FROM REFLECTED LÉVY PROCESSES TO STOCHASTICALLY MONOTONE MARKOV PROCESSES VIA GENERALIZED INVERSES AND SUPERMODULARITY

OFFER KELLA 🆔,* *The Hebrew University of Jerusalem*
MICHEL MANDJES,** *University of Amsterdam*

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 established that the mean value of the reflected process starting from zero is nondecreasing and concave. In the present paper it is shown, by putting them in a common framework, that these results extend to substantially more general settings. Indeed, instead of reflected Lévy processes, we consider a class of more general stochastically monotone Markov processes. In this setup 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 directly follow, but now for the class of Markov processes considered rather than just reflected Lévy processes. 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 frequently used Markovian models, as illustrated by a series of examples.

Keywords: Stochastic storage process; Lévy-driven queues; Skorokhod problem; monotone and convex autocorrelation; concave mean

2020 Mathematics Subject Classification: Primary 60J99
Secondary 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 two-sided (Skorokhod) reflection map. One could regard [3] as the endpoint of a long-lasting research effort which began over four decades ago. The nonnegativity, nonincreasingness, and convexity of the correlation function of the stationary process was proven in [19] for the case

where the Lévy process under consideration is compound Poisson. The more recent contributions [10, 11] deal with the spectrally positive and negative cases, respectively. Finally, [3] removed the spectral restrictions on the Lévy process assumed in [10, 11]. 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 Skorokhod reflection, when the driving process has stationary increments and starts from zero, the mean of the reflected process (as a function of time) is 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]; we refer to [15, Theorem 11] for a multivariate analogue. The nondecreasingness 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. We have done so by considering stochastically monotone Markov processes (in both discrete and continuous time), and deriving properties of the expected value of bivariate supermodular functions of coordinates of the process.

Importantly, we discovered a neat and quite simple approach to extend a broad range of existing results to a substantially broader class of processes and more general functional setups. We strongly feel that this particular approach gets to the heart of the matter, and also helps in giving a much clearer understanding of earlier results. More specifically, our findings directly imply the type of monotonicity results for 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 needs to be imposed on the underlying Markov transition kernel; this is Condition 1, to follow. However, notably, the monotonicity of the variance established in [1] is *not* valid under the conditions imposed in the current paper, as we will show by means of a counterexample.

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, which states that if $\{X_n \mid n \geq 0\}$ is a stationary stochastically monotone time-homogeneous Markov chain (on a real-valued state space) and f is nondecreasing, then $\mathrm{Cov}(f(X_0), f(X_n))$ (whenever it exists and is finite) is nonnegative and nonincreasing in n. As it turns out, this result also is a special case of the results established in our current paper. Importantly, quite a few frequently used stochastic processes are stochastically monotone Markov processes, including for example birth–death processes and diffusions [16], as well as certain Lévy dams and (state-dependent) random walks, besides the above-mentioned reflected processes.

In our proofs we use the notion of a generalized inverse of a distribution function and some of its properties, conditioning arguments, and the application of the concept of supermodularity and its relationship to comonotonicity. 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 $0 \leq s \leq t$ and $\delta > 0$, where h is a supermodular function. 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].

As so often in mathematics, once we are in the right framework, proofs can be highly compact and seemingly straightforward. It is, however, typically far from trivial to identify this best 'lens' through which one should look at the problem. This phenomenon also applies in the context of the properties derived in the present paper. Indeed, in previous works the focus has been on specific models and specific properties, with proofs that tend to be ad hoc, lengthy, and involved, reflecting the lack of an overarching framework. With the general approach that we develop in this paper, we manage to bring a wide class of existing results under a common denominator, with the underlying proofs becoming clean and insightful. In addition, because we have found the right angle to study this class of problems, we succeed in shedding light on the question of the extent to which these results can be further generalized. Our objective was to present the framework as cleanly as possible; our paper is self-contained in the sense that it does not require any previous knowledge of stochastically monotone Markov processes.

The paper is organized as follows. In Section 2 the formal setup (including Condition 1), the main results, and their proofs will be given. In Section 3 we provide a series of appealing examples of frequently used stochastically monotone Markov processes that satisfy Condition 1. Section 4 concludes.

Throughout we write $a \wedge b = \min(a, b)$, $a \vee b = \max(a, b)$, $a^+ = a \vee 0$, $a^- = -a \wedge 0 = (-a)^+$. In addition, *a.s.* stands for *almost surely* (i.e., *with probability one*), and *CDF* stands for *cumulative distribution function*. Also, $=_d$ means 'distributed' or 'distributed like', and $X \leq_{st} Y$ means $\mathbb{P}(X > t) \leq \mathbb{P}(Y > t)$ for all t (stochastic order).

2. General theory

This section presents our main results. Section 2.1 treats our general theory, whereas Section 2.2 further reflects on general questions that can be dealt with relying on our results, as well as the connection with supermodular functions.

2.1. Main results

For $x \in \mathbb{R}$ and A Borel (one-dimensional), we 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, as was discussed earlier and will be demonstrated later, is a natural property across a broad range of frequently used stochastic 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., queues and other storage systems, as will be pointed out in Section 3.

Condition 1. *For each y, $p(x, (x + y, \infty))$ is nonincreasing in x.*

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

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

to be the *generalized inverse function* associated with the CDF $F_x(y) = p(x, (-\infty, y])$, and similarly let $G_n(x, u)$ be the generalized inverse function associated with p_n. We recall (see e.g. the paper [9], among many others) that $G(x, u)$ is nondecreasing and left-continuous in u (on $(0, 1)$), and that $G(x, u) \leq y$ if and only if $u \leq p(x, (-\infty, y])$. Thus, if $U =_d U(0, 1)$, where

U(0, 1) is the uniform distribution on (0, 1), then $\mathbb{P}(G(x, U) \leq y) = p(x, (-\infty, y])$ and thus $\mathbb{P}(G(x, U) \in A) = p(x, A)$. Similar reasoning applies to $G_n(x, u)$ for every $n \geq 1$.

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

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

$$\{y \mid p(x_1, (-\infty, y]) \geq u\} \supset \{y \mid p(x_2, (-\infty, y]) \geq u\}, \tag{2}$$

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

$$\{y \mid p(x_1, (-\infty, x_1 + y]) \geq u\} \subset \{y \mid p(x_2, (-\infty, x_2 + y]) \geq u\}, \tag{3}$$

and (iii) that $G(x, u) - x = \inf\{y \mid p(x, (-\infty, x + y]) \geq u\}$. $\square$

Now, let $g_k^k(x, u) = G_k(x, u)$, and for $n \geq k + 1$ let

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

It immediately follows by induction that if $p_k, \ldots, p_n$ are stochastically monotone, then $g_k^n(x, u_1, \ldots, u_{n-k+1})$ is nondecreasing in x. Assume that $U_1, U_2, \ldots$ are independent and identically distributed (i.i.d.) with distribution U(0, 1); then with $X_0' = x$ and

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

for $n \geq 1$, $\{X_n' \mid n \geq 0\}$ is a real-valued (possibly time-inhomogeneous) Markov chain with possibly time-dependent transition kernels $p_1, p_2, \ldots$. Let us now set $p_k^{k-1}(x, A) = 1_A(x)$ and, for $n \geq k$,

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

Lemma 2. *If, for $1 \leq k \leq n$, $p_k, \ldots, p_n$ are stochastically 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 $n = k + 1$. If p_k and p_{k+1} are stochastically monotone, then $g_k^{k+1}(x, U_1, U_2)$ is a random variable having the distribution $p_k^{k+1}(x, \cdot)$. Therefore, the stochastic monotonicity of p_k^{k+1} is a consequence of the fact that $g_k^{k+1}(x, U_1, U_2)$ 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 $g_k^{k+1}(x, U_1, U_2) - x$ is nonincreasing in x, which implies that p_k^{k+1} 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-homogeneous 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 $x_1 \leq x_2$ and $y_1 \leq y_2$ we have that

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

If X and Y have CDFs F_X and F_Y, then (X, Y) will be called *comonotone* if $\mathbb{P}(X \leq x, Y \leq y) = \mathbb{P}(X \leq x) \wedge \mathbb{P}(Y \leq y)$ 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, 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 $\mathbb{E}h(X, Y) \leq \mathbb{E}h(X', Y')$. In particular, if X, Y are identically distributed, then $\mathbb{E}h(X, Y) \leq \mathbb{E}h(Y, Y)$, 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 is not made 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 $X =_{\mathrm{d}} Y$. 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*

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

Moreover, if h is supermodular and f_1, f_2 are nondecreasing, then $h(f_1(x), f_2(y))$ is supermodular, and in particular, since $h(x, y) = xy$ is supermodular, $f_1(x)f_2(y)$ 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 $h: \mathbb{R}^2 \to \mathbb{R}$,*

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

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

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

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

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

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,

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

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

A time-homogeneous continuous-time Markov process $\{X_t \mid t \geq 0\}$ 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 $0 < t \leq \epsilon$ for some $\epsilon > 0$.

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

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 $(X_{t_2-t_1}, X_{t_2}) =_d (X_0, X_{t_1})$. Therefore

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

Note that, since $X_t =_d X_0$, Lemma 3 implies that

$$\mathbb{E}h(X_0, X_t) \leq \mathbb{E}h(X_0, X_0), \quad (14)$$

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 $T^{-1}(A) = \{\omega \in \Omega \mid T(\omega) \in A\}$. 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 $\int f_i^2 d\mu < \infty$ for $i = 1, 2$, we have that $\int f_1 \cdot T^n f_2 d\mu \to \int f_1 d\mu \cdot \int f_2 d\mu$, 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, $\text{Cov}(f_1(X_0), f_2(X_n)) \to 0$ 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*; see e.g. [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 [7, Theorem 4], 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 $t \geq 0$. Then $\{X_t \mid t \geq 0\}$ 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 $t \to \infty$.   ◇

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 $h : \mathbb{R}^2 \to \mathbb{R}$,*

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

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

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

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 $U_1, U_2, U_3 =_{\mathrm{d}}$ 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 $X_1' =_{\mathrm{d}} X_0'$,

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

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, $\mathbb{E}[X_2 - X_3|X_0]$ 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).   □

Theorem 4. *Consider a stationary stochastically monotone discrete-time or continuous-time time-homogeneous Markov process $\{X_t \mid t \geq 0\}$, 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 $R(t) = \text{Cov}(X_s, X_{s+t})$ (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,

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

where the right equality follows from stationarity. If 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 $\{X_t \mid t \ge 0\}$ be a stochastically monotone discrete-time or continuous-time time-homogeneous Markov process. Assume that the initial distribution can be chosen so that $X_0 \le X_t$ a.s. for every $t \ge 0$. Then the following hold:*

- *(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) If 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) If $\mathbb{E}X_t$ 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 $X_0 \le X_t$ 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 $X_0 \le X_t$ by $X_0 \le_{\text{st}} X_t$.

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 $U_0, U_1, \ldots$ be i.i.d. with $U_i =_{\text{d}} U(0, 1)$. Since $X_0 \le X_\epsilon$ we have with $X_0' = X_0$ that $X_0' \le G_\epsilon(X_0', U_0)$, and thus

$$X_s' \equiv G_s(X_0', U_1) \le G_s(G_\epsilon(X_0', U_0), U_1) \equiv X_{s+\epsilon}'' =_{\text{d}} X_{s+\epsilon}' \equiv G_\epsilon(X_s', U_2), \quad (19)$$

implying stochastic monotonicity.

Taking $X_t' = G_{t-s-\epsilon}(X_{s+\epsilon}', U_3)$, we have that $(X_0', X_s', X_{s+\epsilon}', X_t')$ is distributed like $(X_0, X_s, X_{s+\epsilon}, X_t)$. Since h is nondecreasing in its first variable and $X_s' \le X_{s+\epsilon}''$, it follows that

$$\mathbb{E}h(X_s', X_t') \le \mathbb{E}h(X_{s+\epsilon}'', X_t'). \quad (20)$$

The (by now repetitive) fact that

$$\begin{aligned}
\mathbb{E}[h(X_{s+\epsilon}'', X_t')|U_3] &= \mathbb{E}[h(X_{s+\epsilon}'', G_{t-s-\epsilon}(X_{s+\epsilon}', U_3)|U_3] \\
&\le \mathbb{E}[h(X_{s+\epsilon}', G_{t-s-\epsilon}(X_{s+\epsilon}', U_3)|U_3] = \mathbb{E}[h(X_{s+\epsilon}', X_t')|U_3] \quad (21)
\end{aligned}$$

follows from the supermodularity of $h(x, G_{t-s-\epsilon}(y, u_3))$ in x, y. Taking expected values, together with (20), implies 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

$$h(X_s, X_{s+t}) =_d h(X_s, G_t(X_s, U_0)),$$

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, it clearly follows that $\mathbb{E}X_t$ is nondecreasing. When Condition 1 is met, 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 that it is concave (again, see [4, 24]).   □

We complete this section by noting that although, for the sake of convenience, all the results are 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 instance, for the examples discussed in Section 3, the state space is either $[0, \infty)$ or $\mathbb{Z}_+$.

2.2. Role of stochastic monotone Markov processes and supermodularity

To put our work into perspective, we conclude this section by explaining the train of thought that led us to this research. One result that sparked this line of research was that for a stationary reflected (both one- and two-sided) Lévy process X with $\mathbb{E}X_0^2 < \infty$, $R(t) = \text{Cov}(X_0, X_t)$ is nonnegative, nonincreasing, and convex [3, 10, 11]. Since in this case $\mathbb{E}X_t = \mathbb{E}X_0$, this is equivalent to $\mathbb{E}X_0 X_t$ being nonnegative, nonincreasing, and convex. Another result is that for a stationary discrete-time stochastically monotone Markov process and a nondecreasing function g with $\mathbb{E}g^2(X_0) < \infty$, $\text{Cov}(g(X_0), g(X_n))$ is nonincreasing [7]. This is equivalent to $\mathbb{E}g(X_0)g(X_n)$ being nonincreasing. We observed that xy, $g(x)g(y)$ (when g is nondecreasing) are supermodular functions of x, y. Also, we knew that the one- and two-sided Skorokhod reflection of a process of the form $x + X_t$ is monotone in x. When X is a Lévy process, the reflected process is Markovian, and from the monotonicity in x it must be stochastically monotone.

Therefore, a natural question arose: is it true that for any stationary stochastically monotone Markov process (in discrete or continuous time) and for any supermodular h we have that $\mathbb{E}h(X_0, X_t)$ is nonincreasing? If yes, then this would make the abovementioned results from [3, 7, 10, 11], each of which uses a totally different (often quite involved) approach, simple special cases. We were happy to see that the answer to this question was 'yes'.

The next step was to try to identify what condition ensures convexity of the correlation or even more general function. As seen in Section 2, the condition that ensures convexity of $\mathbb{E}g(X_0)X_t$ for a nondecreasing function g (in particular $g(x) = x$) is Condition 1. We lack a full intuitive understanding of why it is this precise condition that makes it work. As will be seen in Section 3, various natural Markovian models satisfy stochastic monotonicity and Condition 1.

Finally, we also knew that for a (one- and two-sided) reflected Lévy process X starting from the origin (and actually also for more general processes), $\mathbb{E}X_t$ is nonincreasing and concave [1, 12, 14, 15]. We asked ourselves whether this too might be seen as a special case of the more general theory we had discovered, and, as it turned out, the answer to this question was affirmative as well.

3. Examples

The purpose of this section is to give a number of motivating examples for the kinds of processes to which our theory applies. We chose these examples to illustrate the huge potential of the methodology developed in this paper: it can serve to make proofs of existing results significantly more transparent and compact (as in the examples with reflected Lévy processes), and it can serve in the derivation of entirely new structural properties (as in the examples with the Lévy dam, the state-dependent random walk, and the birth-and-death processes).

3.1. Lévy process reflected at the origin

Consider a càdlàg Lévy process $Y = \{Y_t \mid t \geq 0\}$ 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

$$X_t(x) = x + Y_t - \inf_{0 \leq s \leq t}(x + Y_s) \wedge 0 = Y_t + L_t \wedge x, \tag{22}$$

where $L_t = -\inf_{0 \leq s \leq t} Y_s \wedge 0$ 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 the following conditions:

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

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

(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

$$\int_{[0,\infty)} X_s(x) L_{\mathrm{d}s}(x) = 0, \tag{23}$$

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 (one-dimensional) reflected Brownian motion (where the reflection takes place at 0).

It turns out that when $x = 0$, $\mathbb{E}X_t$ is nondecreasing and concave in t [12, 14]. When $\mathbb{E}Y_1 < 0$, this process, which is well known to be Markovian, has a stationary distribution. If W has this stationary distribution and is independent of the process Y, then $X_t^* = X_t(W)$ is a stationary process, and it has been established [10, 11, 19] that when $\mathbb{E}W^2 < \infty$, the autocovariance $R(t) = \mathrm{Cov}(X_s^*, X_{s+t}^*)$ (or autocorrelation $R(t)/R(0)$) is nonnegative, nonincreasing, and convex in t. Where earlier proofs tended to be ad hoc (e.g. dealing only with spectrally one-sided processes, i.e., assuming that the Lévy process has jumps in one direction) and involved (e.g. requiring delicate manipulations with completely monotone functions), with the techniques developed in the present paper this property now follows virtually immediately. The remainder of this subsection illustrates this.

Let us define $p^t(x, A) = \mathbb{P}(X_t(x) \in A)$, so that $p^t(\cdot, \cdot)$ is the transition kernel of the associated Markov process. Since $X_t(x) = Y_t + L_t \wedge x$ is nondecreasing in x and $X_t(x) - x = Y_t - (x - L_t)^+$ is nonincreasing in x (for each $t > 0$), we clearly have that

$$p^t(x, (y, \infty)) = \mathbb{P}(X_t(x) > y) \tag{24}$$

is nondecreasing in x and

$$p^t(x, (x + y, \infty)) = \mathbb{P}(X_t(x) > x + y) = \mathbb{P}(Y_t - (x - L_t)^+ > y) \tag{25}$$

is nonincreasing in x. Therefore stochastic monotonicity and Condition 1 are satisfied.

Remark 2. We note that for the same reasons, a (general) random walk reflected at 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.

3.2. Lévy process with a two-sided reflection

In this subsection we argue that the structural properties discussed in the previous subsection carry over to the case of a two-sided reflection. With Y as defined in Section 3.1, 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 the following:

   (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 $t \geq 0$.

   (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

$$\int_{[0,\infty)} X_t(x) L_{\mathrm{d}t}(x) = \int_{[0,\infty)} (b - X_t(x)) U_{\mathrm{d}t}(x) = 0. \tag{26}$$

For this case it is also known that $\mathbb{E}X_t(0)$ is nondecreasing and concave [1] as well as that for the stationary version the autocovariance $R(t)$ is nonnegative, nondecreasing, and convex [3]. As in the one-sided case, we can apply our newly developed results to establish these facts in a very compact manner provided we verify that the process under consideration is a stochastically monotone Markov process that fulfills Condition 1.

Since Y is a Lévy process, we have that $X_t(x)$ is a time-homogeneous 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 representation (27) to follow).

In order to verify that it satisfies Condition 1, we recall from [17], upon re-denoting by $X_t^0(x)$ the one-sided reflected process described in Section 3.1, that

$$X_t(x) = X_t^0(x) - \sup_{0 \leq s \leq t} \left[\left(X_s^0(x) - b \right)^+ \wedge \inf_{s \leq u \leq t} X_t^0(x) \right]. \tag{27}$$

Since $X_t^0(x)$ is nondecreasing and $X_t^0(x) - x$ is nonincreasing in x (as explained in Section 3.1), it immediately follows that $X_t(x) - x$ is nonincreasing in x, which implies Condition 1.

Regarding the results for $R(t)$ that hold under stationarity, observe that in this two-sided reflected case a stationary distribution always exists and has a bounded support. Therefore, we do not need to impose any additional requirements on Y. This is in contrast to the one-sided case, where it was necessary to assume that $\mathbb{E}Y_1 < 0$ and that the stationary distribution has a finite second moment. As in the case with one-sided reflection, the findings carry over to the discrete-time counterpart: the two-sided reflected random walk.

3.3. Lévy dams with a nondecreasing release rule

In the previous two subsections, dealing with the one- and two-sided reflected Lévy process, we mentioned that existing (and also some previously nonexistent) results may be instantly concluded from the theory that we develop in this paper. The present subsection gives an illustration of our theory's potential to produce similar results for dam processes which, to the best of our knowledge, are completely new for these kinds of processes and were not known earlier. More concretely, it shows that, with the general theory that we developed, the structural results discussed above carry over to more than just reflected Lévy processes.

Let the process $J = \{J_t \mid t \geq 0\}$ 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 and left-continuous on $(0, \infty)$, with $r(0) = 0$. Consider the following *dam process*:

$$X_t(x) = x + J_t - \int_0^t r(X_s(x))\mathrm{d}s. \tag{28}$$

It is well known [6] that, under the stated assumptions, the solution to (28) is unique (pathwise) and belongs to the class of time-homogeneous Markov processes.

As before, we need to check that the process under consideration is stochastically monotone and fulfills Condition 1. For $x < y$ we have that

$$X_t(y) - X_t(x) = y - x - \int_0^t (r(X_s(y)) - r(X_s(x))\mathrm{d}s. \tag{29}$$

Denote by τ the first time (if it exists) for which the right side is zero. Because $x < y$, 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 $h \geq 0$,

$$X_{\tau+h}(x) = X_\tau(x) + J_{\tau+h} - J_\tau + \int_0^h r(X_{\tau+s}(x))\mathrm{d}s$$
$$= X_\tau(y) + J_{\tau+h} - J_\tau + \int_0^h r(X_{\tau+s}(x))\mathrm{d}s = X_{\tau+h}(y), \tag{30}$$

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

$$Z_h = z + J_{\tau+h} - J_\tau - \int_0^h r(Z_s)\mathrm{d}s. \tag{31}$$

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

$$(X_t(x) - x) - (Y_t(y) - y) = \int_0^t (r(X_s(y)) - r(X_s(x))\mathrm{d}s, \tag{32}$$

and thus, since r is assumed to be nondecreasing, we also have that $X_t(x) - x \geq X_t(y) - y$. 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. As a consequence, the structural properties on $\mathbb{E}X_t$ and $R(t)$, as we discussed above, apply to this process as well. Regarding $R(t)$, 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).

To perform a sanity check, we note that if we choose $r(x) = rx$, then the resulting process is a (generalized) shot-noise process. In this case it is well known that we can explicitly write

$$X_t(x) = xe^{-rt} + \int_{(0,t]} e^{-r(t-s)} J_{ds}. \tag{33}$$

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, and convex in t (for any distribution of X_0 having a finite second moment), and which 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 $\int_{(1,\infty)} x^2 \nu(dx) < \infty$, where ν is the associated Lévy measure.

Remark 3. We note that if we replace J by any Lévy process and $-r(\cdot)$ by $\mu(\cdot)$ where μ is nonincreasing (not necessarily negative), then whenever the process described by (28) is well defined, we have (for the same reasons as for the dam process) stochastic monotonicity and the validity of Condition 1. In particular, this is the case when $J_t = \sigma B_t$ and B is a Wiener process, resulting in a diffusion with constant diffusion coefficient and a nonincreasing drift.

3.4. Discrete-time (state-dependent) random walks

Consider a discrete-time random walk on $\mathbb{Z}_+$ with the following transition probabilities:

$$p_{ij} = \begin{cases} q_i, & i \geq 1, \ j = i-1, \\ r_i, & i \geq 0, \ j = i, \\ p_i, & i \geq 0, \ j = i+1, \\ 0 & \text{otherwise}, \end{cases} \tag{34}$$

where $q_i + r_i + p_i = 1$ for $i \geq 1$ and $r_0 + p_0 = 1$. It is a trivial exercise to check that $p(i, A) = \sum_{j \in A} p_{ij}$ is stochastically monotone if and only if $p_{i-1} \leq 1 - q_i$ for $i \geq 1$, and that it satisfies Condition 1 if and only if q_i is nondecreasing in i and p_i is nonincreasing in i. From Lemma 2 it may be concluded that these are also the respective conditions for this Markov chain to be stochastically monotone and to satisfy Condition 1. Therefore, the condition that both are satisfied is that for each $i \geq 1$

$$p_i \leq p_{i-1} \leq 1 - q_i \leq 1 - q_{i-1}, \tag{35}$$

where $q_0 \equiv 0$.

From this it immediately follows that if $r_i = 0$ for all $i \geq 1$, then the stochastic monotonicity together with Condition 1 is equivalent to the assumption that $p_i = p$ for $i \geq 0$ and $r_0 = q_i = 1 - p$ for $i \geq 1$. This is precisely the reflected process associated with $\sum_{k=1}^n Y_k$ (zero for $n = 0$) where the Y_k are i.i.d. with $\mathbb{P}(Y = 1) = 1 - \mathbb{P}(Y = -1) = p$, which is a simple special case of Remark 2.

3.5. Birth-and-death processes

Consider a right-continuous birth-and-death process on $\mathbb{Z}_+$ with birth rates λ_i, $i \geq 0$, and death rates $\mu_i \geq 1$. If the process is explosive we assume that the 'cemetery' state is $+\infty$ and that from a possible time of explosion the process remains there forever (a minimal construction, that is).

As observed in [16], this Markov process is stochastically monotone. To see this (for the sake of ease of reference), simply start two such independent processes from different initial states. By independence they do not jump at the same time (a.s.), and thus as long as they do not meet in a given state one remains strictly above the other. If and when they do meet (stopping time + strong Markov property) we let them continue together. If the higher process explodes then it remains at infinity forever and thus the bottom process never overtakes the top one. This coupling immediately implies stochastic monotonicity.

Therefore it remains to identify when Condition 1 holds. For the random walk of Subsection 3.4 the condition was that p_i is nonincreasing and q_i is nondecreasing in i. Therefore it is conceivable that the condition here should be that λ_i is nonincreasing and μ_i is nondecreasing in i. This indeed turns out to be correct. In order to see this we will first need the following lemma.

Lemma 4. *For $k = 1, 2$, let $X^k = \{X_t^k \,|\, t \geq 0\}$ be a right-continuous birth-and-death process on $\mathbb{Z}_+$ with birth rates λ_i^k for $i \geq 0$ and death rates μ_i^k for $i \geq 1$. Assume that when the process is explosive, the 'cemetery' is $+\infty$, and that the process remains there from the possible point of explosion onward. Then the following two conditions are equivalent:*

(i) $X_t^1 \leq_{\mathrm{st}} X_t^2$ for every $t \geq 0$ for any choice of initial distributions such that $X_0^1 \leq_{\mathrm{st}} X_0^2$.

(ii) $\lambda_i^1 \leq \lambda_i^2$ for every $i \geq 0$ and $\mu_i^1 \geq \mu_i^2$ for every $i \geq 1$.

Proof. We use the standard notation $\mathbb{P}_i$ to indicate the probability measure when the initial state is i. Also recall that a right-continuous birth-and-death process is *orderly*, in the sense that the probability that there are two or more jumps in the interval $[0, t]$ starting from any initial state is $o(t)$. This implies that $\lim_{t\downarrow 0} t^{-1}\mathbb{P}_i(|X_t^k - i| \geq 2) = 0$ and thus

$$\lim_{t\downarrow 0} \frac{1}{t}\mathbb{P}_i(X_t^k \geq i + 1) = \lim_{t\downarrow 0} \frac{1}{t}\mathbb{P}_i(X_t^k = i + 1) = \lambda_i \tag{36}$$

for $i \geq 0$. Similarly, $\lim_{t\downarrow 0} t^{-1}\mathbb{P}_i(X_t^k \leq i - 1) = \mu_i$ for $i \geq 1$.

To show (i) $\Rightarrow$ (ii), suppose that $X_0^1 = X_0^2 = i$ a.s. (clearly $X_0^1 \leq_{\mathrm{st}} X_0^2$). From $X_t^1 \leq_{\mathrm{st}} X_t^2$ we have that $\mathbb{P}_i(X_t^1 \geq i + 1) \leq \mathbb{P}_i(X_t^2 \geq i + 1)$ for $i \geq 0$ and $t > 0$. Dividing by t and letting $t \downarrow 0$ gives $\lambda_i^1 \leq \lambda_i^2$. Similarly, since $\mathbb{P}_i(X_t^1 \geq i) \leq \mathbb{P}_i(X_t^2 \geq i)$ we have that $\mathbb{P}_i(X_t^1 \leq i - 1) \leq \mathbb{P}_i(X_t^2 \geq i - 1)$ for $i \geq 1$ and $t > 0$. Dividing by t and letting $t \downarrow 0$ gives $\mu_i^1 \geq \mu_i^2$.

It remains to show (ii) $\Rightarrow$ (i). If we can find a coupling such that $X_t^1 \leq X_t^2$ a.s. for each $t \geq 0$, then the lemma instantly follows. It is well known that if $X_0^1 \leq_{\mathrm{st}} X_0^2$, then there are $(\tilde{X}_0^1, \tilde{X}_0^2)$ such that $\tilde{X}_0^k =_{\mathrm{d}} X_0^k$ and $\tilde{X}_0^1 \leq \tilde{X}_0^2$ a.s. Now assume that (X^1, X^2) is a right-continuous Markov process on $\{(i, j) \,|\, 0 \leq i \leq j\}$ with the following transition rates (with $\mu_0^1 = \mu_0^2 = 0$): for $i < j$ we have

$$(i, j + 1)$$

$$\lambda_j^2 \Big\uparrow$$

$$(i - 1, j) \xleftarrow{\mu_i^1} \quad (i, j) \quad \xrightarrow{\lambda_i^1} \quad (i + 1, j) \tag{37}$$

$$\mu_j^2 \Big\downarrow$$

$$(i, j - 1)$$

whereas for $i = j$

$$(i, i+1) \qquad\qquad (i+1, i+1)$$

$$\lambda_i^2 - \lambda_i^1 \uparrow \quad \lambda_i^1 \nearrow \qquad\qquad\qquad (38)$$

$$(i-1, i) \quad \overset{\mu_i^1 - \mu_i^2}{\longleftarrow} \quad (i, i)$$

$$\mu_i^2 \swarrow$$

$$(i-1, i-1)$$

All other transition rates are zero. Since the process lives on $\{(i, j) \mid 0 \leq i \leq j\}$, it is clear that $X_t^1 \leq X_t^2$ a.s. for each t. It is easy to check that for each $k = 1, 2$, X^k is a birth-and-death process with birth rates λ_i^k and death rates μ_i^k. $\qquad\square$

Note that the processes suggested in the proof run independently as long as they do not meet, and only if and when they meet does something else happen that ensures that the bottom process does not overtake the top one.

Theorem 6. *Assume that $X = \{X_t \mid t \geq 0\}$ is a right-continuous birth-and-death process with birth rates λ_i for $i \geq 0$ and death rates μ_i for $i \geq 1$. Also assume that the 'cemetery' (when explosive) is $+\infty$ and that the process remains there forever from such an epoch onward. Then X satisfies Condition 1 if and only if λ_i is nonincreasing in i and μ_i is nondecreasing in i.*

Proof. Let us define $X(i)$ to be a birth-and-death process with the above parameters which starts from i. We need to show that $X_t(i+1) - (i+1) \leq_{\mathrm{st}} X_t(i) - i$, or, equivalently, that $X_t(i+1) \leq_{\mathrm{st}} X_t(i) + 1$. Now $X_t(i) + 1$ is a birth-and-death process on $\mathbb{Z}_+$ starting from $i+1$, with arrival rates $\tilde{\lambda}_i = \lambda_{i-1}$ for $i \geq 1$, and $\tilde{\lambda}_0$ may be chosen arbitrarily in $[\lambda_0, \infty)$. The departure rates of $X_t(i) + 1$ are $\tilde{\mu}_i = \mu_{i-1}$ for $i \geq 2$ and $\tilde{\mu}_1 = 0$.

From Lemma 4 it follows that the necessary and sufficient condition that we seek is $\tilde{\lambda}_i \geq \lambda_i$ for $i \geq 0$ and $\tilde{\mu}_i \leq \mu_i$, which is equivalent to the condition that λ_i is nonincreasing and μ_i is nondecreasing (in i). $\qquad\square$

Remark 4. Note that pure birth and pure death processes are special cases of birth-and-death processes, and thus Theorem 4 and Corollary 6 hold for these processes as well. Also note that the state space does not need to be $\mathbb{Z}_+$, but can also be any finite or infinite subinterval of $\mathbb{Z}$.

4. Some concluding remarks

In this paper have focused on techniques to prove structural properties of Markov processes. We have developed a framework that succeeds in bringing various branches of the literature under a common umbrella. As indicated earlier, we strongly believe that the results presented in this paper open up the opportunity to produce compact proofs of existing results, as well as to efficiently derive new properties. We conclude this paper with a couple of remarks.

Recalling Remark 1, when there exists a stationary distribution, for all of the examples discussed in Section 3 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 the findings of Sections 3.1 and 3.2, [3, Theorems 1 and 2] (as well as the earlier [10, Theoremm 3.1] and

[11, Theorem 2.2]) are special cases of our Theorems 2 and 4. In addition, [14, Theorem 3.1] restricted to the Lévy case (and the earlier [12, Theorem 3.3]), as well as the mean (not variance) parts of [1, Theorems 4.6 and 7.5], are special cases of (iv) in our Theorem 5 (upon taking $X_0 = 0$ in [1], that is).

One could wonder whether the monotonicity of the variance, as was discovered in [1] in a reflected Lévy context, would carry over to any stochastically monotone Markov process satisfying Condition 1. However, as it turns out, this particular result from [1] essentially follows from the specific properties of reflected Lévy processes (or, in the discrete-time case, reflected random walks) and is not true in general. One elementary counterexample is the following. Let $\{N_t \mid t \geq 0\}$ be a Poisson process (starting at 0) and take $X_t = (k + N_t) \wedge m$. Then $\{X_t \mid t \geq 0\}$ is a Markov process with state space $\{i \mid i \leq m\}$, with an initial value k and an absorbing barrier m. On $k \leq m$ 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 $t \to \infty$, by bounded convergence ($k \leq X_t \leq m$) the variance vanishes as $t \to \infty$. Since the variance is strictly positive for all $0 < t < \infty$, it cannot be monotone in t.

Funding information

Competing interests

There were no competing interests to declare which arose during the preparation or publication process of this article.

References

[1] ANDERSEN, L. N. AND MANDJES, M. (2009). Structural properties of reflected Lévy processes. *Queueing Systems* **63**, 301–322.

[2] ATHREYA, K. B. AND PANTULA, S. G. (1986). Mixing properties of Harris chains and autoregressive processes. *J. Appl. Prob.* **23**, 880–892.

[3] BERKELMANS, W., CICHOCKA, A. AND MANDJES, M. (2019). The correlation function of a queue with Lévy and Markov additive input. *Stoch. Process. Appl.* **130**, 1713–1734.

[4] BLUMBERG, H. (1919). Convex functions. *Trans. Amer. Math. Soc.* **20**, 40–44.

[5] CAMBANIS, S., SIMONS, G. AND STOUT, W. (1976). Inequalities for $\mathbb{E}k(X, Y)$ when the marginals are fixed. *Z. Wahrscheinlichkeitsth.* **36**, 285–294.

[6] ÇINLAR, E. AND PINSKY, M. (1972). On dams with additive inputs and a general release rule. *J. Appl. Prob.* **9**, 422–429.

[7] DALEY, D. J. (1968). Stochastically monotone Markov chains. *Z. Wahrscheinlichkeitsth.* **10**, 305–317.

[8] DĘBICKI, K. AND MANDJES, M. (2015). *Queues and Lévy Fluctuation Theory.* Springer, New York.

[9] EMBRECHTS, P. AND HOFERT, M. (2013). A note on generalized inverses. *Math. Meth. Operat. Res.* **77**, 423–432.

[10] ES-SAGHOUANI, A. AND MANDJES, M. (2008). On the correlation structure of a Lévy-driven queue. *J. Appl. Prob.* **45**, 940–952.

[11] GLYNN, P. AND MANDJES, M. (2011). Simulation-based computation of the workload correlation function in a Lévy-driven queue. *J. Appl. Prob.* **48**, 114–130.

[12] KELLA, O. Concavity and reflected Lévy process. (1992). *J. Appl. Prob.* **29**, 209–215.

[13] KELLA, O. (2006). Reflecting thoughts. *Statist. Prob. Lett.* **76**, 1808–1811.

[14] KELLA, O. AND SVERCHKOV, M. (1994). On concavity of the mean function and stochastic ordering for reflected processes with stationary increments. *J. Appl. Prob.* **31**, 1140–1142.

[15] KELLA, O. AND WHITT, W. (1996). Stability and structural properties of stochastic storage networks. *J. Appl. Prob.* **33**, 1169–1180.

[16] KEILSON, J. AND KESTER, A. (1977). Monotone matrices and monotone Markov processes. *Stoch. Process. Appl.* **5**, 231–241.

[17] KRUK, L., LEHOCZKY, J., RAMANAN, K. AND SHREVE, S. (2007). An explicit formula for the Skorokhod map on [0, *a*]. *Ann. Prob.* **35**, 1740–1768.

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

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

[20] PUCCETTI, G. AND SCARSINI, M. (2010). Multivariate comonotonicity. *J. Multivariate. Anal.* **101**, 291–304.

[21] PUCCETTI, G. AND WANG, R. (2015). Extremal dependence concepts. *Statist. Sci.* **30**, 385–517.

[22] RÜSCHENDORF, L. (2008). On a comparison result for Markov processes. *J. Appl. Prob.* **45**, 279–286.

[23] SIEGMUND, D. (1976). The equivalence of absorbing and reflecting barrier problems for stochastically monotone Markov processes. *Ann. Prob.* **4**, 914–924.

[24] SIERPINSKI, W. (1920). Sur les fonctions convexes mesurables. *Fund. Math.* **1**, 125–129.