ages in the second system.} \end{aligned}$$

Throughout this section, we shall only encounter pairs of states satisfying $(i, \mathbf{a}) \preccurlyeq (i', \mathbf{a}')$. We assume that the ages in $\mathbf{a}'$ are enumerated according to the matching requirement of (18); of course, if $n < n'$, then there are $n' - n$ customers in the second system whose labels $n+1, \ldots, n'$ are irrelevant.

We let the number of customers at time $t \geq 0$ be denoted by Y_t and Y'_t for the two systems; $X_t = (Y_t, A_t)$ and $X'_t = (Y'_t, A'_t)$, where A_t and A'_t are the respective age configurations at $t \geq 0$.

We assume, to begin with, that the customer arrival process is deterministic, with arrivals at time points $s_j, j \geq 1$, such that

$$0 < s_1 < s_2 < \cdots$$

and

$$s_j \to \infty \quad \text{as } j \to \infty. \tag{19}$$

For the crucial definitions (20) and (21), we have in mind that the stations in the two systems are numbered, from 1 to k, and that a customer with label j is served at station j. Since a customer may get a new label at an s_j time point or when a service is completed, she may also change station during her service. However, that is of no significance.

Let $\mathbf{a} = [a_1, \ldots, a_n]$ and $\mathbf{a}' = [a'_1, \ldots, a'_{n'}] \in \mathcal{A}$. Set, for $x > 0$,

$$\begin{aligned} \rho_j(x) &= r_{a_j}(x), \quad 1 \leq j \leq n, \\ \rho_j(x) &= 0, \quad n+1 \leq j \leq k, \\ \rho'_j(x) &= r_{a'_j}(x), \quad 1 \leq j \leq n', \\ \rho'_j(x) &= 0, \quad n'+1 \leq j \leq k, \end{aligned} \tag{20}$$

and

$$\rho^*_j(x) = \max(\rho_j(x), \rho'_j(x)), \quad 1 \leq j \leq k.$$

https://doi.org/10.1017/S0269964802161018 Published online by Cambridge University Press

Recall ξ from Section 3. In order to use (15) and the subsequent comments, define subsets of $\{(x,y); x>0, y\ge 0\}$ as follows, for $1\le j\le k$:

$$B_j^* = \left\{\sum_{i=1}^{j-1}\rho_m^*(x)\le y<\sum_{i=1}^{j}\rho_m^*(x)\right\},$$
$$B_j = \left\{\sum_{i=1}^{j-1}\rho_m^*(x)\le y<\sum_{i=1}^{j-1}\rho_m^*(x)+\rho_j(x)\right\}, \tag{21}$$

and

$$B_j' = \left\{\sum_{i=1}^{j-1}\rho_m^*(x)\le y<\sum_{i=1}^{j-1}\rho_m^*(x)+\rho_j'(x)\right\}.$$

The sets B_{jt}^*, B_{jt}, and B_{jt}', for $1\le j\le k$ and $t>0$, are

$$B_{jt}^* = \{(x,y)\in B_j^*; x\le t\}$$

and so on.

Now, take $(i_0,\mathbf{a}_0)$ and $(i_0',\mathbf{a}_0')$ such that $(i_0,\mathbf{a}_0)\preccurlyeq(i_0',\mathbf{a}_0')$; they will be the initial states of $(X_t)_0^\infty$ and $(X_t')_0^\infty$, respectively. We shall prove that

$$X_t\preccurlyeq X_t' \quad \text{for all } t\ge 0. \tag{22}$$

To that end, we observe, first, that we need to pay attention only to the jump points (arrivals and times for completed services) because, obviously, the inequality is retained between such time points.

Note that if $\mathbf{a}=[a_1,\dots,a_n]$ and $\mathbf{a}'=[a_1',\dots,a_{n'}']$ are such that $n\le n'$ and $a_m\le a_m'$ for $m\le n$, then the sets B_j^* and B_j are equal for $j\le n$ since $r(x), x>0$, is decreasing.

To prove (22), set $\mathbf{a}$ and $\mathbf{a}'$ preceding (20) to equal $\mathbf{a}_0$ and $\mathbf{a}_0'$, respectively. Let

$$\sigma = \inf\{t>0; \xi(\cup_1^k B_{jt}^*)=1\} \tag{23}$$

and

$$\sigma^0 = \min(\sigma, s_1).$$

It turns out that there are four possibilities ((i)–(iv)) for jumps, where (i)–(iii) appear when $\sigma^0<s_1$, and (iv) when $\sigma^0=s_1$. Now,

$$\text{locate the station } j \text{ such that } \xi(B_{j\sigma}^*)=1. \tag{24}$$

(i) $\xi(B_{j\sigma})=1$, $\xi(B_{j\sigma}')=1$: finished service in both systems.
We set

$$Y_\sigma = i_0-1,$$
$$Y_\sigma' = i_0'-1$$

https://doi.org/10.1017/S0269964802161018 Published online by Cambridge University Press

and

$$A_\sigma = [a_{01} + \sigma, \ldots, a_{0n} + \sigma] \backslash [a_{0j} + \sigma],$$
$$A'_\sigma = [a'_{01} + \sigma, \ldots, a'_{0n'} + \sigma] \backslash [a'_{0j} + \sigma].$$

(ii) $\xi(B_{j\sigma}) = 1$, $\xi(B'_{j\sigma}) = 0$: finished service in the first but not in the second system.
We set

$$Y_\sigma = i_0 - 1,$$
$$Y'_\sigma = i'_0$$

and

$$A_\sigma = [a_{01} + \sigma, \ldots, a_{0n} + \sigma] \backslash [a_{0j} + \sigma],$$
$$A'_\sigma = [a'_{01} + \sigma, \ldots, a'_{0n'} + \sigma].$$

(iii) $\xi(B_{j\sigma}) = 0$, $\xi(B'_{j\sigma}) = 1$: finished service in the second but not in the first system.
We set

$$Y_\sigma = i_0,$$
$$Y'_0 = i'_0 - 1$$

and

$$A_\sigma = [a_{01} + \sigma, \ldots, a_{0n} + \sigma],$$
$$A'_\sigma = [a'_{01} + \sigma, \ldots, a'_{0n'} + \sigma] \backslash [a'_{0j} + \sigma].$$

(iv) We set

$$Y_{s_1} = i_0 + 1,$$
$$Y'_{s_1} = i'_0 + 1$$

and

$$A_{s_1} = \begin{cases} [a_{01} + s_1, \ldots, a_{0n} + s_1] \cup [0] & \text{if } n < k \\ [a_{01} + s_1, \ldots, a_{0n} + s_1] & \text{if } n = k, \end{cases}$$
$$A'_{s_1} = \begin{cases} [a'_{01} + s_1, \ldots, a'_{0n'} + s_1] \cup [0] & \text{if } n' < k \\ [a'_{01} + s_1, \ldots, a'_{0n'} + s_1] & \text{if } n' = k. \end{cases}$$

It is easy to check that $(Y_{\sigma^0}, A_{\sigma^0}) \preccurlyeq (Y'_{\sigma^0}, A'_{\sigma^0})$ in all four cases (i)–(iv); note that (iii) may appear only if $n < n'$, and then it is a station with a number greater than n that completes a service.

https://doi.org/10.1017/S0269964802161018 Published online by Cambridge University Press

For $0 \le t < \sigma^0$, we define X_t and X'_t as follows:

$$Y_t = i_0,$$
$$Y'_t = i'_0 \tag{25}$$

and

$$A_t = [a_{01} + t, \ldots, a_{0n} + t],$$
$$A'_t = [a'_{01} + t, \ldots, a'_{0n'} + t].$$

We content ourselves with a sketch of how to proceed in the production of X_t and X'_t for $t > \sigma^0$:

(26)
(i) Relabel the customers: enumerate the ages in A_{σ^0} and A'_{σ^0} according to (17) and the comment after (18).
(ii) Make σ^0 to a zero point, consider ξ restricted to $(\sigma^0, \infty) \times [0, \infty)$ with that new zero point.
(iii) Set $\mathbf{a}$ and $\mathbf{a}'$ preceding (20) to equal A_{σ^0} and A'_{σ^0} respectively, update (20) and (21) accordingly, and carry out (23)–(25).

Repeated use of (26) yields $(X_t)_0^\infty$ and $(X'_t)_0^\infty$ satisfying (22).

A standard construction of a probability space is now made in order to allow random initial values and arrival process; that is carried out in such a way that the arrival process satisfies (19) a.s. and is independent of the service times. We omit the details of that construction.

Note that if the first system is idle at time 0, then its initial state $(0, \phi)$ is trivially dominated by any initial state $(i', \mathbf{a}')$ of the second system.

It is due time to sum up our findings so far.

THEOREM 1: *Suppose the arrival process satisfies (19) a.s. and is independent of the initial values X_0 and X'_0 as well as the service times of the two systems. Further, suppose that*

$$X_0 \preccurlyeq X'_0,$$

which holds if the first system is idle at time 0. Then, in the DFR case, there exists a coupling $((X_t)_0^\infty, (X'_t)_0^\infty)$ such that

$$X_t \preccurlyeq X'_t \quad \text{for all } t \ge 0.$$

PROOF: Indeed, the proof is complete already, but in order to point out exactly where the DFR condition comes in, recall the comment between (22) and (23) and that preceding (25). ■

https://doi.org/10.1017/S0269964802161018 Published online by Cambridge University Press

COROLLARY 1: *Suppose the arrival process and the failure rate function r allow the existence of a stationary version. Then, if the first system is idle at time 0 and the second one is stationary, we have*

$$X_t \overset{\mathcal{D}}{\preccurlyeq} \pi \quad \text{for all } t \geq 0,$$

where π is the distribution of X_0'.

PROOF: The theorem yields that $X_t \overset{\mathcal{D}}{\preccurlyeq} X_t'$ for all $t \geq 0$. However, X_t' has the same distribution π for all $t \geq 0$. ■

COROLLARY 2: *Suppose that the first system is idle at time 0. Then, in the DFR case,*

$$Y_t \leq Y_t' \quad \text{for all } t \geq 0$$

regardless of the initial value Y_0'.

PROOF: Use the fact that $(i, \mathbf{a}) \preccurlyeq (i', \mathbf{a}')$ implies $i \leq i'$. ■

Corollary 2 should be seen as a natural example of what could be read directly from the theorem. However, this result holds for any service distribution, as long as the services are independent of the initial values and the arrival process. To prove that, we use a coupling as follows: Let the arrivals be the same for the two systems, as earlier. For each arrival, let the service times be identical for the two customers (one in the first system, one in the second system) that turn up at that time point. Then, it is a matter of a straightforward but careful study of the paths to prove that $Y_t \leq Y_t'$ for all $t \geq 0$ if $Y_0 = 0$. For that, one splits the time range into the intervals $[0, s_1), [s_1, s_2), \ldots$, where s_i, $i = 1, 2, \ldots$, are the arrival time points, and one considers the number of customers and the residual service times (more convenient than the ages) of the customers at service, in those time intervals, successively.

Remarks:

1. We made assumption (19) for convenience; also, it suffices for the $M/G/k$ results to follow. However, our methods work for any sequence $(s_j)_0^\infty$ such that $s_j \to \infty$ as $j \to \infty$; hence, we may allow batch arrivals.

   We could do without our special definition of age configurations if we are restricted to arrival processes satisfying (19); but such a restriction would be unfortunate.
2. The condition $X_0 \preccurlyeq X_0'$ may be weakened to $X_0 \overset{\mathcal{D}}{\preccurlyeq} X_0'$, due to Strassen's theorem on stochastic domination (cf. Lindvall [2, Sect. IV.1]).
3. Svensson [5] has demonstrated that Poisson imbedding works well for simulations of processes at least as complicated as the one we studied. Note that our proof of Theorem 1 is carried out in terms of an algorithm which could be used as a base for a simulation program.
4. The word "failure" is, of course, peculiar in a number of contexts where the term "failure rate" is used, but it and the abbreviations DFR and IFR are so well established that we employ them here too.

https://doi.org/10.1017/S0269964802161018 Published online by Cambridge University Press

For the rest of this section, assume that the customers arrive according to a Poisson process on $[0,\infty)$. We then have an $M/G/k$ system, and the purpose is to reveal its monotonicity properties. We observe that X_t, $t \geq 0$, is a Markov process; hence, the theory of Section 2 is at our disposal. Since Theorem 1 tells us that (8) is satisfied, we obtain the following result from (9).

THEOREM 2: *In the DFR case, we have that X_t is stochastically increasing in t if the system is idle at time 0. In particular, Y_t is stochastically increasing in t under that condition.*

PROOF: The latter result follows from the fact that $g : E \to \mathbb{Z}_+$ defined by $g(i, \mathbf{a}) = i$ is increasing. ■

Remarks:

1. The reader is referred to Lindvall [2, Sect. IV, Thm. 9.3], which supports a generalization of Theorem 2. Indeed, if D_E is the state space for $(X_t)_0^\infty$ and $\psi : D_E \to \mathbb{R}$ is increasing in the obvious sense, then

$$\psi(\theta_t X) \text{ is stochastically increasing in } t, \tag{27}$$

   where $(\theta_t x)(\cdot) = x(t + \cdot)$ defines the shift θ_t of $x = (x(s))_{s=0}^\infty \in D_E$. Of course, (27) is under the condition that the system is idle at time 0.
2. Recall (12) and (13). The concavity in (3) follows easily if we can prove that $(Y_t)_0^\infty$ has the SDI property; and it has. Indeed, the birth and death intensities of an $M/M/k$ system are such that we can produce a coupling satisfying

$$Y_t' - Y_t \text{ is decreasing in } t \tag{28}$$

   if $Y_0 \leq Y_0'$. However, that does not hold in the DFR case, at least with the coupling we have produced, which seems to be a natural one. The explanation is obvious: If a_j and a_j' are the ages of two matched customers, which means that $a_j \leq a_j'$, then the customer in the first system may complete service at some instant, but that does not occur for the matched customer in the second system. Hence, the topic of "stochastic concavity" for an $M/G/k$ system, idle at time 0 and with G of DFR type, is left open.

5. THE IFR CASE

Now, let r be increasing. Recall the assumption that the service time distribution G has unbounded support; in the IFR case, that holds if and only if r is bounded on finite intervals.

It is suitable to change the principle of labeling of customers; now, that is done in decreasing order of age. Except for this modification, the machinery of Section 4 is left intact.

If the key word for the DFR case was "monotonicity," it is now "alternation." A reader familiar with recent renewal theory is not surprised by this (cf. Lindvall [2, Sect. V.5]). Indeed, if N_t and N_t', $t \geq 0$, are the numbers of renewals in $[0, t]$ for

https://doi.org/10.1017/S0269964802161018 Published online by Cambridge University Press

a zero-delayed and, for example, a stationary renewal process, respectively, then if the life-length distribution F is of DFR type, we can produce a coupling such that

$$N_t - N_t' \text{ is increasing in } t.$$

That is not at all possible in the IFR case. However, there, we can achieve

$$N_t - N_t' \text{ equals 0 or 1 for } t \geq 0.$$

This motivates us to find bounds for $Y_t' - Y_t$, $t \geq 0$. To that end, let us concentrate on the case where the first system is idle at time 0, for clarity.

For the main result of this section, let $k = 1$: the single-server case. We simplify notation to a and a' for the ages of the customers in the first and second systems, if busy. One aim is to prove that $Y_t' - Y_t \geq 0$ for all $t \geq 0$. In the following, we pay attention to the jumps of $(X_t)_0^\infty$ and $(X_t')_0^\infty$ only at the time points of completed services; an arrival does not affect $(Y_t' - Y_t)_0^\infty$.

Now, consider the following types of state for X_t and X_t':

	First System	Second System	
(1)	(i, a)	$(i+1, a')$	$i \geq 1, a \leq a'$
(2)	(i, a)	$(i+1, a')$	$i \geq 1, a > a'$
(3)	$(0, \phi)$	$(1, a')$	
and			
(4)	$(j, 0)$	$(j+1, 0)$	$j \geq 1$
(5)	$(0, \phi)$	$(1, 0)$	
(6)	(j, a)	$(j, 0)$	$j \geq 1$
(7)	$(0, \phi)$	$(0, \phi)$	
(8)	$(1, a)$	$(1, 0)$	
(9)	$(j, 0)$	$(j+2, a')$	$j \geq 1$
(10)	$(0, \phi)$	$(2, a')$	

(29)

The process $Y_t' - Y_t$, $t \geq 0$, makes jumps ± 1. Suppose, now, that we have reached a state where $i' = i + 1$ [i.e., of type (1), (2), or (3)]. We find that at a time point where a service is completed, the transitions are

From	To
(1)	(4), (5), (6), or (8)
(2)	(4), (5), (9), or (10)
(3)	(7)

(30)

https://doi.org/10.1017/S0269964802161018 Published online by Cambridge University Press

However, all of states of type (4)–(10) are such that a jump to a state of type $(i, i-1)$ is impossible.

Now, suppose $X_0 = (0, \phi)$ and $X_0' = (i_0', a_0')$. With an analysis similar to but easier than that using (29) and (30) to prove $Y_t' - Y_t \geq 0$, we find that $Y_t' - Y_t \leq i_0'$ for all $t \geq 0$.

THEOREM 3: *For single-server systems, there exists a coupling* $((X_t)_0^\infty, (X_t')_0^\infty)$ *in the IFR case such that*

$$0 \leq Y_t' - Y_t \leq Y_0' \quad \text{for all } t \geq 0$$

if the first system is idle at time 0.

PROOF: The result is immediate, from the preceding arguments and a randomization of i_0', a_0', and $(s_j)_1^\infty$. ■

Remarks:

1. We have not come to grips with the case $k \geq 2$, but believe that the theorem is valid in general.
2. The case with infinitely many stations, and hence no waiting lines, is easier to handle because we are always able to match arriving customers immediately, in an obvious way. A moment's thought convinces the reader that Theorems 1–3 hold. She finds that the infinitely many-server case is also conceptually simpler; we can use $\mathcal{A}$ as state space and concentrate on the processes A_t and A_t', $t \geq 0$.

6. COMPARING *M/M/k* AND *M/M/k'*

Now, we have k and k' stations in the first and second systems, respectively, where $k < k'$. A first glance at the problem of comparing the efficiency of the systems, under some natural condition, makes the smaller but quicker system appealing, since we disapprove of idle stations.

Let $k = 1$ and suppose that

$$B \overset{\mathcal{D}}{=} B'/k', \tag{31}$$

where B and B' are typical service times in the first and second systems, respectively. (Let B and B' have that meaning throughout.) Then, in the $GI/G/\cdot$ case, one can prove pathwise domination of the queue length processes if both systems are idle at time 0 (cf. Asmussen [1, Chap. XI, Thm. 1.2]).

However, the single-server case is an exception: The server is busy to a maximal extent. If $k \geq 2$, one should not hope for strong domination results under general conditions.

Wolff [6, p. 258] hints, after an examination of the case $k = 1$ and $k' = 2$, that "it is easy to generalise the number of servers" in the $M/M/\cdot$ case and mentions a

https://doi.org/10.1017/S0269964802161018 Published online by Cambridge University Press

stochastic domination result for $k \geq 2$. The coupling we shall present is a special case of Lindvall [2, Sect. V.4, (4.19)].

Recall condition (4): $k\delta = k'\delta'$, where δ and δ' are the service intensities. Let the Poisson process of arrivals have intensity β. Then, $(Y_t)_0^\infty$ and $(Y_t')_0^\infty$ are birth and death processes, respectively, with birth intensities $\beta_i = \beta_i' = \beta$ for all $i \geq 0$, and death intensities $\delta_i = (k \wedge i)\delta$ and $\delta_i' = (k' \wedge i)\delta'$ for $i \geq 1$. We note that $\delta_i' \leq \delta_i$ for all $i \geq 1$, and $\delta_i' < \delta_i$ for $1 \leq i \leq k$, due to (4).

Now, let (Y_t, Y_t'), $t \geq 0$, be governed by the following intensity table:

From	To	With Intensity	
(i,j)	$(i+1, j+1)$	β	for $0 \leq i \leq j$,
	$(i-1, j)$	δ_i	
	$(i, j-1)$	δ_j'	for $1 \leq i < j$,
(i,i)	$(i-1, i-1)$	δ_i'	
	$(i-1, i)$	$\delta_i - \delta_i'$	for $i \geq 1$.

(32)

The transition intensities from states (i,j) such that $i > j$ are omitted; they are irrelevant here. For any initial values such that $Y_0 \leq Y_0'$, the intensities in (32) work to the effect that $Y_t \leq Y_t'$ for all t; no jumps to the set of states $\{(i,j); i > j\}$ ever take place.

7. FURTHER OBSERVATIONS: $k < k'$

Now, recall condition (5), $kr(x) = k'r'(x)$, $x > 0$, for the failure rate functions r and r' to be used now. To begin, consider the DFR case. Suppose that the arrival process is Poisson with intensity $\beta > 0$. The traffic intensities of the systems are

$$\beta E[B]/k \quad \text{and} \quad \beta E[B']/k', \tag{33}$$

respectively. For a comparison, put $R(x) = \int_0^x r(s)\,ds$ and $R'(x) = \int_0^x r'(s)\,ds$ and $\alpha = k/k'$. Recall that $\mathbb{P}(B > x) = e^{-R(x)}$ for $x \geq 0$. We get

$$\begin{aligned}
\mathbb{E}[B'] &= \int_0^\infty \mathbb{P}(B' > x)\,dx \\
&= \int_0^\infty e^{-R'(x)}\,dx \\
&= \int_0^\infty e^{-\alpha R(x)}\,dx \\
&\geq \int_0^\infty e^{-R(\alpha x)}\,dx \\
&= (1/\alpha)\int_0^\infty e^{-R(x)}\,dx \\
&= (1/\alpha)\mathbb{E}[B]
\end{aligned} \tag{34}$$

https://doi.org/10.1017/S0269964802161018 Published online by Cambridge University Press

since R is concave. Equality holds if and only if r is constant, a case we exclude for now. From (34), concerning the traffic intensities, we deduce that

$$\beta E[B]/k < \beta E[B']/k', \tag{35}$$

which speaks in favor of the first system. Recall that an $M/G/\cdot$ system is stable (has a stationary distribution) if and only if the traffic intensity is less than 1 (cf. Asmussen [1, Sect. XI.2]). This means that for certain values of β, the first system is stable while the second one is not.

In the IFR case, the inequality in (34) is reversed and so is the one in (35); this speaks in favor of the second system.

We are now interested in whether the methods of Sections 4 and 5 yield any domination results that support these observations. However, we should not be optimistic because the traffic intensities concern expectations whose comparisons seem hard to read off the paths.

We omit the details concerning the necessary modification of the setup in Section 4. However, note that the matching idea underlying (21) is not relevant now. We redefine the sets $B_j, 1 \le j \le k$, and $B'_j, 1 \le j \le k'$, as follows:

$$B_j = \left\{ \sum_{i=1}^{j-1} \rho_m(x) \le y < \sum_{i=1}^{j} \rho_m(x) \right\},$$
$$B'_j = \left\{ \sum_{i=1}^{j-1} \rho_m(x) \le y < \sum_{i=1}^{j} \rho_m(x) \right\}, \tag{36}$$

and the sets B_{jt} and B'_{jt} accordingly. To make (23)–(26) work, let

$$B^* = \bigcup_{j=1}^{k} B_j \cup \bigcup_{1}^{k'} B'_j \tag{37}$$

and

$$B^*_t = \{(x, y) \in B^*; x \le t\}.$$

Now, σ is defined by

$$\sigma = \inf\{t; \xi(B^*_t) = 1\}. \tag{38}$$

Then, we use (24)–(26) in an obvious way; note that services can be completed at the same time at stations with different numbers in the two systems.

Now, consider the DFR case, with $k = 1$ and $k' = 2$. We assume that

$$r(x) = \begin{cases} C, & 0 < x \le 1 \\ c, & x > 1, \end{cases} \tag{39}$$

where C is very large and c is very small. Let $i_0 = i'_0 = 1$ and $a_{01} = a'_{01} = 1.1$. It is quite clear how the two systems react to an intense stream of customers; in the first one, they queue for a long time, whereas in the second one, they are rapidly served by

https://doi.org/10.1017/S0269964802161018 Published online by Cambridge University Press

station 2. This means that $Y_t - Y_t'$ grows large to begin with; we realize that it is impossible to support, by a simple pathwise inequality, that the first system is more efficient than the second one.

Our coupling reveals nothing of interest in the IFR case. However, we remind about the possibility of using Poisson imbedding for simulations; a program based on (35)–(37) would be rather simple to write and well suited for comparisons of systems with different numbers of stations.

Acknowledgment

I am indebted to the referee for several valuable comments.

References

1. Asmussen, S. (1987). *Applied probability and queues.* Chichester: Wiley.
2. Lindvall, T. (1992). *Lectures on the coupling method.* New York: Wiley.
3. Lindvall, T. (1997). Stochastic monotonicities in Jackson queueing networks. *Probability in the Engineering and Informational Sciences* 11: 1–9.
4. Stoyan, D. (1983). *Comparison methods for queues and other stochastic models.* Chichester: Wiley.
5. Svensson, D. (2000). *A class of renewal processes in random environments.* [to appear] Ph.D. thesis, Chalmers University of Technology and Göteborg University, Göteborg, Sweden.
6. Wolff, R.W. (1989). *Stochastic modeling and the theory of queues.* Englewood Cliffs, NJ: Prentice-Hall.

https://doi.org/10.1017/S0269964802161018 Published online by Cambridge University Press