

## ON ASSOCIATION AND OTHER FORMS OF POSITIVE DEPENDENCE FOR FELLER PROCESSES

EDDIE TU,\* *Dickinson College*

### Abstract

We characterize various forms of positive dependence, such as association, positive supermodular association and dependence, and positive orthant dependence, for jump-Feller processes. Such jump processes can be studied through their state-space dependent Lévy measures. It is through these Lévy measures that we will provide our characterization. Finally, we present applications of these results to stochastically monotone Feller processes, including Lévy processes, the Ornstein–Uhlenbeck process, pseudo-Poisson processes, and subordinated Feller processes.

*Keywords:* Association; supermodular dependence; supermodular association; orthant dependence; Feller process; Lévy process; integro-differential operator; symbols

2010 Mathematics Subject Classification: Primary 60E07; 60E15; 60J25; 60J35; 60J75

### 1. Introduction

Multi-dimensional Feller processes have been useful for modelling the evolution of dynamical systems that are spatially inhomogeneous. These processes have been important models in finance and physics [3]. Of particular interest is the study of the dependence between the marginal processes. Some different notions of positive dependence include association (A), positive supermodular association (PSA), positive supermodular dependence (PSD), and positive orthant dependence (POD). If a process exemplifies a certain notion of positive dependence between the marginals, then one can better study the evolution of the process.

It is known that Lévy processes in  $\mathbb{R}^d$  can be characterized by their characteristic triplet  $(b, \Sigma, \nu)$ , where  $b \in \mathbb{R}^d$  is the nonrandom linear drift,  $\Sigma$  is the covariance matrix of the (continuous) Brownian motion, and  $\nu$  is the Lévy measure which characterizes the jump behaviour of the process. Feller processes have behaviour that is ‘locally Lévy’, that is, for a Feller process  $(X_t^x)_{t \geq 0}$  that starts at point  $x$  ( $X_0^x = x$  a.s.), there exists a Lévy process  $(Y_t)_{t \geq 0}$  such that, over a short time, the behaviour of  $(X_t^x)_{t \geq 0}$  can be approximated by the behaviour of  $(Y_t + x)_{t \geq 0}$  [5, p. 46]. This idea is related to the notion that, if the domain  $\mathcal{D}(\mathcal{A})$  is ‘rich’, i.e. contains  $C_c^\infty(\mathbb{R}^d)$ , the space of smooth functions with compact support, then the Feller process can be described by a characteristic triplet  $(b(x), \Sigma(x), \nu(x, dy))$ , where the function  $b: \mathbb{R}^d \rightarrow \mathbb{R}^d$  represents the nonrandom component,  $\Sigma: \mathbb{R}^d \rightarrow \mathbb{R}^{d \times d}$  represents the continuous diffusion-like behaviour, and  $x \mapsto \nu(x, dy)$  is a measurable kernel representing the jump behaviour of the process. Unlike the Lévy process, the Feller process triplet has dependence on  $x$ , the state variable of the process, representing its spatial inhomogeneity. We will characterize the different notions of positive dependence via these triplets.

Received 11 May 2018; revision received 6 April 2019.

\* Postal address: Department of Mathematics and Computer Science, Dickinson College, PO Box 1773, Carlisle, PA 17013, USA.

Association, the strongest form of positive dependence that we will examine, has been well studied for infinitely divisible distributions. Infinitely divisible random vectors  $X$  also have a characteristic triplet  $(b, \Sigma, \nu)$  by the famous Lévy–Khintchine formula, where  $b$  represents the nonrandom component,  $\Sigma$  is covariance of the Gaussian component, and  $\nu$  is the Lévy measure of the Poisson component. Pitt [21] characterized association for Gaussian distributions  $(b, Q, 0)$  under the condition that the entries  $\Sigma_{ij}$  of  $\Sigma$  are nonnegative. Resnick [22] proved that a sufficient condition for association of Poisson distributions  $(0, 0, \nu)$  is that  $\nu$  be concentrated on the positive and negative orthants  $\mathbb{R}_+^d$  and  $\mathbb{R}_-^d$ , i.e.

$$\nu((\mathbb{R}_+^d \cup \mathbb{R}_-^d)^c) = 0. \quad (1.1)$$

These results lead to the characterization of association between the marginal processes of a Lévy process, since, for a Lévy process  $Y = (Y_t)_{t \geq 0}$ ,  $Y_t$  is infinitely divisible for each  $t \geq 0$ , and the process can be described by its characteristic triplet  $(b, \Sigma, \nu)$ . Herbst and Pitt [13] extended Pitt's result in [21] to Brownian motion with covariance matrix  $\Sigma$ . For jump-Lévy processes  $Y \sim (0, 0, \nu)$ , Samorodnitsky [25] showed that condition (1.1) is a sufficient and necessary condition for the association of each  $Y_t$ . This result was also proved by Houdré *et al.* [14] using a covariance identity. Bäuerle *et al.* [2] extended Samorodnitsky's results for jump-Lévy processes to association in time, and showed that condition (1.1) is also equivalent to PSD and POD. Liggett [18] proved a necessary and sufficient condition for association of stochastically monotone Markov processes on compact state spaces based on the generator of the process. Szekli [29, Chapter 3.7] and Rüschemdorf [23, Corollary 3.1] extended this result to more general state spaces. Rüschemdorf [23, Corollary 3.4] also extended the Liggett condition for PSA of the Markov process.

In this paper, we want to characterize various forms of positive dependence for stochastically monotone Feller process. Those forms of dependence include association, weak association (WA), PSA, PSD, POD, positive upper orthant dependence (PUOD), and positive lower orthant dependence (PLOD). The association of diffusion processes, i.e.  $(b(x), \Sigma(x), 0)$ , has been characterized by Chen [7], so we will only focus on jump-Feller process  $(b(x), 0, \nu(x, dy))$ . Association of jump-Feller processes, i.e.  $(b(x), 0, \nu(x, dy))$ , was given by Wang [32, Theorem 1.4], but under certain continuity and integrability conditions on the characteristic triplet (see Remark 2.2). Here, we will relax those conditions, allowing us to consider a larger class of Feller processes. Additionally, we characterize WA, PSA, PSD, POD, PUOD, and PLOD for jump-Feller processes. Our techniques extend the ideas of Liggett, Szekli, and Rüschemdorf to the extended generator of the process, an integro-differential operator. We use ideas of the probabilistic symbol  $p(x, \xi)$  of the process developed by Jacob and Schilling [5, pp. 57–58]. Furthermore, for proving the necessary condition of association, WA, PSA, PSD, POD, PUOD, PLOD, we use the technique of small-time asymptotics of the Feller process [16], which will allow us to avoid the use of the (extended) generator and use solely the state-space dependent Lévy measure  $\nu(x, dy)$ . Finally, we provide examples of Feller processes satisfying the conditions of our main results.

In a concurrent paper of ours, entitled 'Association and other forms of positive dependence for Feller evolution systems' [31], we characterize dependence structures for Feller evolution processes (FEP), which are time-inhomogeneous Markov processes having strongly continuous Markov evolutions and Lévy-type behaviour. These FEPs are more general than the Feller processes (time-homogeneous) in this paper, but we need the results of this paper in order to characterize dependence structures of FEPs. We utilize Böttcher's transformation of time-inhomogeneous FEPs into time-homogeneous Feller processes (see [4]) and,

in a nontrivial way, use our results in this paper to prove characterizations of positive dependence for FEPs. This yields positive dependence characterizations for interesting time-inhomogeneous processes, such as additive processes. For a more comprehensive overview of time-inhomogeneous Markov processes, we recommend that the reader explore the paper by Rüschenendorf *et al.* [24], which also discusses comparison theorems of time-inhomogeneous Markov processes.

The present paper is organized in the following way. In Section 2, we give some background on the positive dependence structures, association, WA, PSA, PSD, POD, PUOD, and PLOD, along with definitions of various stochastic orderings. We also provide background on Lévy processes, Feller processes, and the different tools we use to analyse them. In Section 3, we state and prove our main results about the positive dependence structures of jump-Feller processes. Finally, in Section 4, we give a collection of interesting examples of multi-dimensional Feller processes to which we can apply these results.

## 2. Background

### 2.1. Dependence and stochastic orderings

Let  $X = (X_1, \dots, X_d)$  be a random vector in  $\mathbb{R}^d$ . We say  $X$  is *positively correlated* (PC) if  $\text{Cov}(X_i, X_j) \geq 0$  for all  $i, j \in \{1, \dots, d\}$ . This is one of the weakest forms of positive dependence, and we are interested in stronger forms of positive dependence which will be of greater use in our study of stochastic processes. *Association* is the strongest form of positive dependence that we will study.

**Definition 2.1.**  $X = (X_1, \dots, X_d)$  is *associated* (A) if we have

$$\text{Cov}(f(X), g(X)) \geq 0,$$

for all  $f, g: \mathbb{R}^d \rightarrow \mathbb{R}$  nondecreasing in each component, such that  $\text{Cov}(f(X), g(X))$  exists.

We will also study other forms of positive dependence that are weaker than association but stronger than positive correlation. We list them below.

**Definition 2.2.** A random vector  $X = (X_1, \dots, X_d)$  is *weakly associated* (WA) if, for any pair of disjoint subsets  $I, J \subseteq \{1, \dots, d\}$ , with  $|I| = k, |J| = n$ ,

$$\text{Cov}(f(X_I), g(X_J)) \geq 0,$$

where  $X_I := (X_i : i \in I)$ ,  $X_J := (X_j : j \in J)$ , for any  $f: \mathbb{R}^k \rightarrow \mathbb{R}$ ,  $g: \mathbb{R}^n \rightarrow \mathbb{R}$  nondecreasing, such that  $\text{Cov}(f(X_I), g(X_J))$  exists.

**Definition 2.3.**  $X$  is *positive supermodular associated* (PSA) if  $\text{Cov}(f(X), g(X)) \geq 0$  for all  $f, g \in \mathcal{F}_{ism} := \{h: \mathbb{R}^d \rightarrow \mathbb{R}, \text{nondecreasing, supermodular}\}$ . Here, *f supermodular* means, for all  $x, y \in \mathbb{R}^d$ ,  $f(x \wedge y) + f(x \vee y) \geq f(x) + f(y)$ , where  $x \wedge y$  is the component-wise minimum, and  $x \vee y$  is the component-wise maximum.

Now let  $\hat{X} = (\hat{X}_1, \dots, \hat{X}_d)$  be a random vector such that for all  $i$ ,  $\hat{X}_i \stackrel{d}{=} X_i$  and  $\hat{X}_i$  are mutually independent.

**Definition 2.4.**  $X$  is *positive supermodular dependent* (PSD) if, for all  $f: \mathbb{R}^d \rightarrow \mathbb{R}$  supermodular,  $\mathbb{E}f(\hat{X}) \leq \mathbb{E}f(X)$ .



FIGURE 1: Implication map of various positive dependence structures.

**Definition 2.5.**  $X$  is *positive upper orthant dependent* (PUOD) if, for all  $t_1, \dots, t_d \in \mathbb{R}$ ,

$$\mathbb{P}(X_1 > t_1, \dots, X_d > t_d) \geq \mathbb{P}(X_1 > t_1) \cdots \mathbb{P}(X_d > t_d).$$

**Definition 2.6.**  $X$  is *positive lower orthant dependent* (PLOD) if, for all  $t_1, \dots, t_d \in \mathbb{R}$ ,

$$\mathbb{P}(X_1 \leq t_1, \dots, X_d \leq t_d) \geq \mathbb{P}(X_1 \leq t_1) \cdots \mathbb{P}(X_d \leq t_d).$$

**Definition 2.7.**  $X$  is *positive orthant dependent* (POD) if  $X$  is PUOD and PLOD.

One can state another equivalent definition to PUOD (PLOD). For  $i = 1, \dots, d$ , let  $f_i: \mathbb{R} \rightarrow \mathbb{R}_+$  be nondecreasing (nonincreasing) functions. Then  $X = (X_1, \dots, X_d)$  PUOD (PLOD) if and only if

$$\mathbb{E} \left( \prod_{i=1}^d f_i(X_i) \right) \geq \prod_{i=1}^d \mathbb{E} f_i(X_i).$$

**Note.** Definition 2.1 first appeared in [11], and as Definition 2.2 in [6], Definition 2.3 in [23, p. 284], Definition 2.4 in [15], and Definitions 2.5–2.7 in [17]. Definitions 2.4–2.7 can also be stated in terms of stochastic orderings. For more on this, we refer the reader to Müller and Stoyan’s book [19, Chapter 3]. It is useful to see the relationship between these different forms of positive dependence. We state the relationships in the following proposition.

**Proposition 2.1.** *The implications in Figure 1 hold.*

*Proof.* Proofs for these implications can be found in Müller and Stoyan’s book [19, Chapter 3], and implications involving PSD can be found in [8]. □

These notions of dependence can be extended from random vectors to stochastic processes. Let  $X = (X_t)_{t \geq 0}$  be a stochastic process in  $\mathbb{R}^d$ .

**Definition 2.8.** (a) Process  $X$  is *associated in space* or *spatially associated* if, for every  $t \geq 0$ , the random vector  $X_t = (X_t^{(1)}, \dots, X_t^{(d)})$  is associated.

(b) Process  $X$  is *associated in time* or *temporally associated* if, for all  $0 \leq t_1 < \dots < t_n$ , the random vector  $(X_{t_1}, \dots, X_{t_n})$  in  $\mathbb{R}^{dn}$  is associated.

**Remark 2.1.** (i) Clearly, (b) is stronger than (a) in the above definition.

(ii) We can define other forms of positive dependence in stochastic processes if we replace ‘associated’ in Definitions 2.8 (a) and (b) with ‘WA’, ‘PSA’, ‘PSD’, ‘POD’, ‘PUOD’, ‘PLOD’.

(iii) Definition (a) is equivalent to the statement that the ‘process preserves positive correlations’, as given in [18, p. 80] and [7].

**2.2. Feller processes, extended generators, small-time asymptotics**

2.2.1. *Feller process.* Consider a time-homogeneous Markov process  $X = (X_t)_{t \geq 0}$  on the space  $(\Omega, \mathcal{G}, (\mathcal{G}_t)_{t \geq 0}, \mathbb{P}^x)_{x \in \mathbb{R}^d}$  on state space  $\mathbb{R}^d$ . Here  $(\mathcal{G}_t)_{t \geq 0}$  is the filtration, and the index ‘ $x$ ’ indicates the starting point of the process:  $\mathbb{P}^x(X_0 = x) = 1$ . We associate with a Markov process  $X$  a positivity-preserving, contraction semigroup of bounded operators  $(T_t)_{t \geq 0}$  defined by

$$T_t f(x) := \mathbb{E}^x f(X_t), \quad x \in \mathbb{R}^d,$$

where  $f \in B_b(\mathbb{R}^d)$ , the space of bounded measurable functions on  $\mathbb{R}^d$ . Let  $(C_0(\mathbb{R}^d), \|\cdot\|_\infty)$  be the Banach space of continuous functions that vanish at infinity, i.e.  $\lim_{|x| \rightarrow \infty} f(x) = 0$ , where  $\|\cdot\|_\infty$  is the sup-norm. Define  $\mathcal{F}_i := \{f: \mathbb{R}^d \rightarrow \mathbb{R}, \text{ nondecreasing in each component}\}$ . A Markov process is *stochastically monotone* if  $T_t f \in \mathcal{F}_i$  for all  $f \in \mathcal{F}_i$ . We define the *generator*  $\mathcal{A}$  of the process  $X$  to be

$$\mathcal{A}f := \lim_{t \searrow 0} \frac{T_t f - f}{t}, \tag{2.1}$$

for all  $f \in \mathcal{D}(\mathcal{A})$ , where  $\mathcal{D}(\mathcal{A})$  is the *domain of the generator* defined to be

$$\mathcal{D}(\mathcal{A}) = \{u \in C_0(\mathbb{R}^d): \text{ limit on right-hand side of (2.1) exists uniformly}\}.$$

The Markov process is a *Feller process* if the semigroup  $(T_t)_{t \geq 0}$  satisfies the following properties:

- (i)  $T_t: C_0(\mathbb{R}^d) \rightarrow C_0(\mathbb{R}^d)$ ,
- (ii)  $\lim_{t \rightarrow 0} \|T_t u - u\|_\infty = 0$ .

If, additionally, the domain of the generator contains smooth functions with compact support, i.e.  $\mathcal{D}(\mathcal{A}) \supset C_c^\infty(\mathbb{R}^d)$ , we call the process  $X$  a *rich Feller process*. It follows from Courrège’s theorem [9] that  $-\mathcal{A}$  becomes a pseudo-differential operator  $p(x, D)$  on the space of  $C_c^\infty(\mathbb{R}^d)$ :  $\mathcal{A}|_{C_c^\infty(\mathbb{R}^d)} = -p(x, D)$ , where  $p(x, D)$  is defined to be

$$\mathcal{A}f(x) = -p(x, D)f(x) = (2\pi)^{-d/2} \int_{\mathbb{R}^d} e^{i\xi \cdot x} p(x, \xi) \hat{f}(\xi) \, d\xi, \quad f \in C_c^\infty(\mathbb{R}^d). \tag{2.2}$$

The function  $-p(x, \cdot)$  is a continuous negative definite function, in the sense of Schoenberg, for all  $x \in \mathbb{R}^d$ , which yields a Lévy–Khintchine representation for each  $x$ :

$$-p(x, \xi) = -ib(x) \cdot \xi + \frac{1}{2} \xi \cdot \Sigma(x) \xi - \int_{\mathbb{R}^d \setminus \{0\}} (e^{i\xi \cdot y} - 1 - i\xi \cdot y \chi(y)) \nu(x, dy), \tag{2.3}$$

where  $\chi: \mathbb{R}^d \rightarrow \mathbb{R}$  is a *cut-off function*. In this paper, unless otherwise mentioned, we will assume  $\chi(y) = \mathbf{1}_{(0,1)}(|y|)$ . For each  $x$ ,  $(b(x), \Sigma(x), \nu(x, dy))$  is the (*Lévy characteristic triplet*), where  $b(x) \in \mathbb{R}^d$ ,  $\Sigma(x) \in \mathbb{R}^{d \times d}$  a symmetric positive definite matrix, and  $\nu(x, dy)$ , the *Lévy measure*, is a  $\sigma$ -finite measure on  $\mathbb{R}^d \setminus \{0\}$  satisfying  $\int_{\mathbb{R}^d \setminus \{0\}} (1 \wedge |y|^2) \nu(x, dy) < \infty$ . We call the function  $p(x, \xi)$  the *symbol* of the process. We also write  $X \sim (b(x), \Sigma(x), \nu(x, dy))$  to signify that  $X$  is a Feller process with that characteristic triplet.

When the symbol and the corresponding triplet are constant in  $x$ , i.e.  $p(x, \xi) = p(\xi)$  and triplet  $(b(x), \Sigma(x), \nu(x, dy)) = (b, \Sigma, \nu)$ , then process  $X$  is a *Lévy process*, i.e. a stochastically continuous Markov process with stationary and independent increments. The symbol  $p(\xi)$  is also the Lévy symbol of the process, with characteristic function  $\phi_{X_t}(\xi) = e^{tp(\xi)}$ . In the

Lévy case,  $b$  is the nonrandom linear drift,  $\Sigma$  is covariance of the Brownian motion, and  $\nu$  is a measure representing the jumps of the process.

Continuous negative definite functions  $p(x, \xi)$  which are associated with a Feller process have a form of local boundedness in the first argument. In other words, we say the symbol  $p(x, \xi)$  is locally bounded if, for all  $K \subset \mathbb{R}^d$  compact, there exists  $c_K > 0$  such that

$$\sup_{x \in K} |p(x, \xi)| \leq c_K(1 + |\xi|^2). \tag{2.4}$$

We say the *symbol is bounded* if (2.4) holds for  $K = \mathbb{R}^d$ . The local boundedness (or boundedness) of the symbol corresponds to the local boundedness (boundedness) of the characteristics  $(b(x), \Sigma(x), \nu(x, dy))$  (see [27, Lemma 2.1]).

**2.2.2. Integro-differential operator.** For any rich Feller process, the triplet  $(b(x), \Sigma(x), \nu(x, dy))$  characterizes the behaviour of the process, with  $b(x)$  representing nonrandom continuous behaviour,  $\Sigma(x)$  representing the diffusion-like continuous behaviour, and  $\nu(x, dy)$  representing the jump behaviour. To analyse the process, one of the crucial tools we will use is the extended generator. For the case of rich Feller processes, when we substitute (2.3) into the right-hand side of (2.2), by elementary Fourier analysis, we get an *integro-differential operator*  $I(p)$ ,

$$I(p)f(x) = b(x) \cdot \nabla f(x) + \frac{1}{2} \nabla \cdot \Sigma(x) \nabla f(x) + \int_{y \neq 0} (f(x+y) - f(x) - y \cdot \nabla f(x) \chi(y)) \nu(x, dy), \tag{2.5}$$

where

$$\nabla \cdot \Sigma(x) \nabla f(x) = \sum_{j,k=1}^d \Sigma_{jk}(x) \partial_j \partial_k f(x).$$

Clearly, the operator  $I(p)$  is defined on  $C_b^2(\mathbb{R}^d)$ , the space of continuous twice-differentiable bounded functions. When the symbol  $p(x, \xi)$  is bounded,  $I(p)$  is an extension of  $-p(x, D)$ ,

$$I(p)|_{C^\infty(\mathbb{R}^d)} = -p(x, D) = \mathcal{A}|_{C^\infty(\mathbb{R}^d)},$$

and an extension of generator  $\mathcal{A}$ :  $I(p)|_{\mathcal{D}(\mathcal{A})} = \mathcal{A}$ , as shown by Schilling [27, Lemma 2.3]. Our interest in this integro-differential operator  $I(p)$  comes with wanting to use the idea of Liggett’s characterization of association via the generator.

**Theorem 2.1.** (Liggett [18], p. 80.) *Let  $X = (X_t)_{t \geq 0}$  be a Feller process on state space  $E$  with generator  $(\mathcal{A}, \mathcal{D}(\mathcal{A}))$  and semigroup  $(T_t)_{t \geq 0}$ . If  $X$  is stochastically monotone, then*

$$\mathcal{A}fg \geq g\mathcal{A}f + f\mathcal{A}g \quad \text{for all } f, g \in \mathcal{F}_i \cap \mathcal{D}(\mathcal{A}) \tag{2.6}$$

*if and only if  $X_t$  is associated for all  $t \geq 0$  with respect to  $\mathbb{P}^x$  for all  $x \in E$ .*

Liggett proved this for  $E$  compact and  $\mathcal{A}$  bounded. This was extended by Szekli and Rüschemdorf to more general Polish spaces  $E$  and  $\mathcal{A}$  unbounded ([29, Chapter 3.7], [23, Corollary 3.1]). For the Feller processes we consider in the above setting, particularly those of the jump variety, the domain  $\mathcal{D}(\mathcal{A})$  is often defined to be a dense subspace of  $C_0(\mathbb{R}^d)$ , and thus  $\mathcal{D}(\mathcal{A}) \cap \mathcal{F}_i = \{f \equiv 0\}$ . Hence, in that case, inequality (2.6) would always hold. Thus, we would like to extend Theorem 2.1 to the extended generator  $I(p)$ .

2.2.3. *Small-time asymptotics.* The (extended) generator gives us a connection between the notion of association and the Lévy characteristics  $(b(x), \Sigma(x), \nu(x, dy))$  due to the representation of the integro-differential operator. Thus, to characterize association for Feller processes using the Lévy characteristics, an extension of Theorem 2.1 becomes quite useful. However, under weaker conditions on the symbol  $p(x, \xi)$ , such as local boundedness, it is useful to avoid using the generator (as we will show in Section 3) and show a more direct connection between the Lévy characteristics and the notion of association. We will establish such a connection by looking at small-time asymptotics of a Feller process. Additionally, this notion will allow us to characterize weaker forms of positive dependence under the Lévy characteristics.

The classical results of small-time asymptotics have been primarily established for Lévy processes. For a given Lévy process  $L = (L_t)_{t \geq 0}$ , it is known that, for all  $f \in C_c(\mathbb{R}^d \setminus \{0\})$ ,

$$\lim_{t \searrow 0} \frac{1}{t} \mathbb{E}^0 f(L_t) = \int_{\mathbb{R}^d \setminus \{0\}} f(y) \nu(dy) \tag{2.7}$$

(see [16, p. 2] for a reference). Thus, by the Portmanteau Theorem, (2.7) implies

$$\lim_{t \searrow 0} \frac{1}{t} \mathbb{P}^0(L_t \in A) = \nu(A)$$

for all  $A \in \mathcal{B}(\mathbb{R}^d \setminus \{0\})$  with  $0 \notin \bar{A}$  and  $\nu(\partial A) = 0$ . This result naturally extends to a general starting point  $x$ : for every  $x \in \mathbb{R}^d$ ,  $\lim_{t \searrow 0} \frac{1}{t} \mathbb{P}^x(L_t - x \in A) = \nu(A)$  by translation invariance of a Lévy process. Until recently, an analogous statement of the above for Feller processes was not known. However, Kühn and Schilling [16] proved such a statement for these processes.

**Theorem 2.2.** (Kühn, Schilling (2016) [16], Corollary 3.3.) *Let  $X = (X_t)_{t \geq 0}$  be a rich Feller process with symbol  $p(x, \xi)$  and characteristics  $(b(x), \Sigma(x), \nu(x, dy))$ . If  $f \in C_0(\mathbb{R}^d)$  and  $f|_{B(0, \delta)} = 0$  for some  $\delta > 0$ , then*

$$\lim_{t \searrow 0} \frac{1}{t} \mathbb{E}^x f(X_t - x) = \int_{\mathbb{R}^d \setminus \{0\}} f(y) \nu(x, dy).$$

Additionally, by the Portmanteau Theorem,

$$\lim_{t \searrow 0} \frac{1}{t} \mathbb{P}^x(X_t - x \in A) = \nu(x, A)$$

for all  $A \in \mathcal{B}(\mathbb{R}^d \setminus \{0\})$  such that  $0 \notin \bar{A}$  and  $\nu(x, \partial A) = 0$ .

The small-time asymptotics given by Theorem 2.2 give us a direct connection between the Lévy measure and the Feller process, avoiding the representation of the generator. Also, note that the result holds for more general, locally bounded symbols.

Our interest focuses on jump-Feller processes, i.e.  $X \sim (b(x), 0, \nu(x, dy))$ , since the association of diffusion processes  $X \sim (b(x), \Sigma(x), 0)$  was completed by Mu-Fa Chen [7]. In the following section, we will prove a sufficient and necessary condition for the jump-Feller process to be associated, WA, PSA, PSD, POD, PUOD, and PLOD in space, where the condition is

$$\nu(x, (\mathbb{R}_+^d \cup \mathbb{R}_-^d)^c) = 0 \quad \text{for all } x \in \mathbb{R}^d. \tag{2.8}$$

**Remark 2.2.** We note that Jie Ming Wang [32, Theorem 1.4] proved that spatial association is equivalent to (2.8) under certain continuity and integrability conditions (unknown to the author at the time). These assumptions include the following:

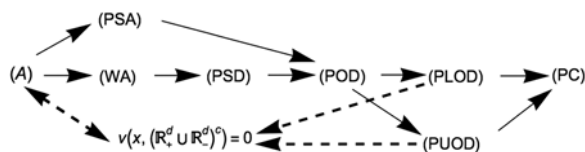


FIGURE 2: Equivalence of dependences under condition (2.8) for Feller processes.

- $b_i, \Sigma_{ij} \in C(\mathbb{R}^d)$ , for all  $i, j$ ,
- $\int h_i(z)(\nu(\cdot, dz) - \nu(\cdot, d(-z))) \in C(\mathbb{R}^d)$ , where  $h: \mathbb{R}^d \rightarrow \mathbb{R}^d$  is defined by

$$h_i(z) = \text{sgn}(z_i)(1 \wedge |z_i|),$$

- $\int_A |h(z)|^2 \nu(\cdot, dz) \in C(\mathbb{R}^d)$  for all  $A \in \mathcal{B}(\mathbb{R}^d)$ ,
- $\int g(z)\nu(\cdot, dz) \in C(\mathbb{R}^d)$  for any  $g \in C_b(\mathbb{R}^d)$  that is 0 near the origin.

We relax these conditions, and furthermore our work includes characterizations of the other dependence structures mentioned in Definitions 2.1–2.7.

### 3. Main results

Consider a rich Feller process  $X = (X_t)_{t \geq 0}$  on the space  $(\Omega, \mathcal{G}, (\mathcal{G}_t)_{t \geq 0}, \mathbb{P}^x)_{x \in \mathbb{R}^d}$  with Lévy characteristics  $(b(x), 0, \nu(x, dy))$ . If we assume that  $X$  is stochastically monotone, then condition (2.8) is a necessary and sufficient condition for the association, WA, PSA, PSD, POD, PUOD, and PLOD in the space of the process  $X$ . These equivalences can be illustrated in the implication map in Figure 2. The dashed arrows are the implications we will prove.

To show these equivalences, we first give a proof that, under stochastic monotonicity, condition (2.8) is equivalent to association in space. We show this in Section 3.1. Then, in Section 3.2, we show that PUOD in space (and, similarly, PLOD) implies condition (2.8).

#### 3.1. Association is equivalent to condition (2.8)

**Theorem 3.1.** *Let  $X = (X_t)_{t \geq 0}$  be a rich Feller process with stochastically monotone transition semigroup  $(T_t)_{t \geq 0}$ , a generator  $(\mathcal{A}, \mathcal{D}(\mathcal{A}))$ , bounded symbol  $p(x, \xi)$ , and  $(b(x), 0, \nu(x, dy))$ . Then  $X_t$  is associated for all  $t \geq 0$  if and only if (2.8),  $\nu(x, (\mathbb{R}_+^d \cup \mathbb{R}_-^d)^c) = 0$ , is satisfied.*

We prove this by first showing that association of  $X_t$  is equivalent to a Liggett-type inequality for the extended generator, the statement of which is in the following theorem.

**Theorem 3.2.** *Let  $X = (X_t)_{t \geq 0}$  be a rich Feller processes with stochastically monotone transition semigroup  $(T_t)_{t \geq 0}$ , a generator  $(\mathcal{A}, \mathcal{D}(\mathcal{A}))$ , bounded symbol  $p(x, \xi)$ , and an (extended) integro-differential operator  $I(p)$ . Assume  $x \mapsto p(x, 0)$  is continuous. Then*

$$I(p)fg \geq fI(p)g + gI(p)f \quad \text{for all } f, g \in C_b^2(\mathbb{R}^d) \cap \mathcal{F}_t \tag{3.1}$$

if and only if, for all  $t \geq 0$ ,

$$T_t fg \geq T_t f \cdot T_t g \quad \text{for all } f, g \in C_b(\mathbb{R}^d) \cap \mathcal{F}_t. \tag{3.2}$$

Inequality (3.2) in Theorem 3.2 is another way to formulate that  $X_t$  is associated for all  $t \geq 0$ . Since (3.2) means, for all  $x \in \mathbb{R}^d$ ,  $\mathbb{E}^x f(X_t)g(X_t) \geq \mathbb{E}^x f(X_t)\mathbb{E}^x g(X_t)$ , then  $X_t$  is associated



with respect to  $\mathbb{P}^x$ . Inequality (3.1) intuitively means that the process moves either up or down, which, in multidimensional Euclidean space, means that if the process is currently at point  $x$ , then it can only move to another point  $y$  if  $y \geq x$  or  $y \leq x$  component-wise.

Note that in Theorem 3.2 we are using the extended generator  $I(p)$ . In previous statements of Liggett's characterization, the generator  $\mathcal{A}$  is used, but we need to use  $I(p)$  for the reasons given in the comments after Theorem 2.1. Hence, it is necessary to show the Liggett-type inequality as a characterization of association for rich Feller processes. Such an extension has not been seen by the author of this paper. We first need the following lemmas to prove Theorem 2.1. We will often assume Setting 3.1 throughout this section.

**Setting 3.1.** Let  $X = (X_t)_{t \geq 0}$  be a rich Feller process, semigroup  $(T_t)_{t \geq 0}$ , generator  $(\mathcal{A}, \mathcal{D}(\mathcal{A}))$ , symbol  $p(x, \xi)$ , (extended) integro-differential operator  $I(p)$ , and characteristics

$$(a(x), b(x), \Sigma(x), \nu(x, dy)),$$

where  $b, \Sigma, \nu$  are the same as before, except that we have an additional characteristic  $a: \mathbb{R}^d \rightarrow \mathbb{R}_+$  which represents the 'killing rate'.

**Remark 3.1.** With the additional characteristic  $a(x)$ , function  $-p(x, \xi)$  would look like  $a(x) +$  the right-hand side of equation (2.3). Also,  $I(p)f(x)$  would look like  $-a(x)f(x) +$  the right-hand side of equation (2.5). Unless stated otherwise, we will assume that  $a(x) \equiv 0$ . For more on the case when  $a(x) \neq 0$ , see the paper by Schnurr [28], which discusses such processes satisfying  $a(x) \neq 0$  and their connection to the symbol.

**Lemma 3.1.** Assume Setting 3.1 and that  $p(x, \xi)$  is bounded. Then  $I(p)$  generates the semigroup  $(T_t)_{t \geq 0}$  locally uniformly, i.e.

$$I(p)f = \lim_{t \searrow 0} \frac{1}{t}(T_t f - f), \quad f \in C_b^2(\mathbb{R}^d),$$

where the convergence is locally uniform.

For a detailed proof, see the appendix.

**Lemma 3.2.** Assume Setting 3.1 and the symbol  $p(x, \xi)$  is bounded. For all  $f \in C_b^2(\mathbb{R}^d)$ ,

$$\frac{d}{dt} T_t f = I(p)T_t f = T_t I(p)f,$$

where the derivative is defined based on locally uniform convergence.

For a detailed proof, see the appendix. Finally, we can extend Liggett's solution to a Cauchy problem [18, Theorem 2.15, p. 19] to integro-differential operators that generate a semigroup locally uniformly.

**Lemma 3.3.** (Cauchy problem.) Let  $(\mathcal{A}, \mathcal{D}(\mathcal{A}))$  be a (rich) Feller generator of a semigroup  $(T_t)_{t \geq 0}$  with bounded Lévy characteristics and symbol  $p(x, \xi)$ . Let  $I(p)$  be the extended generator on  $C_b^2(\mathbb{R}^d)$ . Suppose  $F, G: [0, \infty) \rightarrow C_b(\mathbb{R}^d)$  such that

- (a)  $F(t) \in \mathcal{D}(I(p))$  for all  $t \geq 0$ ,
- (b)  $G(t)$  is continuous on  $[0, \infty)$  (locally uniformly),
- (c)  $F'(t) = I(p)F(t) + G(t)$  for all  $t \geq 0$ .

Then  $F(t) = T_t f(0) + \int_0^t T_{t-s} G(s) \, ds$ .

For a detailed proof, see the appendix. We are now ready to prove the main theorems of this section.

*Proof of Theorem 3.2.* ( $\Leftarrow$ ) Assume  $T_t f g \geq T_t f T_t g$  for all  $f, g \in C_b^2(\mathbb{R}^d) \cap \mathcal{F}_i$ . This implies

$$T_t f g - f g \geq T_t f T_t g - f g = T_t f T_t g - f g + g T_t f - g T_t f = T_t f [T_t g - g] + g [T_t f - f].$$

Hence, for all  $t > 0$ ,

$$\frac{1}{t}(T_t f g - f g) \geq T_t f \frac{T_t g - g}{t} + g \frac{T_t f - f}{t}.$$

Therefore,

$$\begin{aligned} I(p) f g &= \lim_{t \searrow 0} \frac{1}{t}(T_t f g - f g) \geq \lim_{t \searrow 0} \left( T_t f \frac{T_t g - g}{t} + g \frac{T_t f - f}{t} \right) \\ &= \left( \lim_{t \searrow 0} T_t f \right) \left( \lim_{t \searrow 0} \frac{T_t g - g}{t} \right) + g \left( \lim_{t \searrow 0} \frac{T_t f - f}{t} \right) \\ &= f I(p) g + g I(p) f, \end{aligned}$$

where the convergence is locally uniform.

( $\Rightarrow$ ) Assume  $I(p) f g \geq f I(p) g + g I(p) f$  for all  $f, g \in C_b^2(\mathbb{R}^d) \cap \mathcal{F}_i$ . By monotonicity,  $T_t f, T_t g \in C_b^2(\mathbb{R}^d) \cap \mathcal{F}_i$ , which implies

$$I(p)(T_t f)(T_t g) \geq T_t f [I(p) T_t g] + T_t g [I(p) T_t f]. \tag{3.3}$$

Define  $F(t) := T_t f g - T_t f T_t g$ . Then, by Lemma 3.2, we have

$$\begin{aligned} F'(t) &= I(p) T_t f g - (T_t f [I(p) T_t g] + T_t g [I(p) T_t f]) \geq I(p) T_t f g - (I(p) T_t f T_t g) \\ &= I(p)(T_t f g - T_t f T_t g) \\ &= I(p) F(t), \end{aligned}$$

where the inequality comes from (3.3). Define  $G(t) := F'(t) - I(p) F(t) \geq 0$ . Then, by Lemma 3.3, the solution to the Cauchy problem  $F'(t) = G(t) + I(p) F(t)$  is given by

$$F(t) = T_t f(0) + \int_0^t T_{t-s} G(s) \, ds = \int_0^t T_{t-s} G(s) \, ds$$

since  $F(0) = 0$ . Since  $G(s) \geq 0$  for all  $s$ , and  $T_{t-s}$  is a positivity-preserving linear operator,  $F(t) \geq 0$  for all  $t \geq 0$ . Thus,  $T_t f g \geq T_t f \cdot T_t g$  for all  $f, g \in C_b^2(\mathbb{R}^d) \cap \mathcal{F}_i$ . This inequality also holds for all  $f, g \in C_b(\mathbb{R}^d) \cap \mathcal{F}_i$ , since we can approximate nondecreasing, continuous, bounded functions by nondecreasing smooth, bounded functions, and then use a dominated convergence argument.  $\square$

**Remark 3.2.** For the necessary condition, we did not need stochastic monotonicity.

*Proof of Theorem 3.1.* ( $\Leftarrow$ ) Fix  $x \in \mathbb{R}^d$ . Assume  $\nu(x, (\mathbb{R}_+^d \cup \mathbb{R}_-^d)^c) = 0$ . Then, for all  $f, g \in C_b^2(\mathbb{R}^d) \cap \mathcal{F}_i$ ,

$$\begin{aligned} & I(p)fg(x) - g(x)I(p)f(x) - f(x)I(p)g(x) \\ &= b(x) \cdot \nabla fg(x) + \int_{y \neq 0} (f(x+y)g(x+y) - f(x)g(x) - y \cdot \nabla fg(x)\mathbf{1}_{(0,1)}(|y|))\nu(x, dy) \\ &\quad - b(x) \cdot g(x)\nabla f(x) - \int_{y \neq 0} (f(x+y)g(x) - f(x)g(x) - y \cdot g(x)\nabla f(x)\mathbf{1}_{(0,1)}(|y|))\nu(x, dy) \\ &\quad - b(x) \cdot f(x)\nabla g(x) - \int_{y \neq 0} (f(x)g(x+y) - f(x)g(x) - y \cdot f(x)\nabla g(x)\mathbf{1}_{(0,1)}(|y|))\nu(x, dy) \\ &= \int_{y \neq 0} (f(x+y)g(x+y) - f(x+y)g(x) - f(x)g(x+y) + f(x)g(x))\nu(x, dy) \\ &= \int_{y \neq 0} (f(x+y) - f(x))(g(x+y) - g(x))\nu(x, dy) \\ &= \int_{\mathbb{R}_+^d} (f(x+y) - f(x))(g(x+y) - g(x))\nu(x, dy) \\ &\quad + \int_{\mathbb{R}_-^d} (f(x+y) - f(x))(g(x+y) - g(x))\nu(x, dy) \\ &\geq 0, \end{aligned}$$

where the drift terms and the cut-off term in the integrand vanish because

$$\nabla fg(x) = f(x)\nabla g(x) + g(x)\nabla f(x).$$

Additionally, we get positivity at the end there because, for all  $y \in \mathbb{R}_+^d$ ,  $f(x+y) - f(x) \geq 0$ , and  $g(x+y) - g(x) \geq 0$ , so  $(f(x+y) - f(x))(g(x+y) - g(x)) \geq 0$  on  $\mathbb{R}_+^d$ . A similar result holds on  $\mathbb{R}_-^d$ . By Theorem 3.2, this implies  $T_t fg(x) \geq T_t f(x)T_t g(x)$ , where  $f, g \in C_b^2(\mathbb{R}^d) \cap \mathcal{F}_i$ . Now, to obtain association of  $X_t$ , this inequality needs to hold for all  $f, g \in C_b(\mathbb{R}^d) \cap \mathcal{F}_i$ . However, we can use an approximation of a function  $f \in C_b(\mathbb{R}^d) \cap \mathcal{F}_i$  by  $f_n \in C_b^\infty(\mathbb{R}^d) \cap \mathcal{F}_i$ , which gives us the desired result.

( $\Rightarrow$ ) Assume that  $X_t$  is associated for all  $t \geq 0$ . This means  $T_t fg(x) \geq T_t f(x)T_t g(x)$  for all  $x \in \mathbb{R}^d$ , for all  $f, g \in C_b(\mathbb{R}^d) \cap \mathcal{F}_i$ . So this inequality of course holds for  $f, g \in C_b^2(\mathbb{R}^d) \cap \mathcal{F}_i$ , which yields  $I(p)fg \geq gI(p)f + fI(p)g$  for such  $f, g$  by Theorem 3.2. This implies, by a similar calculation in the ( $\Leftarrow$ ) direction, that

$$\int_{y \neq 0} (f(x+y) - f(x))(g(x+y) - g(x))\nu(x, dy) \geq 0.$$

For simplicity, assume  $d = 2$ , but note that we can easily generalize this result to higher dimensions using correction functions. Fix  $x = (x_1, x_2) \in \mathbb{R}^2$ . Assume for contradiction that Resnick’s condition is not satisfied. Without loss of generality, let us say  $\nu(x, (0, \infty) \times (-\infty, 0)) > 0$ . By continuity of measure, there exists  $a > 0$  such that  $\nu(x, (a, \infty) \times (-\infty, a)) > 0$ . Let  $\varepsilon \in (0, 1)$ , and define  $f, g \in C_b^\infty(\mathbb{R}^2) \cap \mathcal{F}_i$  by

$$f(y_1, y_2) = \begin{cases} 0 & \text{if } y_1 \leq x_1 + \varepsilon a, \\ 1 & \text{if } y_1 \geq x_1 + a, \end{cases} \quad g(y_1, y_2) = \begin{cases} 0 & \text{if } y_2 \geq x_2 - \varepsilon a, \\ -1 & \text{if } y_2 \leq x_2 - a. \end{cases}$$

This implies  $f(x) = g(x) = 0$ . Hence,

$$\begin{aligned} 0 &\leq \int_{y \neq 0} (f(x+y) - f(x))(g(x+y) - g(x))\nu(x, dy) \\ &= \int_{y \neq 0} f(x+y)g(x+y)\nu(x, dy) \\ &= \int_{(a, \infty) \times (-\infty, -a)} f(x+y)g(x+y)\nu(x, dy) + \int_{(a, \infty) \times [-a, -\varepsilon a]} f(x+y)g(x+y)\nu(x, dy) \\ &\quad + \int_{[\varepsilon a, a] \times (-\infty, -a)} f(x+y)g(x+y)\nu(x, dy) + \int_{[\varepsilon a, a] \times [-a, -\varepsilon a]} f(x+y)g(x+y)\nu(x, dy) \\ &= -\nu(x, (a, \infty) \times (-\infty, -a)) - \int_{(a, \infty) \times [-a, -\varepsilon a]} g(x+y)\nu(x, dy) \\ &\quad + \int_{[\varepsilon a, a] \times (-\infty, -a)} f(x+y)\nu(x, dy) + \int_{[\varepsilon a, a] \times [-a, -\varepsilon a]} f(x+y)g(x+y)\nu(x, dy) \\ &\leq -\nu(x, (a, \infty) \times (-\infty, -a)), \end{aligned}$$

implying  $\nu(x, (a, \infty) \times (-\infty, -a)) \leq 0$ . Hence,  $\nu(x, (a, \infty) \times (-\infty, -a)) = 0$ , a contradiction. □

### 3.2. PUOD implies condition (2.8)

**Lemma 3.4.** *If  $Y = (Y_1, \dots, Y_d)$  is PUOD, then  $(Y_{k_1}, \dots, Y_{k_n})$  is PUOD for all multi-indices  $\{k_j\}_{j=1}^n \subset \{1, \dots, d\}$ .*

*Proof.* If  $Y$  PUOD, then we know  $\mathbb{E}(\prod_{i=1}^d f_i(Y_i)) \geq \prod_{i=1}^d \mathbb{E} f_i(Y_i)$ , where  $f_i: \mathbb{R} \rightarrow \mathbb{R}_+$  non-decreasing. So for all  $i \in \{1, \dots, d\} \setminus \{k_j\}_{j=1}^n$ , set  $f_i = \mathbf{1}_{\mathbb{R}}$ . Then the above inequality becomes

$$\mathbb{E}\left(\prod_{j=1}^n f_j(Y_{k_j})\right) \geq \prod_{j=1}^n \mathbb{E} f_j(Y_{k_j}).$$

Thus,  $(Y_{k_1}, \dots, Y_{k_n})$  is PUOD. □

**Theorem 3.3.** *Let  $X = (X_t)_{t \geq 0}$  be a rich Feller process with symbol  $p(x, \xi)$  and triplet*

$$(b(x), 0, \nu(x, dy)).$$

*If  $X_t$  is PUOD for each  $t \geq 0$ , then condition (2.8) is satisfied, i.e.*

$$\nu(x, (\mathbb{R}_+^d \cup \mathbb{R}_-^d)^c) = 0.$$

*Proof.* Assume  $X_t$  is PUOD (with respect to  $\mathbb{P}^x$ ) for each  $t \geq 0$ . Fix  $x = (x_1, \dots, x_d) \in \mathbb{R}^d$ . Since  $X_t$  is PUOD, then  $X_t - x$  is PUOD for all  $t \geq 0$ . Assume for contradiction that  $\nu$  is not concentrated on  $\mathbb{R}_+^d \cup \mathbb{R}_-^d$ . Without loss of generality, say  $\nu(x, (0, \infty)^{d-1} \times (-\infty, 0)) > 0$ . By continuity of measure, there exists  $a > 0$  such that

$$\nu(x, (a, \infty)^{d-1} \times (-\infty, -a)) > 0$$

and

$$\nu(x, \partial[(a, \infty)^{d-1} \times (-\infty, -a)]) = \nu(x, \partial[(a, \infty) \times \mathbb{R}^{d-1}]) = 0.$$

Then, by Theorem 2.2,

$$\lim_{t \rightarrow 0} \frac{1}{t} \mathbb{P}^x(X_t - x \in (a, \infty)^{d-1} \times (-\infty, -a)) = v(x, (a, \infty)^{d-1} \times (-\infty, -a)).$$

Hence,

$$\begin{aligned} & 0 < v(x, (a, \infty)^{d-1} \times (-\infty, -a)) \\ &= \lim_{t \rightarrow 0} \frac{1}{t} \mathbb{P}^x(X_t - x \in (a, \infty)^{d-1} \times (-\infty, -a)) \\ &= \lim_{t \rightarrow 0} \frac{1}{t} \mathbb{P}^x(X_t^{(1)} - x_1 > a, \dots, X_t^{(d-1)} - x_{d-1} > a, X_t^{(d)} - x_d < -a) \\ &\leq \lim_{t \rightarrow 0} \frac{1}{t} \mathbb{P}^x(X_t^{(1)} - x_1 > a, \dots, X_t^{(d-1)} - x_{d-1} > a, X_t^{(d)} - x_d \leq -a) \\ &= \lim_{t \rightarrow 0} \frac{1}{t} \mathbb{P}^x(\{X_t^{(1)} - x_1 > a\} \setminus [\{X_t^{(1)} - x_1 > a\} \cap \{X_t^{(2)} - x_2 > a, \dots, X_t^{(d)} - x_d \leq -a\}^c]) \\ &= \lim_{t \rightarrow 0} \frac{1}{t} [\mathbb{P}^x(X_t^{(1)} - x_1 > a) \\ &\quad - \mathbb{P}^x(\{X_t^{(1)} - x_1 > a\} \cap \{X_t^{(2)} - x_2 > a, \dots, X_t^{(d)} - x_d \leq -a\}^c)] \\ &= \lim_{t \rightarrow 0} \frac{1}{t} [\mathbb{P}^x(X_t^{(1)} - x_1 > a) \\ &\quad - \mathbb{P}^x(\{X_t^{(1)} - x_1 > a\} \cap [\{X_t^{(2)} - x_2 \leq a\} \cup \dots \cup \{X_t^{(d)} - x_d > -a\}])] \\ &= \lim_{t \rightarrow 0} \frac{1}{t} [\mathbb{P}^x(X_t^{(1)} - x_1 > a) \\ &\quad - \mathbb{P}^x(\{X_t^{(1)} - x_1 > a, X_t^{(2)} - x_2 \leq a\} \cup \dots \cup \{X_t^{(1)} - x_1 > a, X_t^{(d)} - x_d > -a\}])] \\ &\leq \lim_{t \rightarrow 0} \frac{1}{t} [\mathbb{P}^x(X_t^{(1)} - x_1 > a) - \mathbb{P}^x(X_t^{(1)} - x_1 > a, X_t^{(d)} - x_d > -a)] \\ &\leq \lim_{t \rightarrow 0} \frac{1}{t} [\mathbb{P}^x(X_t^{(1)} - x_1 > a) - \mathbb{P}^x(X_t^{(1)} - x_1 > a) \mathbb{P}^x(X_t^{(d)} - x_d > -a)] \\ &= \lim_{t \rightarrow 0} \frac{1}{t} [\mathbb{P}^x(X_t^{(1)} - x_1 > a)(1 - \mathbb{P}^x(X_t^{(d)} - x_d > -a))] \\ &= \lim_{t \rightarrow 0} \frac{1}{t} \mathbb{P}^x(X_t^{(1)} - x_1 > a) \mathbb{P}^x(X_t^{(d)} - x_d \leq -a) \\ &= \left[ \lim_{t \rightarrow 0} \frac{1}{t} \mathbb{P}^x(X_t^{(1)} - x_1 > a) \right] \left[ \lim_{t \rightarrow 0} \mathbb{P}^x(X_t^{(d)} - x_d \leq -a) \right] \\ &= v(x, (a, \infty) \times \mathbb{R}^{d-1}) \mathbb{P}^x(X_0^{(d)} - x_d \leq -a) \\ &= 0. \end{aligned}$$

We obtain lines 4, 9 by set containment, line 5 by the fact that  $A \cap B = A \setminus (A \cap B^c)$ , line 10 by Lemma 3.4, and line 14 by Theorem 2.2. This contradiction gives us the desired result.  $\square$

**Remark 3.3.** (i) We could also have shown that PLOD implies condition (2.8) using techniques similar to those above.

(ii) Symbol  $p(x, \xi)$  in the above theorem need not be bounded, only locally bounded.

**Corollary 3.1.** For stochastically monotone jump-Feller processes, i.e.  $X \sim (b(x), 0, \nu(x, dy))$  with bounded symbols  $p(x, \xi)$ , then condition (2.8),  $\nu(x, (\mathbb{R}_+^d \cup \mathbb{R}_-^d)^c) = 0$ , is equivalent to  $X$  being associated, WA, PSA, PSD, POD, PUOD, and PLOD in space.

*Proof.* This holds by Theorems 3.1 and 3.3. □

### 3.3. Association in time

Our results can also be used to study the temporal association of Feller processes. We first examine the case of Lévy processes, a subclass of Feller processes with constant characteristic triplet  $(b, Q, \nu)$ . For Lévy processes, spatial association is equivalent to temporal association.

**Theorem 3.4.** Let  $X = (X_t)_{t \geq 0}$  be a stochastic process in  $\mathbb{R}^d$  with independent and stationary increments, i.e.  $X_t - X_s \perp\!\!\!\perp X_s - X_r$ , for all  $0 \leq r < s < t$ , and  $X_t - X_s \stackrel{d}{=} X_{t-s}$  for all  $0 \leq s < t$ . Then  $X$  is associated in time if and only if  $X$  is associated in space.

*Proof.* The forward direction is trivial by definition. We only need to prove the backward direction. Assume  $X_t$  is associated in  $\mathbb{R}^d$  for every  $t \geq 0$ . Choose  $0 \leq t_1 < \dots < t_n$ . Then

$$\begin{aligned} (X_{t_1}, \dots, X_{t_n}) &= (X_{t_1}, X_{t_1} + (X_{t_2} - X_{t_1}), \dots, X_{t_1} + (X_{t_2} - X_{t_1}) + \dots + (X_{t_n} - X_{t_{n-1}})) \\ &= (X_{t_1}, \dots, X_{t_1}) + (0, X_{t_2} - X_{t_1}, \dots, X_{t_2} - X_{t_1}) + \dots + (0, \dots, 0, X_{t_n} - X_{t_{n-1}}). \end{aligned}$$

Now observe that by stationary increments,  $X_{t_{k+1}} - X_{t_k} \stackrel{d}{=} X_{t_{k+1}-t_k}$  and  $X_{t_{k+1}-t_k}$  is associated, which makes  $X_{t_{k+1}} - X_{t_k}$  associated (association is preserved under equality in distribution), for all  $k \in \{1, \dots, n-1\}$ . Further, if  $\hat{X}$  is associated in  $\mathbb{R}^d$ , then each block  $(0, \dots, 0, \hat{X}, \dots, \hat{X})$  is associated in  $\mathbb{R}^{dn}$ , where there are  $k$  0 vectors and  $(n-k)$   $\hat{X}$  vectors. Therefore, each block  $(0, \dots, 0, X_{t_{k+1}} - X_{t_k}, \dots, X_{t_{k+1}} - X_{t_k})$  is associated, for each  $k \in \{1, \dots, n-1\}$ . By independent increments, each block is independent. Therefore, since the sum of independent random vectors, each of which is associated, is associated, then  $(X_{t_1}, \dots, X_{t_n})$  is associated. □

**Corollary 3.2.** Any Lévy process  $X$  that is associated in space is also associated in time. Additionally, if  $X$  has triplet  $(b, 0, \nu)$ , then  $X$  is associated in time if and only if  $\nu((\mathbb{R}_+^d \cup \mathbb{R}_-^d)^c) = 0$ .

*Proof.* Any Lévy process has independent and stationary increments, thus the result holds by Theorem 3.4. □

We would also like to consider conditions for temporal association of general Feller processes. Early work on this has been done by Harris [12, Corollary 1.2] and Liggett [18, p. 82] for Feller processes with a countable state space. This can be extended to more general state spaces, as given in the following theorem.

**Theorem 3.5.** Let  $X = (X_t)_{t \geq 0}$  be a time-homogeneous, stochastically monotone Feller process on  $\mathbb{R}^d$ . If  $X$  is spatially associated, and  $X_0 \sim \mu$ , where  $\mu$  satisfies

$$\int fg \, d\mu \geq \int f \, d\mu \int g \, d\mu, \quad f, g \in B_b(\mathbb{R}^d) \cap \mathcal{F}_i,$$

then  $X$  is temporally associated.

The proof is similar to Liggett’s proof found in [18, p. 82]. For details on the proof, we refer the reader to the author’s dissertation [30, p. 59]. Theorem 3.5 yields the following corollary about jump-Feller processes.

**Corollary 3.3.** *Let  $X = (X_t)_{t \geq 0}$  be a stochastically monotone Feller process with characteristics  $(b(x), 0, \nu(x, dy))$ . Assume  $X_0 \sim \mu \in \mathcal{M}_a$ . Then  $\nu(x, (\mathbb{R}_+^d \cup \mathbb{R}_-^d)^c) = 0$  if and only if  $X$  is associated in time.*

*Proof.* The proof follows from Theorems 3.1 and 3.5. □

### 4. Examples

We give a collection of interesting Feller processes that satisfy stochastic monotonicity.

#### 4.1. Lévy processes

Any Lévy process satisfies stochastic monotonicity. Let  $(T_t)_{t \geq 0}$  be a semigroup of a Lévy process. Then, for  $f \in \mathcal{F}_i$ , we have

$$T_t f(x) = \mathbb{E}^x f(X_t) = \mathbb{E}^0 f(X_t + x).$$

Thus monotonicity of function  $f$  and of the expectation  $\mathbb{E}^0$  gives us  $T_t f \in \mathcal{F}_i$ .

Let  $X = (X_t)_{t \geq 0}$  be a jump-Lévy process whose Lévy characteristics look like  $(b, 0, \nu)$ , where there is no state-space dependence. Then  $\nu((\mathbb{R}_+^d \cup \mathbb{R}_-^d)^c) = 0$  is equivalent to  $X_t$  being associated, WA, PSA, PSD, POD, PUOD, and PLOD since all Lévy processes are stochastically monotone. This was proved in [2] for association, PSD, and POD, but not for the other dependence structures. Furthermore, the technique in [2] to prove condition (1.1) is equivalent to PSD and POD, and required Lévy copulas. Our method of short-time asymptotics avoids Lévy copulas altogether, and solely uses the Lévy measure. Additionally, condition (1.1) is equivalent to temporal association of  $X$ , by Corollary 3.3.

#### 4.2. Ornstein–Uhlenbeck process

An Ornstein–Uhlenbeck (OU) process  $X = (X_t)_{t \geq 0}$  in  $\mathbb{R}^d$  is the solution to the general Langevin equation

$$\begin{aligned} dX_t &= -\lambda X_t dt + dL_t, \\ X_0 &= x \text{ a.s.,} \end{aligned}$$

where  $\lambda > 0$ ,  $L = (L_t)_{t \geq 0} \sim (b_L, \Sigma_L, \nu_L)$  is a Lévy process in  $\mathbb{R}^d$ , and  $x \in \mathbb{R}^d$ . Then the OU process looks like

$$X_t = e^{-\lambda t} x + \int_0^t e^{-\lambda(t-s)} dL_s$$

The semigroup  $(T_t)_{t \geq 0}$  of this process is called the Mehler semigroup and is given by

$$T_t f(x) = \int_{\mathbb{R}^d} f(e^{\lambda t} x + y) \mu_t(dy), \quad L_t \sim \mu_t.$$

**Claim 4.1.** The OU process is stochastically monotone.

*Proof.* Let  $f \in B_b(\mathbb{R}^d)$  be an increasing function. Assume  $x < y$ , and fix some  $t \geq 0$ . Then  $e^{-\lambda t} x < e^{-\lambda t} y$ . This implies  $f(e^{-\lambda t} x + z) \leq f(e^{-\lambda t} y + z)$  for all  $z \in \mathbb{R}^d$ . Hence,

$$T_t f(x) = \int_{\mathbb{R}^d} f(e^{-\lambda t} x + z) \mu_t(dz) \leq \int_{\mathbb{R}^d} f(e^{-\lambda t} y + z) \mu_t(dz) = T_t f(y).$$

Thus,  $T_t f$  is an increasing function on  $\mathbb{R}^d$ . □

Process  $X$  has a characteristic triplet,  $(b_L - \lambda x, \Sigma_L, \nu_L)$  [1]. Thus, the characterization of positive dependence (association, WA, PSA, PSD, POD, PUOD, PLOD) is equivalent to  $\nu_L((\mathbb{R}_+^d \cup \mathbb{R}_-^d)^c) = 0$  when  $\Sigma = 0$ .

### 4.3. Feller’s pseudo-Poisson process

Here we construct a stochastically monotone pseudo-Poisson process. Let  $S = (S(n))_{n \in \mathbb{N}}$  be a homogeneous Markov process taking values in  $\mathbb{R}^d$ . Let  $(q^{(n)})_{n \in \mathbb{N}}$  define the  $n$ -step transition probabilities,

$$q^{(n)}(x, B) = \mathbb{P}(S(n) \in B | S(0) = x),$$

for all  $B \in \mathcal{B}(\mathbb{R}^d)$ . Let  $Q$  be the transition operator of  $S$ , defined by

$$(Qf)(x) = \int_{\mathbb{R}^d} f(y)q(x, dy)$$

for all  $f \in B_b(\mathbb{R}^d)$ ,  $x \in \mathbb{R}^d$ . Note that  $Q^n f(x) = \int_{\mathbb{R}^d} f(y)q^{(n)}(x, dy)$ . Let  $N = (N_t)_{t \geq 0}$  be a Poisson process with rate  $\lambda$  that is independent of  $S$ . Define  $X = (X_t)_{t \geq 0}$  by subordination:

$$X_t := S(N_t) \quad \text{for all } t \geq 0.$$

Process  $X$ , called *Feller’s pseudo-Poisson process*, is a Feller process.

The semigroup  $(T_t)_{t \geq 0}$  and generator  $\mathcal{A}$  of  $X$  are given by

$$T_t f(x) = e^{t[\lambda(Q-I)]} f(x) = e^{-\lambda t} \sum_{n=0}^{\infty} \frac{(\lambda t)^n}{n!} Q^n f(x),$$

$$\mathcal{A}f(x) = \lambda(Q - I)f(x) = \int_{\mathbb{R}^d} [f(y) - f(x)]\lambda q(x, dy).$$

**Claim 4.2.** If  $S$  is a stochastically monotone Markov process, then  $X$  is stochastically monotone.

*Proof.* We will show that if  $f \in \mathcal{F}_i$  then  $T_t f \in \mathcal{F}_i$ . Observe that, since  $S$  is stochastically monotone, we deduce that  $q(x, B)$  is a monotone function in  $x$  for any monotone set  $B \in \mathcal{B}(\mathbb{R}^d)$ . Additionally, for  $f \in B_b(\mathbb{R}^d) \cap \mathcal{F}_i$ ,  $Qf(x) = \int_{\mathbb{R}^d} f(y)q(x, dy)$  is a monotone function. We show, by induction, that for all  $n$

$$G_n := e^{-\lambda t} \frac{(\lambda t)^n}{n!} Q^n f$$

is a nondecreasing function.

For the base case,  $n = 0$ ,  $G_0(x) = e^{-\lambda t} f(x)$  is nondecreasing. For  $n = 1$ ,

$$G_1(x) = e^{-\lambda t} \lambda t Qf(x) = e^{-\lambda t} \lambda t \int_{\mathbb{R}^d} f(z)q(x, dz)$$

is nondecreasing.

*Induction hypothesis.* Assume that

$$G_n(x) = e^{-\lambda t} \frac{(\lambda t)^n}{n!} Q^n f(x) = e^{-\lambda t} \frac{(\lambda t)^n}{n!} \int_{\mathbb{R}^d} f(z)q^{(n)}(x, dz)$$

is a nondecreasing function.



*Inductive step.* We have

$$\begin{aligned} G_{n+1}(x) &= e^{-\lambda t} \frac{(\lambda t)^{n+1}}{(n+1)!} Q^{n+1} f(x) \\ &= e^{-\lambda t} \frac{(\lambda t)^{n+1}}{(n+1)!} \int_{\mathbb{R}^d} f(z) q^{(n+1)}(x, dz) \\ &= e^{-\lambda t} \frac{(\lambda t)^{n+1}}{(n+1)!} \int_{\mathbb{R}^d} \left( \int_{\mathbb{R}^d} f(z) q^{(n)}(y, dz) \right) q(x, dy) \\ &=: e^{-\lambda t} \frac{(\lambda t)^{n+1}}{(n+1)!} \int_{\mathbb{R}^d} H(y) q(x, dy), \end{aligned}$$

where  $H(y) = \int_{\mathbb{R}^d} f(z) q^{(n)}(y, dz)$  is a nondecreasing function in  $y$  by induction hypothesis, and we have used the Chapman–Kolmogorov equations. Thus, by the base case, the integral  $\int_{\mathbb{R}^d} H(y) q(x, dy)$  is nondecreasing in  $x$ . Hence we deduce that  $G_n$  is a nondecreasing function for all  $n$ . Hence,  $Tf$  is nondecreasing, giving us our desired result.  $\square$

Now, to find the characteristic triplet  $(b(x), \Sigma(x), \nu(x, dy))$ , we consider the generator:

$$\begin{aligned} \mathcal{A}f(x) &= \int_{\mathbb{R}^d} (f(z) - f(x)) \lambda q(x, dz) \\ &= \int_{\mathbb{R}^d} (f(x+z) - f(x)) \lambda q(x, dz+x) \\ &= \int_{\mathbb{R}^d} (f(x+z) - f(x)) \lambda \hat{q}(x, dz), \quad \text{where } \hat{q}(x, B) := q(x, B+x), \\ &= \int_{\mathbb{R}^d} (f(x+z) - f(x) - \nabla f(x) \cdot z \chi(z)) \lambda \hat{q}(x, dz) + \int_{\mathbb{R}^d} \nabla f(x) \cdot z \chi(z) \lambda \hat{q}(x, dz) \\ &= \int_{\mathbb{R}^d} (f(x+z) - f(x) - \nabla f(x) \cdot z \chi(z)) \lambda \hat{q}(x, dz) + \nabla f(x) \cdot \left( \int_{\mathbb{R}^d} z \chi(z) \lambda \hat{q}(x, dz) \right). \end{aligned}$$

Thus, the Lévy triplet will be  $(b(x), \Sigma(x), \nu(x, dy))$ , where

$$b(x) = \int_{\mathbb{R}^d} z \chi(z) \lambda \hat{q}(x, dz), \quad \Sigma(x) = 0, \quad \nu(x, A) = \lambda \hat{q}(x, A) = \lambda q(x, A+x).$$

**4.4. Bochner’s subordination of a Feller process**

Consider a continuous-time Feller process  $Y = (Y(t))_{t \geq 0}$  with semigroup  $(T_t)_{t \geq 0}$  and generator  $(\mathcal{A}, \mathcal{D}(\mathcal{A}))$ . Let  $N = N_t$  be a subordinator independent of  $Y$  with Lévy characteristics  $(b, \lambda)$ , i.e. it has Lévy symbol  $\eta(u) = ibu + \int_0^\infty (e^{iuy} - 1) \lambda(dy)$ , where  $\mathbb{E} e^{iuN_t} = e^{t\eta(u)}$ . Additionally, we can attain a Laplace transform of the subordinator,  $\mathbb{E} e^{-uN_t} = e^{-t\psi(u)}$ , where

$$\psi(u) := -\eta(iu) = bu + \int_0^\infty (1 - e^{-uy}) \lambda(dy).$$

Function  $\psi$  is called the *Laplace symbol* or *Bernstein function* of the subordinator. The following is a theorem of Phillips.

**Theorem 4.1.** (Phillips [20].) *Let  $X = (X_t)_{t \geq 0}$  be given by the prescription  $X_t = Y(N_t)$ . Then  $X$  is a Feller process with semigroup  $(T_t^X)_{t \geq 0}$  and generator  $(\mathcal{A}^X, \mathcal{D}(\mathcal{A}^X))$ , given by*

$$T_t^X f = \int_0^\infty (T_{sf}) \mu_{N_t}(ds), \quad \mathcal{A}^X f = b\mathcal{A}f + \int_0^\infty (T_{sf} - f) \lambda(ds).$$

**Claim 4.3.** If  $Y$  is a stochastically monotone Feller process with semigroup  $(T_t)_{t \geq 0}$ , i.e.  $T_t f \in \mathcal{F}_i$  for  $f \in \mathcal{F}_i$ , and  $N = (N_t)_{t \geq 0}$  is a subordinator, then  $X = (X_t)_{t \geq 0}$  given by  $X_t = Y(N_t)$  is a stochastically monotone Feller process.

*Proof.* We already know that  $X$  is Feller with semigroup  $(T_t^X)_{t \geq 0}$ . So choose  $f \in \mathcal{F}_i \cap C_b(\mathbb{R}^d)$ . Then  $T_s f \in \mathcal{F}_i \cap C_b(\mathbb{R}^d)$  for all  $s \geq 0$ . Choose  $x < y$ . Then  $T_s f(x) \leq T_s f(y)$  for all  $s \geq 0$ . Hence,

$$T_t^X f(x) = \int_0^\infty (T_s f)(x) \mu_{N_t}(ds) \leq \int_0^\infty (T_s f)(y) \mu_{N_t}(ds) = T_t^X f(y).$$

Thus,  $T_t^X f \in \mathcal{F}_i$ . □

Let  $Y$  have symbol  $p(x, \xi)$ . Then  $X = Y(N)$  is a Feller process with symbol  $p_X(x, \xi)$  that is given by

$$p_X(x, \xi) = \psi(p(x, \xi)) + \text{lower-order perturbation.}$$

This ‘perturbation’ is ‘measured in a suitable scale of anisotropic function spaces’ [5, p. 104]. Particularly interesting examples are when  $N$  is an  $\alpha$ -stable subordinator, inverse Gaussian subordinator, and Gamma subordinator, and  $Y$  is a diffusion process  $Y \sim (b(x), Q(x), 0)$ .

**Example 4.1.** Let  $Y$  be a stochastically monotone diffusion process in  $\mathbb{R}^d$ . This means that  $Y$  has Lévy characteristics  $(b(x), Q(x), 0)$ . Mu-Fa Chen and Feng-yu Wang [7] proved that such a process is stochastically monotone if and only if  $q_{ij}(x)$  only depends on  $x_i$  and  $x_j$ , and  $b_i(x) \leq b_i(y)$  whenever  $x \leq y$  with  $x_i = y_i$ . The generator of  $Y$  is given by

$$\mathcal{A}^Y f(x) = b(x) \cdot \nabla f(x) + \frac{1}{2} \nabla \cdot Q(x) \nabla f(x).$$

Let  $N$  be an  $\alpha$ -stable subordinator with Lévy characteristics  $(0, \lambda)$ , where

$$\lambda(dy) = \frac{\alpha}{\Gamma(1-\alpha)} \frac{1}{y^{1+\alpha}} dy.$$

The generator  $\mathcal{A}^X$  of process  $X = Y(N)$  looks like

$$\mathcal{A}^X f(x) = \int_0^\infty (T_s f(x) - f(x)) \lambda(ds) = \int_0^\infty (T_s f(x) - f(x)) \frac{\alpha}{\Gamma(1-\alpha)} \frac{1}{s^{1+\alpha}} ds.$$

### Appendix A. Proofs of some lemmas from Section 3

Throughout this appendix, we assume Setting 3.1.

**Lemma A.1.** (Schilling [26], Theorem 4.3.) *Assume  $p(x, \xi)$  is bounded. If  $x \mapsto p(x, 0)$  is continuous, then  $(T_t)_{t \geq 0}$  extends to a  $C_b$ -Feller semigroup, i.e. it satisfies*

- (a)  $T_t : C_b(\mathbb{R}^d) \rightarrow C_b(\mathbb{R}^d)$ ,
- (b)  $\lim_{h \searrow 0} \|T_{t+hu} - T_t u\|_{\infty, K} = 0$  for all  $K \subset \mathbb{R}^d$  compact,  $u \in C_b(\mathbb{R}^d)$ ,  $t \geq 0$ , where  $\|u\|_{\infty, K} := \sup_{y \in K} |u(y)|$ , i.e. locally uniformly continuous.

*Proof.* For a proof, see [26, p. 247]. □

*Proof of Lemma 3.1.* The process

$$M_t^f := f(X_t) - f(X_0) - \int_0^t I(p) f(X_{s-}) ds$$

is, for every  $f \in C_b^2(\mathbb{R}^d)$ , a martingale with respect to  $\mathbb{P}^x$ , for all  $x$  (see [27, Lemma 3.2, p. 579]). This implies

$$\begin{aligned} 0 &= \mathbb{E}^x f(X_t) - \mathbb{E}^x f(X_0) - \mathbb{E}^x \int_0^t I(p)f(X_{s-}) \, ds \\ &= T_t f(x) - f(x) - \int_0^t \mathbb{E}^x I(p)f(X_{s-}) \, ds \\ &= T_t f(x) - f(x) - \int_0^t T_s I(p)f(x) \, ds \end{aligned}$$

for every  $x \in \mathbb{R}^d, t \geq 0$ . Note that we can switch integrals in line 2 because  $I(p)f \in C_b(\mathbb{R}^d)$  by Remark 4.5(ii) in [26]. This implies

$$\frac{1}{t}(T_t f - f) = \frac{1}{t} \int_0^t T_s I(p)f \, ds.$$

We argue that when taking the limit as  $t \searrow 0$ , the right-hand side converges locally uniformly to  $I(p)f$ . Note that since  $I(p)f \in C_b(\mathbb{R}^d)$ , then  $(T_s I(p)f)\mathbf{1}_K$  is continuous in  $s$  for every compact set  $K$  by the  $C_b$ -Feller property, i.e.

$$\|(T_{s+h} I(p)f)\mathbf{1}_K - (T_s I(p)f)\mathbf{1}_K\|_\infty = \sup_{x \in K} |T_{s+h} I(p)f(x) - T_s I(p)f(x)| \rightarrow 0.$$

So, the function  $T_{(\cdot)} I(p)f\mathbf{1}_K$  is the integrand of a Bochner-type integral that is continuous in  $s$  and integrable on any closed interval  $[a, b]$ . Therefore, by the Fundamental Theorem of Calculus for Bochner integrals [10, pp. 21–22],

$$\lim_{t \searrow 0} \frac{1}{t}(T_t f - f)\mathbf{1}_K = \lim_{t \searrow 0} \frac{1}{t} \int_0^t (T_s I(p)f)\mathbf{1}_K \, ds = (I(p)f)\mathbf{1}_K$$

for all  $K \subset \mathbb{R}^d$  compact. Hence,  $I(p)f = \lim_{t \searrow 0} \frac{1}{t}(T_t f - f)$ , where convergence is locally uniform. □

*Proof of Lemma 3.2.* By Lemma A.1, our semigroup  $(T_t)_{t \geq 0}$  satisfies the  $C_b$ -Feller property. Choose  $f \in C_b^2(\mathbb{R}^d)$ . Observe that for all  $x \in \mathbb{R}^d$ ,

$$\begin{aligned} T_{t+h} f(x) - T_t f(x) &= T_t(T_h f(x) - f(x)) \\ &= T_t \int_0^h T_s I(p)f(x) \, ds \\ &= \mathbb{E}^x \int_0^h T_s I(p)f(X_t) \, ds \\ &= \int_0^h \mathbb{E}^x T_s I(p)f(X_t) \, ds, \quad \text{by Fubini's theorem,} \\ &= \int_0^h T_t T_s I(p)f(x) \, ds \\ &= \int_0^h T_s T_t I(p)f(x) \, ds. \end{aligned}$$

Thus,

$$\lim_{h \rightarrow 0} \frac{1}{h} (T_{t+h}f - T_t f) = \lim_{h \rightarrow 0} \frac{1}{h} \int_0^h T_s T_t I(p) f \, ds = T_t I(p) f$$

because  $T_t I(p) f \in C_b(\mathbb{R}^d)$  by the  $C_b$ -Feller property, thus making  $T_s T_t I(p) f \mathbf{1}_K$  continuous in  $s$  for every compact  $K$ . Once again, by the Fundamental Theorem of Calculus for Bochner integrals (see [10, pp. 21–22]), we get the convergence shown above.

Finally, we want to show  $I(p)T_t f = T_t I(p) f$ . We choose  $(\phi_n)_{n \in \mathbb{N}} \subset C_c^\infty(\mathbb{R}^d)$  such that  $\mathbf{1}_{B(0,n)} \leq \phi_n \leq 1$  for all  $n$ . Hence,  $f\phi_n \in C_c^2(\mathbb{R}^d) \subset \mathcal{D}(\mathcal{A})$ , the domain of generator  $\mathcal{A}$ , and we have  $I(p)T_t f \phi_n = T_t I(p) f \phi_n$ . By an approximation argument, we get our desired result.  $\square$

*Proof of Lemma 3.3.* Observe that all limits (and corresponding derivatives) we take here are with respect to locally uniform convergence. Note that by the assumption that  $x \mapsto p(x, 0)$  is continuous, our semigroup  $(T_t)_{t \geq 0}$  satisfies the  $C_b$ -Feller property by Lemma A.1. Also, by Lemma 3.1, we have  $\lim_{t \searrow 0} \frac{1}{t} (T_t u - u) = I(p)u$  for all  $u \in C_b^2(\mathbb{R}^d)$ . Observe that we will define the derivative  $F'(s)$  by

$$F'(s) = \lim_{h \rightarrow 0} \frac{F(s+h) - F(s)}{h},$$

where the limit is under locally uniform convergence. Also, our statement of (b) is different from that of Liggett.

Liggett’s: if  $t_n \rightarrow t$ , then  $\|G(t_n) - G(t)\|_\infty \rightarrow 0$  as  $n \rightarrow \infty$ .

Ours: if  $t_n \rightarrow t$ , then  $\|G(t_n) - G(t)\|_{\infty, K} \rightarrow 0$  as  $n \rightarrow \infty$  for all  $K$  compact.

Although Liggett’s assumption would be sufficient, we do not need something that strong in our setting, and our  $G$  will satisfy locally uniform continuity. Choosing any compact set  $K \subset \mathbb{R}^d$ , we have

$$\begin{aligned} & \frac{T_{t-s-h}F(s+h) - T_{t-s}F(s)}{h} \cdot \mathbf{1}_K \\ &= \frac{T_{t-s-h}F(s+h)}{h} \cdot \mathbf{1}_K - \frac{T_{t-s}F(s)}{h} \cdot \mathbf{1}_K \\ & \quad + [T_{t-s-h} - T_{t-s}]F'(s) \cdot \mathbf{1}_K - [T_{t-s-h} - T_{t-s}]F'(s) \cdot \mathbf{1}_K \\ & \quad + \frac{T_{t-s-h}F(s)}{h} \cdot \mathbf{1}_K - \frac{T_{t-s-h}F(s)}{h} \cdot \mathbf{1}_K \\ & \quad + \frac{T_{t-s}F(s+h)}{h} \cdot \mathbf{1}_K - \frac{T_{t-s}F(s+h)}{h} \cdot \mathbf{1}_K \\ & \quad + \frac{T_{t-s}F(s)}{h} \cdot \mathbf{1}_K - \frac{T_{t-s}F(s)}{h} \cdot \mathbf{1}_K \\ &= (1) + (2) + (3) + (4) + (5) + (6) + (7) + (8) + (9) + (10) \\ &= [(2) + (7)] + [(5) + (10)] + [(3)] + [(4) + (1) + (9) + (8) + (6)] \\ &= T_{t-s} \left[ \frac{F(s+h) - F(s)}{h} \right] \cdot \mathbf{1}_K + \left[ \frac{T_{t-s-h} - T_{t-s}}{h} \right] F(s) \cdot \mathbf{1}_K \\ & \quad + [T_{t-s-h} - T_{t-s}]F'(s) \cdot \mathbf{1}_K + [T_{t-s-h} - T_{t-s}] \left[ \frac{F(s+h) - F(s)}{h} - F'(s) \right] \cdot \mathbf{1}_K \\ &= (I) + (II) + (III) + (IV). \end{aligned}$$

Now we consider the limits as  $h$  goes to 0 for each of these four terms.

(I)

$$\begin{aligned} \lim_{h \searrow 0} T_{t-s} \left[ \frac{F(s+h) - F(s)}{h} \right] \cdot \mathbf{1}_K &= T_{t-s} \lim_{h \searrow 0} \left[ \frac{F(s+h) - F(s)}{h} \right] \cdot \mathbf{1}_K \\ &= T_{t-s} F'(s) \cdot \mathbf{1}_K \end{aligned}$$

because  $T_{t-s}$  is a bounded operator, which means it is a continuous operator.

(II) Let  $u = t - s$ . Then  $s = t - u$  and  $ds = -du$ . For a function  $f \in C_b(\mathbb{R}^d)$ ,

$$\begin{aligned} \lim_{h \searrow 0} \left[ \frac{T_{t-s-h} - T_{t-s}}{h} \right] f \cdot \mathbf{1}_K &= \frac{d}{ds} T_{t-s} f \cdot \mathbf{1}_K \\ &= -\frac{d}{du} T_u f \cdot \mathbf{1}_K \\ &= -I(p) T_u f \cdot \mathbf{1}_K \\ &= -I(p) T_{t-s} f \cdot \mathbf{1}_K. \end{aligned}$$

Therefore,

$$\lim_{h \searrow 0} \left[ \frac{T_{t-s-h} - T_{t-s}}{h} \right] F(s) \cdot \mathbf{1}_K = -I(p) T_{t-s} F(s) \cdot \mathbf{1}_K = -T_{t-s} I(p) F(s) \cdot \mathbf{1}_K.$$

(III) By the  $C_b$ -Feller property, since  $F'(s) \in C_b(\mathbb{R}^d)$ ,  $\lim_{h \searrow 0} [T_{t-s-h} - T_{t-s}] F'(s) \cdot \mathbf{1}_K = 0$  uniformly.

(IV) Observe that  $T_{t-s-h}$  and  $T_{t-s}$  are both contractions. Hence,

$$\begin{aligned} &\left\| [T_{t-s-h} - T_{t-s}] \left[ \frac{F(s+h) - F(s)}{h} - F'(s) \right] \right\|_{\infty, K} \\ &\leq \left\| T_{t-s-h} - T_{t-s} \right\| \cdot \left\| \left[ \frac{F(s+h) - F(s)}{h} - F'(s) \right] \right\|_{\infty, K} \\ &\leq 2 \left\| \left[ \frac{F(s+h) - F(s)}{h} - F'(s) \right] \right\|_{\infty, K} \rightarrow 0 \end{aligned}$$

as  $h \rightarrow 0$ . Thus, we have, for  $0 < s < t$ ,

$$\begin{aligned} \frac{d}{ds} T_{t-s} F(s) \cdot \mathbf{1}_K &= \lim_{h \searrow 0} \frac{T_{t-(s+h)} F(s+h) - T_{t-s} F(s)}{h} \cdot \mathbf{1}_K \\ &= \lim_{h \searrow 0} [(I) + (II) + (III) + (IV)] \\ &= T_{t-s} F'(s) \cdot \mathbf{1}_K - T_{t-s} I(p) F(s) \cdot \mathbf{1}_K \\ &= T_{t-s} [F'(s) - I(p) F(s)] \cdot \mathbf{1}_K \\ &\stackrel{(c)}{=} T_{t-s} G(s) \cdot \mathbf{1}_K. \end{aligned}$$

The right-hand side is a continuous function of  $s$  because  $G$  is a continuous function of  $s$  and the semigroup is uniformly continuous on  $K$  by the  $C_b$ -Feller property. We justify this as follows.

Let  $\varepsilon > 0$ . Then there exists  $N$  such that  $\|G(s_n) - G(s)\|_{\infty, K} < \varepsilon/2$  for all  $n \geq N$ . Also, there exists  $N'$  such that  $\|T_{t-s_n}G(s_N) - T_{t-s}G(s_N)\|_{\infty, K} = \|(T_{t-s_n} - T_{t-s})G(s_N)\|_{\infty, K} < \varepsilon/2$  for all  $n \geq N'$ , since the semigroup operator is uniformly continuous on compact sets. Let  $M = \max(N, N')$ . Then,

$$\begin{aligned} & \|T_{t-s_M}G(s_M) - T_{t-s}G(s)\|_{\infty, K} \\ &= \|T_{t-s_M}G(s_M) - T_{t-s}G(s_M) + T_{t-s}G(s_M) - T_{t-s}G(s)\|_{\infty, K} \\ &\leq \|T_{t-s_M}G(s_M) - T_{t-s}G(s_M)\|_{\infty, K} + \|T_{t-s}G(s_M) - T_{t-s}G(s)\|_{\infty, K} \\ &\leq \|T_{t-s_M}G(s_M) - T_{t-s}G(s_M)\|_{\infty, K} + \|G(s_M) - G(s)\|_{\infty, K} \\ &< \frac{\varepsilon}{2} + \frac{\varepsilon}{2} \\ &= \varepsilon. \end{aligned}$$

Therefore we can integrate these functions with respect to  $s$  from 0 to  $t$ . By the Fundamental Theorem of Calculus for Bochner integrals (see [10, pp. 21–22]), we obtain

$$\begin{aligned} \int_0^t T_{t-s}G(s) \, ds \cdot \mathbf{1}_K &= \int_0^t \frac{d}{ds} T_{t-s}F(s) \, ds \cdot \mathbf{1}_K \\ &= (T_{t-t}F(t) - T_t f(0))\mathbf{1}_K \\ &= (F(t) - T_t f(0))\mathbf{1}_K. \end{aligned}$$

Since  $K$  was an arbitrary compact set, we have established our desired result:  $F(t) = T_t f(0) + \int_0^t T_{t-s}G(s) \, ds$ . □

### Acknowledgement

The author would like to thank Dr Jan Rosinski for his helpful advice and guidance regarding the ideas of this paper.

### References

- [1] APPLEBAUM, D. (2007). On the infinitesimal generators of Ornstein–Uhlenbeck processes with jumps in Hilbert space. *Potential Anal.* **26**, 79–100.
- [2] BAUERLE, N., BLATTER, A. AND MÜLLER, A. (2008). Dependence properties and comparison results for Lévy processes. *Math. Meth. Oper. Res.* **67**, 161–186.
- [3] BÖTTCHER, B. (2010). Feller processes: The next generation in modeling: Brownian motion, Lévy processes and beyond. *PLoS ONE* **5** (12), 1–8.
- [4] BÖTTCHER, B. (2014). Feller evolution systems: Generators and approximation. *Stoch. Dyn.* **14** (3), 1–15.
- [5] BÖTTCHER, B., SCHILLING, R. AND WANG, J. (2013). *Lévy Matters III*. Springer.
- [6] BURTON, R., DABROWSKI, A. AND DEHLING, H. (1986). An invariance principle for weakly associated random vectors. *Stoch. Proc. Appl.* **23**, 301–306.
- [7] CHEN, M. AND WANG, F. (1993). On order-preservation and positive correlations for multidimensional diffusion processes. *Prob. Theory Rel. Fields* **95**, 421–428.
- [8] CHRISTOFIDES, T. AND VAGGELATOU, E. (2004). A connection between supermodular ordering and positive/negative association. *J. Multivar. Anal.* **88**, 138–151.
- [9] COURRÈGE, P. (1965). Sur la forme intégrable-différentielle des opérateurs de  $C_k^\infty$  dans  $C$  satisfaisant au principe du maximum. *Sém. Théorie du Potentiel* **2**, 1–38.
- [10] DYNKIN, E. (1965). *Markov Processes*. Springer.
- [11] ESARY, J., PROSCHAN, F. AND WALKUP, D. (1967). Association of random variables, with applications. *Ann. Math. Statist.* **38** (5), 1466–1474.
- [12] HARRIS, T. (1977). A correlation inequality for Markov processes in partially ordered state spaces. *Ann. Prob.* **5** (3), 451–454.

- [13] HERBST, I. AND PITT, L. (1991). Diffusion equation techniques in stochastic monotonicity and positive correlations. *Prob. Theory Rel. Fields* **87**, 275–312.
- [14] HOUDRÉ, C., PÉREZ-ABREU, V. AND SURGAILIS, D. (1998). Interpolation, correlation identities, and inequalities for infinitely divisible variables. *Fourier Anal. Appl.* **4**, 651–668.
- [15] HU, T. (2000). Negatively superadditive dependence of random variables with applications. *Chinese J. Appl. Probab. Statist.* **16** (2), 133–144.
- [16] KÜHN, F. AND SCHILLING, R. (2019). On the domain of fractional Laplacians and related generators of Feller processes. *J. Funct. Anal.* **276** (8), 2397–2439.
- [17] LEHMANN, E. (1966). Some concepts of dependence. *Ann. Math. Statist.* **37** (5), 1137–1153.
- [18] LIGGETT, T. (1985). *Interacting Particle Systems*. Springer.
- [19] MÜLLER, A. AND STOYAN, D. (2002). *Comparison Methods for Stochastic Models and Risks*. Wiley.
- [20] PHILLIPS, R. (1952). On the generation of semigroups of linear operators. *Pacific J. Math.* **2**, 343–369.
- [21] PITT, L. (1982). Positive correlated normal random variables are associated. *Ann. Prob.* **10**, 496–499.
- [22] RESNICK, S. (1988). Association and extreme value distributions. *Austral. J. Statist.* **30A**, 261–271.
- [23] RÜSCHENDORF, L. (2008). On a comparison result for Markov processes. *J. Appl. Prob.* **45**, 279–286.
- [24] RÜSCHENDORF, L., SCHNURR, A. AND WOLF, V. (2016). Comparison of time-inhomogeneous Markov processes. *Adv. Appl. Prob.* **48** (4), 1015–1044.
- [25] SAMORODNITSKY, G. (1995). Association of infinitely divisible random vectors. *Stoch. Proc. Appl.* **55**, 45–55.
- [26] SCHILLING, R. (1998). Conservativeness and extensions of Feller semigroups. *Positivity* **2**, 239–256.
- [27] SCHILLING, R. (1998). Growth and Hölder conditions for the sample paths of Feller processes. *Prob. Theory Rel. Fields* **112**, 565–611.
- [28] SCHNURR, A. (2017). The fourth characteristic of a semimartingale. Available at [arXiv:1709.06756v3](https://arxiv.org/abs/1709.06756v3).
- [29] SZEKLI, R. (1995). *Stochastic Ordering and Dependence in Applied Probability*. Springer.
- [30] TU, E. (2017). Dependence structures in Lévy-type Markov processes. Doctoral thesis, University of Tennessee, Knoxville.
- [31] TU, E. (2018). Association and other forms of positive dependence for Feller evolution systems. Available at [arXiv:1805.03080](https://arxiv.org/abs/1805.03080).
- [32] WANG, J. M. (2009). Stochastic comparison and preservation of positive correlations for Lévy-type processes. *Acta Math. Sinica* **25** (5), 741–758.