Stochastic Orders and Decision under Risk
IMS Lecture Notes – Monograph Series (1991)

STOCHASTIC ORDERING FOR MARKOV PROCESSES ON PARTIALLY ORDERED SPACES WITH APPLICATIONS TO QUEUEING NETWORKS

BY WILLIAM A. MASSEY

AT&T Bell Laboratories

The state spaces for queueing networks are intrinsically multidimensional. As a result, a theory of stochastic ordering for Markov processes on partially ordered spaces is a natural setting for the formulation of comparison theorems of queueing networks. The partially ordered structure of the state space gives rise to a variety of stochastic orders, that are distinct only when the space is *not* totally ordered. In particular, we can define stochastic orderings that are not equivalent to sample path comparisons. For all these orderings, we give a unified theory that allows us to compare Markov processes on such spaces, given their initial distributions and infinitesimal generators. Examples will be given to show how these results apply to queueing networks. In addition, we will give a simplified proof for a special case of Strassen's theorem as it applies to a sample path comparison of random variables.

0. Introduction. Queueing network theory has applications to such diverse areas as data communications, manufacturing, highway traffic, and population biology. Abstractly, we can view a queueing network as a collection of *sites* and *particles*. Sites can be thought of as nodes in a network, and particles as customers for the system. Each object has its own set of instructions as to how to move from one site to another, and how long to stay at any given site. The theory that develops is the method by which we analyze the flows between sites. In particular, we may have interest in how long any given object may take to traverse these sites, which is better known as the sojourn time. We may also want to know how many particles may accumulate at various sites, which we call the *queue length vector process*. It is the latter on which we will focus our attention.

Since we never have precise information on the arrival, service, or transfer patterns for particles in these networks, we are naturally led to the use of

---

AMS 1980 Subject Classification: Primary 60J99, 60K25; Secondary 60J80.

Key words and phrases: Monotone Markov processes, strong and weak orderings, Jackson network, Strassen's theorem.

stochastic processes to model these systems. Due to the complexity of queueing networks, the theory of continuous time Markov processes plays a larger role in their analysis than they do for single node systems. We start by modeling a network in terms of its initial configuration or distribution. How the system evolves in time is governed by a differential equation that is uniquely specified by an *infinitesimal generator*. Such a generator is characterized by the arrival, service, and transfer rates of the network as well as the routing and service disciplines. The basic paradigm for Markovian queueing network theory is to analyze the time evolution of the queue length process, given its initial distribution and infinitesimal generator. In terms of closed form solutions, a time dependent or transient analysis for even these Markovian systems can be difficult. Most results deal with the *relatively* easier task of finding the equilibrium distribution for the system. Prominent examples are the work of Jackson [5] and Kelly [9]. However, even an equilibrium analysis of networks becomes difficult when we want to treat networks with priority classes of finite buffers.

Formally, a *stochastic ordering* is a partial ordering for probability measures on the same state space. It provides a tool for doing a transient and equilibrium analysis of these networks, through the derivation of bounds and limit theorems. In our Markovian content, useful results would be ones where a stochastic comparison can be established purely in terms of initial distributions and generators. Results of this nature were first developed by Kalmykov [6] for one-dimensional Markov processes. Daley [3] identified the critical role that monotone Markov processes play in such theorems. He also gave necessary and sufficient conditions for monotonicity in terms of the transition functions on totally ordered spaces. Keilson and Kester [10] and Kester [11] gave many examples of monotone Markov processes, for the one-dimensional case, and derived other properties for them such as whether a given process is stochastically increasing in time. This approach is summarized in Chapter 4 of Stoyan [27]. For queueing networks however, our state space is multidimensional. Such spaces have, at best, a natural partial ordering relation. The fundamental result in stochastic sample path comparisons of random variables on partially ordered spaces, follows from a special case of Strassen's theorem [28]. These results were generalized by Kamae, Krengel, and O'Brien [8] for comparing stochastic processes on partially ordered Polish spaces.

The purpose of this paper is first to show that out of the partially ordered structure of the state space, a wider variety of stochastic orderings can be defined here than for totally ordered spaces. The appropriate analogy here is that of topological spaces. A topology can be defined by collection of open sets, and a weaker topology is obtained by restricting the type of open sets used. In this same manner, we can construct different stochastic orderings.

The strongest of these orderings corresponds to a sample path comparison of the processes. Moreover, the various orderings presented here differ only for partially ordered spaces that are *not* totally ordered. Thus these orderings are useful precisely for the case of queueing networks with multidimensional state spaces. For all these orderings, we develop a unified comparison theory for Markov processes on countable, partially ordered spaces. Such a theory makes it possible to establish a stochastic ordering between two Markov processes where no sample path comparison exists. This is indeed the case as we will show in the later examples.

The paper will summarize the results of [21]. In Section 1, we define a stochastic order and identify three natural candidates for orderings, the strong, weak, and weak*. Section 2 introduces the notion of a monotone Markov process, the main comparison theorems, and methods for constructing these monotone processes. We also generalize these results to time-inhomogeneous Markov processes. In Section 3 we show what additional results can be developed for the strong and weak orderings. In Sections 4 and 5 we give examples of how stochastic orderings can be applied to Jackson networks as well as birth-death-migration processes. Section 6 serves as an appendix for a proof of Strassen's theorem. The proof in Strassen [28] is far more general than is needed for the case of stochastic ordering, so we give a simplified argument for our special case of interest.

Finally, we want to emphasize that the work summarized here is only one part of a larger body of work on stochastic ordering for queueing networks. The papers of Baccelli, Shanthikumar, Whitt, and Yao explore different aspects of these issues.

1. Stochastic Orderings. Given a countably infinite, partially ordered set E, let $I(E)$ be a family of subsets of E that includes E itself and the empty set. We can then induce a transitive relation for probability measures of E. If P and Q are two such measures, we say that $P \leq_I Q$ whenever

$$P(\Gamma) \leq Q(\Gamma) \text{ for all } \Gamma \text{ in } I(E). \tag{1}$$

For all x and y in E, let δ_x be the point mass measure on E at x and similarly define δ_y. We will say that $\leq_I$ is a *stochastic ordering* on E if

1. $x \leq y$ if and only if $\delta_x \leq_I \delta_y$,
2. the relation $\leq_I$ is a partial order on the space of probability measures.

If X and Y are two E-valued random variables, we will say that $X \leq_I Y$ whenever their induced measures can be so ordered by $\leq_I$.

For any subset A of E, we borrow the following notation from Kamae and Krengel [7].

$$A^{\uparrow} = \{y \mid y \geq x \text{ for some } x \text{ in } A\} \tag{2}$$

$$A^{\downarrow} = \{y \mid y \leq x \text{ for some } x \text{ in } A\}. \tag{3}$$

We will define a subset A of E to be an *increasing set* if $A = A^{\uparrow}$. A family of increasing sets is said to be *strongly separating* if for all $x \not\leq y$, the family contains a set Γ such that $x \in \Gamma$ and $y \notin \Gamma$. Adapting the notion of *determining class* used in O'Brien [22] to a family of sets by using their indicator functions, we can show that $I(E)$ induces a stochastic order, if and only if $I(E)$ is a strongly separating family of increasing sets that form a determining class. Note that we are emphasizing only the partially ordered structure of the state space. In this framework we will not be discussing stochastic orderings such as convex orderings (see Baccelli, Massey and Towsley [1] and Kirstein [12] for examples).

While many stochastic orderings can be formulated, there are three natural ones that can be formed. Let $I_{st}(E)$, $I_{wk}(E)$, and $I_{wk^*}(E)$ denote respectively the strong, weak, and weak* orderings where

$$I_{st}(E) = \{\text{all increasing sets in } E\} \tag{4}$$

$$I_{wk}(E) = \{\{x\}^{\uparrow} \mid x \in E\} \cup \{E, \emptyset\} \tag{5}$$

$$I_{wk^*}(E) = \{(\{x\}^{\downarrow})^c \mid x \in E\} \cup \{E, \emptyset\}. \tag{6}$$

The strong ordering is the one that is equivalent to a sample path ordering of the random variables (see Strassen [28]. Weak orderings are equivalent to comparing tail distribution functions, and weak* orderings serve the same role for distribution functions. Examples of their usefulness can be found in Tong [29] and Stoyan [27]. The similarity of the nomenclature above to the various types of convergence on linear topological spaces is intentional. Just as weaker topologies are defined by restricting the family of open sets used, we can define stochastic orderings weaker than the strong one by restricting the family of increasing sets used. Continuing the analogy, recall that for finite dimensional spaces, all topologies are equivalent. Totally ordered spaces fill the corresponding role for stochastic orderings. It can be shown [17] that if E is a totally ordered space, then all of its stochastic orderings are equivalent.

Now let E and F be two partially ordered spaces, with $I(E)$ and $J(F)$ defining their respective stochastic orderings. We say that a function f from E to F is an *isotone mapping* from $I(E)$ to $J(F)$ if $f^{-1}(J(F)) \subset I(E)$. If $E = F$ and $f^{-1}(I(E)) \subset I(E)$, we say that f is $I(E)$-*isotone*.

PROPOSITION 1.1. *If f is an isotone mapping from $I(E)$ to $J(F)$, then for any two E-valued random variables X and Y we have*

$$X \leq_I Y \Rightarrow f(X) \leq_J f(Y). \tag{7}$$

Any isotone mapping is then an increasing function.

2. Monotone Markov Processes. Let $X(t)$ be a time-homogeneous Markov process with state space E. Its infinitesimal generator $\boldsymbol{A}$ acts as a linear operator on the space of bounded, real-valued functions on E. Given such a function f, the resulting new function will be denoted $\boldsymbol{A}f$. We will say that $\boldsymbol{A}$ is the generator of an *$I(E)$-monotone Markov process* if $X(t)$ and $Y(t)$ both have $\boldsymbol{A}$ as their generator, but

$$X(0) \leq_I Y(0) \text{ implies that } X(t) \leq_I Y(t). \tag{8}$$

In Shanthikumar and Yao [26], monotonicity is the key result necessary to derive their stochastic bounds for closed queueing networks. Now we show how monotone Markov processes can be used in general to determine stochastic bounds.

THEOREM 2.1. *Let $X(t)$ and $Y(t)$ both be Markov processes having E as their state space. If $\boldsymbol{A}$ and $\boldsymbol{B}$ are their respective generators, $\boldsymbol{A} \leq_I \boldsymbol{B}$, and either $X(t)$ or $Y(t)$ is $I(E)$-monotone, then*

$$X(0) \leq_I Y(0) \textit{ implies that } X(t) \leq_I Y(t) \tag{9}$$

where $\boldsymbol{A} \leq_I \boldsymbol{B}$ stands for the relation $\boldsymbol{A}1_\Gamma(x) \leq \boldsymbol{B}1_\Gamma(x)$ for all x in E, and 1_Γ is the indicator function of Γ.

Suppose that $X(t)$ and $Y(t)$ are two Markov processes having E and F, two different posets, as their respective state spaces. Moreover, let f be a mapping from E to F. Using the operator formalism [21], we can establish a similar criterion for comparison between $f(X(t))$ and $Y(t)$. This is a way of extending these results to get stochastic bounds for non-Markovian processes. This topic is pursued in more detail in Whitt [30]. Theorems can also be developed to establish the notion of being *$I(E)$-time increasing (decreasing)*. This is proving that $\Pr\{X(t) \in \Gamma\}$ is an increasing (decreasing) function of time for all Γ in $I(E)$. To construct $I(E)$-monotone generators, we appeal to the theorem below:

THEOREM 2.2. *The class of $I(E)$-monotone infinitesimal generators form a cone that is closed under the strong operator topology.*

We conclude this section by generalizing these results for the time-inhomogeneous Markov processes. The author [18] used these results for an asymptotic analysis of the time-inhomogeneous M/M/1 queue. Let $X(t_0, t)$ denote a time-inhomogeneous Markov process with state space E, and let $\{\boldsymbol{A}(s) \mid t_0 \leq s \leq t\}$ be its family of infinitesimal generators from time t_0 to time t. For simplicity, we let $X(t_0) = X(t_0, t_0)$. We define $\{\boldsymbol{A}(s) \mid t_0 \leq s \leq t\}$ to be the family of infinitesimal generators for an $I(E)$-monotone Markov process on the time interval $[t_0, t]$ if $X(t_0, t)$ and $Y(t_0, t)$ both have the same family of generators, and

$$X(s) \leq_I Y(s) \text{ implies that } X(s,t) \leq_I Y(s,t) \tag{10}$$

for almost all s in $[t_0, t]$.

THEOREM 2.3. *Let $X(t_0, t)$ and $Y(t_0, t)$ both be time-homogeneous Markov processe having E as their state space. If $\{\boldsymbol{A}(s) \mid t_0 \leq s \leq t\}$ and $\{\boldsymbol{B}(s) \mid t_0 \leq s \leq t\}$ are their respective family of generators, $\boldsymbol{A}(s) \leq_I \boldsymbol{B}(s)$ for almost all s in $[t_0, t]$, and either $X(t_0, t)$ or $Y(t_0, t)$ is $I(E)$-monotone on $[t_0, t]$, then*

$$X(t_0) \leq_I Y(t_0) \text{ implies that } X(t_0, t) \leq_I Y(t_0, t). \tag{11}$$

The following theorem gives us a way to construct a monotone family of generators.

THEOREM 2.4. *The family of infinitesimal generators $\{\boldsymbol{A}(s) \mid t_0 \leq s \leq t\}$ is $I(E)$-monotone on $[t_0, t]$ if $\boldsymbol{A}(s)$ is $I(E)$-monotone for almost every s in $[t_0, t]$.*

3. Strong and Weak Orderings. We state here a special case of Strassen's theorem as it applies to sample path comparisons of random variables. This theorem is so fundamental as to sample path stochastic ordering, that it is overlooked that the results apply to *any* binary, closed operation. In the appendix, we will state and prove this more general, but still special case of Strassen's theorem.

THEOREM 3.1 (Strassen) *Let P and Q be two probability measures on E. We have $P \leq_{st} Q$ if and only if there exists a bivariate distribution R on $E \times E$ such that $R(x,y) = 0$ if $x \not\leq y$, and*

$$P(x) = \sum_{y \geq x} R(x,y) \tag{12}$$

$$Q(x) = \sum_{y \leq x} R(y,x). \tag{13}$$

Assuming that the generator for the Markov process is a bounded operator is equivalent to saying probabilistically that the process is *uniformizable.* This means that we can synchronize the sample path behavior of the process to the jumps of a Poisson process. If we let λ be sufficiently large then $\boldsymbol{P}_\lambda(\boldsymbol{A}) = \boldsymbol{I} + \frac{1}{\lambda}\boldsymbol{A}$, is a stochastic matrix. The probabilistic and the analytic stories merge together in the following formula:

$$\exp(t\boldsymbol{A}) = \sum_{n=0}^{\infty} \frac{e^{-\lambda t}(\lambda t)^n}{n!} \boldsymbol{P}_\lambda(\boldsymbol{A})^n. \tag{14}$$

THEOREM 3.2. *Let $X(t)$ and $Y(t)$ both be Markov processes having E as their state space. If $\boldsymbol{A}$ and $\boldsymbol{B}$ are their respective generators, then $X(t) \leq_{st} Y(t)$ if and only if $X(0) \leq_{st} Y(0)$, and for all $x \leq y$ in E and sufficiently large λ, we have*

$$\boldsymbol{P}_\lambda(\boldsymbol{A})1_\Gamma(x) \leq \boldsymbol{P}_\lambda(\boldsymbol{B})1_\Gamma(y), \tag{15}$$

for all Γ in $I(E)$, where $\boldsymbol{P}_\lambda(\boldsymbol{A})1_\Gamma(x) = 1_\Gamma(x) + \frac{1}{\lambda}\boldsymbol{A}1_\Gamma(x)$.

If we let $\boldsymbol{A} = \boldsymbol{B}$, then we have the necessary and sufficient conditions for $\boldsymbol{A}$ to be strongly monotone.

For weak orderings, there is a special class of weakly monotone Markov processes. They were identified by Kester [11] as *Möbius monotone* Markov processes. Letting $1_{\{x\}}$ be the indicator function of the singleton set x, we define an operator $\mathbf{Z}(E)$ as

$$\mathbf{Z}(E)1_{\{x\}} = 1_{\{x\}^\uparrow}. \tag{16}$$

In combinatorial theory (see Rota [25]), this corresponds to the *Zeta function*, and its inverse, when it exists, is called the *Möbius function.* A generator $\boldsymbol{A}$ is said to be Möbius monotone if for some $\lambda > 0$, there exists a positive operator $\boldsymbol{M}$ such that

$$\boldsymbol{P}_\lambda(\boldsymbol{A}) \cdot \mathbf{Z}(E) = \mathbf{Z}(E) \cdot \boldsymbol{M}. \tag{17}$$

Thus using operator techniques, we have a sufficient (but not necessary [21]) condition for weak monotonicity.

4. The Jackson Network. The prototypical example of queueing network is the Jackson network [5]. This is initially an N node collection of M/M/1 queues. The ith queue, $Q_i(t)$, has a Poisson input at rate λ_i, and exponential service at rate μ_i. The output from a given node i is then routed either to some node j with probability p_{ij}, or routed away from the system with probability q_i. Let $\boldsymbol{Q}(t) = (Q_1(t), \cdots, Q_N(t))$ equal the joint queue length process, with $\mathbf{Z}_+^N$ the set of nonnegative integer N-tuples equaling its state space. Jackson's theorem states that if we have equilibrium, then for all $m \in \mathbf{Z}_+^N$ we have

$$\lim_{t\to\infty} P_m\{Q_1(t) = n_1, \cdots, Q_N(t) = n_N\} = \prod_{i=1}^{N}(1-\rho_i)\rho_i^{n_1} \tag{18}$$

where $\rho_i = \theta_i/\mu_i$ and $\theta = (\theta_1, \cdots, \theta_N)$ solves the matrix equation

$$\theta = \theta\, \boldsymbol{P} + \lambda \tag{19}$$

where $\boldsymbol{P} = \{p_{ij} \mid 1 \le i, j \le N\}$ and $\lambda = (\lambda_1, \cdots, \lambda_N)$, with the constraint that $\theta_i < \mu_i$.

Suppose we want to know what happens when the total system does not attain steady state. Some nodes may become unstable, but there may still be a maximal subnetwork of stable nodes. We can formulate a solution heuristically. Let θ now be the solution of

$$\theta = (\theta \wedge \mu)\boldsymbol{P} + \lambda \tag{20}$$

where $\theta \wedge \mu$ is the componentwise minimum of the vectors θ and $\mu = (\mu_1, \cdots, \mu_N)$. It can be shown [14], that (20) has a unique solution. Moreover, $I = \{i \mid \theta_i < \mu_i\}$ is a good candidate for the maximal stable subnetwork. We should have for all n in $\mathbf{Z}_+^N$

$$\lim_{t\to\infty} P_{\boldsymbol{m}(\boldsymbol{I})}\{Q_i(t) = n_i;\ i \in I\} = \prod_{i \in I}(1-\rho_i)\rho_i^{n_i} \tag{21}$$

where $\rho_i = \theta_i/\mu_i$, and for all $j \notin I$

$$\lim_{t\to\infty} \Pr\{Q_j(t) = n_j\} = 0. \tag{22}$$

The problem that arises is that $\{Q_i(t) \mid i \in I\}$, the subnetwork, is not in general a Markov process. As such, there is no immediate ergodic theorem to quote that says that (21) *is* the limiting distribution. In Goodman and Massey [4], we instead proved (21) and (22) by stochastic ordering. This allowed us to bound the non-Markovian system above and below by Markovian processes whose limiting distributions were known.

Another natural question for Jackson networks concerns their transient behavior. No general solution is known for these models. Using stochastic ordering [15], we are able to establish the following bound

$$P_m(Q_1(t) \ge n_1, \cdots, Q_N(t) \ge n_N) \le \prod_{i=1}^{N} P_{m_i}(X_i(t) \ge n_i) \tag{23}$$

where $X_i(t)$ is an M/M/1 queue with arrival rate $\lambda_i + \sum_{j=1}^N \mu_j p_{ji}$, and exponential service rate μ_i.

At first glance, the result is not surprising. However, we can prove that this inequality is not intuitive! We can show [15] that if $\boldsymbol{X}(t) = (X_1(t), \cdots, X_N(t))$, where the $X_i(t)$'s are independent, then it is *never* the case that $Q(t) \leq \boldsymbol{X}(t)$ (using componentwise ordering) on all sample paths, unless the two processes are identical. This result can be generalized [19] to bound the Jackson network by independent, smaller Jackson networks. The resulting stochastic ordering is one that is not equivalent to either the strong, weak, or weak* orderings.

Since a stochastic ordering is involved here that is not equivalent to sample path comparisons, we must resort to methods that are more analytic than probabilistic. Bounds of the type were derived for the Jackson network, see [15] and [19], and the Kelly network [16] by creating an operator calculus. Primitive operators are defined that correspond to arrivals, departures, or transfers from a network. The infinitesimal generator for the network is then defined by decomposing it into an algebraic function of these primitive operators.

5. Birth-Death-Migration Processes. Here we illustrate how multidimensional processes can be more complex than their one-dimensional counterparts. A birth-death-migration process can be viewed as a multidimensional birth-death process, or a Jackson network where arrivals, departures, and transfer rates for the system are all functions of the total state of the system. For some state $\boldsymbol{m} \in \mathbf{Z}_+^N$, we define the infinitesimal rates as follows:

- The *birth rate* for the ith colony is $\alpha_i(\boldsymbol{m})$.
- The *death rate* for the ith colony is $\beta_i(\boldsymbol{m})$.
- The *migration rate* from the ith colony to the jth colony is $\gamma_{ij}(\boldsymbol{m})$.

Recall that one dimensional birth-death processes are *always* monotone with respect to the strong, or sample path ordering (see Proposition 4.2.10 of Stoyan [27]). Below, we show that the general case is much different.

THEOREM 5.1. *A B-D-M process $Q(t)$ is strongly monotone if and only if for all $\boldsymbol{m}, \boldsymbol{n} \in Z_+^N$ with $\boldsymbol{m} \leq \boldsymbol{n}$, we have*

$$\alpha_i(\boldsymbol{m}) + \sum_{j\in K} \gamma_{ji}(\boldsymbol{m}) \leq \alpha_i(\boldsymbol{n}) + \sum_{j\in K} \gamma_{ji}(\boldsymbol{n}) \tag{24}$$

and

$$\beta_i(\boldsymbol{m}) + \sum_{j\in K} \gamma_{ij}(\boldsymbol{m}) \geq \beta_i(\boldsymbol{n}) + \sum_{j\in K} \gamma_{ij}(\boldsymbol{n}), \tag{25}$$

where $\{1, 2, \cdots, N\}$ *is partitioned into sets* I *and* J, *such that* $\boldsymbol{m}_I = \boldsymbol{n}_I$ *and* $\boldsymbol{m}_J < \boldsymbol{n}_J$, *with* $i \in I$ *and* K *is any superset of* J.

The proof (see [20]) follows from using Theorem 3.2. There, similar stochastic bounds as for the Jackson network can be proved using both the weak and weak* orderings.

6. Appendix (Proof of Strassen's Theorem). In this section, we will assume that E is a Polish space, and it has a relation $\sim$, that is closed with respect to the metric topology. For alternate proofs in the case of partial orderings, see Liggett [13] as well as Preston [23] for the special case where E is a finite set. Let R be a bivariate probability measure on the space $E \times E$. We will consider only *closed relations*. What we mean is that the relation $\sim$ defines a set $E(\sim) \equiv \{(x, y) \mid x \sim y\}$ which is a closed subset of $E \times E$, with respect to the product topology. The set supp(R) is also defined to be a closed subset of $E \times E$, where any open set disjoint with it, is a set of measure zero with respect to the probability measure R. Denoting its first and second marginals by $R^{(1)}$ and $R^{(2)}$, respectively, let Λ equal the following set:

$$\Lambda = \{(P, Q) \mid R^{(1)} = P,\ R^{(2)} = Q,\ R \in M(E \times E), \text{ and } \operatorname{supp}(R) \subset E(\sim)\}. \tag{26}$$

Our goal is to show that $(P, Q) \in \Lambda$ can be shown simply by using inequalities between P and Q. For the theorem below, we will use $\mathbb{E}$ to denote the expectation of a random variable.

THEOREM 6.1 (Strassen). *If* X *and* Y *are* E*-valued random variables, and* $\sim$ *is a closed relation on* E, *then we have that* X *is equivalent in distribution to some* $\widetilde{X}$, *and similarly* Y *to* $\widetilde{Y}$, *with* $\widetilde{X} \sim \widetilde{Y}$ *if and only if for all bounded, real valued, continuous functions* f *on* E, *we have*

$$\mathbb{E}(f(X)) \leq \mathbb{E}(\sup_{x \sim Y} f(x)). \tag{27}$$

PROOF. Observe that Λ, viewed as a subset of $M(E) \oplus M(E)$, the direct sum of the space of countably additive signed measures with itself, is a closed, convex set. The convexity is clear. The closure, with respect to the vague topology, follows from observing that if K_1 and K_2 are two compact subsets of E,then

$$R((K_1 \times K_2)^c) \leq R^{(1)}(K_1^c) + R^{(2)}(K_2^c). \tag{28}$$

So if (P_n, Q_n) is a sequence in Λ that forms a tight family (see Billingsley [2]), then the corresponding family of R_n forms a tight family as well, which makes Λ a closed set under the vague topology.

A continuous, linear functional on $M(E) \oplus M(E)$ is represented by (f, g) where f and g are bounded, continuous, real valued functions on E. Its action on (P, Q) is defined to be $(f, g) \cdot (P, Q) \equiv \int_E f\, dP + \int_E g\, dQ$. If X and Y are E-valued random variables with distributions P and Q, respectively, we will express $\int_E f\, dP$ as the expectation of $f(X)$, or $\mathbb{E}(f(X))$, and similarly set $\int_E g dQ = \mathbb{E}(g(Y))$. Now if (X, Y) does *not* belong to the convex set Λ, then by the Hahn-Banach theorem (see Reed and Simon [24]), we can find a hyperplane (continuous, linear functional), defined by f and g that separates the point from the convex set such that

$$\mathbb{E}(f(X) + g(Y)) > \sup_{(X', Y') \in \Lambda} \mathbb{E}(f(X') + g(Y')). \tag{29}$$

On the contrary, we have for all f and g,

$$\begin{aligned} \mathbb{E}(f(X) + g(Y)) &\leq \mathbb{E}(\sup_{x \sim Y} f(x) + g(Y)) \\ &\leq \sup_{y \in E}(\sup_{x \sim y} f(x) + g(y)) \\ &\leq \sup_{(X', Y') \in \Lambda} \mathbb{E}(f(X') + G(Y')), \end{aligned}$$

hence $(X, Y) \in \Lambda$, and this completes the proof. ∎

REFERENCES

[1] Baccelli, F., Massey, W. A. and Towsley, D. (1989). Acyclic fork join queueing networks. *J. Assoc. Comput. Math.* **36**, 615–642.

[2] Billingsley, P. (1982). *Convergence of Probability Measures*, John Wiley & Sons, Inc., New York.

[3] Daley, D. J. (1968). Stochastically monotone Markov chains. *Z. Wahrsch. verw. Gebiete* **10**, 305–317.

[4] Goodman, J. and Massey, W. A. (1984). The non-ergodic Jackson network. *J. Appl. Probab.* **21**, 860–869.

[5] Jackson, J. R. (1957). Networks of waiting lines. *Oper. Res.* **5**, 518–521.

[6] Kalmykov, G. I. (1962). On the partial ordering of one-dimensional Markov processes. *Theory Probab. Appl.* **7**, 456–459.

[7] Kamae, T. and Krengel, U. (1978). Stochastic partial ordering. *Ann. Probab.* **6**, 1044–1049.

[8] Kamae, T., Krengel, U., and O'Brien, G. L. (1977). Stochastic inequalities on partially ordered spaces. *Ann. Probab.* **5**, 899–912.

[9] Kelly, F. (1976). Networks of queues. *Adv. in Appl. Probab.* **8**, 416–432.

[10] KEILSON, J. and KESTER, A. (1977). Monotone matrices and monotone Markov processes. *Stochastic Proc. Appl.* **5**, 231–241.

[11] KESTER, A. J. M. (1977). Preservation of cone characterizing properties in Markov chains. Ph.D. Thesis, University of Rochester.

[12] KIRSTEIN, B. M. (1976). Monotonicity and comparability of time homogeneous Markov processes with discrete state space. *Math. Operationsforsch. Statist.* **1**, S. 151–168.

[13] LIGGETT, T. M. (1985). *Interacting Particle Systems.* Springer, New York.

[14] MASSEY, W. A. (1981). Non-stationary queues. Ph.D. Thesis, Stanford University.

[15] MASSEY, W. A. (1984). An operator analytic approach to the Jackson network. *J. Appl. Probab.* **21**, 379–393.

[16] MASSEY, W. A. (1984). Open networks of queues: Their algebraic structure and estimating their transient behavior. *Adv. in Appl. Probab.* **16**, 176–201.

[17] MASSEY, W. A. (1984). Tail distribution functions on partially ordered spaces. AT&T Technical Memorandum.

[18] MASSEY, W. A. (1985). Asymptotic analysis of the time-dependent M/M/1 queue. *Math. Oper. Res.* **10**, 305–327.

[19] MASSEY, W. A. (1986). A family of bounds for the transient behaviors of a Jackson network. *J. Appl. Probab.* **23**, 543–549.

[20] MASSEY, W. A. (1987). Stochastic orderings for birth-death-migration processes. Submitted to *Stochastic Proc. Appl.*

[21] MASSEY, W. A. (1987). Stochastic orderings for Markov processes on partially ordered spaces. *Math. Oper. Res.* **12**, 350–367.

[22] O'BRIEN, G. (1982). Determining classes and independence. *Statist. Probab. Lett.* **1**, 57–59.

[23] PRESTON, C. J. (1974). A generalization of the FKG inequalities. *Comm. Math. Phys.* **36**, 233–241.

[24] REED, M. and SIMON, B. (1972). *Functional Analysis. Vol. 1.* Academic Press, New York.

[25] ROTA, G.-C. (1964). On the foundations of combinatorial Theory. I. Theory of Möbius functions. *Z. Wahrsch. verw. Gebiete.* **2**, 340–368.

[26] SHANTHIKUMAR, G. and YAO, D. (1987). Stochastic monotonicity of the queue lengths in closed queueing networks. *Oper. Res.* **35**, 583–588.

[27] STOYAN, D. (1983). *Comparison Methods for Queues and Other Stochastic Models.* John Wiley & Sons, Inc.

[28] STRASSEN, V. (1965). The existence of probability measures with given marginals. *Ann. Math. Statist.* **36**, 423–439.

[29] TONG, Y. L. (1980). *Probability Inequalities in Multivariate Distributions.* Academic Press, New York.

[30] WHITT, W. (1986). Comparing non-Markov processes. *Math. Oper. Res.* **11**, 608–618.

AT&T BELL LABORATORIES
600 MOUNTAIN AVENUE
MURRAY HILL, NEW JERSEY 07974