Dependence Ordering for Markov Processes on Partially Ordered Spaces

H. Daduna and R. Szekli

REPORT No. 16, 2004/2005, fall

ISSN 1103-467X
ISRN IML-R- -16-04/05- -SE+fall

INSTITUT MITTAG-LEFFLER
THE ROYAL SWEDISH ACADEMY OF SCIENCES

Dependence ordering for Markov processes on partially ordered spaces

Hans Daduna *
Hamburg University
Department of Mathematics
Bundesstrasse 55
20146 Hamburg
Germany

Ryszard Szekli †
Wrocław University
Mathematical Institute
pl. Grunwaldzki 2/4
50-384 Wrocław
Poland

March 2, 2005

Abstract

We consider for monotone Markov processes with partially ordered state space the internal dependence structure with respect to concordance order and supermodular order. For two processes to be oredered with respect to these orderings we obtain for discrete time processes conditions on the one step transition probabilities and for continuous time processes we obtain a criterion on the infinitesimal generator. We show by examples that (complete) monotonicity is not necessary for ordering processes.

1 Introduction

A standard technique in analysis related to complicated functions is to use simpler functions bounding them above or below. This requires to find a proper way to compare relevant characteristics of the functions. One application of this idea is to bound distribution functions for a class of probability measures on R^d in which distributions have fixed one dimensional marginal distributions. When two such measures concentrate their mass on finitely many atoms, say that one is *more concordant then the other* if it can be obtained from it by a finite number of mass transformations which add a fixed amount of mass at $min(x, y)$ and $max(x, y)$ while subtracting it at x and y, $x, y \in R^d$, so that large, resp. small, values of selected coordinates are more often associated with large, resp. small, values of other coordinates. Such a comparison leads to the problem of finding bounds on functionals $\int \psi dF$ where F belongs to the Frechet class of distributions on R^d in which distributions have

*Work supported by Deutscher Akademischer Austauschdienst

†Work supported by the by KBN Grant 2P03A02023

[0] *Key Words*:Concordance order, supermodular order, monotonicity, dependence order

AMS (1991) subject classification: 60K25

Short title: Dependence ordering for Markov processes; filename and version number: modularml250205.tex fixed one dimensional marginal distributions and ψ belongs to some class possessing specific monotonicity properties, for example being supermodular, or n-monotone ($\Delta-$monotone). The case of discrete distributions, $d = 2$, $\psi(x_1, x_2) = x_1 x_2$ leads to the Hardy, Littlewood and Polya [HLP52] rearrangement theorem based on Birkhoff's theorem on extreme points of the convex set of doubly stochastic matrices. The rearrangement result was extended to arbitrary d and ψ supermodular by Lorentz [Lor53]. Independently, Hoeffding [Hoe40] studied the case $d = 2$ and $\psi(x_1, x_2) = f(x_1 - x_2)$, for f convex. Whitt [Whi76] and Tchen [Tch80] noted the connection between Hoeffding's and Hardy, Littlewood and Polya's results. Rüschendorf [Rüs80] studied the case of arbitrary d and ψ being $\Delta-$monotone. Functionals of the form $\int \psi dF$ considered for ψ in a specified cone of functions lead in a natural way to stochastic orderings, for example we say that two distributions on R^d, F, G, are supermodular stochastically ordered if $\int \psi dF \leq \int \psi dG$ for all supermodular ψ for which the integrals exist. If F, G are ordered in this ordering then necessarily they have equal one dimensional marginal distributions, and ordered covariances between respective coordinates - therefore this ordering is a *dependence ordering.* Upper and lower bounds in this ordering are of interest in many contexts. For further refinements and review of research in this field we refer to Müller and Stoyan [MS02]. Most recent applications of this theory are for example in modelling of multivariate portfolios and financial risks see eg. [Rüs04] or in obtaining Kolmogorov type, Hajek-Renyi inequalities for negatively associated random variables, see eg. Christofides and Vaggelatou [CV04].

Application of the above described ideas to Markov chains leads to investigate the class of *stochastically monotone* Markov chains introduced by Daley [Dal68]. Property of stochastic monotonicity is very often the explanation underlying successful applications of comparison techniques in specific problems for example in random walks, epidemic processes, genetics processes or queueing processes. If the state space for such a process is a subset of the real line then covariances decrease monotonically in time. We again refer to Müller and Stoyan [MS02] for stochastic monotonicity and related results.

A general question about comparison of stationary sequences of random variables has been studied from different points of view, for example concerning variability, distance between two stationary sequences, maxima and minima, entropy or dependence. A special interest in comparison of dependence in two stationary (ergodic) sequences is motivated by examples where a stationary sequence represents an input to a complex system (eg. queueing or reliability system) for which replacing an input by another one with the same one dimensional stationary distribution but being more dependent results in a dramatic change of basic performance measures of the system (see e.g. Szekli et al. [SDH94] for such an example in a queueing system). A special case - stationary Markov chain with the state space being a subset of the real line - was studied by Hu and Pan [HP00], where stochastic monotonicity of the processes allows to obtain necessary and sufficient conditions for supermodular ordering and concordance ordering for finite dimensional distributions of Markov chains in discrete and continuous time. Dependence ordering of some derived stationary sequences is studied in Kulik and Szekli [KS04].

For applications to complex systems with multidimensional or more general state space the results of Hu and Pan are not adequat due to the restriction to linearly ordered ($\subseteq \mathbb{R}$) state space. The main aim of the present paper is to extend the results from Hu and Pan [HP00] to the context of arbitrary partially ordered Polish spaces. We further elaborate in detail on the problem of necessity of monotonicity assumptions for the processes. Application of this theoretical development to queueing networks will be presented in forthcoming papers. In section 2 we introduce preliminary facts, in section 3 we present results for discrete time and than for continuous time processes.

2 Definitions and preliminary results

We shall consider probability measures on a partially ordered Polish space E endowed with a closed partial order $\prec$ and the Borel $\sigma-$algebra $\mathcal{E}$ denoted by $(E, \mathcal{E}, \prec)$ and random elements $X : (\Omega,\mathcal{F},P) \to (E, \mathcal{E}, \prec)$. In the following we shall always asssume that for all random elements a common underlying probability space $(\Omega,\mathcal{F},P)$ is fixed.

We denote by $\mathcal{I}^* = \mathcal{I}^*(E)$ the set of all real valued increasing measurable bounded functions on E (f increasing means: for all x, y, $x \prec y$ implies $f(x) \leq f(y)$), and $\mathcal{I}(E)$ the set of all increasing sets (i.e. sets for which indicator functions are increasing). The decreasing analogues are denoted by $\mathcal{D}^*(E)$, and $\mathcal{D}(E)$, respectively.

For $A \subseteq E$ we denote $A^\uparrow := \{y \in E : y \succ x \ for\ some\ x \in A\}$, and $A^\downarrow := \{y \in E : y \prec x\ for\ some\ x \in A\}$.

Further we define $\mathcal{I}_{wk}(E) = \{\{x\}^\uparrow : x \in E\}$ and $\mathcal{D}_{wk}(E) = \{\{x\}^\downarrow : x \in E\}$.

For $(E, \mathcal{E}, \prec)$ which is a lattice (i.e. for any $x, y \in E$ there exist a largest lower bound $x \wedge y \in E$ and a smallest upper bound $x \vee y \in E$ uniquely determined) we denote by $\mathcal{L}_{sm}(E)$ the set of all real valued bounded measurable supermodular functions on E, i.e., functions which fulfill for all $x, y \in E$

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

For the product spaces we use the following notation, $E^{(n)} = E_1 \times ... \times E_n$, for E_i partially ordered Polish spaces ($i = 1, \ldots, n$). If $E_i = E$ for all i then we write E^n instead of $E^{(n)}$. Analogously we write $E^{(\infty)}$ and E^∞ for infinite products. Product spaces will be considered with the product topology. Elements of $E^{(n)}$ will be denoted by $x^n = (x_1, \ldots, x_n)$, of $E^{(\infty)}$ by $x^{(\infty)}$. For random elements we use capital letters in this notation. We denote the coordinatewise ordering on $E^{(n)}$ by $\prec^n$.

Definition 2.1 *We say that two random elements X, Y of $(E, \mathcal{E}, \prec)$ are* strongly stochastically ordered *(and write $X \prec_{st} Y$ or $Y \succ_{st} X$) if $Ef(X) \leq Ef(Y)$ for all $f \in \mathcal{I}^*(E)$, for which the expectations exist.*

We say that two random elements X, Y of $(E, \mathcal{E}, \prec)$ are supermodular stochastically ordered *(and write $X \prec_{sm} Y$ or $Y \succ_{sm} X$) if $Ef(X) \leq Ef(Y)$ for all $f \in \mathcal{L}_{sm}(E)$, for which the expectations exist.*

A simple sufficient criterion for $\prec_{sm}$ order for E which is a discrete lattice is given as follows.

Lemma 2.2 *Let P_1 be a probability measure on a discrete lattice ordered space $(E, \mathcal{E}, \prec)$ and assume that for points $x, y \in E$ we have $P_1(x) \geq \alpha$ and $P_1(y) \geq \alpha$ for some $\alpha > 0$. Define a new probability measure P_2 on $(E, \mathcal{E}, \prec)$ by*

$$\begin{aligned} P_2(x) &= P_1(x) - \alpha & \qquad P_2(x \vee y) &= P_1(x \vee y) + \alpha \\ P_2(y) &= P_1(y) - \alpha & \qquad P_2(x \wedge y) &= P_1(x \wedge y) + \alpha \\ P_2(z) &= P_1(z) \qquad \textit{otherwise.} & & \end{aligned} \tag{2.1}$$

Then $P_1 \prec_{sm} P_2$.
If some probability measure P_2 *on* $(E, \mathcal{E}, \prec)$ *can be obtained from* P_1 *by a finite sequence of transformations of the form* (2.1), *then* $P_1 \prec_{sm} P_2$.

Proof. If $x \prec y$ or $y \prec x$ we have $P_1 = P_2$. Assume now, x and y are not comparable. Then for any $f \in \mathcal{L}_{sm}(E)$ the supermodularity of f yields

$$\int_E P_2(dz) f(z) - \int_E P_1(dz) f(z) = \alpha(f(x \wedge y) + f(x \vee y) - f(x) - f(y)) \geq 0. \tag{2.2}$$

The second statement follows because $\prec_{sm}$ is transitive. ⊙

Remark 2.3 If in Lemma 2.2 the state space E is the set of all subsets of a finite set (it is a lattice under set inclusion) than the transformation described in (2.1) is called in [LX00] a *pairwise* g^+ *transform* and our Lemma 2.2 specializes then to Proposition 5.5., [LX00].

We shall use the following dependence orderings.

Definition 2.4 *For arbitrary random elements* X, Y *of* $(E, \mathcal{E}, \prec)$ *we say that* X, Y *are* concordant weakly stochastically ordered, *and write* $X \prec_{cc-wk} Y$ *if*

$$P(X \in A) \leq P(Y \in A)$$

for all $A \in \mathcal{I}_{wk}(E)$ *and for all* $A \in \mathcal{D}_{wk}(E)$.

Another weaker ordering than $\prec_{sm}$ can be defined on product spaces.

Definition 2.5 *Let* $X^{(n)}, Y^{(n)}$ *be random elements of* $E^{(n)}$. *We say that* $X^{(n)}$ *and* $Y^{(n)}$ *are* concordant stochastically ordered *(written as* $X^{(n)} \prec_{cc}^n Y^{(n)}$ *or* $Y^{(n)} \succ_{cc}^n X^{(n)}$*) if*

$$E\left[\prod_{i=1}^n f_i(X_i)\right] \leq E\left[\prod_{i=1}^n f_i(Y_i)\right] \tag{2.3}$$

for all $f_i \in \mathcal{I}^*(E_i)$ *and all* $f_i \in \mathcal{D}^*(E_i)$, $i = 1, \ldots, n$, *such that the expectations exist.*

Note that concordant ordering is defined only for product spaces while concordant weak ordering is defined on arbitrary partially ordered spaces. Definitions of concordant orderings do not require lattice structure on the space, which in turn is required for supermodular ordering.

On product lattices Lemma 2.2 recuces to the concept of concordance order investigated by Joe and can be illustrated as follows (see [Joe90][Ex. 2.2]). Let $(E, \mathcal{E}, \prec)$ be a finite lattice and P_1 and P_2 probability measures on E^2. Let $x, y \in E^2$ be two points which are not comparable under coordinate-wise ordering. If P_1 is obtained from P_2 by adding the probability mass $\alpha > 0$ at points $x \wedge y$ and $x \vee y$ and by subtracting the probability mass $\alpha > 0$ at points x, y, then $P_1 \prec_{cc}^2 P_2$. We can drop the assumption that E is a lattice, and obtain the following.

Corollary 2.6 *Let* P_1 *be a probability measure on* E^2, *where* $(E, \mathcal{E}, \prec)$ *is a discrete partially ordered space. Assume that for* $a, b, c, d \in E$ *we have* $a \prec c$ *and* $b \prec d$ *such that* $(a, d) \in E^2$ *and* $(c, b) \in E^2$ *are not comparable in the coordinate-wise ordering on* E^2. *Assume that we*

have $P_1((a,d)) \geq \alpha$ *and* $P_1((c,b)) \geq \alpha$ *for some* $\alpha > 0$.
Define a new probability measure P_2 *on* E^2 *by*

$$\begin{aligned} P_2(a,d) &= P_1(a,d) - \alpha & P_2(a,b) &= P_1(a,b) + \alpha \\ P_2(c,b) &= P_1(c,b) - \alpha & P_2(c,d) &= P_1(c,d) + \alpha \\ P_2(x,y) &= P_1(x,y) \quad \text{otherwise.} \end{aligned} \tag{2.4}$$

Then $P_1 \prec_{cc}^2 P_2$.
If some probability measure P_2 *on* E^2 *can be obtained from* P_1 *by a finite sequence of transformations of the form that of* (2.4), *then* $P_1 \prec_{cc}^2 P_2$.

In the special case of product spaces $(E^{(n)}, \prec^n)$ concordant weak ordering is characterized in the following way: $X^{(n)} \prec_{cc-wk}^n Y^{(n)}$ iff

$$P(X_1 \in A_1, \ldots, X_n \in A_n) \leq P(Y_1 \in A_1, \ldots, Y_n \in A_n)$$

for all $A_i \in \mathcal{I}_{wk}(E_i)$ and for all $A_i \in \mathcal{D}_{wk}(E_i)$, $i = 1, \ldots, n$. Since $\mathcal{I}_{wk}(E_i) \subset \mathcal{I}(E_i)$ and $\mathcal{D}_{wk}(E_i) \subset \mathcal{D}(E_i)$, $i = 1, \ldots, n$ it is clear that $\prec_{cc}^n$ implies $\prec_{cc-wk}^n$. For products of linearly ordered sets the above inclusion relations can be replaced by equalities and therefore $\prec_{cc}^n$ and $\prec_{cc-wk}^n$ are equivalent then.

In the next example we show that in general $\prec_{cc-wk}^n$ does not imply $\prec_{cc}^n$.

Example 2.7 If we consider $E = \{(0,0), (0,1), (1,0), (1,1)\}$ with the usual partial ordering $(0,0) \prec (0,1) \prec (1,1)$ and $(0,0) \prec (1,0) \prec (1,1)$ and $P_1(\{(1,1)\}) = P_1(\{(0,0)\}) = 1/2$, $P_2(\{(0,1)\}) = P_2(\{(1,0)\}) = 1/2$ then directly from the definition we obtain on E^2 for product measures $P_1 \times P_1 \prec_{cc-wk}^2 P_2 \times P_2$ but not $P_1 \times P_1 \prec_{cc}^2 P_2 \times P_2$.

The proof of the following lemma can be given utilizing similar arguments as Lindqvist [Lin88][Theorem 3.1].

Lemma 2.8 *The following conditions are equivalent for random elements* $X^{(n)}, Y^{(n)}$ *of* $E^{(n)}$.

1. $X^{(n)} \prec_{cc}^n Y^{(n)}$
2. $P(X_1 \in A_1, \ldots, X_n \in A_n) \leq P(Y_1 \in A_1, \ldots, Y_n \in A_n)$
   for all $A_i \in \mathcal{I}(E_i)$ *and for all* $A_i \in \mathcal{D}(E_i)$ $\quad i = 1, \ldots, n$,
3. $P(X_1 \in A_1, \ldots, X_n \in A_n) \leq P(Y_1 \in A_1, \ldots, Y_n \in A_n)$
   for all closed sets $A_i \in \mathcal{I}(E_i)$ *and for all closed sets* $A_i \in \mathcal{D}(E_i)$, $i = 1, \ldots, n$,
4. $P(X_1 \in A_1, \ldots, X_n \in A_n) \leq P(Y_1 \in A_1, \ldots, Y_n \in A_n)$
   for all compact generated sets (for definition see [Lin88]) $A_i \in \mathcal{I}(E_i)$ *and for all compact generated sets* $A_i \in \mathcal{D}(E_i)$, $i = 1, \ldots, n$,

If in addition E_i *are normally ordered (see [Lin88]) then* $X^{(n)} \prec_{cc}^n Y^{(n)}$ *iff (2.3) holds with* f_i *which are in addition continuous.*

The following basic example shows how for concordance orderings products structures can be connected with lattice - non product structures in a natural way.

Example 2.9 Suppose the state space E is the set of all subsets of a finite set $\{e_1, \dots, e_n\}$. E is a lattice under the set inclusion $\prec := \subset$. Let P_1, P_2 be probability measures on E such that $P_1 \prec_{cc-wk} P_2$. It is natural to define on E the following random vector

$$\delta(A) = (I_A(e_1), \dots, I_A(e_n)), \; A \in E.$$

This is a random vector with n coordinates $(\delta_1, \dots, \delta_n)$ taking on values 0 or 1, indicating which elements belong to A. Denote by $P_1^\delta = P_1\delta^{-1}$ and $P_2^\delta = P_2\delta^{-1}$ the distributions of δ on R^n with respect to P_1 and P_2, accordingly. We consider on R^n the usual coordinate-wise ordering $\leq$. It follows than from definitions that $P_1 \prec_{cc-wk} P_2$ iff $P_1^\delta \leq_{cc}^n P_2^\delta$. The ordering $P_1 \prec_{cc-wk} P_2$ for two probability measures on E can be defined by two orderings defined in [LX00][Definition 5.1], that is by *majorization* of P_1 by P_2 *from the roots* and by *majorization* of P_1 by P_2 *from the leaves.*

We shall now proceed to state without proof (usual arguments can be used) the standard properties of dependence orderings for $\prec_{cc}^n$ (for the case of real valued random vectors see [MS02][Theorem 3.8.7]). We skip the superscript n in the notation $\prec_{cc}^n$ here.

Proposition 2.10 *Let $X^{(n)}, Y^{(n)}$ be random elements of $E^{(n)}$. Then the following properties hold.*

P 1 (bivariate concordance)*: $X^{(n)} \prec_{cc} Y^{(n)}$ implies that for any pair of indices $1 \leq i < j \leq n$ the bivariate marginal vectors fulfill $(X_i, X_j) \prec_{cc}^2 (Y_i, Y_j)$, and therefore $Cov(f(X_i), g(X_j)) \leq Cov(f(Y_i), g(Y_j))$ for each pair $f \in \mathcal{I}^*(E_i), g \in \mathcal{I}^*(E_j)$.*

P 2 (transitivity)*: $X^{(n)} \prec_{cc} Y^{(n)}$ and $Y^{(n)} \prec_{cc} Z^{(n)}$ implies $X^{(n)} \prec_{cc} Z^{(n)}$.*

P 3 (reflexivity)*: $X^{(n)} \prec_{cc} X^{(n)}$.*

P 4 (antisymmetry)*: For E_i for which increasing sets are measure determining classes , $X^{(n)} \prec_{cc} Y^{(n)}$ and $Y^{(n)} \prec_{cc} X^{(n)}$ implies that the distributions of $X^{(n)}$ and $Y^{(n)}$ are equal.*

P 5 (weak convergence)*: Let $(X_k^{(n)}, Y_k^{(n)}, k \in \mathbb{N})$, be a sequence of random elements of $E^{(n)}$. If $E^{(n)}$ is normally ordered then for weak convergence denoted by $\xrightarrow{w}$ on $E^{(n)}$: $X_k^{(n)} \prec_{cc} Y_k^{(n)}$ for all $k \in \mathbb{N}$, and $X_k^{(n)} \xrightarrow{w}_{k\to\infty} X^{(n)}, Y_k^{(n)} \xrightarrow{w}_{k\to\infty} Y^{(n)}$ implies $X^{(n)} \prec_{cc} Y^{(n)}$.*

P 6 (invariance with respect to permutation of indices)*: $X^{(n)} \prec_{cc} Y^{(n)}$ implies $X^{\pi(n)} = (X_{\pi(1)}, \dots, X_{\pi(n)}) \prec_{cc} (Y_{\pi(1)}, \dots, Y_{\pi(n)}) = Y^{\pi(n)}$ for all permutations π of $\{1, 2, \dots, n\}$.*

P 7 (invariance with respect to increasing transforms)*: $X^{(n)} \prec_{cc} Y^{(n)}$ implies*

$$(f_1(X_1), \dots, f_n(X_n)) \prec_{cc} (f_1(Y_1), \dots, f_n(Y_n))$$

for all increasing (decreasing) functions $f_i : E_i \to E_i, i = 1, \dots, n$.

P 8 (closure under marginalization)*: If $X^{(n)} \prec_{cc} Y^{(n)}$ then for any marginal vectors $X_M = (X_i : i \in M), Y_M = (Y_i : i \in M)$ with $M \subseteq \{1, \dots, n\}$ it follows $X_M \prec_{cc}^{|M|} Y_M$.*

P9 (equality of one dimensional marginals)*: $X^{(n)} \prec_{cc} Y^{(n)}$ implies that for all $i = 1, \ldots, n$, the distributions of the respective coordinates X_i and Y_i are equal.*

P10 (closure under identical concatenation)*: $X^{(n)} \prec_{cc} Y^{(n)}$, implies for all $K, L \subseteq \{1, \ldots, n\}$ $(X_K, X_L) \prec_{cc}^{|K|+|L|} (Y_K, Y_L)$. Here $K \cap L \neq \emptyset$ is allowed.*

P11 (closure under independent concatenation)*: Let $X_j^{(n)} \prec_{cc} Y_j^{(n)}$, for $j = 1, 2$, and assume that $X_1^{(n)}$ is independent of $X_2^{(n)}$, and $Y_1^{(n)}$ is independent of $Y_2^{(n)}$. Then $(X_1^{(n)}, X_2^{(n)}) \prec_{cc} (Y_1^{(n)}, Y_2^{(n)})$.*

Note that for a completely regular ordered space [Nac65][p. 54] the set formed of the open decreasing and open increasing subsets is a subbase of the topology [Nac65][Proposition 6]. Then the increasing subsets of E_i are measure determining.

3 Main results

The relevant general definition we use to compare stochastic processes is as follows.

Definition 3.1 *Let $T \subseteq \mathbb{R}$ be an index set for stochastic processes $X = (X_t : t \in T)$ and $Y = (Y_t : t \in T)$, $X_t, Y_t : (\Omega, \mathcal{F}, P) \to (E, \mathcal{E}, \prec)$, $t \in T$. We say that X and Y are concordant stochastically ordered (and write $X \prec_{cc} Y$) if for all $n \geq 2$ and all $t_1 < t_2 < \ldots t_n$, we have on E^n*

$$(X_{t_1}, \ldots, X_{t_n}) \prec_{cc} (Y_{t_1}, \ldots, Y_{t_n}).$$

If in addition E is a lattice, we say that X and Y are supermodular stochastically ordered (and write $X \prec_{sm} Y$), if for all $n \geq 2$ and all $t_1 < t_2 < \ldots t_n$, we have on E^n, $(X_{t_1}, \ldots, X_{t_n}) \prec_{sm}^n (Y_{t_1}, \ldots, Y_{t_n})$.

3.1 Discrete time Markov processes

Recall that a stochastic kernel $K : E \times \mathcal{E} \to [0, 1]$ is strongly stochastically monotone if $\int f(x) K(s, dx)$ is increasing (decreasing) in s for each increasing (decreasing) function f.

Let $X = (X_t : t \in \mathbb{Z})$ and $Y = (Y_t : t \in \mathbb{Z})$, $X_t, Y_t : (\Omega, \mathcal{F}, P) \to (E, \mathcal{E}, \prec)$, be stationary homogeneous Markov processes. Assume that π is an invariant (stationary) one–dimensional marginal distribution common for both X and Y, and denote the 1–step transition kernels for X and Y, by $K^X : E \times \mathcal{E} \to [0, 1]$, and $K^Y : E \times \mathcal{E} \to [0, 1]$, respectively. Denote the respective transition kernels for the time reversed processes $\bar{X}$, $\bar{Y}$ by $\bar{K}^X$, $\bar{K}^Y$.

Theorem 3.2 *If either K^Y and $\bar{K}^X$ are strongly stochastically monotone, or K^X and $\bar{K}^Y$ are strongly stochastically monotone then*

$$X \prec_{cc} Y \text{ iff } (X_0, X_1) \prec_{cc}^2 (Y_0, Y_1). \tag{3.1}$$

Proof. We consider the case of K^X and $\bar{K}^Y$ being monotone. We proceed by induction in n. For $n = 1$ (3.1) is the required relation. Suppose for $n - 1 > 0$ we have $(X_0, \ldots, X_{n-1}) \prec_{cc} (Y_0, \ldots, Y_{n-1})$.

Then for $f_i : E \to I\!R \in \mathcal{I}^*(E),\ i = 0,\ldots,n,\ (f_i : E \to I\!R \in \mathcal{D}^*(E),\ i = 0,\ldots,n)$, denoting the joint distribution of $(X_0,\ldots,X_n)$ by $P^{X_0,\ldots,X_n}$ (and using similar notation for conditional distributions)

$$
\begin{aligned}
& E_\pi\left[\prod_{i=0}^{n} f_i(X_i)\right] \\
= & \int_{E^n} P^{X_0,\ldots,X_{n-1}}(d(x_0,\ldots,x_{n-1}))\prod_{i=0}^{n-1} f_i(x_i)\int_E P^{X_n|X_{n-1}=x_{n-1}}(dx_n)f_n(x_n) \\
= & \int_{E^n} P^{X_0,\ldots,X_{n-1}}(d(x_0,\ldots,x_{n-1}))\prod_{i=0}^{n-1} f_i(x_i)\qquad \underbrace{\int_E K^X(x_{n-1},dx_n)f_n(x_n)}_{\text{increasing (decreasing) in } x_{n-1} \text{ because } K^X \text{ is monotone}} \\
\overset{(1)}{\leq} & \int_{E^n} P^{Y_0,\ldots,Y_{n-1}}(d(x_0,\ldots,x_{n-1}))\prod_{i=0}^{n-1} f_i(x_i)\int_E K^X(x_{n-1},dx_n)f_n(x_n) \\
= & \int_E P^{Y_{n-1}}(dx_{n-1})\int_{E^{n-1}} P^{Y_0,\ldots,Y_{n-2}|Y_{n-1}=x_{n-1}}(d(x_0,\ldots,x_{n-2}))\prod_{i=0}^{n-1} f_i(x_i)\int_E K^X(x_{n-1},dx_n)f_n(x_n) \\
= & \underbrace{\int_E \pi(dx_{n-1})\int_E K^X(x_{n-1},dx_n)}_{=P^{X_{n-1},X_n}(d(x_{n-1},x_n))}\prod_{i=n-1}^{n} f_i(x_i) \\
& \qquad\cdot\underbrace{\int_{E^{n-1}} P^{Y_0,\ldots,Y_{n-2}|Y_{n-1}=x_{n-1}}(d(x_0,\ldots,x_{n-2}))\prod_{i=0}^{n-2} f_i(x_i)}_{\text{increasing (decreasing) in } x_{n-1} \text{ because } \bar{K}^Y \text{ is monotone}} \\
\overset{(2)}{\leq} & \int_{E^2} P^{Y_{n-1},Y_n}(d(x_{n-1},x_n))\prod_{i=n-1}^{n} f_i(x_i)\cdot\int_{E^{n-1}} P^{Y_0,\ldots,Y_{n-2}|Y_{n-1}=x_{n-1}}(d(x_0,\ldots,x_{n-2}))\prod_{i=0}^{n-2} f_i(x_i) \\
= & E_\pi\left[\prod_{i=0}^{n} f_i(Y_i)\right].
\end{aligned}
$$

Here $\overset{(1)}{\leq}$ follows from the induction hypotheses and $\overset{(2)}{\leq}$ follows from stationarity and (3.1). ⊙

Theorem 3.3 *Assume that the state space E is linearly ordered. If either K^Y and $\bar{K}^X$ are strongly stochastically monotone, or K^X and $\bar{K}^Y$ are strongly stochastically monotone then*

$$X \prec_{sm} Y \text{ iff } (X_0,X_1) \prec_{sm}^2 (Y_0,Y_1). \tag{3.2}$$

Proof. Analoguously to the case $E = \mathbb{R}$, see Theorem 3.2 in [HP00].

A natural question is whether the monotonicity assumptions on the kernels arising in the above theorems are necessary. The following example indicates that this assumption is not necessary in order to obtain concordance ordering for finite sequences.

Example 3.4 Consider two homogeneous Markov chains X, Y on a finite state space E with a common invariant distribution π Denote by $P = [p(i,j)]$, and $Q = [q(i,j]$, the transition matrices of X, and Y, respectively. Assume that for $a, b, c, d \in E$ we have $a \prec c$ and $b \prec d$ such that $(a, d) \in E^2$ and $(c, b) \in E^2$ are not comparable, and that for $P^{X_0,X_1}((i,j)) := \pi_i p(i,j)$, and $P^{Y_0,Y_1}((i,j)) := \pi_i q(i,j)$ defined for all $(i,j) \in E^2$, we have $P^{X_0,X_1}((a,d)) \geq \alpha$, $P^{X_0,X_1}((c,b)) \geq \alpha$ for some $\alpha > 0$.
Assume moreover that P^{Y_0,Y_1} and P^{X_0,X_1} are related by $P^{Y_0,Y_1}((a,b)) = P^{X_0,X_1}((a,b)) + \alpha$, $P^{Y_0,Y_1}((c,d)) = P^{X_0,X_1}((c,d)) + \alpha$, and $P^{Y_0,Y_1}((a,d)) = P^{X_0,X_1}((a,d)) - \alpha$, $P^{Y_0,Y_1}((c,b)) = P^{X_0,X_1}((c,b)) - \alpha$, and $P^{Y_0,Y_1}((u,v)) = P^{X_0,X_1}((u,v))$ for all other $(u,v) \in E^2$ (see Corollary 2.6). Then the one-dimensional marginal distributions of (X_0, X_1) and (Y_0, Y_1) are π, and the transition matrix $Q = (q(x,y) : x, y \in E)$ is related to $P = (p(x,y) : x, y \in E)$ as follows:

$$\begin{aligned}
q(a,d) &= p(a,d) - \frac{\alpha}{\pi(a)}, & q(c,b) &= p(c,b) - \frac{\alpha}{\pi(c)} \\
q(a,b) &= p(a,b) + \frac{\alpha}{\pi(a)}, & q(c,d) &= p(c,d) + \frac{\alpha}{\pi(c)} \\
q(u,v) &= p(u,v) \quad \text{otherwise.} & &
\end{aligned} \tag{3.3}$$

If $E = \{1,2,3,4\}$, $\prec := \leq$ (natural ordering), $(a,d) = (2,4), (c,b) = (4,2)$ and $\alpha = \frac{1}{12}$ then for the transition matrix

$$P = \frac{1}{36} \cdot \begin{pmatrix} 18 & 0 & 18 & 0 \\ 6 & 12 & 6 & 12 \\ 7 & 11 & 7 & 11 \\ 5 & 13 & 5 & 13, \end{pmatrix}$$

the rows $P(i,\cdot), i = 1,2,3,4$ treated as distributions on E fulfill

$$P(1,\cdot) \leq_{st} P(3\cdot) \leq_{st} P(2,\cdot) \leq_{st} P(4,\cdot)$$

and especially $P(2,\cdot)$ is not equal to $P(3,\cdot)$, so P is not stochastically monotone.
Because P is doubly stochastic, its unique invariant vector is $\pi = (\frac{1}{4}, \frac{1}{4}, \frac{1}{4}, \frac{1}{4})$, and the time reversal $\bar{X}$ of X has transition matrix $\bar{P} = P^T$:

$$\bar{P} = \frac{1}{36} \cdot \begin{pmatrix} 18 & 6 & 7 & 5 \\ 0 & 12 & 11 & 13 \\ 18 & 6 & 7 & 5 \\ 0 & 12 & 11 & 13 \end{pmatrix}.$$

The rows $\bar{P}(i,\cdot), i = 1,2,3,4$ of this matrix treated as distributions on E fulfill

$$\bar{P}(1,\cdot) = \bar{P}(3\cdot) \leq_{st} \bar{P}(2,\cdot) = \bar{P}(4,\cdot)$$

and in particular $\bar{P}(2,\cdot)$ is not equal to $\bar{P}(3,\cdot)$. So $\bar{P}$ is not stochastically monotone.
The stationary two dimensional distribution P^{X_0,X_1} of (X_0, X_1) is given by

$$[P^{X_0,X_1}((i,j))]_{i,j=1,\ldots,4} = \frac{1}{144} \begin{pmatrix} 18 & 0 & 18 & 0 \\ 6 & 12 & 6 & 12 \\ 7 & 11 & 7 & 11 \\ 5 & 13 & 5 & 13 \end{pmatrix}$$

The distribution P^{Y_0,Y_1} of (Y_0, Y_1) is given by

$$[P^{Y_0,Y_1}((i,j))]_{i,j=1,\dots,4} = \frac{1}{144}\begin{pmatrix} 18 & 0 & 18 & 0 \\ 6 & 24 & 6 & 0 \\ 7 & 11 & 7 & 11 \\ 5 & 1 & 5 & 25 \end{pmatrix}$$

with transition matrix Q given by

$$Q = \frac{1}{36}\cdot\begin{pmatrix} 18 & 0 & 18 & 0 \\ 6 & 24 & 6 & 0 \\ 7 & 11 & 7 & 11 \\ 5 & 1 & 5 & 25 \end{pmatrix}$$

which is not stochastically monotone, but doubly stochastic. It can be checked that the time reversal $\bar{Q}$ of Q is not monotone as well. By a direct (lengthy) computation we see that for any triple of functions $f_i, i = 0, 1, 2,$ we have

$$\begin{aligned} R &:= E_\pi[\prod_{i=0}^{2} f_i(Y_i)] - E_\pi[\prod_{i=0}^{2} f_i(X_i)] \\ &= \alpha\,(f_2(d) - f_2(b))\cdot\left(f_1(c)\left[\sum_{x\in E}\bar{p}(c,x)f_0(x))\right] - f_1(a)\left[\sum_{x\in E}\bar{p}(a,x)f_0(x))\right]\right) \\ &+ \alpha\,(f_0(c) - f_0(a))\cdot\left(f_1(d)\left[\sum_{x\in E}p(d,x)f_2(x))\right] - f_1(b)\left[\sum_{x\in E}p(b,x)f_2(x))\right]\right). \end{aligned}$$

Now with $(a,d) = (2,4), (c,b) = (4,2)$ we obtain

$$\begin{aligned} R &= \alpha\,(f_2(4) - f_2(2))\cdot\left(f_1(4)\left[\sum_{x\in E}\bar{p}(4,x)f_0(x))\right] - f_1(2)\left[\sum_{x\in E}\bar{p}(2,x)f_0(x))\right]\right) \\ &+ \alpha\,(f_0(4) - f_0(2))\cdot\left(f_1(4)\left[\sum_{x\in E}p(4,x)f_2(x))\right] - f_1(2)\left[\sum_{x\in E}p(2,x)f_2(x))\right]\right). \end{aligned}$$

Recalling that

$$\bar{P}(2,\cdot) = \bar{P}(4\cdot),\ P(2,\cdot) \le_{st} P(4,\cdot)$$

we conclude that for functions $f_i, i = 0, 1, 2,$ which are monotone in the same direction we have $R \ge 0$, therefore $(X_0, X_1, X_2) \le_{cc}^3 (Y_0, Y_1, Y_2)$.

Proceeding in a similar way and applying Corollary 2.6 to the pair of states $(a,d) = (2,3), (c,b) = (3,2)$ we find that in this case $R \ge 0$ does not hold for general $f_i, i = 0, 1, 2,$ which are monotone in the same direction. So we do not have $(X_0, X_1, X_2) \le_{cc}^3 (Y_0, Y_1, Y_2)$, although for two dimensional marginal distributions we have $(X_0, X_1) \le_{cc} (Y_0, Y_1)$ and $(X_1, X_2) \le_{cc} (Y_1, Y_2)$. A closer look reveals that the essential properties are:

$$\bar{P}(2,\cdot) \le_{st} \bar{P}(4\cdot),\ P(2,\cdot) \le_{st} P(4,\cdot)$$

in the first case, and

$$\bar{P}(3,\cdot) \leq_{st} \bar{P}(2\cdot),\ P(3,\cdot) \leq_{st} P(2,\cdot)$$

in the second case and while in the first case we shift the probability mass α in a direction which is comonotone with the inherent stochastic monotonicity of P and $\bar{P}$, in the second case we shift the probability mass α in a direction which is converse to the inherent stochastic monotonicity of P and $\bar{P}$.

The next example indicates that total monotonicity assumptions are not necessary in order to obtain supermodular ordering for finite Markov sequences.

Example 3.5 Let E be a discrete lattice ordered space with partial order $\prec$ and assume that for $a, b \in E, a \neq b$, we have $b \prec a$. Let $X = (X_n : n = 0, 1, 2)$ be a stationary homogeneous Markov chain with state space E and with transition matrix P, transition matrix $\bar{P}$ for the time reversal $\bar{X}$ of X, and with one-dimensional stationary distribution π. We assume that for the rows of P treated as distributions on E

$$P(b,\cdot) \prec_{st} P(a,\cdot) \quad \text{and} \quad \bar{P}(b,\cdot) \prec_{st} \bar{P}(a,\cdot) \tag{3.4}$$

and that

$$P^{(X_0,X_1)}(a,b) \geq \alpha, \quad P^{(X_0,X_1)}(b,a) \geq \alpha$$

for some $\alpha > 0$. We compare X with stationary homogeneous Markov chain $Y = (Y_n : n = 0, 1, 2)$ with transition matrix Q obtained by:

$$\begin{aligned} q(a,b) &= p(a,b) - \frac{\alpha}{\pi(a)}, & q(b,a) &= p(b,a) - \frac{\alpha}{\pi(b)} \\ q(a,a) &= p(a,a) + \frac{\alpha}{\pi(a)}, & q(b,b) &= p(b,b) + \frac{\alpha}{\pi(b)} \\ q(u,v) &= p(u,v) \quad \text{otherwise.} \end{aligned} \tag{3.5}$$

Note that Y has invariant distribution π as well. Let $f : (E^3, \prec^3) \longrightarrow I\!R$ be a bounded supermodular function. Direct computation yields

$$(E_\pi[f(Y_0,Y_1,Y_2)] - E_\pi[f(X_0,X_1,X_2)]) = \tag{3.6}$$

$$\sum_{x_0\in E} \bar{p}(a,x_0)[f(x_0,a,a) - f(x_0,a,b)] - \sum_{x_0\in E} \bar{p}(b,x_0)[f(x_0,b,a) - f(x_0,b,b)] \tag{3.7}$$

$$+ \sum_{x_2\in E} p(a,x_2)[f(a,a,x_2) - f(b,a,x_2)] - \sum_{x_2\in E} p(b,x_2)[f(a,b,x_2) - f(b,b,x_2)] \tag{3.8}$$

$$+ \frac{\alpha}{\pi(a)}[f(a,a,a) - f(a,a,b) - f(b,a,a) + f(b,a,b)] \tag{3.9}$$

$$+ \frac{\alpha}{\pi(b)}[f(b,b,b) - f(b,b,a) - f(a,b,a) + f(a,b,a)]. \tag{3.10}$$

If we assume now that $f \in \mathcal{L}_{sm}(E)$, then from supermodularity (3.9) and (3.10) are nonnegative. Utilizing the increasing differences property (which follows from supermodularity), for a fixed in the second coordinate $f(x_0,a,a) - f(x_0,a,b)$ is increasing in x_0, and for b in the second coordinate fixed, $f(x_0,b,a) - f(x_0,b,b)$ is increasing in x_0, and further for x_0 fixed $f(x_0,a,a) - f(x_0,a,b) \geq f(x_0,b,a) - f(x_0,b,b)$. So $\bar{P}(b,\cdot) \prec_{st} \bar{P}(a,\cdot)$ implies that (3.7) is nonnegative. With similar arguments and $P(b,\cdot) \prec_{st} P(a,\cdot)$ we obtain that (3.8) is nonnegative. Because this holds for any $f \in \mathcal{L}_{sm}(E)$, we conclude $(X_0,X_1,X_2) \prec_{sm} (Y_0,Y_1,Y_2)$.

3.2 Continuous time Markov processes

Let $X = (X_t : t \in I\!R)$ and $Y = (Y_t : t \in I\!R)$, $X_t, Y_t : (\Omega,\mathcal{F},P) \to (E,\mathcal{E},\prec)$, be stationary homogeneous Markov processes. Assume that π is an invariant distribution common for both X and Y. Denote the corresponding families of transition kernels of X, and Y, by $K^X = (K_t^X : E \times \mathcal{E} \to [0,1] : t \geq 0)$, and $K^Y = (K_t^Y : E \times \mathcal{E} \to [0,1] : t \geq 0)$, respectively, and the respective transition kernels for the time reversed processes $\bar{X}$, $\bar{Y}$ by $\bar{K}^X = (\bar{K}_t^X : E \times \mathcal{E} \to [0,1] : t \geq 0)$, and $\bar{K}^Y = (\bar{K}_t^Y : E \times \mathcal{E} \to [0,1] : t \geq 0)$, respectively. Recall that K^X (K^Y) is strongly stochastically monotone if for each $t > 0$, K_t^X (K_t^Y) is strongly stochastically monotone.

Utilizing similar argument as in the proof of Theorem 3.2 we have

Theorem 3.6 *If either K^Y and $\bar{K}^X$ are strongly stochastically monotone, or K^X and $\bar{K}^Y$ are strongly stochastically monotone then*

$$X \prec_{cc} Y \text{ iff } \quad (X_0, X_t) \prec_{cc}^2 (Y_0, Y_t) \text{ for all } t > 0. \tag{3.11}$$

We will now show sufficient and necessary conditions for the concordance ordering in terms of infinitesimal characteristics beginning with countable state spaces.

3.2.1 Countable state space

Let us assume that the Markov processes under consideration have countable state space $(E, \mathcal{E}, \prec)$, and $\mathcal{E} = 2^E$.

Theorem 3.7 *Suppose that $(E, \mathcal{E}, \prec)$ is countable and chains X and Y have bounded intensity matrices Q_X and Q_Y, respectively. If either K^X and $\bar{K}^Y$ are strongly stochastically monotone, or K^Y and $\bar{K}^X$ are strongly stochastically monotone then the following properties are equivalent*

1. $X \prec_{cc} Y$,
2. $(X_0, X_t) \prec_{cc}^2 (Y_0, Y_t) \quad \forall t > 0$,
3. *for each comonotone (either both increasing or both decreasing) pair of bounded functions f, g*

$$\sum_{x\in E} \pi(x) f(x) \sum_{y\in E} Q_X(x,y) g(y) \leq \sum_{x\in E} \pi(x) f(x) \sum_{y\in E} Q_Y(x,y) g(y),$$

4. *for each comonotone pair of sets $F, G \subseteq E$*

$$\sum_{x\in F} \pi(x) \sum_{y\in G} Q_X(x,y) \leq \sum_{x\in F} \pi(x) \sum_{y\in G} Q_Y(x,y).$$

Proof. That 1. and 2. are equivalent follows from the statement of Theorem 3.6. That 2. implies 3. follows from the statement of Theorem 3.8 below. That 3. and 4. are equivalent follows by standard approximation arguments. In order to finish the proof it is sufficient to show that 3. implies 2. We consider the case $f, g \in \mathcal{I}^*(E)$ and K^X and $\bar{K}^Y$ being strongly stochastically monotone kernels and shall utilize the representation (which follows from assuming bounded intensity matrices)

$$K^X(t;x,\{y\}) = \lim_{n\to\infty} (\mathbb{1} + \frac{t}{n}Q_X)^n(x,y), \quad x,y \in E, \tag{3.12}$$

where $\mathbb{1}$ is the identity operator, and similarly for K^Y.
We first show by induction that for sufficiently small $\eta > 0$ we have

$$\sum_{x\in E}\pi(x)f(x)\sum_{y\in E}(\mathbb{1}+\eta Q_X)^n(x,y)g(y) \le \sum_{x\in E}\pi(x)f(x)\sum_{y\in E}(\mathbb{1}+\eta Q_Y)^n(x,y)g(y). \tag{3.13}$$

For $n = 1$ this is immediate from 3., so assume we have proved (3.13) for all $m \le n$. Then substitute in (3.13)

$$g(x) := \sum_{y\in E}(\mathbb{1}+\eta Q_X)(x,y)g(y) = (\mathbb{1}+\eta Q_X)g(x),$$

which is an increasing function in x because $\mathbb{1} + \eta Q_X$ is a monotone operator as defined in Massey [Mas87] (and by induction that $(\mathbb{1} + \eta Q_X)^n)(g)$ is an increasing function.) We obtain

$$\begin{aligned}
&\sum_{x\in E}\pi(x)f(x)\sum_{y\in E}(\mathbb{1}+\eta Q_X)^{n+1}(x,y)g(y)\\
=\ &\sum_{x\in E}\pi(x)f(x)\sum_{y\in E}(\mathbb{1}+\eta Q_X)^n(x,y)\left[\sum_{z\in E}(\mathbb{1}+\eta Q_X)(y,z)g(z)\right]\\
\overset{(1)}{\le}\ &\sum_{x\in E}\pi(x)f(x)\sum_{y\in E}(1+\eta Q_Y)^n(x,y)\left[\sum_{z\in E}(\mathbb{1}+\eta Q_X)(y,z)g(z)\right]\\
=\ &\sum_{y\in E}\pi(y)\left[\sum_{z\in E}(\mathbb{1}+\eta Q_X)(y,z)g(z)\right]\underbrace{\sum_{x\in E}\frac{\pi(x)}{\pi(y)}f(x)(\mathbb{1}+\eta Q_Y)^n(x,y)}_{=^{(\star)}\sum_{x\in E}(\mathbb{1}+\eta\bar{Q}_Y)^n(y,x)f(x)}\\
=\ &\sum_{y\in E}\pi(y)\sum_{x\in E}(\mathbb{1}+\eta\bar{Q}_Y)^n(y,x)f(x)\left[\sum_{z\in E}(\mathbb{1}+\eta Q_X)(y,z)g(z)\right]\\
\overset{(2)}{\le}\ &\sum_{y\in E}\pi(y)\sum_{x\in E}(\mathbb{1}+\eta\bar{Q}_Y)^n(y,x)f(x)\left[\sum_{z\in E}(\mathbb{1}+\eta Q_Y)(y,z)g(z)\right]\\
=\ &\sum_{y\in E}\pi(y)\sum_{x\in E}\frac{\pi(x)}{\pi(y)}(\mathbb{1}+\eta Q_Y)^n(x,y)f(x)\left[\sum_{z\in E}(\mathbb{1}+\eta Q_Y)(y,z)g(z)\right]\\
=\ &\sum_{x\in E}\pi(x)f(x)\sum_{z\in E}\sum_{y\in E}(\mathbb{1}+\eta Q_Y)^n(x,y)(\mathbb{1}+\eta Q_Y)(y,z)g(z)\\
=\ &\sum_{x\in E}\pi(x)f(x)\sum_{y\in E}(\mathbb{1}+\eta Q_Y)^{n+1}(x,y)g(y)
\end{aligned}$$

Here $\overset{(1)}{\leq}$ follows from the induction hypothesis (3.13) and the monotonicity of $(\mathbb{1}+\eta Q_X)g(x)$. The inequality $\overset{(2)}{\leq}$ is obtained by applying 3. with the substitution

$$f(x) := \sum_{y\in E}(\mathbb{1}+\eta\bar{Q}_Y)^n(x,y)f(y).$$

which is an increasing function. The equality $(\star)$ follows by direct computations from (note: $B^0 = \mathbb{1}$):

$$\pi(x)(\mathbb{1}+\eta Q_Y)^n(x,y) = \sum_{k=0}^{n}\binom{n}{k}\eta^k\pi(x)Q_Y^k(x,y),$$

and $\pi(x)Q_Y^n(x,y) = \sum_{z\in E}\bar{Q}_Y(z,x)\pi(z)Q_Y^{n-1}(z,y)$.

Put now in (3.13) $\eta := t/n$ and insert into

$$\lim_{n\to\infty}\sum_{x\in E}\pi(x)f(x)\sum_{y\in E}(\mathbb{1}+\frac{t}{n}Q_X)^n(x,y)g(y) \quad\leq\quad \lim_{n\to\infty}\sum_{x\in E}\pi(x)f(x)\sum_{y\in E}(\mathbb{1}+\frac{t}{n}Q_Y)^n(x,y)g(y)$$

Since $\pi(x)f(x), x\in E$ is a counting density of a finite measure on E, and the sequences of functions $\sum_{y\in E}(\mathbb{1}+\frac{t}{n}Q_X)^n(x,y)g(y), n\in I\!N$, resp. $\sum_{y\in E}(\mathbb{1}+\frac{t}{n}Q_Y)^n(x,y)g(y), n\in I\!N$, are bounded and uniformly convergent to the bounded functions $\sum_{y\in E}K^X(t;x,\{y\})g(y), n\in I\!N$, resp. $\sum_{y\in E}K^Y(t;x,\{y\})g(y), n\in I\!N$, so by interchanging limit and integration we obtain 2 in the form

$$\int_{E^2}P^{(X_0,X_t)}(d(u,v))f(u)g(v) \leq \int_{E^2}P^{(Y_0,Y_t)}(d(u,v))f(u)g(v), \quad \forall f,g\in\mathcal{I}^*(E).$$

For decreasing functions we repeat the above argument. $\odot$

3.2.2 General state space

Assume that X and Y are Feller processes on a normally ordered space $(E,\mathcal{E},\prec)$ with corresponding semigroups $T^X = (T_t^X : \mathcal{C}_b(E)\to\mathcal{C}_b(E) : t\geq 0)$, resp. $T^Y = (T_t^Y : \mathcal{C}_b(E)\to\mathcal{C}_b(E) : t\geq 0)$ on $\mathcal{C}_b(E)$, (the space of real valued bounded continuous functions) possesing infinitesimal generators respectively A^X with the domain $D_{A^X}\subseteq\mathcal{C}_b(E)$ and A^Y with the domain $D_{A^Y}\subseteq\mathcal{C}_b(E)$. We assume $T_0^X = T_0^Y = \mathbb{1}$ (identity operator). The corresponding families of transition kernels of X, and Y, are $K^X = (K_t^X : E\times\mathcal{E}\to[0,1] : t\geq 0)$, and $K^Y = (K_t^Y : E\times\mathcal{E}\to[0,1] : t\geq 0)$, respectively.

Theorem 3.8 *If for X and Y described above K^X and $\bar{K}^Y$ are monotone and $D_{A^X}\cap D_{A^Y}$ is dense in $\mathcal{C}_b(E)$ then*

(i)

$$X \prec_{cc} Y \tag{3.14}$$

implies

(ii) for each comonotone pair of functions $f\in\mathcal{C}_b(E)$, and $g\in D_{A^X}\cap D_{A^Y}$ it holds

$$\int_E \pi(dx)f(x)A^Xg(x) \leq \int_E \pi(dx)f(x)A^Yg(x). \tag{3.15}$$

Proof. $X \prec_{cc} Y$ implies $(X_0, X_t) \prec_{cc}^2 (Y_0, Y_t)$ for all $t > 0$, therefore for all comonotone pairs $f \in \mathcal{C}_b(E)$, $g \in D_{A^X} \cap D_{A^Y}$,

$$\int_E \pi(dx) f(x) \underbrace{\int_E K_t^X(x, dy) g(y)}_{=(T_t^X g)(x)} \le \int_E \pi(dx) f(x) \underbrace{\int_E K_t^Y(x, dy) g(y)}_{=(T_t^Y g)(x)}. \tag{3.16}$$

Hence for all $h > 0$

$$\int_E \pi(dx) f(x) \frac{1}{h} \left[T_h^X - \mathbb{1}\right] g(x) \le \int_E \pi(dx) f(x) \frac{1}{h} \left[T_h^Y - \mathbb{1}\right] g(x). \tag{3.17}$$

For any $\epsilon > 0$ there exists some sufficiently small $h(\epsilon)$ such that for all $h < h(\epsilon)$, $\|\frac{1}{h}\left[T_h^X - \mathbb{1}\right] g - A^X g\| \le \epsilon$ holds (with the uniform norm). It follows

$$\begin{aligned} & \left| \int_E \pi(dx) f(x) \frac{1}{h} \left[T_h^X - \mathbb{1}\right] g(x) - \int_E \pi(dx) f(x) (A^X g)(x) \right| \\ \le & \int_E \pi(dx) \|f\| \cdot \|\frac{1}{h} \left[T_h^X - \mathbb{1}\right] g - A^X g\| \le \|f\| \cdot \epsilon \end{aligned}$$

Therefore we may take $h \to 0$ in (3.17) and interchange integration to obtain (3.15). ⊙

4 Conclusion

Comparison of random elements having fixed marginals with respect to internal dependence structure has has found much interest in the literature and is useful for determing bounds for complex performance measures in applied probability. The results of Hu and Pan [HP00] and others recently opened the way to compare stationary Markov processes with common equlibrium over an infinite horizon with respect to supermodular and concordance ordering. A drawback of their results is that the state spaces of their systems needed to be subsets of $\mathbb{R}$ to prove the respective theorems.

We have shown in this paper that this restriction can be weakened considerably. Our theorems cover the cases of queueing networks or reliability networks with multistate components, applications will be provided in a forthcoming paper.

We further studied the question how strong the necessity of having monotone processes is to obtain the respective dependence order. Our result is that weakening in this direction may be possible but this is part of our still ongoing research.

Acknowledgement. Thanks are due to Mittag-Leffler Institute and particularly to prof I. Kaj who invited R. Szekli to spend some time at Mittag-Leffler Institute in Fall 2004, where a part of the present research was done.

References

[CV04] T. C. Christofides and E. Vaggelatou. A connection between supermodular ordering and positive/negative association. *J. multivariate Analysis*, 88:138–151, 2004.

[Dal68] D. J. Daley. The correlation structure of the output process of some single server queueing system. *The Annals of Mathematical Statistics*, 39:1007–1019, 1968.

[HLP52] G.H. Hardy, J.E. Littlewood, and G. Polya. *Inequalities.* Cambridge University Press, London, 2 edition, 1952.

[Hoe40] W. Hoeffding. Maßstabinvariante Korrelationstheorie. *Schr. math. Inst. u. Inst. angew. Math. Univ. Berlin*, 5:181–233, 1940.

[HP00] T. Hu and X. Pan. Comparisons of dependence for stationary Markov processes. *Probability in the Engineering and Informational Sciences*, 14:299–315, 2000.

[Joe90] H. Joe. Multivariate concordance. *Journal of Multivariate Analysis*, 35:12–30, 1990.

[KS04] R. Kulik and R. Szekli. Dependence orderings for some functionals of multivariate point processes. *Journal of Multivariate Analysis*, page to appear, 2004.

[Lin88] B. H. Lindqvist. Association of probability measures. *Journal of Multivariate Analysis*, 26:111–132, 1988.

[Lor53] G.G. Lorentz. An inequality for rearrangements. *American mathmatical Monthly*, 60:176–179, 1953.

[LX00] H. Li and S. H. Xu. Stochastic bounds and dependence properties of survival times in a multicomponent shock model. *Journal of Applied Probability*, 37:1020–1043, 2000.

[Mas87] W. A. Massey. Stochastic ordering for Markov processes on partially ordered spaces. *Mathematics of Operations Research*, 12:350–367, 1987.

[MS02] A. Müller and D. Stoyan. *Comparison Methods for Stochastic Models and Risks.* Wiley, Chichester, 2002.

[Nac65] L. Nachbin. *Topology and Order*, volume 4 of *Van Nostrand Mathematical Series.* Van Nostrand, Princeton, 1965.

[Rüs80] L. Rüschendorf. Inequalities for the expectation of δ monotone functions. *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete*, 54:341–349, 1980.

[Rüs04] L. Rüschendorf. Comparison of multivariate risks and positive dependence. *Advances in Applied Probability*, 41:391–406, 2004.

[SDH94] R. Szekli, R.L. Disney, and S. Hur. MR/GI/1 queues with positively correlated arrival streams. *Journal of Applied Probability*, 31:497–514, 1994.

[Tch80] A. Tchen. Inequalities for distributions with given marginals. *Annals of Probability*, 8:811–827, 1980.

[Whi76] W. Whitt. Bivariate distributions with given marginals. *Annals of Statistics*, 4:1280–1289, 1976.