Monotonicity properties for bivariate supermodular functions of coordinates of stochastically monotone Markov processes

Kella, O.; Mandjes, M.R.H.

Publication date
2020

Document Version
Submitted manuscript

Published in
Journal of Applied Probability

Monotonicity properties for bivariate supermodular functions of coordinates of stochastically monotone Markov processes

Offer Kella*† and Michel Mandjes‡§

August 10, 2020

Abstract

It was recently proven that the correlation function of the stationary version of a reflected Lévy process is nonnegative, nonincreasing and convex. In another branch of the literature it was shown that the mean value of the reflected process is nonnegative, nondecreasing and concave. The main objective of this paper is to explore whether these two types of results can be covered by a common framework, and to investigate to what level they can be generalized. To this end, instead of a reflected Lévy process, a more general stochastically monotone Markov process, in addition satisfying a certain condition, is considered. We show monotonicity results associated with a supermodular function of two coordinates of our Markov process from which the above-mentioned monotonicity and convexity/concavity results follow for any such Markov process. In addition, various results for the transient case (when the Markov process is not in stationarity) are provided. The conditions imposed are natural, in that they are satisfied by various relevant processes. We specifically consider a class of dam processes as well as reflected Lévy processes (with both one- and two-sided reflection).

Keywords: stochastically monotone Markov processes, supermodular function, stochastic storage process, Lévy-driven queues, Skorokhod problem, monotone and convex correlation.

AMS Subject Classification (MSC2010): 60J99, 60G51, 90B05.

1 Introduction

In the context of Lévy-driven queues [8] and Lévy storage processes [18], it was recently shown [3] that, whenever the stationary distribution exists and has a finite second moment, the correlation function associated with the stationary version of the reflected process is nonnegative, nonincreasing and convex. Here, a Lévy-driven queue is to be interpreted as the one-sided (Skorokhod) reflection map applied to a Lévy process. Notably, the results in [3] show that the mentioned structural properties carry over to the finite-buffer Lévy-driven queue, i.e., the double-sided (Skorokhod) reflection map. One could regard [3] as the endpoint of a long-lasting research effort: the nonnegativity, nonincreasingness and convexity of the stationary correlation function was proven over four decades ago in [19] for the case the Lévy process under consideration is of compound Poisson type, whereas the more recent contributions [10] and [11] deal with the spectrally-positive and negative cases,

*Department of Statistics and Data Science, the Hebrew University of Jerusalem, Jerusalem 9190501, Israel (offer.kella@huji.ac.il).

†Supported in part by grant No. 1647/17 from the Israel Science Foundation and the Vigevani Chair in Statistics.

‡Korteweg-de Vries Institute for Mathematics, University of Amsterdam, Science Park 904, 1098 XH Amsterdam, The Netherlands (m.r.h.mandjes@uva.nl).

§Partly funded by the NWO Gravitation Programme Networks (Grant Number 024.002.003) and an NWO Top Grant (Grant Number 613.001.352).

respectively. Whereas [10, 11, 19] rely on the machinery of completely-monotone functions, [3] uses a direct conditioning argument in combination of elementary properties of the reflection map.

A second strand of research that we would like to mention concerns structural properties of the mean value (and related quantities) of the reflected process. It was found [14] that for a one-sided Skorohod reflection, when the driving process has stationary increments and starts from zero, the mean of the reflected process (as a function of time, that is) is nonnegative, nondecreasing and concave. In particular this holds when the driving process also has independent increments (i.e., the Lévy case), which for the spectrally-positive case had been discovered earlier [12], where we refer to [15, Thm. 11] for a multivariate analogue. The nonnegativity, nonincreasingness and concavity of the mean was proven to extend to the two-sided reflection case in [1], where it was also shown that for the one- and two-sided reflection cases the variance is nondecreasing.

The main objective of this paper is to explore to what level of generality the results from the above two branches of the literature can be extended, and whether they could be somehow brought under a common umbrella. Importantly, in our attempt to understand the above-mentioned structural properties better, we discovered that they are covered by a substantially broader framework. Indeed, we will consider stochastically monotone Markov processes (in both discrete and continuous time), and check what can be said about the expected value of bivariate supermodular functions of coordinates of the process. This turns out to be sufficient to imply the type of monotonicity results of the covariance that were found in [3, 10, 11] and [12, 14]. For the convexity results of [3, 10, 11] and the concavity results of [14] (restricted to Lévy processes) and [12] a further, rather natural, condition is needed – this is Condition 1 to follow. We conclude that our new findings directly cover a broad range of existing results. However, notably, the monotonicity of the variance established in [1] does *not* follow from the results of the current paper; a counterexample is provided.

The area of stochastically monotone Markov processes is vast. Without aiming at giving a full overview, we would like to mention [7, 16, 22, 23]. In particular, in [7] a main result is Theorem 4, stating that if $\{X_n \,|\, n \geqslant 0\}$ is a stationary stochastically monotone time-homogeneous Markov chain (real valued state space) and f is nondecreasing, then $\mathrm{Cov}(f(X_0), f(X_n))$ (whenever exists and is finite) is nonnegative and nonincreasing in n. As it turns out, this result as well is a special case the results established in our current paper.

In our proofs we use conditioning arguments similar to those relied upon in [3] in combination with the application of the concept of supermodularity. More concretely, it will be important to study the properties of $h(X_s, X_t)$ or $h(X_s, X_t - X_{t+\delta})$ (and others) for $0 \leqslant s \leqslant t$ and $\delta > 0$, where h is a supermodular function. These are well known, but we will state them in our notation (for the special case of bivariate functions of real variables) for ease of reference. This will be done for both the stationary case and the transient case (under various conditions). For background on results associated with supermodular functions, and in particular the relationship with comonotone random variables, which we will need several times, we refer to [5, 20, 21].

The paper is organized as follows. In Section 2 the setup, the main results and their proofs will be given. In Section 3 we provide three examples of stochastically monotone Markov processes that satisfy Condition 1: a Lévy dam process with left continuous release function $r(\cdot)$ with $r(0) = 0$, a Lévy process reflected at zero (one-sided reflection) and a Lévy process reflected at 0 and $b > 0$ (two-sided reflection). Our paper is self-contained in the sense that it does not require previous knowledge of stochastically monotone Markov processes.

2 Structural results

In what follows we write $a \wedge b = \min(a, b)$, $a \vee b = \max(a, b)$, $a^+ = a \vee 0$, $a^- = -a \wedge 0 = (-a)^+$. In addtion, *a.s.* abbreviates *almost surely* (i.e., *with probability one*), and *cdf* abbreviates *cumulative* *distribution function.*

For $x \in \mathbb{R}$ and A Borel (one-dimensional), let $p(x, A)$ be a Markov transition kernel. By this we mean that for every Borel A, $p(\cdot, A)$ is a Borel function and for each $x \in \mathbb{R}$, $p(x, \cdot)$ is a probability measure. We will say that p is *stochastically monotone* if $p(x, (y, \infty))$ is nondecreasing in x for each $y \in \mathbb{R}$, which is obviously a natural property across a broad range of frequently used models.

The following condition plays a crucial role in our results. Whenever it is satisfied, it allows us to establish highly general results. The condition is natural in the context of e.g. storage systems, as pointed out in Section 3.

Condition 1. $p(x, (x + y, \infty))$ is nonincreasing in x for each $y \in \mathbb{R}$.

Now, for $n \geqslant 1$, let p and p_n, for $n \geqslant 1$, be transition kernels. Define

$$G(x, u) = \inf\{y \,|\, p(x, (-\infty, y]) \geqslant u\} \tag{1}$$

the *generalized-inverse function* associated with the cdf $F_x(y) = p(x, (-\infty, y])$, and let similarly $G_n(x, u)$ be the generalized-inverse function associated with p_n. We recall (*e.g.,* [9], among many others) that $G(x, u)$ is nondecreasing and left continuous in u (on $(0, 1)$), and that $G(x, u) \leqslant y$ if and only if $u \leqslant p(x, (-\infty, y])$. Thus, if $U \sim \mathrm{U}(0, 1)$, with $\mathrm{U}(0, 1)$ denoting a standard uniform random variable, then $P(G(x, U) \leqslant y) = p(x, (-\infty, y])$ and thus $P(G(x, U) \in A) = p(x, A)$. A similar reasoning applies to $G_n(x, u)$ for every $n \geqslant 1$.

Lemma 1. *p is stochastically monotone if and only if, for each $u \in (0, 1)$, $G(x, u)$ is nondecreasing in x, and Condition 1 is satisfied if and only if, for each u, $G(x, u) - x$ is nonincreasing in x.*

Proof. Follows from the facts (i) that for $x_1 < x_2$

$$\{y \,|\, p(x_1, (-\infty, y] \geqslant u\} \supset \{y \,|\, p(x_2, (-\infty, y] \geqslant u\}\,, \tag{2}$$

(ii) that under Condition 1 for $x_1 < x_2$

$$\{y \,|\, p(x_1, (-\infty, x_1 + y] \geqslant u\} \subset \{y \,|\, p(x_2, (-\infty, x_2 + y] \geqslant u\}, \tag{3}$$

and (iii) that $G(x, u) - x = \inf\{y \,|\, p(x, (-\infty, x + y] \geqslant u\}$. ☐

Now, denote $g_k^{k+1}(x, u) = G_k(x, u)$ and, for $n \geqslant k + 2$,

$$g_k^n(x, u_1, \ldots, u_{n-k}) = G_n(g_k^{n-1}(x, u_1, \ldots, u_{n-k-1}), u_{n-k})\,. \tag{4}$$

It immediately follows by induction that in case $p_k, \ldots, p_n$ are stochastically monotone, it holds that $g_k^n(x, u_1, \ldots, u_{n-k})$ is nondecreasing in x. Assuming that $U_1, U_2, \ldots$ are i.i.d. and distributed $\mathrm{U}(0, 1)$, then with $X'_0 = x$ and

$$X'_n = g_1^n(x, U_1, \ldots, U_n) \tag{5}$$

for $n \geqslant 1$, $\{X'_n \,|\, n \geqslant 0\}$ is a real valued (possibly time-inhomogenous) Markov chain with possibly time-dependent transition kernels $p_1, p_2, \ldots$. Let us now denote $p_k^k(x, A) = 1_A(x)$ and, for $n \geqslant k+1$,

$$p_k^n(x, A) = \int_{\mathbb{R}} p_k^{n-1}(y, A) p_n(x, \mathrm{d}y)\,. \tag{6}$$

Lemma 2. *If, for $1 \leqslant k \leqslant n$, $p_k, \ldots, p_n$ are stocastically monotone Markov kernels (resp., and in addition satisfy Condition 1), then p_k^n is stochastically monotone (resp., and in addition satisfies Condition 1).*

Proof. By induction, it suffices to show this for the case $n = k+2$. If p_k and p_{k+1} are stochastically monotone, then $g_k^{k+2}(x, U_1, U_2)$ is a random variable having the distribution $p_k^{k+2}(x, \cdot)$. Therefore, the stochastic monotonicity of p_k^{k+2} is a consequence of the fact that $g_k^{k+2}(x, U_1, U_2)$ is nondecreasing in x. Now, if stochastic monotonicity and Condition 1 hold then $G_k(x, U_1) - x$ and $G_{k+1}(G_k(x, U_1), U_2) - G_k(x, U_1)$ are nonincreasing in x and thus so is their sum. This implies that $g_k^{k+2}(x, U_1, U_2) - x$ is nonincreasing in x, which implies that p_k^{k+2} satisfies Condition 1. $\square$

A (possibly time-inhomogeneous) Markov chain with stochastically monotone transition kernels will be called a *stochastically monotone Markov chain* (see, *e.g.*, [7] for the time-homogenous case). Lemma 2 immediately implies the following.

Corollary 1. *Any subsequence of a stochastically monotone Markov chain (resp., in addition satisfying Condition 1) is also a stochastically monotone Markov chain (resp., in addition satisfying Condition 1).*

Therefore, a subsequence of a time-homogeneous stochastically monotone Markov chain (resp., in addition satisfying Condition 1) may no longer be time-homogeneous, but is always a stochastically monotone Markov chain (resp., in addition satisfying Condition 1).

Recall that $h : \mathbb{R}^2 \to \mathbb{R}$ is called *supermodular* if whenever $x_1 \leqslant x_2$ and $y_1 \leqslant y_2$ we have that

$$h(x_1, y_2) + h(x_2, y_1) \leqslant h(x_1, y_1) + h(x_2, y_2) \ . \tag{7}$$

If X and Y have cdfs F_X and F_Y, then (X, Y) will be called *comonotone* if $\mathbb{P}(X \leqslant x, Y \leqslant y) = \mathbb{P}(X \leqslant x) \wedge \mathbb{P}(Y \leqslant y)$ for all x, y. There are various equivalent definitions for comonotonicity. In particular it is worth mentioning that when X and Y are identically distributed, then they are comonotone if and only if $\mathbb{P}(X = Y) = 1$. It is well known that if (X', Y') is comonotone and has the same marginals as (X, Y), then for any Borel supermodular h for which $\mathbb{E}h(X, Y)$ and $\mathbb{E}h(X', Y')$ exist and are finite, we have that $\mathbb{E}h(X, Y) \leqslant \mathbb{E}h(X', Y')$. In particular, when X, Y are identically distributed then $\mathbb{E}h(X, Y) \leqslant \mathbb{E}h(Y, Y)$, which is a property that we will need later in this paper. For such results and much more see, *e.g.*, [21] and references therein, where the Borel assumption was missing, but is actually needed as there are non-Borel supermodular functions for which $h(X, Y)$ is not necessarily a random variable. We write down what we will need later as a lemma. Everything in this lemma is well known.

Lemma 3. *Let (X, Y) be a random pair such that $X \sim Y$. Then for every Borel supermodular function $h : \mathbb{R}^2 \to \mathbb{R}$ for which $\mathbb{E}h(X, Y)$ and $\mathbb{E}h(Y, Y)$ exist and are finite, we have that*

$$\mathbb{E}h(X, Y) \leqslant \mathbb{E}h(Y, Y) \ . \tag{8}$$

Moreover, if h is supermodular and f_1, f_2 are nondecreasing, then $h(f_1(x), f_2(x))$ is supermodular and in particular, since $h(x, y) = xy$ is supermodular, $f_1(x)f_2(x)$ is supermodular as well.

As usual, we call π *invariant* for a Markov kernel p if, for every Borel A, $\int_{\mathbb{R}} p(x, A)\pi(\mathrm{d}x) = \pi(A)$. We proceed by stating and proving our first main result.

Theorem 1. *Assume that X_0, X_1, X_2 is a stochastically monotone Markov chain where p_1 has an invariant distribution π_1 and X_0 is π_1 distributed. Then for every Borel supermodular $h : \mathbb{R}^2 \to \mathbb{R}$,*

$$\mathbb{E}h(X_0, X_2) \leqslant \mathbb{E}h(X_1, X_2) \tag{9}$$

whenever the means exist and are finite. In particular, for any nondecreasing f_1, f_2 for which the means of $f_1(X_0)$, $f_2(X_2)$, $f_1(X_0)f_2(X_2)$ and $f_1(X_1)f_2(X_2)$ exist and are finite, we have that

$$0 \leqslant \mathrm{Cov}(f_1(X_0), f_2(X_2)) \leqslant \mathrm{Cov}(f_1(X_1), f_2(X_2)) \ . \tag{10}$$

Proof. Let X_0', U_1, U_2 be independent with $U_1, U_2 \sim \mathrm{U}(0,1)$ and $X_0' = X_0$. Then with $X_1' = G_1(X_0', U_1)$ and $X_2' = G_2(X_1', U_2)$ we have that $(X_0', X_1', X_2') \sim (X_0, X_1, X_2)$. Now, we note that since $G_2(y, u_2)$ is nondecreasing in y, then (due to Lemma 3) $h(x, G_2(y, u_2))$ is supermodular in x, y for every fixed u_2. Since $X_0' \sim X_1'$, then it follows from Lemma 3 that

$$\mathbb{E}(h(X_0', G_2(X_1', U_2)|U_2) \leqslant \mathbb{E}(h(X_1', G_2(X_1', U_2)|U_2)\,. \tag{11}$$

Taking expected values on both sides gives (9). Noting that $\mathbb{E}f_1(X_0)\mathbb{E}f_2(X_2) = \mathbb{E}f_1(X_1)\mathbb{E}f_2(X_2)$ and that $f_1(x)f_2(y)$ is supermodular gives the right inequality in (10). To show the left inequality, note that $G_2(G_1(x, u_1), u_2)$ is a nondecreasing function of x and thus, so is $\gamma(x) = \mathbb{E}f_2(G_2(G_1(x, U_1, U_2)))$. Now,

$$\mathrm{Cov}(f_1(X_0), f_2(X_2)) = \mathrm{Cov}(f_1(X_0), \mathbb{E}[f_2(X_2)|X_0]) = \mathrm{Cov}(f_1(X_0), \gamma(X_0)) \tag{12}$$

and it is well known that the covariance of comonotone random variables (whenever well defined and finite) must be nonnegative. □

A time-homogenous continuous-time Markov process $\{X_t \,|\, t \geqslant 0\}$ will be called stochastically monotone, whenever $p^t(x, A) = \mathbb{P}_x(X(t) \in A)$ is a stochastically monotone kernel for each $t > 0$. We will say that it satisfies Condition 1 whenever p^t satisfies this condition for every $t > 0$. Note that by Corollary 1 this is equivalent to the assumption that these conditions are satisfied for $0 < t \leqslant \epsilon$ for some $\epsilon > 0$.

Theorem 2. *Consider a stationary stochastically monotone discrete-time or continuous-time time-homogenous Markov process $\{X_t \,|\, t \geqslant 0\}$. Then for every supermodular h for which the following expectations exist and are finite, $\mathbb{E}h(X_0, X_t)$ is nonincreasing in $t \geqslant 0$ where t is either nonnegative integer or nonnegative real valued. In particular, when f_1, f_2 are nondecreasing and the appropriate expectations exist,* $\mathrm{Cov}(f_1(X_0), f_2(X_t))$ *is nonnegative and nonincreasing in $t \geqslant 0$.*

Proof. For every $0 < t_1 < t_2$ we have that $X_0, X_{t_2 - t_1}, X_{t_2}$ satisfy the conditions and hence the conclusions of Theorem 1 (for the discrete time case, recall Corollary 1). By stationarity we have that $(X_{t_2 - t_1}, X_{t_2}) \sim (X_0, X_{t_1})$. Therefore

$$\mathbb{E}h(X_0, X_{t_2}) \leqslant \mathbb{E}h(X_{t_2 - t_1}, X_{t_2}) = \mathbb{E}h(X_0, X_{t_1})\,. \tag{13}$$

Note that, since $X_t \sim X_0$, Lemma 3 implies that

$$\mathbb{E}h(X_0, X_t) \leqslant \mathbb{E}h(X_0, X_0) \tag{14}$$

so that $\mathbb{E}h(X_0, X_t)$ is nonincreasing on $[0, \infty)$ and not just on $(0, \infty)$. Since $\mathbb{E}f_1(X_0)\mathbb{E}f_2(X_t) = \mathbb{E}f_1(X_0)\mathbb{E}f_2(X_0)$, the result for the covariance follows by taking $h(x, y) = f_1(x)f_2(y)$. □

Remark 1. The following is a standard and very simple exercise in ergodic theory. Let T be a measure-preserving transformation on $(\Omega, \mathcal{F}, \mu)$, where μ is a σ-finite measure. This means that $\mu(T^{-1}(A)) = \mu(A)$ for every $A \in \mathcal{F}$, where $T^{-1}(A) = \{\omega \in \Omega \,|\, T(\omega) \in A\}$. Then T is mixing (in the sense that $\mu(A \cap T^{-n}B) \to \mu(A)\mu(B)$ as $n \to \infty$, for every $A, B \in \mathcal{F}$) if and only if for every $f_1, f_2 : \Omega \to \mathbb{R}$ such that $\int f_i^2 \mathrm{d}\mu < \infty$ for $i = 1, 2$ we have that $\int f_1 \cdot T^n f_2 \mathrm{d}\mu \to \int f_1 \mathrm{d}\mu \cdot \int f_2 \mathrm{d}\mu$, where $T^n f_2(\omega) = f_2(T^n \omega)$.

The implication of this under the assumptions of Theorem 2 is that with such mixing conditions $\mathrm{Cov}(f_1(X_0), f_2(X_n)) \to 0$ for every Borel f_1, f_2 such that $\mathbb{E}f_i^2(X_0) < \infty$ for $i = 1, 2$. This in particular holds when in addition f_i, $i = 1, 2$, are nondecreasing. This also implies that the same would hold in the continuous-time case as the covariance is nonincreasing (when f_i are nondecreasing) and thus it suffices that it vanishes along any subsequence (such as $t_n = n$). In particular this will hold for any Harris-recurrent Markov process for which there exists a stationary distribution.

In this case an even stronger form of mixing is known to hold (called *strong mixing* or α-mixing, *e.g.*, see [2]). All the examples discussed in Section 3 for which a stationary distribution exists are in fact Harris-recurrent and even have a natural regenerative state. Alternatively, the same holds whenever the stationary distribution is unique and f_1, f_2 are nondecreasing with $\mathbb{E}f_i(X_0)^2 < \infty$ for $i = 1, 2$. This may be shown by adapting the proof of Theorem 4 of [7] in which $f_1 = f_2$. For all of our examples in which a stationary distribution exists, it is also unique.

Of course, we cannot expect that the covariance will vanish without such mixing conditions or uniqueness of the stationary distribution. For example, if ξ is some random variable having a finite second moment and variance $\sigma^2 > 0$, set $X_t = \xi$ for all $t \geqslant 0$. Then $\{X_t \,|\, t \geqslant 0\}$ is trivially a stochastically monotone, stationary Markov process (and also trivially satisfies Condition 1), but (taking $f_1(x) = f_2(x) = x$) $\mathrm{Cov}(X_0, X_t) = \sigma^2$ clearly does not vanish as $t \to \infty$. $\diamond$

We continue with two theorems in which Condition 1 is imposed.

Theorem 3. *Assume that X_0, X_1, X_2, X_3 is a stochastically monotone Markov chain satisfying Condition 1, where p_1 has an invariant distribution π_1 and X_0 is π_1 distributed. Then for every Borel supermodular $h : \mathbb{R}^2 \to \mathbb{R}$,*

$$\mathbb{E}h(X_0, X_2 - X_3) \leqslant \mathbb{E}h(X_1, X_2 - X_3) \tag{15}$$

whenever the expectations exist and are finite. In particular, for f nondecreasing and the appropriate expectations exist and are finite, then

$$0 \leqslant \mathrm{Cov}(f(X_0), X_2) - \mathrm{Cov}(f(X_0), X_3) \leqslant \mathrm{Cov}(f(X_1), X_2) - \mathrm{Cov}(f(X_1), X_3)\,. \tag{16}$$

Proof. The proof is very similar to the proof of Theorem 1. That is, we let $X_0' = X_0$, $X_n' = G_n(X_{n-1}', U_n)$ for $n = 1, 2, 3$ where X_0', U_1, U_2, U_3 are independent and $U_1, U_2, U_3 \sim \mathrm{U}(0, 1)$. From the stochastic monotonicity and Condition 1 it follows that $G_2(x, u_2) - G_3(G_2(x, u_2), u_3)$ is nondecreasing in x. Therefore, by Lemma 3 we have that, since $h(x, G_2(y, u_2) - G_3(G_2(y, u_2), u_3))$ is supermodular in x, y and $X_1' \sim X_0'$, we have that

$$\begin{aligned}
\mathbb{E}[h(X_0', X_2' - X_3')|U_2, U_3] &= \mathbb{E}[h(X_0', G_2(X_1', U_2) - G_3(G_2(X_1', U_2), U_3))|U_2, U_3] \\
&\leqslant \mathbb{E}[h(X_1', G_2(X_1', U_2) - G_3(G_2(X_1', U_2), U_3))|U_2, U_3] \\
&= \mathbb{E}[h(X_1', X_2' - X_3')|U_2, U_3]\ ,
\end{aligned} \tag{17}$$

and taking expected values establishes (15). Taking $h(x, y) = f(x)y$ gives the right inequality of (16). The left inequality is obtained via comonotonicity, by observing that since $G_2(G_1(x, u_1), u_2) - G_3(G_2(G_1(x, u_1), u_2), u_3)$ is nondecreasing in x, then $\mathbb{E}[X_2 - X_3|X_0]$ is a nondecreasing function of X_0, so that this inequality follows from the comonotonicity of $f(X_0)$ and $\mathbb{E}[X_2 - X_3|X_0]$ as in the proof of the left inequality of (10). $\square$

Theorem 4. *Consider a stationary stochastically monotone discrete-time or continuous-time time-homogenous Markov process $\{X_t \,|\, t \geqslant 0\}$, satisfying Condition 1. Then for every $s > 0$ and every supermodular h for which the following expectations exist and are finite, $\mathbb{E}h(X_0, X_t - X_{t+s})$ is nonincreasing in t where t is either nonnegative integer or nonnegative real valued. In particular, when f is nondecreasing and the appropriate expectations exist, $\mathrm{Cov}(f(X_0), X_t)$ is nonnegative, nonincreasing and convex in t.*

In particular, note that when choosing $f(x) = x$ and assuming that $\mathbb{E}X_0^2 < \infty$, we see that, under the conditions of Theorem 4, the auto-covariance $R(t) = \mathrm{Cov}(X_s, X_{s+t})$ (or auto-correlation $R(t)/R(0)$ when X_0 is not a.s. constant) is nonnegative, nonincreasing and convex in t.

Proof. Let $t_1 < t_2$ then X_0, $X_{t_2-t_1}$, $X_{t_2-t_1+s}$, X_{t_2+s} satisfy the conditions and hence the conclusion of Theorem 3. Therefore,

$$\mathbb{E}h(X_0, X_{t_2} - X_{t_2+s}) \leqslant \mathbb{E}h(X_{t_2-t_1}, X_{t_2} - X_{t_2+s}) = \mathbb{E}h(X_0, X_{t_1} - X_{t_1+s}) \tag{18}$$

where the right equality follows from stationarity. When f is nondecreasing then $h(x, y) = f(x)y$ is supermodular and thus $\mathbb{E}f(X_0)X_{t+s} - \mathbb{E}f(X_0)X_t$ is nonincreasing in t for every $s > 0$. Therefore $\mathbb{E}f(X_0)X_t$ is midpoint convex and since by Theorem 2 it is nonnegative and nonincreasing (hence Borel), it must be convex (see [4, 24]). □

Can anything be said for the case where the initial distribution is not invariant? Here is one possible answer.

Theorem 5. *Let $\{X_t \,|\, t \geqslant 0\}$ be a stochastically monotone discrete-time or continuous-time time-homogenous Markov process. Assume that the initial distribution can be chosen so that $X_0 \leqslant X_t$ a.s. for every $t \geqslant 0$. Then,*

(i) X_t is stochastically increasing in t.

(ii) For every Borel supermodular function which is nondecreasing in its first variable and for which the expectations exist and are finite, $\mathbb{E}h(X_s, X_t)$ is nondecreasing in s on $[0, t]$. When in addition Condition 1 is satisfied, the same is true for $\mathbb{E}h(X_s, X_t - X_{t+\delta})$ for every $\delta > 0$ (whenever the expectations exist and are finite).

(iii) When h is nondecreasing in both variables (not necessarily supermodular) and expected values exist and are finite, then $\mathbb{E}h(X_s, X_{t+s})$ is nondecreasing in s. When Condition 1 is satisfied, the same is true for $\mathbb{E}h(X_s, X_{s+t} - X_{s+t+\delta})$ for $\delta > 0$.

(iv) When $\mathbb{E}X_t$ exists and is finite then it is nondecreasing and under Condition 1 it is also concave.

We note that it would suffice to assume that $X_0 \leqslant X_t$ for $t \in (0, \epsilon]$ for some $\epsilon > 0$, or in discrete time for $t = 1$. Also, we note that for (i) we can replace $X_0 \leqslant X_t$ by $X_0 \leqslant_{\text{st}} X_t$.

Proof. For any $s < t$ take $\epsilon \in (0, t - s]$ and let $G_v(x, u)$ be the generalized-inverse with respect to the kernel p^v. Let $U_0, U_1, \dots$ be i.i.d. with $U_i \sim \text{U}(0, 1)$. Since $X_0 \leqslant X_\epsilon$ we have with $X'_0 = X_0$, that $X'_0 \leqslant G_h(X'_0, U_0)$ and thus

$$X'_s \equiv G_s(X'_0, U_1) \leqslant G_s(G_\epsilon(X'_0, U_0), U_1) \equiv X''_{s+\epsilon} \sim X'_{s+\epsilon} \equiv G_\epsilon(X'_s, U_2)\,, \tag{19}$$

implying stochastic monotonicity.

Taking $X'_t = G_{t-s-\epsilon}(X'_{s+\epsilon}, U_3)$, then $(X'_0, X'_s, X'_{s+\epsilon}, X'_t)$ is distributed like $(X_0, X_s, X_{s+\epsilon}, X_t)$. Since h is nondecreasing in its first variable and $X'_s \leqslant X''_{s+\epsilon}$, it follows that

$$\mathbb{E}h(X'_s, X'_t) \leqslant \mathbb{E}h(X''_{s+\epsilon}, X'_t)\ . \tag{20}$$

The (by now, repetitive) fact that

$$\begin{aligned}\mathbb{E}[h(X''_{s+\epsilon}, X'_t)|U_3] &= \mathbb{E}[h(X''_{s+\epsilon}, G_{t-s-\epsilon}(X'_{s+\epsilon}, U_3)|U_3] \\ &\leqslant \mathbb{E}[h(X'_{s+\epsilon}, G_{t-s-\epsilon}(X'_{s+\epsilon}, U_3)|U_3] = \mathbb{E}[h(X'_{s+\epsilon}, X'_t)|U_3]\end{aligned} \tag{21}$$

follows from the supermodularity of $h(x, G_{t-s-\epsilon}(y, u_3))$ in x, y. Taking expected values implies, together with (20), that $\mathbb{E}h(X_s, X_t)$ is nondecreasing in s on $[0, t]$. The proof of the fact that, under Condition 1, $\mathbb{E}h(X_s, X_t - X_{t+\delta})$ is nondecreasing in s on $[0, t]$ is similar, once we define

$X'_{t+\delta} = G_\delta(X'_t, U_4)$ and observe that $X'_t - X'_{t+\delta} = G_{t-s-\epsilon}(X'_{s+\epsilon}, U_3) - G_\delta(G_{t-s-\epsilon}(X'_{s+\epsilon}, U_3), U_4)$ is nondecreasing in $X'_{s+\epsilon}$.

When h is nondecreasing in both variables we have that $h(X_s, X_{s+t}) \sim h(X_s, G_t(X_s, U_0))$ so that by stochastic monotonicity $\mathbb{E}(h(X_s, G_t(X_s, U_0))|U_0)$ is nondecreasing in s and hence also $\mathbb{E}h(X_s, X_{s+t})$. The proof for $\mathbb{E}h(X_s, X_{s+t} - X_{s+t+\delta})$, under Condition 1, is similar.

Finally, since X_t is stochastically increasing then it clearly follows that $\mathbb{E}X_t$ is nondecreasing. When Condition 1 is met, then taking $h(x, y) = y$ (nondecreasing in both variables) we have that $\mathbb{E}X_{s+t} - \mathbb{E}X_{s+t+\delta}$ is nondecreasing. This implies midpoint concavity, so that since $\mathbb{E}X_t$ is monotone (hence Borel) it follows by that it is concave (again, see [4, 24]). $\square$

We complete this section by noting that although, for the sake of convenience, all the results were written for the case where the state space is $\mathbb{R}$, they hold whenever the state space is any Borel subset of $\mathbb{R}$ as was assumed in [7]. For the three examples that will be covered in Section 3, the state space is $[0, \infty)$.

3 Examples

In this section we discuss some examples of stochastically monotone time-homogeneous Markov processes satisfying Condition 1. We start by providing a number of general observations.

For each of the examples contained in this section there exists a huge body of literature. If we were only interested in stochastic monotonicity, then standard examples are birth-death processes and diffusions, *e.g.*, [16]. Recalling Remark 1, when there exists a stationary distribution, for all of these examples we have that $\mathrm{Cov}(f_1(X_0), f_2(X_t))$ vanishes as $t \to \infty$ whenever f_1, f_2 are nondecreasing with $\mathbb{E}f_i^2(X_0) < \infty$ for $i = 1, 2$. We note that in light of Subsections 3.2 and 3.3, [3, Thms. 1 and 2] (as well as the earlier [10, Thm. 3.1] and [11, Thm. 2.2]) are special cases of the Theorems 2 and 4 here. In addition, [14, Thm. 3.1] (which holds for any reflected process with stationary, not necessarily independent, increments, hence not necessarily Markov) restricted to the Lévy case (and the earlier [12, Thm. 3.3]) as well as the mean (not variance) parts of [1, Thms. 4.6 and 7.5] are special cases of (iv) of Theorem 5 here (upon taking $X_0 = 0$ in [1], that is).

It would have been nice if the monotonicity of the variance discovered in [1] would hold for any stochastically monotone Markov process satisfying Condition 1. However, it turns out that this particular result from [1] essentially follows due to the specific properties of reflected Lévy processes (or, in the discrete time case, reflected random walks) and, unfortunately, it is not true in general. One very simple counterexample is the following. Let $\{N_t \,|\, t \geqslant 0\}$ be a Poisson process (starting at 0) and take $X_t = (k + N_t) \wedge m$. Then $\{X_t \,|\, t \geqslant 0\}$ is a Markov process with state space $\{i \,|\, i \leqslant m\}$, with an initial value k and an absorbing barrier m. On $k \leqslant m$ we have that $(k + N_t) \wedge m$ is nondecreasing in k and $X_t - X_0 = N_t \wedge (m - k)$ is nonincreasing in k and thus this is a stochastically monotone Markov process satisfying Condition 1. Clearly, $\mathrm{Var}(X_0) = 0$ and since $X_t \to m$ (a.s.) as $t \to \infty$ then by bounded convergence ($k \leqslant X_t \leqslant m$) the variance vanishes as $t \to \infty$. Since the variance is strictly positive for all $0 < t < \infty$, it cannot be monotone in t.

We continue by discussing three settings that obey the desired stochastic monotonicity as well as Condition 1, so that our results apply. All three may be referred to either Lévy-driven queues [8] or Lévy storage processes [18] or both.

3.1 Lévy dams with nondecreasing left continuous general release rule

Let the process $J = \{J_t \,|\, t \geqslant 0\}$ be a right continuous subordinator (nondecreasing Lévy process) with $\mathbb{P}(J_0 = 0) = 1$ and let $r : [0, \infty) :\to [0, \infty)$ be nondecreasing, left continuous on $(0, \infty)$, with

$r(0) = 0$. Consider the following *dam process*:

$$X_t(x) = x + J_t - \int_0^t r(X_s(x))\mathrm{d}s\,. \tag{22}$$

It is well known [6] that, under the stated assumptions, the solution to (22) is unique (pathwise) and belongs to the class of time-homogeneous Markov processes. For $x < y$ we have that

$$X_t(y) - X_t(x) = y - x - \int_0^t (r(X_s(y) - r(X_s(x))\mathrm{d}s\ . \tag{23}$$

Denote τ to be first time (if it exists) for which the right side is zero. Because $x < y$, then for every $t < \tau$ we have that $X_t(x) < X_t(y)$. On $\tau < \infty$ we clearly have that $X_\tau(x) = X_\tau(y)$ and therefore we also have that for any $h \geqslant 0$,

$$\begin{aligned} X_{\tau+h}(x) &= X_\tau(x) + J_{\tau+h} - J_\tau + \int_0^h r(X_{\tau+s}(x))\mathrm{d}s \\ &= X_\tau(y) + J_{\tau+h} - J_\tau + \int_0^h r(X_{\tau+s}(x))\mathrm{d}s = X_{\tau+h}(y)\,, \end{aligned} \tag{24}$$

where the right inequality follows from the uniqueness of the solution Z to the equation

$$Z_h = z + J_{\tau+h} - J_\tau - \int_0^h r(Z_s)\mathrm{d}s\,. \tag{25}$$

Therefore, we have that $X_t(x) \leqslant X_t(y)$ for every $t \geqslant 0$. Moreover, note that

$$(X_t(x) - x) - (Y_t(y) - y) = \int_0^t (r(X_s(y)) - r(X_s(x))\mathrm{d}s \tag{26}$$

and thus, since r is assumed to be nondecreasing, we also have that $X_t(x) - x \geqslant X_t(y) - y$. The conclusion is that $X_t(x)$ is nondecreasing in x and $X_t(x) - x$ is nonincreasing in x. In other words, the process considered is a stochastically monotone Markov process satisfying Condition 1. We note that here a stationary distribution exists whenever $\mathbb{E}J_1 < r(x)$ for some $x > 0$ (recalling that $r(\cdot)$ is nondecreasing).

So as to perform a sanity check, we note that when choosing $r(x) = rx$ the resulting process is a (generalized) shot-noise process. In this case we can explicitly write

$$X_t(x) = xe^{-rt} + \int_{(0,t]} e^{-r(t-s)} J_{\mathrm{d}s}. \tag{27}$$

In this setting it is immediately clear that $X_t(x)$ is nondecreasing and $X_t(x)-x$ is nonincreasing in x. We observe that here, if $\mathbb{E}X_0^2 < \infty$, then $R(t) = \mathrm{Cov}(X_0, X_t) = \mathrm{Var}(X_0)e^{-rt}$, so that $R(t)/R(0) = e^{-rt}$, which is, as expected, nonnegative, nonincreasing, convex in t (for any distribution of X_0) and also converges to zero as $t \to \infty$. It is well known that in this particular case, the stationary distribution has a finite second moment if and only if J_1 has a finite second moment. This is equivalent to requiring that $\int_{(1,\infty)} x^2\nu(\mathrm{d}x) < \infty$, where ν is the associated Lévy measure.

3.2 Lévy process reflected at the origin

Consider a càdlàg Lévy process $Y = \{Y_t \,|\, t \geqslant 0\}$ with $\mathbb{P}(Y_0 = 0) = 1$ (not necessarily spectrally one-sided). For every x, the one-sided (Skorokhod) reflection map, with reflection taking place at level 0, is defined through

$$X_t(x) = x + Y_t - \inf_{0\leqslant s\leqslant t}(x + Y_s) \wedge 0 = Y_t + L_t \wedge x \tag{28}$$

where $L_t = -\inf_{0\leqslant s\leqslant t} Y_s \wedge 0$ with $L_t(x) = (L_t - x)^+$ (so, in particular, $L_t = L_t(0)$). The pair $(L_t(x), X_t(x))$ is known to be the unique process satisfying

(i) $L_t(x)$ is right continuous, nondecreasing in t, with $L_0 = 0$.

(ii) $X_t(x)$ is nonnegative for every $t \geqslant 0$.

(iii) For every $t > 0$ such that $L_s(x) < L_t(x)$ for every $s < t$ we have that $X_t(x) = 0$.

It is known [13] that (iii) is equivalent to the condition that

$$\int_{[0,\infty)} X_s(x) L_{\mathrm{d}s}(x) = 0, \tag{29}$$

or alternatively to the condition that $L_t(x)$ is the minimal process satisfying (i) and (ii). Special cases of such processes are the workload process in an M/G/1 queue and the (single dimensional) reflected Brownian motion (where the reflection takes place at 0).

Clearly $X_t(x) = Y_t + L_t \wedge x$ is nondecreasing in x and $X_t(x) - x = Y_t + (L_t - x)^-$ is nonincreasing in x. Since Y is a Lévy process, then $X_t(x)$ is well known to be a time-homogeneous Markov process starting at x. Therefore, it is necessarily a stochastically monotone Markov process satisfying Condition 1. Here a stationary distribution exists whenever $\mathbb{E}Y_1 < 0$.

We note that for the same reasons, a (general) random walk reflected a the origin is a discrete time version of the process featuring in the above setup. As a consequence, it is also stochastically monotone and satisfies Condition 1. In particular, this applies to the consecutive waiting times upon arrivals of customers in a GI/GI/1 queue.

3.3 Lévy process with a two-sided reflection

With Y defined in Subsection 3.2, a two-sided (Skorokhod) reflection in $[0, b]$ for $b > 0$ (and similarly in $[a, b]$ for any $a < b$) is defined as the unique process $(X_t(x), L_t(x), U_t(x))$, with $X_t(x) = x + Y_t + L_t(x) - U_t(x)$, satisfying

(i) $L_t(x)$, $U_t(x)$ are right continuous and nondecreasing with $L_0(x) = U_0(x) = 0$.

(ii) $X_t(x) \in [0, b]$ for all $t \geqslant 0$.

(iii) For every $t > 0$ such that $L_s(x) < L_t(x)$ (resp., $U_s(x) < U_t(x)$) for every $s < t$, $X_t(x) = 0$ (resp., $X_t(x) = b$).

Also here, (iii) is equivalent to

$$\int_{[0,\infty)} X_t(x) L_{\mathrm{d}t}(x) = \int_{[0,\infty)} (b - X_t(x)) U_{\mathrm{d}t}(x) = 0. \tag{30}$$

Since Y is a Lévy process, we have that $X_t(x)$ is a time-homogenous Markov process starting at x. The driving process Y being the same for both $X_t(x)$ and $X_t(y)$, we find that choosing $x < y$ means that $X_t(x)$ can never overtake $X_t(y)$. Consequently, $X_t(x)$ is nondecreasing in x and thus the Markov chain is stochastically monotone (with some effort, this can also be shown directly from (31) to follow). In order to verify that it satisfies Condition 1, we recall from [17], upon re-denoting by $X_t^0(x)$ the single sided reflected process described in Subsection 3.2, that

$$X_t(x) = X_t^0(x) - \sup_{0 \leqslant s \leqslant t} \left[\left(X_s^0(x) - b \right)^+ \wedge \inf_{s \leqslant u \leqslant t} X_t^0(x) \right] . \tag{31}$$

Since $X_t^0(x)$ is nondecreasing and $X_t^0(x) - x$ is nonincreasing in x (as explained in Subsection 3.2), it immediately follows that $X_t(x) - x$ is nonincreasing in x, which implies Condition 1. In this doubly reflected case, a stationary distribution always exists. Again the findings carry over to the process' discrete time counterpart.

References

[1] Andersen, L. N. and M. Mandjes. 2009. Structural properties of reflected Lévy processes. *Queueing Syst.* **63**, 301-322.

[2] Athreya, K. B. and S. G. Pantula. 1986. Mixing properties of Harris chains and autoregressive processes. *J. Appl. Probab.* **23**, 880-892.

[3] Berkelmans, W., Cichocka, A. and M. Mandjes. 2019. The correlation function of a queue with Lévy and Markov additive input. *Stoch. Proc. Appl.* **130**, 1713-1734.

[4] Blumberg, H. 1919. Convex functions. *Trans. Amer. Math. Soc.* **20**, 40-44.

[5] Cambanis, S., Simons, G. and W. Stout. 1976. Inequalities for $\mathbb{E}k(X,Y)$ when the marginals are fixed. *Z. Wahrscheinlichkeitstheorie verw. Geb.* **36**, 285-294.

[6] Çinlar, E. and M. Pinsky. 1972. On dams with additive inputs and a general release rule. *J. Appl. Probab.* **9**, 422-429.

[7] Daley, D. J. 1968. Stochastically monotone Markov chains. *Z. Wahrscheinlichkeitstheorie verw. Geb.* **10**, 305-317.

[8] Dębicki, K. and M. Mandjes. 2015. *Queues and Lévy Fluctuation Theory.* Springer, New York, NY.

[9] Embrechts, P. and M. Hofert. 2013. A note on generalized inverses. *Math. Meth. Oper. Res.* **77,** 423-432.

[10] Es-Saghouani, A. and M. Mandjes. 2008. On the correlation structure of a Lévy-driven queue. *J. Appl. Probab.* **45**, 940-952.

[11] Glynn, P. and M. Mandjes. 2011. Simulation-based computation of the workload correlation function in a Lévy-driven queue. *J. Appl. Probab.* **48**, 114-130.

[12] Kella, O. Concavity and reflected Lévy process. *J. Appl. Probab.* **29**, 209-215.

[13] Kella, O. 2006. Reflecting thoughts. *Stat. Prob. Let.* **76**, 1808-1811.

[14] Kella, O. and M. Sverchkov. 1994. On concavity of the mean function and stochastic ordering for reflected processes with stationary increments. *J. Appl. Probab.* **31**, 1140-1142.

[15] Kella, O. and W. Whitt. 1996. Stability and structural properties of stochastic storage networks. *J. Appl. Probab.* **33**, 1169-1180.

[16] Keilson, J. and A. Kester. 1977. Monotone matrices and monotone Markov processes. *Stoch. Proc. Appl.* **5**, 231-241.

[17] Kruk, L., Lehoczky, J., Ramanan, K. and S. Shreve. 2007. An explicit formula for the Skorokhod map on $[0,a]$. *Ann. Probab.* **35**, 1740-1768.

[18] Kyprianou, A. 2006. *Introductory Lectures on Fluctuations of Lévy Processes with Applications.* Springer, New York, NY.

[19] Ott, T. 1977. The covariance function of the virtual waiting-time process in an M/G/1 queue. *Adv. Appl. Probab.* **9**, 158-168.

[20] Puccetti, G. and M. Scarsini. 2010. Multivariate comonotonicity. *J. Multivar. Anal.* **101**, 291-304.

[21] Puccetti, G. and R. Wang 2015. Extremal dependence concepts. *Stat. Sci.* **30**, 385-517.

[22] Rüschendorf, L. 2008. On a comparison results for Markov processes. *J. Appl. Probab.* **45**, 279-286.

[23] Siegmund, D. 1976. The equivalence of absorbing and reflecting barrier problems for stochastically monotone Markov processes. *Ann. Probab.* **4**, 914-924.

[24] Sierpinski, W. 1920. Sur les fonctions convexes mesurables. *Fund. Math.* **1**, 125-129.