Order relations of measures when avoiding decreasing sets[☆]

Pierre Collet[a], F. Javier López[b,*], Servet Martínez[c]

[a] *Centre de Physique Théorique, CNRS UMR 7644 Ecole Polytechnique, F-91128 Palaiseau Cedex, France*
[b] *Depto. Métodos Estadísticos, Facultad de Ciencias, Universidad de Zaragoza, C/Pedro Cerbuna 12, Zaragoza 50009, Spain*
[c] *CMM and DIM, UMR UCHILE-CNRS 2071, Universidad de Chile, Casilla 170-3 Correo 3, Santiago, Chile*

Abstract

We consider a discrete-time ergodic Markov chain on a partially ordered state space and study the stochastic comparison between its invariant measure and some measures related with the behaviour of the chain conditioned to avoid a decreasing subset of the state space. We also study the situation when several decreasing sets are avoided.

Keywords: Markov chain; Stochastic comparison; Quasistationary measure

1. Introduction

Stochastic comparison is one of the main tools in the study of stochastic processes (see Müller and Stoyan, 2002, and references therein). In many situations, it is useful to study the comparison of random elements conditioned to some particular event, rather than on the whole space or the comparison between the conditional and unconditional distribution of the element.

We consider a discrete-time ergodic Markov chain (Z_n) on a partially ordered finite state space I. We are interested in the comparison of the behaviour of the original chain and the chain conditioned

☆ Research supported by CONSID + D project P119/2001, project BFM 2001-2449 of MCYT, Project ECOS-Conicyt E99C09 and Nucleus Millenium Information and Randomness, ICM P01-005

* Corresponding author.

E-mail addresses: pierre.collet@cpht.polytechnique.fr (P. Collet), javier.lopez@posta.unizar.es (F.J. López), smartine@dim.uchile.cl (S. Martínez).

to avoid certain subsets B of the state space. With that aim, we study the stochastic comparison of four measures related to the long-term behaviour of the chain. We briefly describe these measures now (the formal definitions will be given below).

The first measure we consider is π, the invariant measure of the chain (Z_n), which reflects the limit behaviour of the (unrestricted) chain. The three other measures are related to the behaviour of the chain conditioned to stay in $\tilde{I} = I \setminus B$. The measure r is the limit measure of the chain defined on $\tilde{I}$, obtained by forbidding the passage to B and renormalizing the entries of the transition matrix.

To define the measure v, we make the states of B absorbing and take v as the quasistationary measure of the chain. That is, v verifies $P_v(Z_n{=}i \mid \tau_{\mathrm{B}} > n){=}v_i$ for all $i \in \tilde{I}$, $n \geqslant 1$, where τ_{B} is the time of absorption in B. The theory of quasistationary distribution is an important and well-established topic (see Darroch and Seneta, 1965; van Doorn 1991; Ferrari et al., 1995; or Moler et al., 2000, and references therein for results and applications). For absorbing discrete-time Markov chains on a finite state space which are aperiodic and irreducible on the set of transient states, the (unique) quasistationary measure coincides with the limiting conditional measure, defined by

$$\lim_{n\to\infty} \frac{P_\eta(Z_n = i)}{P_\eta(\tau_{\mathrm{B}} > n)}$$

for any initial measure η on the transient states. In other words, the measure v represents the limit probability of staying in state i at time n given that absorption has not occurred by that time.

The last measure we study is ρ, to be defined below, which corresponds to the limiting conditional mean ratio (see Darroch and Seneta, 1965):

$$\rho_i = \lim_{n\to\infty} E_\eta \left(\frac{\text{number of visits to } i \text{ between 1 and } n}{n} \,\middle|\, \tau_{\mathrm{B}} > n \right)$$

for any initial probability measure η. That is, ρ represents the limit proportion of time that the chain spends in each state up to time n, given that absorption has not occurred by that time.

We give conditions under which these measures are stochastically ordered. As $\mu_1 \preccurlyeq_{\mathrm{st}} \mu_2$ means $\int f \,\mathrm{d}\mu_1 \leqslant \int f \,\mathrm{d}\mu_2$ for f increasing, this can be used to compare different characteristics of the behaviour of the conditional and unconditional chain and to give bounds on some quantities related to v and ρ (which are usually difficult to compute) in terms of π and r. In the case where I is totally ordered, the stochastic comparison between π, v and ρ is studied in Section 3.7 of Kijima (1997).

It is also interesting to compare the behaviour of the chain conditioned to avoid different subsets of I. We study the stochastic comparison between the limiting conditional mean ratios ρ^i, obtained by removing successive subsets B_i. The conditions for the comparison are given in terms of the original chain. We also compare the measures r^i.

The paper is organized as follows: in the rest of this section, the definitions and some preliminary results are given. In Section 2, the main result, giving conditions for comparability is shown. Section 3 is devoted to the situation where successive subsets B_i are avoided.

Let I be a finite set. $\mathscr{M}(I)$ denotes the set of positive measures and $\mathscr{P}(I)$ the set of probability measures on I. For $f : I \to \mathbf{R}$, $\mu \in \mathscr{M}(I)$, we write $\mu(f) = \int f \,\mathrm{d}\mu$, $|f| = \sum_{x\in I} |f(x)|$, $|\mu| = \mu(1)$.

We assume I is endowed with a partial order $\preccurlyeq$. Denote by $\mathscr{I}(I) = \{f : I \to \mathbf{R} : f(x) \leqslant f(y), \forall x \preccurlyeq y\}$ the set of increasing functions defined on I and $\mathscr{I}_+(I) = \{f \in \mathscr{I}(I) : f \geqslant 0\}$. A set $A \subseteq I$ is *increasing* if $1_A \in \mathscr{I}_+(I)$, that is if $x \in A$, $y \in I$, $x \preccurlyeq y \Rightarrow y \in A$; and A is *decreasing* if $I \setminus A$ is increasing. We also write $A \in \mathscr{I}(I)$ if $A \subseteq I$ is increasing.

The *stochastic* order $\preccurlyeq_{\rm st}$ in $\mathscr{M}(I)$ is defined as $\mu \preccurlyeq_{\rm st} \nu$ if $\mu(A) \leqslant \nu(A)$ for all increasing $A \subseteq I$. As an example $\delta_x \preccurlyeq_{\rm st} \delta_y$ when $x \preccurlyeq y$, where δ_z is the Dirac measure at z.

Lemma 1.1. *Let $\mu \preccurlyeq_{\rm st} \nu$ and $f \in \mathscr{I}_+(I)$ then $\mu(f) \leqslant \nu(f)$.*

Proof. If $\mu(I) = \nu(I)$ the result follows from the classic results on stochastic orders for probability measures. Otherwise, as I is increasing, we have $\mu(I) < \nu(I)$. Define I' as $I \cup \{a\}$ where a is a point not belonging to I and extend the order in I to I' by setting $a \prec x$, for all $x \in I$. Let μ' and ν' be defined on I' as

$$\mu'(x) = \mu(x)1_I(x) + (\nu(I) - \mu(I))1_{\{a\}}(x), \ \ \nu'(x) = \nu(x)1_I(x).$$

Note that $\mu' \preccurlyeq_{\rm st} \nu'$ in I' because, if $A \subseteq I'$ is increasing then, either $a \notin A$ and then $\mu'(A) = \mu(A) \leqslant \nu(A) = \nu'(A)$, or $a \in A$ but this implies, by the definition of increasing set that $A = I'$ and, in this case, $\mu'(I) = \nu'(I)$. Now, extend f to I' by $f(a) = 0$. Obviously, f is increasing in I' and, as $\mu'(I) = \nu'(I)$, we have $\mu(f) = \mu'(f) \leqslant \nu'(f) = \nu(f)$ and the lemma is proved. □

A stronger order than the stochastic order is the *hazard rate* order, where for $\mu, \nu \in \mathscr{M}(I)$, we put $\mu \preccurlyeq_{\rm hr} \nu$ if

$$\mu(A)\nu(A') \leqslant \mu(A')\nu(A) \quad \forall A, A' \in \mathscr{I}(I), A \subseteq A'.$$

This order extends the usual hazard rate order defined on totally ordered spaces (see, for instance, 1.B in Shaked and Shanthikumar, 1994). By taking $A' = I$, it is direct that $\mu \preccurlyeq_{\rm hr} \nu$ and $\mu(I) \leqslant \nu(I)$ imply $\mu \preccurlyeq_{\rm st} \nu$.

A measure $\mu \in \mathscr{M}(I)$ is said to be *associated* if

$$\mu(1)\mu(fg) \geqslant \mu(f)\mu(g) \quad \forall f, g \in \mathscr{I}_+(I).$$

This definition trivially extends the classical definition of association (also called positive correlations) for probability measures (see Definition II.2.11 of Liggett, 1985).

If I is totally ordered, then any $\mu \in \mathscr{P}(I)$ is associated. In fact, if $A, B \subseteq I$ increasing then, as I is totally ordered, $A \subseteq B$ or $B \subseteq A$. Therefore, say $A \subseteq B$,

$$\mu(1_A 1_B) = \mu(A) \geqslant \mu(A)\mu(B) = \mu(1_A)\mu(1_B).$$

Then, as any $f \in \mathscr{I}_+(I)$ can be written as $\sum_{k=1}^{\mathrm{card}(I)} a_k 1_{A_k}$, with $a_k \geqslant 0$, $A_k \subseteq I$ increasing, a direct computation yields that μ is associated. (An alternative proof of this fact, using a coupling argument, can be seen in Theorem 3.10.5 in Müller and Stoyan, 2002.) As a consequence, for I totally ordered, any $\mu \in \mathscr{M}(I)$ is associated.

In the case $I = X_1 \times \cdots \times X_n$, where all the X_i are totally ordered and I is endowed with the coordinatewise order $((x_1, \ldots, x_n) \preccurlyeq (x_1', \ldots, x_n')$ if $x_i \preccurlyeq x_i'$ for all i), another relation $\preccurlyeq_{\rm tp}$ in $\mathscr{M}(I) \backslash \{0\}$ is given by $\mu \preccurlyeq_{\rm tp} \nu$ whenever

$$\mu(x \wedge x')\nu(x \vee x') \geqslant \mu(x)\nu(x') \quad \forall x, x' \in I,$$

where $\vee$ and $\wedge$ stand for the coordinatewise maximum and minimum, respectively. From definition $\mu \preccurlyeq_{\rm tp} \nu \Leftrightarrow \mu/|\mu| \preccurlyeq_{\rm tp} \nu/|\nu|$. This relation $\preccurlyeq_{\rm tp}$ is also known as (multivariate) likelihood ratio order (see 1.C and 4.E in Shaked and Shanthikumar, 1994).

In general, the relation $\preccurlyeq_{\text{tp}}$ is not an order. Now, for any $\mu \in \mathscr{M}(I)$ it holds

$$\mu \preccurlyeq_{\text{tp}} \mu \Rightarrow \mu \text{ is associated.}$$

This follows from the classic results of Fortuin, Kasteleyn and Ginibre (the well-known FKG inequalities) see, e.g., (1.1) in Lindqvist (1988).

From Theorem 2.2 of Karlin and Rinott (1980) we get

$$\mu, \nu \in \mathscr{M}(I), \mu \preccurlyeq_{\text{tp}} \nu \Rightarrow \mu/|\mu| \preccurlyeq_{\text{st}} \nu/|\nu|.$$

When I is totally ordered $\mu \preccurlyeq_{\text{tp}} \nu$ is equivalent to

$$\mu(x)/\nu(x) \geqslant \mu(y)/\nu(y) \quad \forall x \preccurlyeq y.$$

Remark 1.2. Let $I = X_1 \times \cdots \times X_n$, where all the X_i are totally ordered. If $A' \in \mathscr{I}(I)$ is also a product of totally ordered spaces, then for $\mu, \nu \in \mathscr{P}(I)$,

$$\mu \preccurlyeq_{\text{tp}} \nu \Rightarrow \frac{\mu(\cdot)}{\mu(A')} \preccurlyeq_{\text{tp}} \frac{\nu(\cdot)}{\nu(A')} \text{ on } A' \Rightarrow \frac{\mu(\cdot)}{\mu(A')} \preccurlyeq_{\text{st}} \frac{\nu(\cdot)}{\nu(A')} \text{ on } A'$$

and we get $\mu(A)\nu(A') \leqslant \mu(A')\nu(A)$ for all $A \subseteq A'$, $A \in \mathscr{I}(I)$. In particular, in totally ordered spaces, every increasing set satisfies the above condition so $\mu \preccurlyeq_{\text{tp}} \nu$ implies $\mu \preccurlyeq_{\text{hr}} \nu$, recovering Theorem 1.C.1 of Shaked and Shanthikumar (1994).

Let P and Q be nonnegative matrices on I. We put $P \preccurlyeq_{\text{st}} Q$ if $P(x,\cdot) \preccurlyeq_{\text{st}} Q(y,\cdot)$ for all $x \preccurlyeq y$. If P, Q are stochastic this condition is equivalent to $\mu P^n \preccurlyeq_{\text{st}} \nu Q^n$ for $n \geqslant 1$ and every $\mu, \nu \in \mathscr{P}(I)$ such that $\mu \preccurlyeq_{\text{st}} \nu$.

If $P \preccurlyeq_{\text{st}} P$ then P is said to be *monotone*. In this case $Pf \in \mathscr{I}_+(I)$ for any $f \in \mathscr{I}_+(I)$. In fact, for $x \preccurlyeq y$ we have $P(x,\cdot) \preccurlyeq_{\text{st}} P(y,\cdot)$ and Lemma 1.1 implies $Pf(x) \leqslant Pf(y)$.

We remove a subset B from I and denote $\tilde{I} = I \setminus B$. For a matrix P defined on I we denote $\tilde{P} = P|_{\tilde{I}\times\tilde{I}}$ its restriction to $\tilde{I}$.

Lemma 1.3. *Let P be monotone nonnegative. If $\tilde{I}$ is increasing*, then *$\tilde{P}$ is monotone.*

Proof. We just have to check that if $A \in \mathscr{I}(\tilde{I})$ then $A \in \mathscr{I}(I)$. Let $x \in A$ and $x \preccurlyeq y$ with $y \notin A$; then, since A is increasing in $\tilde{I}$ then $y \notin \tilde{I}$ but this contradicts $\tilde{I}$ increasing. $\square$

From now on P will be a nonnegative matrix with nonvanishing rows. We associate to it the stochastic matrix $R = R(P)$ given by

$$R(x,y) = P(x,y)/P(x,I), \quad x, y \in I.$$

If P is also irreducible, by Perron–Frobenius theorem, it has a simple real eigenvalue $\alpha = \alpha(P) > 0$ which is greater in modulus than any other eigenvalue of P. Besides, the corresponding left and right eigenvectors denoted, respectively, by $\eta = \eta(P)$ and $\varphi = \varphi(P)$, can be chosen to be strictly positive. We associate to P the stochastic matrix $Q = Q(P)$ defined on I by

$$Q(x,y) = P(x,y)\varphi(y)/\alpha\varphi(x), \quad x, y \in I.$$

If P is stochastic then $Q=P$ because $\alpha=1$, $\varphi=1$. Denote by $\pi=(\pi(x):x\in I)\in\mathscr{P}(I)$ the invariant measure associated to Q. It verifies $\pi(A)=\eta(\varphi 1_A)/\eta(\varphi)$ for $A\subseteq I$.

It is well known (see, for instance, Seneta, 1981) that η and φ verify

$$\eta=\lim_{n\to\infty}\mu P^n/|\mu P^n| \text{ for } \mu\in\mathscr{M}(I)\backslash\{0\} \quad \text{and} \quad \varphi=\lim_{n\to\infty}P^n f/|P^n f| \text{ for } f\in\mathscr{I}_+(I)\backslash\{0\}.$$

We assume P is monotone. Then $P(\mathscr{I}_+)\subseteq\mathscr{I}_+$ and $\varphi=\lim_{n\to\infty}P^n 1/|P^n 1|$. We deduce that $\varphi\in\mathscr{I}_+$.

We will suppose that $\tilde{P}=P|_{\tilde{I}\times\tilde{I}}$ is also irreducible and aperiodic. Denote $\tilde{\alpha}=\alpha(\tilde{P})>0$ its Perron–Frobenius eigenvalue and by $\tilde{\eta}=\eta(\tilde{P})>0$ and $\tilde{\varphi}=\varphi(\tilde{P})>0$ the corresponding left and right eigenvectors. From Lemma 1.3 we get that $\tilde{P}(\mathscr{I}_+(\tilde{I}))\subseteq\mathscr{I}_+(\tilde{I})$ and then $\tilde{\varphi}$ is also increasing.

Denote by $\tilde{Q}=Q(\tilde{P})$, $\tilde{R}=R(\tilde{P})$, $\lambda=\alpha(\tilde{P})\in(0,1)$, $v=\eta(\tilde{P})$, $h=\varphi(\tilde{P})$. The matrices $\tilde{Q}$ and $\tilde{R}$ verify

$$\tilde{Q}(x,y)=P(x,y)h(y)/\lambda h(x) \quad \text{and} \quad \tilde{R}(x,y)=P(x,y)/P(x,\tilde{I}) \text{ for } x,y\in\tilde{I}.$$

The stationary probability measures of $\tilde{Q}$ and $\tilde{R}$ are denoted respectively by ρ and r. ρ verifies $\rho(x)=v(x)h(x)/v(h)$.

Assume v is normalized, $|v|=1$. If P is stochastic then the matrix $\tilde{Q}^v$ defined on $\tilde{I}$ by $\tilde{Q}^v(x,y)=P(x,y)+P(x,B)v(y)$ for $x,y\in\tilde{I}$ is also stochastic. A direct computation shows v is the stationary measure associated to $\tilde{Q}^v$.

2. Comparison of measures

We will assume P is stochastic, so $Q=P$, and monotone. To compare the stochastic ordering between π and the measures v, ρ and r, we extend these last measures to I, in such a way that they vanish at B and we denote such extensions by v', ρ' and r'.

Theorem 2.1. *Assume P is stochastic and monotone and that $\tilde{I}$ is an increasing set.*

(a) *If $P(x,\cdot)$ is associated for all $x\in\tilde{I}$, then $\pi\preccurlyeq_{\text{st}} r'$ and $\pi\preccurlyeq_{\text{st}}\rho'$.*

(b) *If $\tilde{I}$ verifies $x\notin\tilde{I}, y\in\tilde{I}\Rightarrow x\preccurlyeq y$, then $\pi\preccurlyeq_{\text{st}} v'$.*

(c) *If v is associated, then $v\preccurlyeq_{\text{st}}\rho$.*

(d) *If $\tilde{P}(x,\cdot)$ is associated for all $x\in\tilde{I}$, then $\tilde{R}(x,\cdot)\preccurlyeq_{\text{st}}\tilde{Q}(x,\cdot)$ for every $x\in\tilde{I}$. Moreover, if $P(x,\cdot)\preccurlyeq_{\text{hr}}P(y,\cdot)$ for any $x\preccurlyeq y$, $x,y\in\tilde{I}$, then $\tilde{R}$ is monotone and $r\preccurlyeq_{\text{st}}\rho$.*

Proof. (a) Let $A\subseteq I$ be increasing.

Let us show $\pi\preccurlyeq_{\text{st}} r'$. Let $x\in\tilde{I}$. Since $P(x,\cdot)$ is associated and the sets A and $\tilde{I}$ are increasing, we get $P1_A(x)P1_{\tilde{I}}(x)\leqslant P(1_A 1_{\tilde{I}})(x)$ which is equivalent to

$$P(x,A)\leqslant P(x,A\cap\tilde{I})/P(x,\tilde{I}). \tag{1}$$

We will show that (1) is sufficient to imply the result. Define the matrix R' as

$$R'(x,y)=P(x,y)1_B(x)+\tilde{R}(x,y)1_{\tilde{I}}(x)1_{\tilde{I}}(y).$$

It is clear that R' has r' as its unique invariant measure. Let us see that $P\preccurlyeq_{\text{st}}R'$, that is for any $x\preccurlyeq y$ it holds $P(x,\cdot)\preccurlyeq_{\text{st}}R'(y,\cdot)$. Since P is monotone, it is enough to check that $P(x,\cdot)\preccurlyeq_{\text{st}}R'(x,\cdot)$

for $x \in \tilde{I}$ (if $x \in B$ it is trivial). Let $A \subseteq I$ be increasing, from (1) we get

$$R'(x,A) = P(x, A \cap \tilde{I})/P(x,\tilde{I}) \geqslant P(x,A).$$

Therefore, for any $\mu \in \mathscr{P}(I)$ it holds $\pi = \lim_{n\to\infty} \mu P^n \preccurlyeq_{\text{st}} \lim_{n\to\infty} \mu R'^n = r'$ and the result is proved.

Let us prove $\pi \preccurlyeq_{\text{st}} \rho'$. We denote h' the extension of h to I with h' vanishing at B. Then $P(h'g)(x) = \tilde{P}(hg)(x)$ for $x \in \tilde{I}$ and g any function defined on I. Define the matrix Q' on I by

$$Q'(x,y) = P(x,y)1_B(x) + \tilde{Q}(x,y)1_{\tilde{I}}(x)1_{\tilde{I}}(y).$$

Q' has ρ' as its unique invariant measure. Since h' is increasing we get for $x \in \tilde{I}$ and any increasing A

$$Ph'(x)P1_A(x) \leqslant P(h'1_A)(x). \tag{2}$$

Since $Ph'(x) = \lambda h'(x)$ for $x \in \tilde{I}$ we get that (2) is equivalent to

$$P(x,A) \leqslant \sum_{y \in A \cap \tilde{I}} P(x,y)h(y)/\lambda h(x) = Q'(x,A) \text{ for } x \in \tilde{I}.$$

We have shown that $P(x,\cdot) \preccurlyeq_{\text{st}} Q'(x,\cdot)$ for any $x \in I$ (because for $x \in B$ it is trivial). Since P is monotone we get $P \preccurlyeq_{\text{st}} Q'$, and we deduce $\pi \preccurlyeq_{\text{st}} \rho'$.

(b) First, let us show that the following condition holds for all $x \in \tilde{I}$ and $A \subseteq I$ increasing in I:

$$P(x, A \backslash \tilde{I}) \leqslant v(A \cap \tilde{I})P(x,B). \tag{3}$$

Let A be increasing. If $A \setminus \tilde{I}$ is empty then (3) holds trivially. If $A \setminus \tilde{I} \neq \emptyset$ then, from the hypothesis and A increasing, we have $A \cap \tilde{I} = \tilde{I}$ and $v(A \cap \tilde{I}) = 1$. Hence (3) holds.

Now, (3) suffices to get the result. To prove it, consider the following stochastic matrix S' on I,

$$S'(x,y) = P(x,y)1_B(x) + (P(x,y) + P(x,B)v(y))1_{\tilde{I}}(x)1_{\tilde{I}}(y).$$

It is clear that S' has v' as its unique invariant measure.

Let us see now that $P \preccurlyeq_{\text{st}} S'$. As P is monotone, we just have to see that $P(x,\cdot) \preccurlyeq_{\text{st}} S'(x,\cdot)$ for all $x \in I$. The inequality is trivial if $x \in B$, so take $x \in \tilde{I}$ and A increasing. From condition (3) it is verified

$$S'(x,A) = P(x, A \cap \tilde{I}) + P(x,B)v(A \cap \tilde{I}) \geqslant P(x,A)$$

and we conclude $P \preccurlyeq_{\text{st}} S'$. Then $\pi \preccurlyeq_{\text{st}} v'$.

(c) Let $A \subseteq \tilde{I}$ be increasing. Then, as v is associated and h and 1_A are increasing we have $v(1_A)v(h) \leqslant v(h1_A)$, which is equivalent to $v(A) \leqslant \rho(A)$. The result follows.

(d) Let A be an increasing subset in $\tilde{I}$. Since h is increasing we get

$$(\tilde{P}h)(x)(\tilde{P}1_A)(x) \leqslant \tilde{P}(h1_A)(x)(\tilde{P}1_{\tilde{I}})(x). \tag{4}$$

We have $(\tilde{P}1_A)(x) = P(x,A)$ and $\tilde{P}h = \lambda h$. Then (4) is equivalent to

$$P(x,A)/P(x,\tilde{I}) \leqslant \sum_{y \in A} P(x,y)h(y)/\lambda h(x).$$

The first part is shown.

For the second part, it is immediate to check that $\tilde{R}(x,\cdot) \preccurlyeq_{\text{hr}} \tilde{R}(y,\cdot)$ for any $x \preccurlyeq y$, so $\tilde{R}$ is monotone. This implies (by the first part) $\tilde{R} \preccurlyeq_{\text{st}} \tilde{Q}$, and then $r \preccurlyeq_{\text{st}} \rho$. □

From Theorem 2.1(b) we deduce the following result for monotone matrices.

Corollary 2.2. *Assume P is nonnegative, $Q(P)$ is monotone and $\tilde{I}$ verifies $x \notin \tilde{I}$, $y \in \tilde{I} \Rightarrow x \preccurlyeq y$. Then $\eta(\varphi 1_A)/\eta(\varphi) \leqslant \tilde{\eta}(\varphi 1_A)/\tilde{\eta}(\varphi)$ for every increasing set $A \subseteq \tilde{I}$.*

Proof. From $v\widetilde{Q(P)} = \lambda v$ and the definition of Q we get

$$\sum_{x \in \tilde{I}} v(x)P(x,y)/\varphi(x) = \alpha\lambda v(y)/\varphi(y) \text{ for } y \in \tilde{I}.$$

We deduce $v = \tilde{\eta}\ \varphi/\tilde{\eta}(\varphi)$ (and $\tilde{\alpha} = \alpha\lambda < \alpha$). From Theorem 2.1(b) we get $\pi \preccurlyeq_{\mathrm{st}} v'$. Since $\pi(A) = \eta(\varphi 1_A)/\eta(\varphi)$ we find the result. □

In the following remark we analyze some interesting situations for the application of Theorem 2.1.

Remark 2.3. (i) Assume I is totally ordered. The hypothesis of (b) is trivially fulfilled. Besides as any $\mu \in \mathscr{M}(I)$ is associated, conditions (a), (c) and the first part of (d) are also verified. For the second condition of part (d) it suffices, by Remark 1.2, that $P(x,\cdot) \preccurlyeq_{\mathrm{tp}} P(y,\cdot)$ for $x \preccurlyeq y$, $x, y \in \tilde{I}$.

(ii) Assume $I = X_1 \times \cdots \times X_n$, where all the X_i are totally ordered. Then $P(x,\cdot) \preccurlyeq_{\mathrm{tp}} P(x,\cdot)$ for $x \in \tilde{I}$ suffices for (a). Condition (b) means that B is the point $(m_1, \ldots, m_n)$ with m_i the minimum of X_i.

Concerning (c), the condition $\tilde{P} \preccurlyeq_{\mathrm{tp}} \tilde{P}$ established by Karlin and Rinott (1980) is sufficient to insure that v is associated. The condition $\tilde{P} \preccurlyeq_{\mathrm{tp}} \tilde{P}$ means

$$\tilde{P}(x \wedge x', y \wedge y')\tilde{P}(x \vee x', y \vee y') \geqslant \tilde{P}(x,y)\tilde{P}(x',y'). \quad \forall x, y, x', y' \in \tilde{I}.$$

In Theorem 2.4 of Karlin and Rinott (1980) it is shown that $\mu\tilde{P} \preccurlyeq_{\mathrm{tp}} \mu\tilde{P}$ for all $\mu \in \mathscr{P}(\tilde{I})$ such that $\mu \preccurlyeq_{\mathrm{tp}} \mu$. Since $\delta_a \preccurlyeq_{\mathrm{tp}} \delta_a$, for $a \in \tilde{I}$ (since both sides of the inequality, when applied to $x \neq x'$, are zero), we get by iteration $v \preccurlyeq_{\mathrm{tp}} v$, so v is associated. By the same argument than the one given above, it holds $h \preccurlyeq_{\mathrm{tp}} h$ (and is associated) and therefore, $\rho \preccurlyeq_{\mathrm{tp}} \rho$ (and is associated).

For condition (d), when $\tilde{I}$ is a product of subsets of X_i, it is enough that $P(x,\cdot) \preccurlyeq_{\mathrm{tp}} P(y,\cdot)$ for $x \preccurlyeq y$, $x, y \in \tilde{I}$ (see Remark 1.2).

(iii) Observe that the hypotheses of Theorem 2.1 imply that $P(x,\tilde{I})$ is increasing with x. An extreme case is given by the following condition:

$$P(x,\tilde{I}) \text{ does not depend on } x \in \tilde{I}. \tag{5}$$

In that case, $\tilde{R}$ is proportional to $\tilde{P}$ and no further condition is needed to insure the monotonicity of $\tilde{R}$ in Theorem 2.1(d). Concerning the association of v, when (5) holds it suffices that $\tilde{P}$ is monotone and $\tilde{P}(x,\cdot)$ is associated for all $x \in \tilde{I}$. To see this, let us prove that this condition insures that, if $\mu \in \mathscr{M}(\tilde{I})$ is associated then $\mu\tilde{P}$ is also associated. Let $c = P(x,\tilde{I}) = P1_{\tilde{I}}(x) = \mu\tilde{P}(1)$ for any $x \in \tilde{I}$. Let $f, g \in \mathscr{I}_+(\tilde{I})$. Since $\tilde{P}(x,\cdot)$ is associated we have $\tilde{P}(fg)(x) \geqslant (\tilde{P}1(x))^{-1}\tilde{P}f(x)\tilde{P}g(x) = c^{-1}\tilde{P}f(x)\tilde{P}g(x)$. Then

$$\begin{aligned}(\mu\tilde{P})(1_{\tilde{I}})(\mu\tilde{P})(fg) &= (\mu\tilde{P})(1)\mu(\tilde{P}(fg)) \geqslant c\mu(c^{-1}\tilde{P}f\tilde{P}g) \\ &= \mu(\tilde{P}f\tilde{P}g) \geqslant \mu(\tilde{P}f)\mu(\tilde{P}g) = (\mu\tilde{P})(f)(\mu\tilde{P})(g),\end{aligned}$$

the last inequality because μ is associated. Now, if $a \in \tilde{I}$, δ_a is associated and $\nu = \lim_{n\to\infty} \delta_a \tilde{P}^n / |\delta_a \tilde{P}^n|$. We conclude that ν is associated.

Remark 2.4. From the proof above we recover Theorem 5.2.8 in Müller and Stoyan (2002), which states that, if P is a monotone stochastic matrix, then the necessary and sufficient condition for μP^n being associated whenever μ is, is that $P(x,\cdot)$ is associated for all x. The sufficiency follows from our proof and the necessity from the fact that δ_x is associated and $P(x,\cdot) = \delta_x P$. This compares with the continuous-time case, when a monotone Markov chain keeps the association of the initial measures if and only if every jump is 'up' or 'down' (see Harris, 1977).

Remark 2.5. Association plays an important role in most parts of Theorem 2.1 and, in fact, the corresponding results are not valid if that assumption is dropped. For instance, consider part (a) and take $I = \{1,2,3\}$, with the partial order $1 \preccurlyeq 3$, $2 \preccurlyeq 3$, $\tilde{I} = \{2,3\}$ and the matrix

$$P = \begin{pmatrix} 3/10 & 1/10 & 3/5 \\ 1/5 & 2/5 & 2/5 \\ 1/10 & 1/10 & 4/5 \end{pmatrix},$$

which is stochastic and monotone (with respect to $\preccurlyeq$). Note that $P(2,\cdot) = (1/5, 2/5, 2/5)$ is not associated, for take $f(1)=1$, $f(2)=0$, $f(3)=1$ and $g(1)=0$, $g(2)=1$, $g(3)=1$, which are $\preccurlyeq$—increasing and $0.4 = P(2,\cdot)(1)P(2,\cdot)(fg) \not\geqslant P(2,\cdot)(f)P(2,\cdot)(g) = 0.48$. We have $\pi = (1/7, 1/7, 5/7)$, $r' = (0, 2/11, 9/11)$ and $\rho' = (0, (2-\sqrt{2})/4, (2+\sqrt{2})/4)$. Take $A = \{1,3\}$ which is $\preccurlyeq$—increasing; then $6/7 = \pi(A) \not\leqslant r'(A) = 9/11$, so $\pi \not\preccurlyeq_{\mathrm{st}} r'$; also $6/7 = \pi(A) \not\leqslant \rho'(A) = (2+\sqrt{2})/4$, so $\pi \not\preccurlyeq_{\mathrm{st}} \rho'$.

Also, for parts (c) and (d), take $I = \{0,1,2,3\}$ with the partial order $0 \preccurlyeq 1 \preccurlyeq 2$ (3 is not comparable with any other point). For $\mu_1, \mu_2 \in \mathscr{P}(I)$, $\mu_1 \preccurlyeq_{\mathrm{st}} \mu_2$ if and only if $\mu_1(1)+\mu_1(2) \leqslant \mu_2(1)+\mu_2(2)$, $\mu_1(2) \leqslant \mu_2(2)$ and $\mu_1(3) = \mu_2(3)$ (this equality follows from the fact that the sets $\{0,1,2\}$, $\{3\}$ are increasing and $\mu_1, \mu_2 \in \mathscr{P}(I)$).

Let

$$P = \begin{pmatrix} 0.1 & 0.1 & 0.1 & 0.7 \\ 0.1 & 0.1 & 0.1 & 0.7 \\ 0.05 & 0.05 & 0.2 & 0.7 \\ 0.8 & 0.1 & 0 & 0.1 \end{pmatrix},$$

which is monotone and take $\tilde{I} = \{1,2,3\}$. Then

$$\tilde{P} = \begin{pmatrix} 0.1 & 0.1 & 0.7 \\ 0.05 & 0.2 & 0.7 \\ 0.1 & 0 & 0.1 \end{pmatrix},$$

whose left Perron eigenvector is $\nu = (0.22208, 0.09724, 0.68068)$. ν is not associated on $\tilde{I}$ since, taking $f(1)=0$, $f(2)=1$, $f(3)=0$, $g(1)=0$, $g(2)=0$, $g(3)=1$, we have $0 = \nu(fg) \not\geqslant \nu(f)\nu(g) = 0.06619$. Besides, $\rho = (0.410197, 0.206952, 0.382851)$, and $\nu \not\preccurlyeq_{\mathrm{st}} \rho$ since $\nu(3) \neq \rho(3)$. Thus, part (c) fails if the assumption of association is dropped. For part (d), note that $\tilde{P}(1,\cdot)$ is not associated (this is checked as above). We have

$$\tilde{R} = \begin{pmatrix} 1/9 & 1/9 & 7/9 \\ 1/19 & 4/19 & 14/19 \\ 1/2 & 0 & 1/2 \end{pmatrix}$$

and

$$\tilde{Q} = \begin{pmatrix} 0.233433 & 0.268976 & 0.49759 \\ 0.1012933 & 0.466866 & 0.4318376 \\ 0.766569 & 0 & 0.2334332 \end{pmatrix}$$

and $\tilde{R}(1,\cdot) \not\preccurlyeq_{\rm st} \tilde{Q}(1,\cdot)$ since $\tilde{R}(1,3) \neq \tilde{Q}(1,3)$.

3. Iteration of the inequalities

For P stochastic and monotone we have given general conditions on P and $\tilde{I}$ for having $\pi \preccurlyeq_{\rm st} \rho'$ and $\pi \preccurlyeq_{\rm st} r'$. As ρ and r are the invariant measures of the transition matrices $\tilde{Q}$ and $\tilde{R}$, respectively, we can remove a new set B' from $\tilde{I}$ and, under suitable conditions on $\tilde{P}$ and B', we will have $\rho' \preccurlyeq_{\rm st} \rho''$ and $r' \preccurlyeq_{\rm st} r''$ where ρ'' and r'' are the new vectors arising from $\tilde{Q}$ and $\tilde{R}$, respectively. This procedure can be applied successively, and under adequate conditions, we will have

$$\pi \preccurlyeq_{\rm st} \rho' \preccurlyeq_{\rm st} \rho'' \preccurlyeq_{\rm st} \cdots \preccurlyeq_{\rm st} \rho^{i'} \preccurlyeq_{\rm st} \cdots \quad \text{and} \quad \pi \preccurlyeq_{\rm st} r' \preccurlyeq_{\rm st} r'' \preccurlyeq_{\rm st} \cdots \preccurlyeq_{\rm st} r^{i'} \preccurlyeq_{\rm st} \cdots .$$

A problem is that the conditions need to be checked for the corresponding matrices and removed subsets at each step. Below we study conditions on the original stochastic matrix P and the removed subsets B to guarantee the above chains of inequalities. For the first case, we will restrict ourselves to the case of I totally ordered, while, for the second one, I will be a product of totally ordered spaces.

We will use the notation $I^1 = I$ and $I^i = I^{i-1} \setminus B^{i-1}$, where B^i is the removed set at step i. The measures ρ^i and r^i will be the measures ρ and r corresponding to step i (although defined on I^{i+1}, we extend them to I by giving them the value 0 at $I \setminus I^{i+1}$). It is important to note that the measure r^i (respectively, ρ^i) so defined coincides with the measure r (ρ) defined by the direct removal of $B^1 \cup \cdots \cup B^i$. Therefore, the result below can be used to compare the behaviour of the chain when avoiding different sets $B \subseteq B'$.

If I is totally ordered, for having $\pi \preccurlyeq_{\rm st} \rho'$, it is enough that P is stochastic monotone and B is decreasing. However, P monotone does not imply $\tilde{Q}$ monotone so, in order to iterate the inequality, some extra conditions are to be imposed on P. If $I = X_1 \times \cdots \times X_n$ with X_i totally ordered, to have $\pi \preccurlyeq_{\rm st} r'$, it is enough that P is monotone, $\tilde{I}$ is increasing and $P(x,\cdot)$ is associated. The key point now is that the structure of I lets us use the order $\preccurlyeq_{\rm tp}$; therefore, for the iteration, it will be important that the successive sets I^i are also products of totally ordered sets.

Theorem 3.1. (a) *Let I be totally ordered and P be stochastic. If $P(x,\cdot) \preccurlyeq_{\rm tp} P(y,\cdot)$ for all $x \preccurlyeq y$, and B^i decreasing in I^i, then*

$$\pi \preccurlyeq_{\rm st} \rho^1 \preccurlyeq_{\rm st} \cdots \preccurlyeq_{\rm st} \rho^i \preccurlyeq_{\rm st} \cdots .$$

(b) *Let* $I = X_1 \times \cdots \times X_n$ *with* X_i *totally ordered and* P *be stochastic. If* $P(x,\cdot) \preccurlyeq_{\mathrm{tp}} P(y,\cdot)$ *for all* $x \preccurlyeq y$ *and* B^i *decreasing in* I^i *such that* $I^i = X_1^i \times \cdots \times X_n^i$, *then*

$$\pi \preccurlyeq_{\mathrm{st}} r^1 \preccurlyeq_{\mathrm{st}} \cdots \preccurlyeq_{\mathrm{st}} r^i \preccurlyeq_{\mathrm{st}} \cdots \ .$$

Proof. (a) As $P(x,\cdot) \preccurlyeq_{\mathrm{tp}} P(y,\cdot)$ for $x \preccurlyeq y$ implies P is monotone, we have $\pi \preccurlyeq_{\mathrm{st}} \rho^1$. Now, to iterate the inequality, we just have to check that the matrix $\tilde{Q}$ inherits the properties imposed to P in the statement of the theorem; that is, we have to check that, if $x \preccurlyeq y$ in I^2, then $\tilde{Q}(x,\cdot) \preccurlyeq_{\mathrm{tp}} \tilde{Q}(y,\cdot)$. This is direct from the definition of Q since, for $z \preccurlyeq z'$, we have $P(x,z)/P(y,z) \geqslant P(x,z')/P(y,z')$ and then

$$\frac{\tilde{Q}(x,z)}{\tilde{Q}(y,z)} = \frac{P(x,z)h(y)}{P(y,z)h(x)} \geqslant \frac{P(x,z')h(y)}{P(y,z')h(x)} = \frac{\tilde{Q}(x,z')}{\tilde{Q}(y,z')}.$$

Then (a) follows.

(b) We have P is monotone. Besides, since $P(x,\cdot) \preccurlyeq_{\mathrm{tp}} P(x,\cdot)$, then $P(x,\cdot)$ is associated and the first inequality is proved.

Now, by the definition of R it is evident that $\tilde{R}(x,\cdot) \preccurlyeq_{\mathrm{tp}} \tilde{R}(y,\cdot)$ for $x \preccurlyeq y$ and (b) follows. □

Remark 3.2. The conditions of Theorem 3.1(b) on the sets B^i are satisfied, for instance, if they are unions of sets of the form

$$X_1 \times \cdots \times [x_j \preccurlyeq\ ] \times \cdots \times X_n,$$

where $[x_j \preccurlyeq\ ] = \{y \in X_j : y \preccurlyeq x_j\}$; that is, we remove from I any point which has at least one of its $j_1,\ldots,j_k$ coordinates smaller than $x_{j_1},\ldots,x_{j_k}$.

Remark 3.3. Although all the results of this work are given for finite I, they hold with the same proof for countable I, as long as the measures η and φ can be written as $\eta = \lim_{n\to\infty} \delta_a \tilde{P}^n / |\delta_a \tilde{P}^n|$ for some $a \in I$, and $\varphi = \lim_{n\to\infty} \tilde{P}^n f / |\tilde{P}^n f|$ for some increasing f.

References

Darroch, J.N., Seneta, E., 1965. On quasi-stationary distributions in absorbing discrete-time finite Markov chains. J. Appl. Probab. 2, 88–100.

Ferrari, P., Kesten, H., Martínez, S., Picco, P., 1995. Existence of quasi-stationary distributions. A Renewal Approach. Ann. Probab. 23, 501–521.

Harris, T., 1977. A correlation inequality for Markov processes in partially ordered state spaces. Ann. Probab. 5, 451–454.

Karlin, S., Rinott, Y., 1980. Classes of orderings of measures and related correlation inequalities. I Multivariate totally positive distributions. J. Multivariate Anal. 10, 467–498.

Kijima, M., 1997. Markov Processes for Stochastic Modeling. Chapman & Hall, London.

Liggett, T.M., 1985. Interacting Particle Systems. Springer, New York.

Lindqvist, B.H., 1988. Association of probability measures on partially ordered spaces. J. Multivariate Anal. 26, 111–132.

Moler, J.A., Plo, F., San Miguel, M., 2000. Minimal quasi-stationary distributions under null R-recurrence. Test 9, 455–470.

Müller, A., Stoyan, D., 2002. Comparison Methods for Stochastic Models and Risks. Wiley, Chichester.
Seneta, E., 1981. Non-negative Matrices and Markov Chains. Springer, New York.
Shaked, M., Shanthikumar, J.G., 1994. Stochastic Orders and Applications. Academic Press, New York.
van Doorn, E.A., 1991. Quasi-stationary distributions and convergence to quasi-stationarity of birth-death processes. Adv. Appl. Probab. 23, 683–700.