Markov-modulated Affine Processes

Kevin Kurt[1]   Rüdiger Frey[2]

October 2, 2021

Abstract

We study Markov-modulated affine processes (abbreviated MMAPs), a class of Markov processes that are created from affine processes by allowing some of their coefficients to be a function of an exogenous Markov process. MMAPs allow for richer models in various applications. At the same time MMAPs largely preserve the tractability of standard affine processes, as their characteristic function has a computationally convenient functional form. Our setup is a substantial generalization of earlier work, since we consider the case where the generator of the exogenous process X is an unbounded operator (as is the case for diffusions or jump processes with infinite activity). We prove existence of MMAPs via a martingale problem approach, we derive the formula for their characteristic function and we study various mathematical properties of MMAPs. The paper closes with a discussion of several applications of MMAPs in finance.

1. Introduction

A standard affine process Y in the sense of [17] can be characterized as a strong Markov process whose generator is given in terms of so-called *admissible* parameters that form affine functions of the state y of the process, cf. [37]. Markov-modulated affine processes or MMAPs are a natural extension where the constant (non y-dependent) part of these affine functions depends on some exogenous Markov process X. A simple example is provided by a CIR process with Markov modulated mean reversion level. Formally, we introduce Markov-modulated affine processes as a class of Markov processes, which we define in terms of the generator $\mathfrak{L}$: we assume that $\mathfrak{L}$ is a linear operator of the form

$$\mathfrak{L}f(x,y) = \mathfrak{L}^X f(x,y) + \mathfrak{L}^{Y|X}(x)f(x,y), \quad f \in \mathrm{Dom}(\mathfrak{L}) \subset C_\infty(\mathcal{D}^X \times \mathcal{D}^Y).$$

Here, $\mathfrak{L}^X$ (acting on $x \mapsto f(x,y)$) is the generator of a Feller process X with state space $\mathcal{D}^X \subset \mathbb{R}^d$, and for fixed x the operator $\mathfrak{L}^{Y|X}(x)$ (acting on $y \mapsto f(x,y)$) is the generator of an affine process Y with state space $\mathcal{D}^Y \subset \mathbb{R}^n$. Thus, conditional on the path of X, the process Y can be regarded as a time-inhomogeneous affine process.

Markov-modulation adds flexibility for (financial) modelling to the class of affine processes. At the same time MMAPs largely preserve the tractability of the latter class, as the characteristic

---

[1]Vienna University of Economics and Business (WU Wien), Institute for Statistics and Mathematics, kevin.kurt@wu.ac.at

[2]Vienna University of Economics and Business (WU Wien), Institute for Statistics and Mathematics, ruediger.frey@wu.ac.at

function of their marginal distributions can be computed via an extension of the Riccati equations for standard affine processes. More precisely, for an MMAP it holds that

$$(1.1) \qquad \mathbb{E}_z \left[e^{\langle u, Y_t \rangle} \right] = \varphi(t,x,u) e^{\langle \psi(t,u), y \rangle}, \quad z = (x,y) \in \mathcal{D}^X \times \mathcal{D}^Y.$$

The function ψ is characterized by the same system of generalized Riccati equations as in the standard affine case, whereas φ is determined by means of a Cauchy problem involving ψ and the generator $\mathfrak{L}^X$ of X.

Simple *regime-switching* MMAPs where X is a finite state Markov chain have been used previously in finance. For instance, [25] consider bond pricing in affine short rate models with regime switching; [31] use a portfolio credit risk model with default intensities given by a CIR process with regime switching to analyze securitization products backed by European sovereign bonds, and [24] study option pricing for regime switching pure jump Lévy processes. The recent contribution [56] unifies the previous examples and provides a theory of MMAPs for the case where the generator $\mathfrak{L}^X$ is a *bounded* operator so that X is a finite-activity jump process.

The present paper is a substantial extension of [56]: we make only very mild assumptions on $\mathfrak{L}^X$ and we allow for discontinuities in the coefficients of $\mathfrak{L}^{Y|X}(x)$. In particular, $\mathfrak{L}^X$ may be unbounded, so that models where X is a (jump-)diffusion or a jump process with infinite activity fall within the scope of our analysis. This is relevant for financial applications. There X often represents state variables such as market sentiment or the economic environment, and it may be more natural to model the dynamics of these state variables by continuous processes and not by processes with piecewise constant trajectories; see also the financial applications discussed in Section 6. The extension to unbounded generators $\mathfrak{L}^X$ is also interesting from a mathematical viewpoint. In that case the generator $\mathfrak{L}$ of (X, Y) does not satisfy the regularity conditions commonly imposed in the construction of Markov processes via perturbation arguments, so that classical results from semigroup theory (cf. [48, Section 3.3] or [27, Section 1.7]) are not readily applicable. To deal with this issue we choose a probabilistic approach and use weak convergence results to construct solutions to the martingale problem associated with $\mathfrak{L}$. If the Cauchy problem for the function φ from (1.1) admits a classical solution, relation (1.1) determines the characteristic function and hence the marginal distributions of (X, Y). Classical results from martingale problems thus guarantee the Markov property and consequently the existence of MMAPs. We go on and analyze further mathematical properties of MMAPs such as the Feller property, the existence of real exponential moments and the semimartingale characteristics of MMAPs. Finally, in order to illustrate the wide range of modelling possibilities offered by MMAPs we discuss several applications of MMAPs in finance.

Our analysis builds on the formal treatment of affine processes provided in [17]. Moreover, we make extensive use of the comprehensive treatment of Markov processes and martingale problems in [27]. We further contribute to the list of extensions to affine processes that has started to grow ever since the seminal work of [17]. [28] introduces time-inhomogeneous affine processes. [12] consider matrix-valued affine processes and [13] subsequently generalize the state space by considering affine processes on symmetric cones. Another generalization of affine semimartingales are polynomial processes, which are treated in [14] and [30]. More recently, [38] study a setup where the jump times of affine processes are allowed to be predictable (so that the processes no longer obey stochastic continuity) and [54] discuss infinite dimensional affine diffusions.

The rest of the paper is organized as follows. Section 2 introduces the necessary technical preliminaries and the formal notion of MMAPs. In particular, we discuss the specifics of the operator supposed to act as the generator of our class of processes. In Section 3 we solve the martingale problem associated with this operator. Section 4 is concerned with the Fourier transformation of MMAPs and the associated existence result. Further mathematical properties of MMAPs are discussed in Section 5. Finally, in Section 6 we highlight the applicability of MMAPs by discussing novel models from mathematical finance.

2. Setup

In this section we introduce the necessary notation and we give a formal definition of a Markov modulated affine process (abbreviated MMAP) via its generator.

2.1. Notation and basic concepts

Analysis. Let E be either an open set or the closure of an open set in $\mathbb{C}^k$. Then, $B_b(E)$ is the space of bounded Borel measurable functions, $C(E)$ $(C_b(E))$ denotes the Banach space (equipped with supremum norm $\|\cdot\|_\infty$) of continuous (bounded) functions on E. The complete subspace

$$C_\infty(E) = \{f \in C(E) \,:\, \forall \epsilon > 0 \, \exists K \subset E \text{ compact} : |f| < \epsilon \text{ on } E \setminus K\}$$

is the space of continuous functions vanishing at infinity. We use $C^k(E)$ to denote the space of k times continuously differentiable functions on E with $C^\infty(E) = \cap_{k \in \mathbb{N}} C^k(E)$ and $\mathscr{S}^k \subset C^\infty(\mathbb{R}^k)$ denotes the Schwartz space of rapidly decreasing functions on $\mathbb{R}^k$. The space $C_b^k(E)$ consists of bounded $C^k(E)$-functions (where the derivatives are bounded as well). The space of k times differentiable functions with compact support is $C_c^k(E)$. The set of càdlàg functions on $[0, \infty)$ with values in some space A is denoted by D_A. When there is no ambiguity with respect to dimensionality we write $\mathbf{1}$ for a vector of ones. Similarly, we use Id to denote the identity. We write bp-$\lim_{r \to a} f_r$ to indicate bounded and pointwise convergence of $\{f_r\}_{r \in I} \subset C_b(E)$ for some index set I.

In this paper, a MMAP is defined as a $(d + n)$-dimensional process $Z = (X, Y)$ on the state space $\mathcal{D} = \mathcal{D}^X \times \mathcal{D}^Y$, where $\mathcal{D}^X$ is an open (or the closure of an open) set of $\mathbb{R}^d$ and $\mathcal{D}^Y = \mathbb{R}_+^m \times \mathbb{R}^{n-m}$. Throughout we use x to denote a point in $\mathcal{D}^X$ and $y = (y^+, y^*)$ denotes a point in $\mathcal{D}^Y$ with $y^+ \in \mathbb{R}_+^m$ and $y^* \in \mathbb{R}^{n-m}$. An element in $\mathbb{C}^n = \mathbb{C}^m \times \mathbb{C}^{n-m}$ is denoted by $u = (u^+, u^*)$. We follow the notational conventions established in [17] and introduce $\mathcal{I} := \{1, \ldots, m\}$ and $\mathcal{J} := \{m+1, \ldots, n\}$ along with $\mathcal{I}(i) := \mathcal{I} \setminus \{i\}$ and $\mathcal{J}(i) := \{i\} \cup \mathcal{J}$. For a generic $k \times k$-matrix $\alpha = (\alpha_{ij})$ and a k-tuple $\beta = (\beta_1, \ldots, \beta_k)$, we write $\alpha_{IJ} := (\alpha_{ij})_{i \in I, j \in J}$ and $\beta_I := (\beta_i)_{i \in I}$ for indices $I, J \subset \{1, \ldots, k\}$. For $u \in \mathbb{C}^n$, we write $f_u(x) := e^{\langle u, x \rangle}$. Throughout we fix a continuous truncation function $\chi : \mathbb{R}^n \to [-1, 1]^n$ with $\chi(\xi) = \xi$ on a neighborhood around the origin. The Borel σ fields on $\mathcal{D}^X$, $\mathcal{D}^Y$ and $\mathcal{D}$ are denoted by $\mathcal{B}(\mathcal{D}^X)$, $\mathcal{B}(\mathcal{D}^Y)$ and $\mathcal{B}(\mathcal{D})$.

Probability. In order to avoid ambiguities with respect to different notions of *Feller* semigroups in the literature, we collect the definitions used in this paper in Appendix A. Since affine semigroups are in general *not conservative*, we throughout work on the path space $\Omega := D_{\mathcal{D}^\Delta}$, where $\mathcal{D}^\Delta := \mathcal{D} \cup \{\Delta\}$ is the one-point compactification of $\mathcal{D}$ with $\Delta \notin \mathcal{D}$ describing the point at infinity. We extend any function f on $\mathcal{D}$ to $\mathcal{D}^\Delta$ via $f(\Delta) = 0$ and we endow Ω with the *Skorohod* topology $\mathbb{S}$. Considering the product structure of the state space, we denote the canonical process by $Z_t(\omega) := (X_t(\omega), Y_t(\omega)) := \omega_t$, $t \geq 0$ for $\omega \in \Omega$. Note that $(\Omega, \mathbb{S})$ is a Polish space and that the corresponding Borel σ-algebra is generated by the evaluation maps, see [34, Section VI.1b] for further details. We use $(\mathcal{F}_t^0)_{t \geq 0}$ to denote the filtration generated by Z and we set $\mathcal{F}^0 := \bigvee_{t \geq 0} \mathcal{F}_t^0$. Note that working on $D_{\mathcal{D}^\Delta}$ is no restriction as [15] show that an affine process always has a càdlàg version. We write $\mathfrak{P}(\Omega)$ for the set of probability measures on Ω and $\mathbb{P}_z \in \mathfrak{P}(\Omega)$ denotes a measure with $\mathbb{P}_z(Z_0 = z) = 1$.

2.2. Markov-modulated Affine Processes

We aim to construct a family of probability measures $(\mathbb{P}_z)_{z \in \mathcal{D}} \subset \mathfrak{P}(\Omega)$, for which $(Z, (\mathbb{P}_z))$ is a Markov process with the following properties: X modulates the coefficients of the generator of Y and conditional on the paths of X the process Y is a time-inhomogeneous affine process. For

the characterization of (X, Y) we rely on the martingale problem approach associated to a linear operator of the form

$$\mathfrak{L}f(x,y) = \mathfrak{L}^X f(x,y) + \mathfrak{L}^{Y|X}(x)f(x,y), \quad f \in \mathrm{Dom}(\mathfrak{L}) \subset C_\infty(\mathcal{D}),$$

where $\mathfrak{L}^X$ acts on $x \mapsto f(x,y)$ and is the generator of X, and $\mathfrak{L}^{Y|X}(x)$ acts on $y \mapsto f(x,y)$ and for fixed x is the generator of an affine process. In the sequel, we outline the specifics of the operators $\mathfrak{L}^X$ and $\mathfrak{L}^{Y|X}(x)$ as well as of the operator domain $\mathrm{Dom}(\mathfrak{L})$.

The generator of X. As suggested above, we characterize the process X via the linear operator $\mathfrak{L}^X$ acting on functions in $\mathrm{Dom}(\mathfrak{L}^X)$ which is assumed to be a subset of $C_\infty(\mathcal{D}^X)$. The restrictions we impose on the modulating process are fairly weak as we allow the generator of X to be unbounded, extending the results of [56] to the case of diffusions and infinitely active jump processes. Throughout we work with the following

Assumption 2.1. The operator $(\mathfrak{L}^X, \mathrm{Dom}(\mathfrak{L}^X))$ with $\mathrm{Dom}(\mathfrak{L}^X) \subset C_\infty(\mathcal{D}^X)$ generates a Feller semigroup, denoted by $(P_t^X)_{t\geq 0}$. Moreover, $(P_t^X)_{t\geq 0}$ is conservative and $C_c^2(\mathcal{D}^X)$ is a core of $(\mathfrak{L}^X, \mathrm{Dom}(\mathfrak{L}^X))$.

Assumption 2.1 is fulfilled by a large set of different Markov processes, for instance by conservative affine processes. Note that the operator $\mathfrak{L}^X$ is not time-dependent, i.e. the Markov process X is time-homogeneous. Nevertheless, the extension to the inhomogeneous case is obvious by extending the process via $(t, X_t)_{t\geq 0}$. The restriction that $(P_t^X)_{t\geq 0}$ is conservative could be relaxed, but in the applications we have in mind there is no need for the modulating process to have finite life time.

Remark 2.2. [51] shows that under the above assumption, the semigroup $(P_t^X)_{t\geq 0}$ has an extension to $C_b(\mathcal{D}^X)$, which is C_b-Feller (cf. Definition A.3). We avoid confusions by denoting the corresponding semigroup on $C_b(\mathcal{D}^X)$ with $(\widetilde{P}_t^X)_{t\geq 0}$. Moreover, we associate with the C_b-Feller semigroup a *weak generator* $\widetilde{\mathfrak{L}}^X$ via

$$(2.1) \qquad \mathrm{Dom}(\widetilde{\mathfrak{L}}^X) = \left\{ f \in C_b(\mathcal{D}^X) \mid \exists g \in C_b(\mathcal{D}^X) \text{ such that } \underset{t\to 0+}{\mathrm{bp\text{-}lim}} \frac{\widetilde{P}_t^X f - f}{t} = g \right\},$$

$$(2.2) \qquad \widetilde{\mathfrak{L}}^X f = \underset{t\to 0+}{\mathrm{bp\text{-}lim}} \frac{\widetilde{P}_t^X f - f}{t}, \quad f \in \mathrm{Dom}(\widetilde{\mathfrak{L}}^X).$$

Note that $\widetilde{\mathfrak{L}}^X|_{\mathrm{Dom}(\mathfrak{L}^X)} = \mathfrak{L}^X$, that $\mathrm{Dom}(\widetilde{\mathfrak{L}}^X)$ is weakly dense in $C_b(\mathcal{D}^X)$ and that $\widetilde{\mathfrak{L}}^X$ is weakly closed, cf. [19, Chapter I]. We frequently resort to the operator $(\widetilde{\mathfrak{L}}^X, \mathrm{Dom}(\widetilde{\mathfrak{L}}^X))$, when dealing with functions which are not in $C_\infty(\mathcal{D}^X)$.

In [17], it is shown that the coefficients of the infinitesimal generator of a generic affine process Y are affine functions of y that can be described by a set of *admissible* parameters. For MMAPs the constant part of these affine functions may depend on the state x of the modulating process. This specification ensures the tractability of the characteristic function (see Section 4 below). In the following definition we introduce the Markov-modulated analogue of admissible parameters; their precise role will be clear in conjunction with Equation (2.5) below.

Definition 2.3. *We say that the parameters*

$$(a, \alpha, b, \beta, c, \gamma, m, \mu) = (a(x), \alpha, b(x), \beta, c(x), \gamma, m(x), \mu), \quad x \in \mathcal{D}^X,$$

are x-admissible, if for each fixed $x \in \mathcal{D}^X$

- $a(x) \in \mathrm{Sem}^n$ *with* $a_{\mathcal{I}\mathcal{I}}(x) = 0$, *and we have* $a_{ij}(\cdot) \in B_b(\mathcal{D}^X)$ *for all* $i, j \in \mathcal{I} \cup \mathcal{J}$,

- $\alpha = (\alpha_1, \ldots, \alpha_m)$ *with each* $\alpha_i \in \text{Sem}^n$ *and* $\alpha_{i;\mathcal{I}(i)\mathcal{I}(i)} = 0$ *for all* $i \in \mathcal{I}$,

- $b(x) \in \mathcal{D}^Y$ *with* $b_i(\cdot) \in B_b(\mathcal{D}^X)$ *for all* $i \in \mathcal{I} \cup \mathcal{J}$

- $\beta \in \mathbb{R}^{n \times n}$ *such that* $\beta_{\mathcal{I}\mathcal{J}} = 0$ *and* $\beta_{i\mathcal{I}(i)} \in \mathbb{R}_+^{m-1}$ *for all* $i \in \mathcal{I}$,

- $c(x) \in \mathbb{R}_+$ *with* $c(\cdot) \in B_b(\mathcal{D}^X)$,

- $\gamma \in \mathbb{R}_+^m$,

- $m(x, \cdot)$ *is a Borel measure on* $\mathcal{D}^Y \backslash \{0\}$ *which satisfies* $m(\cdot, d\xi) \in B_b(\mathcal{D}^X)$ *and* $M(x) := M(x, \mathcal{D}^Y \backslash \{0\}) < \infty$ *for*

$$M(x, d\xi) := (\langle \chi_{\mathcal{I}}(\xi), \mathbf{1} \rangle + \|\chi_{\mathcal{J}}(\xi)\|^2) m(x, d\xi),$$

- $\mu = (\mu_1, \ldots, \mu_m)$ *where* μ_i *is a Borel measure on* $\mathcal{D}^Y \backslash \{0\}$ *with* $\mathcal{M}_i := \mathcal{M}_i(\mathcal{D}^Y \backslash \{0\}) < \infty$ *where*

$$\mathcal{M}_i(d\xi) := (\langle \chi_{\mathcal{I}(i)}(\xi), \mathbf{1} \rangle + \|\chi_{\mathcal{J}(i)}(\xi)\|^2) \mu_i(d\xi).$$

The parameters are called strongly x-admissible, *if in addition* $(a(x), b(x), c(x))$ *are continuous in x, and if $m(x, d\xi)$ is weakly continuous on $\mathcal{D}^Y \backslash \{0\}$ in x.*

[28] provides an accessible illustration of the above parameter conditions and their implications.

In this paper we define a MMAP (X, Y) in terms of the generator of its associated semigroup. The following definition is needed to specify the domain of that operator. For $f \in C^2(\mathcal{D})$ we set

(2.3) $$[f]_1(x, y) := (1 + \|y^+\|) \left(|f(x, y)| + \|\nabla_y f(x, y)\| + \sum_{k,l=1}^n \left| \frac{\partial^2 f(x, y)}{\partial y_k \partial y_l} \right| \right),$$

(2.4) $$[f]_2(x, y) := |\langle y^*, \beta^* \nabla_{\mathcal{J}} f(x, y) \rangle|,$$

where $\beta^* := (\beta^\top)_{\mathcal{J}\mathcal{J}} \in \mathbb{R}^{(n-m) \times (n-m)}$. The advantage of the functions $[f]_1$ and $[f]_2$ in handling affine semigroups is best seen in Lemma 6.1 of [28] and in Lemma 8.1 of [17].

The "generator" of (X, Y). Consider for x-admissible parameters $(a, \alpha, b, \beta, c, \gamma, m, \mu)$ the following partial integro differential operator acting on $g \in C_b^2(\mathcal{D}^Y)$,

(2.5)
$$\begin{aligned}
\mathfrak{L}^{Y|X}(x)g(y) = &\sum_{k,l=1}^n a_{kl}(x) \frac{\partial^2 g(y)}{\partial y_k \partial y_l} + \langle b(x), \nabla g(y) \rangle - c(x)g(y) \\
&+ \int_{\mathcal{D}^Y \backslash \{0\}} (g(y + \zeta) - g(y) - \langle \nabla_{\mathcal{J}} g(y), \chi_{\mathcal{J}}(\zeta) \rangle) m(x, d\zeta) \\
&+ \sum_{k,l=1}^n \langle \alpha_{\mathcal{I},kl}, y^+ \rangle \frac{\partial^2 g(y)}{\partial y_k \partial y_l} + \langle \beta y, \nabla g(y) \rangle - \langle \gamma, y^+ \rangle g(y) \\
&+ \sum_{i=1}^m \int_{\mathcal{D}^Y \backslash \{0\}} (g(y + \zeta) - g(y) - \langle \nabla_{\mathcal{J}(i)} g(y), \chi_{\mathcal{J}(i)}(\zeta) \rangle) y_i^+ \mu_i(d\zeta).
\end{aligned}$$

Throughout we consider the linear operator $(\mathfrak{L}, \text{Dom}(\mathfrak{L}))$ with

(2.6a) $$\mathfrak{L}f(x, y) := \mathfrak{L}^X f(x, y) + \mathfrak{L}^{Y|X}(x)f(x, y)$$

where f is an element of

(2.6b) $$\text{Dom}(\mathfrak{L}) := \left\{ f \in C^2(\mathcal{D}) \cap C_\infty(\mathcal{D}) \,\middle|\, \begin{array}{l} f(\cdot, y) \in \text{Dom}(\mathfrak{L}^X) \text{ for all } y \in \mathcal{D}^Y, \\ [f]_1, [f]_2 \in C_\infty(\mathcal{D}) \end{array} \right\}$$

Definition 2.4. *A Markov process* $(Z, (\mathbb{P}_{z \in \mathcal{D}}))$ *with sub-Markov semigroup* $(P_t)_{t \geq 0}$ *and generator* $(L, \mathrm{Dom}(L))$ *is called a* Markov-modulated affine process (MMAP) *if* $L|_{\mathrm{Dom}(\mathfrak{L})} = \mathfrak{L}$. *Moreover, we call* $(P_t)_{t \geq 0}$ *a* Markov-modulated affine semigroup.

If the parameters underlying $\mathfrak{L}$ *are strongly* x-*admissible, we refer to* $(Z, (\mathbb{P}_z)_{z \in \mathcal{D}})$ *and* $(P_t)_{t \geq 0}$ *as a* strongly regular Markov-modulated affine process *and as a* strongly regular Markov-modulated affine semigroup, *respectively.*

Definition 2.4 differs from the classical definition of affine processes in [17]. There, a Markov process Y is called affine if its characteristic function for fixed $t \geq 0$ is of the form $e^{\phi(t,u) + \langle \psi(t,u), y \rangle}$, where the functions ϕ and ψ are assumed to satisfy a crucial regularity assumption. This technical condition is both necessary and sufficient for the existence of an infinitesimal generator of an affine semigroup (see e.g. Example 1.25.g and the surrounding discussion of [5] in the context of Feller processes). [37] show that this regularity assumption is in fact superfluous, and further prove that stochastically continuous affine processes are Feller, cf. [37, Theorem 3.5]. Thus, the results of [37] show that an affine process can equivalently be characterized in terms of its generator, so that MMAPs as defined in this paper are a genuine extension of standard affine processes.

In general, $(\mathfrak{L}, \mathrm{Dom}(\mathfrak{L}))$ is *not* the generator of a strongly continuous semigroup on $C_\infty(\mathcal{D})$. The possible discontinuities in the x-admissible parameters prevent us from applying the theory of Hille-Yosida (cf. [27, Section 1.2]) to construct MMAPs via analytical semigroup arguments. In fact, a priori it is not even clear if for a given generator $\mathfrak{L}^X$ satisfying Assumption 2.1 and a given set of x-admissible parameters a MMAP in the sense of Definition 2.4 exists. We address this issue via the associated martingale problem.

Definition 2.5. *Consider a linear operator* $(\mathfrak{A}, \mathrm{Dom}(\mathfrak{A}))$ *with* $\mathrm{Dom}(\mathfrak{A}) \subset C_\infty(\mathcal{D})$. *We say that* $(\mathbb{P}_z)_{z \in \mathcal{D}} \subset \mathfrak{P}(\Omega)$ *solves the martingale problem associated with* $\mathfrak{A}$, *in short* $\{\mathbb{P}_z\} \in \mathbf{MP}(\mathfrak{A}, \mathrm{Dom}(\mathfrak{A}))$, *if for* $f \in \mathrm{Dom}(\mathfrak{A})$ *and for all* $z \in \mathcal{D}$ *the process*

$$f(Z_t) - f(z) - \int_0^t \mathfrak{A}f(Z_s)ds, \quad t \geq 0$$

is a $\{\mathcal{F}_t\}$-*martingale, where* $\{\mathcal{F}_t\}$ *is the* $\mathbb{P}_z$-*completion of* $\{\mathcal{F}_t^0\}$. *Moreover, the martingale problem is said to be* well-posed, *if for all* z *any two solutions* $\mathbb{P}_z^1, \mathbb{P}_z^2 \in \mathbf{MP}(\mathfrak{A}, \mathrm{Dom}(\mathfrak{A}))$ *have the same finite-dimensional distributions.*

The next example shows that the class of MMAPs comprises also Markov processes with quite irregular behavior.

Example 2.6. Suppose $m = 0$, $n = 1$ and $\mathcal{D}^X = \mathbb{R}$. We let the modulating process X be a Brownian motion so that

$$\mathfrak{L}^X f = \partial_{xx} f, \quad f \in C_c^2(\mathcal{D}^X) \subset \mathrm{Dom}(\mathfrak{L}^X).$$

Concerning x-admissible parameters, we set

$$a(x) := \mathbb{1}_{\{x \geq 0\}} 1/2, \qquad m(x, d\zeta) := \mathbb{1}_{\{x < 0\}}(1/x^2)\delta_x(d\zeta),$$

where $\delta_x(d\zeta)$ is the Dirac measure centred at x; and all the remaining x-admissible parameters vanish. Intuitively, upon X crossing 0, the modulated process Y switches infinitely often between being a diffusion and a pure jump process (as a Brownian motion has infinitely many crossings of zero). The existence of a MMAP with this generator follows from the results of Sections 3 and 4; see Example 4.8 for details.

3. The associated Martingale Problem

In this section we discuss the existence of solutions to the martingale problem associated with the linear operator $\mathfrak{L} = \mathfrak{L}^X + \mathfrak{L}^{Y|X}(x)$ defined in (2.6). For fixed $x \in \mathcal{D}^X$, it follows from [17] that an appropriate extension of $\mathfrak{L}^{Y|X}(x)|_{C_c^\infty(\mathcal{D}^Y)}$ generates an affine semigroup. Thus, we can regard $\mathfrak{L}^X$ as a perturbation of $\mathfrak{L}^{Y|X}(x)$. Standard results on martingale problems related to perturbations typically fall into two categories (cf. [27, Section 4.10.]): either the perturbation is bounded (corresponding to $\mathfrak{L}^X$ being a bounded operator), or the two operators are independent, that is $\mathfrak{L}^{Y|X}(x)$ is independent of x. Hence, the main difficulties in dealing with $\mathbf{MP}(\mathfrak{L}, \mathrm{Dom}(\mathfrak{L}))$ stem from the fact that our setup falls in neither of the two categories. Moreover, the popular approach of showing that $(\mathfrak{L}, \mathrm{Dom}(\mathfrak{L}))$ satisfies the *positive maximum principle* (cf. [27, Theorem 4.5.4]) to solve $\mathbf{MP}(\mathfrak{L}, \mathrm{Dom}(\mathfrak{L}))$ is also not applicable in our setup since it would require $\mathfrak{L}$ to map into $C(\mathcal{D})$.

We tackle these issues via an approximation argument: we approximate $\mathfrak{L}^X$ by a sequence of bounded operators $\{\mathfrak{L}_k^X\}_{k \in \mathbb{N}}$ and we show that the sequence of solutions to the martingale problem for the operator $\mathfrak{L}_k + \mathfrak{L}^{Y|X}(x)$, $k \in \mathbb{N}$, is tight and that every limit point solves the martingale problem for $\mathfrak{L}^X + \mathfrak{L}^{Y|X}(x)$.

We begin with two auxiliary results.

Lemma 3.1. $\mathrm{Dom}(\mathfrak{L})$ *is a dense subset of* $C_\infty(\mathcal{D})$.

Proof. As in the proof of [28, Proposition 6.3.], we introduce the following set

$$\Theta_0 := \left\{ h \in C^\infty(\mathcal{D}^Y) \,\middle|\, \begin{array}{l} h(y^+, y^*) = \int_{\mathbb{R}^{n-m}} e^{\langle (v, iq), (y^+, y^*) \rangle} g(q) dq, \\ v \in \mathbb{C}_{--}^m, \, g \in C_c^\infty(\mathbb{R}^{n-m}) \end{array} \right\}.$$

We further set $\overline{\Theta}_0 := \{gh \mid g \in C_c^2(\mathcal{D}^X), h \in \Theta_0\}$. Pick an arbitrary $f = gh \in \overline{\Theta}_0$. As $C_c^2(\mathcal{D}^X) \subset \mathrm{Dom}(\mathfrak{L}^X)$, we have $f(\cdot, y) \in \mathrm{Dom}(\mathfrak{L}^X)$ for all $y \in \mathcal{D}^Y$.

Note that $\Theta_0 \subset \mathscr{S}^n$. Thus, we have (with slight abuse of notation) for an arbitrary multiindex $\alpha \in \mathbb{N}_0^n$ that $(1 + \|y^+\|)h \in \mathscr{S}^n$ and $\nabla^\alpha h \in \mathscr{S}^n$ (cf. [58, Section V.2]). Since g has compact support we infer that $[f]_1, [f]_2 \in C_\infty(\mathcal{D})$ and thus, $\overline{\Theta}_0 \subset \mathrm{Dom}(\mathfrak{L})$. Since the linear span of $\overline{\Theta}_0$ is dense in $C_\infty(\mathcal{D})$ (see the proof of [28, Proposition 6.3.]), we arrive at the result. $\square$

Lemma 3.2. $\mathrm{Dom}(\mathfrak{L})$ *is an algebra.*

Proof. Let $f, g \in \mathrm{Dom}(\mathfrak{L})$. Obviously, $fg \in C^2(\mathcal{D}) \cap C_\infty(\mathcal{D})$. Moreover, using the product rule we get the estimates (see also [17, Section 8])

$$\|[fg]_1\|_\infty \leq K \|[f]_1\|_\infty \|[g]_1\|_\infty,$$

for some constant K. Thus, we have $[fg]_1, [fg]_2 \in C_\infty(\mathcal{D})$.

It remains to consider $\mathfrak{L}^X(fg)(\cdot, y)$ for arbitrary $y \in \mathcal{D}^Y$. Since $C_c^2(\mathcal{D})$ is a core of $\mathrm{Dom}(\mathfrak{L}^X)$, we can find sequences $(f_k), (g_k) \in C_c^2(\mathcal{D}^X)$ converging to $f(\cdot, y)$ and $g(\cdot, y)$ w.r.t. $\|\cdot\|_{\mathfrak{L}}^X$. Since the f_k and g_k are bounded it holds that $\lim_{k \to \infty} \|f_k g_k - f(\cdot, y)g(\cdot, y)\|_\infty = 0$ with $\{f_k g_k\} \subset C_c^2(\mathcal{D}^X)$ as $C_c^2(\mathcal{D}^X)$ is an algebra. And since the $\|\cdot\|_{\mathfrak{L}}^X$-closure of $C_c^2(\mathcal{D}^X)$ is $\mathrm{Dom}(\mathfrak{L}^X)$, the result follows. $\square$

Next, we introduce an approximating operator sequence $\{\mathfrak{L}_k^X\}_{k \in \mathbb{N}}$ and we study the martingale problem associated with the operator $\mathfrak{L}_k = \mathfrak{L}_k^X + \mathfrak{L}^{Y|X}(x)$. Since $(\mathfrak{L}^X, \mathrm{Dom}(\mathfrak{L}^X))$ generates a strongly continuous semigroup by Assumption 2.1, the Hille-Yosida theorem implies first that $\mathrm{Dom}(\mathfrak{L}^X)$ is dense in $C_\infty(\mathcal{D}^X)$, second that $\mathrm{Range}(\alpha \, \mathrm{Id} - \mathfrak{L}^X) = C_\infty(\mathcal{D}^X)$ for any $\alpha > 0$ and third that $\mathfrak{L}^X$ is dissipative, i.e. $\|\alpha f - \mathfrak{L}^X f\| \geq \alpha \|f\|$ for any $\alpha > 0$ and $f \in \mathrm{Dom}(\mathfrak{L}^X)$. Moreover, we consider for each k the following *Yosida approximation*

$$(3.1) \qquad\qquad\qquad \mathfrak{L}_k^X := k\mathfrak{L}^X G_k,$$

where $G_k = (k\,\mathrm{Id} - \mathfrak{L}^X)^{-1}$. Note that the operator G_k, the resolvent of the semigroup $(P_t^X)_{t \geq 0}$, exists by standard results from semigroup theory, see e.g. Section VII.4 of [58]. Then, $\{\mathfrak{L}_k^X\}_{k \in \mathbb{N}}$ is a sequence of *bounded* operators, and it holds that $\lim_{k \to \infty} \mathfrak{L}_k^X f = \mathfrak{L}^X f$ (where the limit is pointwise for $f \in \mathrm{Dom}(\mathfrak{L}^X)$), since $(\mathfrak{L}^X, \mathrm{Dom}(\mathfrak{L}^X))$ generates a strongly continuous semigroup, see for instance [27, Chapter 1] for details.

Lemma 3.3. *Consider the operator* $(\mathfrak{L}_k, \mathrm{Dom}(\mathfrak{L}))$*, where*

$$\mathfrak{L}_k f(x, y) := \mathfrak{L}_k^X f(x, y) + \mathfrak{L}^{Y|X}(x) f(x, y), \quad f \in \mathrm{Dom}(\mathfrak{L}).$$

Then, there exists a solution to the martingale problem associated with $(\mathfrak{L}_k, \mathrm{Dom}(\mathfrak{L}))$*.*

Proof. To prove the lemma we resort to [27, Proposition 4.10.2]. Accordingly, we have to show that for each k there exists a kernel π_k^X such that we obtain the following representation for any $g \in \mathrm{Dom}(\mathfrak{L}^X)$,

$$(3.2) \qquad \mathfrak{L}_k^X g(x) = k \int_{\mathcal{D}^X} (g(\zeta) - g(x)) \pi_k^X(x, d\zeta), \quad \forall x \in \mathcal{D}^X.$$

First, note that the operator G_k from (3.1) is a bijection from $C_\infty(\mathcal{D}^X)$ onto $\mathrm{Dom}(\mathfrak{L}^X)$. Additionally, since G_k is a positive linear operator, we know by the Riesz representation theorem that there exists a kernel π_k^X such that for any $f \in C_\infty(\mathcal{D}^X)$

$$kG_k f(x) = \int_{\mathcal{D}^X} f(\zeta) \pi_k^X(x, d\zeta), \quad \forall x \in \mathcal{D}^X.$$

Pick an arbitrary $g \in \mathrm{Dom}(\mathfrak{L}^X)$. Consequently, since $(k\,\mathrm{Id} - \mathfrak{L}^X)g \in C_\infty(\mathcal{D}^X)$ by the Hille-Yosida theorem, it holds that

$$kg(x) = (k\,\mathrm{Id} - \mathfrak{L}^X) k G_k g(x) = \int_{\mathcal{D}^X} (k\,\mathrm{Id} - \mathfrak{L}^X) g(\zeta) \pi_k^X(x, d\zeta), \quad x \in \mathcal{D}^X.$$

We further have that $\int_{\mathcal{D}^X} \mathfrak{L}^X g(\zeta) \pi_k^X(\cdot, d\zeta) = \mathfrak{L}_k^X g$, and so overall we obtain the representation (3.2) for any $g \in \mathrm{Dom}(\mathfrak{L}^X)$. $\qquad\square$

Note that the proof of the above lemma shows that in probabilistic terms, the approximation of $\mathfrak{L}^X$ by $\{\mathfrak{L}_k^X\}_{k \in \mathbb{N}}$ corresponds to the approximation of the Feller process X by a sequence of pure jump processes.

Theorem 3.4. *Consider the linear operator* $(\mathfrak{L}, \mathrm{Dom}(\mathfrak{L}))$ *given by* (2.6)*, where the underlying parameters* $(a, \alpha, b, \beta, c, \gamma, m, \mu)$ *are x-admissible and where $\mathfrak{L}^X$ satisfies Assumption 2.1. Then, there exists a solution to the martingale problem associated with* $(\mathfrak{L}, \mathrm{Dom}(\mathfrak{L}))$*.*

Proof. For a closed subset $U \times V \subset \mathcal{D}^X \times \mathcal{D}^Y$ and a function $f \in \mathrm{Dom}(\mathfrak{L})$, we introduce

$$\|f\|_{\sharp; U \times V} := \sup_{(x,y) \in U \times V} \{[f]_1(x, y) + [f]_2(x, y)\},$$

where $[f]_1$ and $[f]_2$ were introduced in (2.3), (2.3). The proof is divided into two steps. First, we show that the measures $\mathbb{P}_z^k \in \mathbf{MP}(\mathfrak{L}_k, \mathrm{Dom}(\mathfrak{L}))$, $k \in \mathbb{N}$ form a tight collection. In the second step, we collect several classical results to conclude that there exist limit points in $(\mathbb{P}_z^k)_{k \in \mathbb{N}}$ that solve $\mathbf{MP}(\mathfrak{L}, \mathrm{Dom}(\mathfrak{L}))$.

Step 1: Tightness. Fix an arbitrary $z \in \mathcal{D}$. By Lemma 3.3 there is for each k a measure $\mathbb{P}_z^k \in \mathbf{MP}(\mathfrak{L}_k, \mathrm{Dom}(\mathfrak{L}))$ with the operator $\mathfrak{L}_k$ acting on $f \in \mathrm{Dom}(\mathfrak{L})$ via

$$\mathfrak{L}_k f(x,y) = \mathfrak{L}_k^X f(x,y) + \mathfrak{L}^{Y|X}(x) f(x,y).$$

To keep the notation lean, we omit the subscript z in the sequel, i.e. we write $\mathbb{P}^k$ and $\mathbb{E}^k$ instead of $\mathbb{P}_z^k$ and $\mathbb{E}_z^k$, respectively. We want to show that $\{\mathbb{P}^k\}_{k \in \mathbb{N}}$ is tight.

Fix an arbitrary $f \in \mathrm{Dom}(\mathfrak{L})$. We use $\mathbb{P}_f^k$ to denote the distribution of the paths of $f(X,Y)$ (with corresponding expectation $\mathbb{E}_f^k$) and let $(\xi_t)_{t \geq 0}$ be the coordinate process of real-valued càdlàg functions, where the σ-algebra $\mathcal{F}^{k,\xi}$ is the $\mathbb{P}_f^k$-completion of $\sigma(\xi_t : t \geq 0)$. We denote the corresponding filtration with $\{\mathcal{F}^{k,\xi}\}_t$. Since $\mathrm{Dom}(\mathfrak{L})$ is dense in $C_\infty(\mathcal{D})$, we know from [27, Theorem 3.9.1] that tightness of $(\mathbb{P}^k)_{k \in \mathbb{N}}$ is equivalent to tightness of $(\mathbb{P}_f^k)_{k \in \mathbb{N}}$. We introduce the following function,

$$g \,:\, \Omega \to D_{\mathbb{R}} \,;\, \omega \mapsto f \circ \omega.$$

Let $\epsilon \in (0,1]$. Set $q(e,d) := |e-d| \wedge 1$, for any $e, d \in \mathbb{R}$. Then, for every $A \in \mathcal{F}_t^{k,\xi}, 0 \leq t < \infty$ and $0 \leq h \leq \epsilon$ we have

$$
\begin{aligned}
\mathbb{E}_f^k \left[\mathbb{1}_A q(\xi_{t+h}, \xi_t)^2 \right] &\leq \int_A (\xi_{t+h} - \xi_t)^2 d\mathbb{P}_f^k \\
&= \int_{g^{-1}(A)} (f(X_{t+h}, Y_{t+h}) - f(X_t, Y_t))^2 d\mathbb{P}^k \\
(3.3) \qquad\qquad &= \int_{g^{-1}(A)} \mathbb{E}^k \left[(f(X_{t+h}, Y_{t+h}) - f(X_t, Y_t))^2 \mid \mathcal{F}_t \right] d\mathbb{P}^k.
\end{aligned}
$$

Since $\mathbb{P}^k \in \mathbf{MP}(\mathfrak{L}_k, \mathrm{Dom}(\mathfrak{L}_k))$ and $f, f^2 \in \mathrm{Dom}(\mathfrak{L}_k)$ by Lemma 3.2, we get

$$
\begin{aligned}
&\mathbb{E}^k \left[(f(X_{t+h}, Y_{t+h}) - f(X_t, Y_t))^2 \mid \mathcal{F}_t \right] \\
&= \mathbb{E}^k \left[f^2(X_{t+h}, Y_{t+h}) - f^2(X_t, Y_t) \mid \mathcal{F}_t \right] \\
&\qquad - 2f(X_t, Y_t)\mathbb{E}^k \left[f(X_{t+h}, Y_{t+h}) - f(X_t, Y_t) \mid \mathcal{F}_t \right] \\
&= \mathbb{E}^k \left[\int_t^{t+h} \mathfrak{L}_k f^2(X_s, Y_s)ds \mid \mathcal{F}_t \right] - 2f(X_t, Y_t)\mathbb{E}^k \left[\int_t^{t+h} \mathfrak{L}_k f(X_s, Y_s)ds \mid \mathcal{F}_t \right]
\end{aligned}
$$

Next, we handle the integrands. For that recall that the specific form of $\mathfrak{L}_k f(x,y)$ equals

$$
\mathfrak{L}_k^X f(x,y) + \sum_{k,l=1}^n a_{kl}(x) \frac{\partial^2 f(x,y)}{\partial y_k \partial y_l} + \langle b(x), \nabla_y f(x,y) \rangle - c(x) f(x,y)
$$

$$
+ \int_{\mathcal{D}^Y \setminus \{0\}} (f(x, y+\zeta) - f(x,y) - \langle \nabla_{\mathcal{J}} f(x,y), \chi_{\mathcal{J}}(\zeta) \rangle) m(x, d\zeta)
$$

$$
(3.4) \qquad + \sum_{k,l=1}^n \langle \alpha_{\mathcal{I},kl}, y^+ \rangle \frac{\partial^2 f(x,y)}{\partial y_k \partial y_l} + \langle \beta y, \nabla_y f(x,y) \rangle - \langle \gamma, y^+ \rangle f(x,y)
$$

$$
(3.5) \qquad + \sum_{i=1}^m \int_{\mathcal{D}^Y \setminus \{0\}} (f(x, y+\zeta) - f(x,y) - \langle \nabla_{\mathcal{J}(i)} f(x,y), \chi_{\mathcal{J}(i)}(\zeta) \rangle) y_i^+ \mu_i(d\zeta).
$$

We compactly write the expressions (3.4) to (3.5) as $\tilde{\mathfrak{L}}^Y f(x,y)$ and note that $\tilde{\mathfrak{L}}^Y$ is the generator of an affine semigroup with admissible parameters $(0, \alpha, 0, \beta, 0, \gamma, 0, \mu)$. Thus, by Lemma 8.1 of [17] we know that for $(x,y) \in \mathcal{D}$ there exists a constant C such that

$$
\tilde{\mathfrak{L}}^Y f(x,y) \leq C \left(\|\alpha\| + \|\beta\| + \|\gamma\| + \sum_{i \in \mathcal{I}} \mathcal{M}_i \right) \|f(x, y+\cdot)\|_{\sharp; \mathcal{D}^Y},
$$

and since $f \in C_\infty(\mathcal{D})$ this in particular implies that there is a constant C_f^Y such that

$$(3.6) \qquad \tilde{\mathfrak{L}}^Y f(x,y) \leq C_f^Y,$$

with the constant C_f^Y depending on f, but not on k. Furthermore, with $\mathfrak{L}_k^X$ being a bounded operator, there exists a constant $C_{k,f}^X$ (depending both on f and k) such that

$$(3.7) \qquad \mathfrak{L}_k^X f(x,y) \leq C_{k,f}^X.$$

For the remaining terms we yet again apply Lemma 8.1 of [17] and exploit the structure of $\mathrm{Dom}(\mathfrak{L}_k)$ to arrive at an inequality of the form

$$(3.8) \quad \mathfrak{L}_k f(x,y) - (\mathfrak{L}_k^X f(x,y) + \tilde{\mathfrak{L}}^Y f(x,y))$$
$$\leq C \left(\left\| (\|a_{kl}\|_\infty)_{k,l=1}^n \right\| + \left\| (\|b_k\|_\infty)_{k=1}^n \right\| + \|c\|_\infty + \|M\|_\infty \right) \|f(x, y+\cdot)\|_{\sharp;\mathcal{D}^Y}$$
$$(3.9) \qquad \leq C_f^{Y|X},$$

for a constant $C_f^{Y|X}$ independent of k. Similar considerations also apply for $\mathfrak{L}_k f^2$ and so the inequalities (3.6) to (3.9) imply

$$\mathbb{E}^k \left[(f(X_{t+h}, Y_{t+h}) - f(X_t, Y_t))^2 \mid \mathcal{F}_t \right]$$
$$(3.10) \qquad \leq \epsilon (C_{k,f^2}^X + C_{f^2}^{Y|X} + C_{f^2}^Y) + 2\epsilon \|f\|_\infty (C_{k,f}^X + C_f^{Y|X} + C_f^Y) \quad \mathbb{P}^k\text{-a.s.}$$

It remains to consider the behavior of the constant $C_{k,f}^X$ when taking the supremum over k. However, for a Yosida approximation it holds that $\mathfrak{L}_k^X f$ converges uniformly to $\mathfrak{L}^X f$, which in turn implies the finiteness of $\sup_k \left\| \mathfrak{L}_k^X f \right\|_\infty$ (uniformly convergent bounded functions are uniformly bounded). Consequently, we can find a constant independent of k and we have thus shown that $\lim_{t \to 0} \sup_{k \in \mathbb{N}} \mathbb{E}_f^k \left[q(\xi_0, \xi_t)^2 \right] = 0$. By defining γ_ϵ as the right-hand side of (3.10), we can regard $(\gamma_\epsilon)_{\epsilon \in (0,1]}$ as a family of functions $\gamma_\epsilon : D_\mathbb{R} \to [0, \infty)$ with the following properties

$$\sup_{k \in \mathbb{N}} \mathbb{E}_f^k [\gamma_\epsilon] \to 0 \quad \text{as } \epsilon \to 0,$$
$$\mathbb{E}_f^k \left[q(\xi_{t+h}, \xi_t)^2 \, q(\xi_t, \xi_{t-v})^2 \mid \mathcal{F}_t^{k,\xi} \right] \leq \mathbb{E}_f^k \left[\gamma_\epsilon \mid \mathcal{F}_t^{k,\xi} \right] \quad \mathbb{P}_f^k\text{-a.s.}$$

for all $0 \leq t \leq T$, $0 \leq h \leq \epsilon \leq 1$ and $0 \leq v \leq \epsilon \wedge t$. We work on $D_{\mathcal{D}^\Delta}$, so that the compact containment condition is trivially fulfilled. Hence, by Theorem 3.8.6 of [27] we conclude that $(\mathbb{P}_f^k)_{k \in \mathbb{N}}$ is tight for all $f \in \mathrm{Dom}(\mathfrak{L})$. By Theorem 3.9.1 of [27] together with Lemma 3.1 this implies that $(\mathbb{P}^k)_{k \in \mathbb{N}}$ is tight.

Step 2: Limit. Since $(\Omega, \mathbb{S})$ is a Polish space, by Prokhorov's Theorem the set $(\mathbb{P}^k)_{k \in \mathbb{N}} \subset \mathfrak{P}(\Omega)$ is relatively compact. So, there exists an accumulation point of $\{\mathbb{P}^k\}$, which we denote by $\mathbb{P}$ (with corresponding expectation $\mathbb{E}$). In the sequel we show that $\mathbb{P} \in \mathbf{MP}(\mathfrak{L}, \mathrm{Dom}(\mathfrak{L}))$, which is equivalent to showing that for all $f \in \mathrm{Dom}(\mathfrak{L})$ it holds that

$$(3.11) \qquad \mathbb{E} \left[\left(f(X_{t_2}, Y_{t_2}) - f(X_{t_1}, Y_{t_1}) - \int_{t_1}^{t_2} \mathfrak{L} f(X_s, Y_s) ds \right) \prod_{l=1}^m h_l(X_{s_l}, Y_{s_l}) \right] = 0,$$

for all $0 \leq s_1 < \dots s_m \leq t_1 \leq t_2$ and all $h_1, \dots, h_m \in B_b(\mathcal{D})$. By Lemma 3.7.7 of [27] the following dense subset of $[0, \infty)$,

$$T_\mathbb{P} := \{ t \geq 0 \mid \mathbb{P}((X_{t-}, Y_{t-}) = (X_t, Y_t)) = 1 \},$$

has an at most countable complement in $[0, \infty)$. We pick a subsequence $\{\mathbb{P}^{k(j)}\}$ for which $\lim_{j \to \infty} \mathbb{P}^{k(j)} = \mathbb{P}$. Then, by [27, Theorem 3.7.8] the finite dimensional distributions of

$((X_{t_1}, Y_{t_1}), \ldots, (X_{t_m}, Y_{t_m}))$ under $\mathbb{P}^{k(j)}$ converge weakly to the corresponding distributions under $\mathbb{P}$ for $t_1, \ldots, t_m \in T_\mathbb{P}$. Thus, it is enough to show (3.11) for $s_1, \ldots, s_m, t_1, t_2 \in T_\mathbb{P}$. Since $\mathbb{P}^{k(j)} \in \mathbf{MP}(\mathfrak{L}_{k(j)}, \mathrm{Dom}(\mathfrak{L}))$, it is left to prove that

$$(3.12) \qquad \lim_{j \to \infty} \mathbb{E}^{k(j)} \left[\int_{t_1}^{t_2} \mathfrak{L}_{k(j)} f(X_s, Y_s) ds \prod_{l=1}^{m} h_l(X_{s_l}, Y_{s_l}) \right]$$

$$= \mathbb{E} \left[\int_{t_1}^{t_2} \mathfrak{L} f(X_s, Y_s) ds \prod_{l=1}^{m} h_l(X_{s_l}, Y_{s_l}) \right]$$

Given the uniform convergence of $\{\mathfrak{L}_{k(j)} f\}$ we have

$$(3.13) \qquad \lim_{j \to \infty} \mathbb{E}^{k(j)} \left[(\mathfrak{L}_{k(j)} f(X_s, Y_s) - \mathfrak{L} f(X_s, Y_s)) \prod_{l=1}^{m} h_l(X_{s_m}, Y_{s_m}) \right] = 0,$$

for all $s \in T_\mathbb{P}$. The weak convergence of $\{\mathbb{P}^{k(j)}\}$ further implies that

$$(3.14) \qquad \lim_{j \to \infty} \mathbb{E}^{k(j)} \left[\mathfrak{L} f(X_s, Y_s) \prod_{l=1}^{m} h_l(X_{s_l}, Y_{s_l}) \right] = \mathbb{E} \left[\mathfrak{L} f(X_s, Y_s) \prod_{l=1}^{m} h_l(X_{s_l}, Y_{s_l}) \right],$$

and so Equations (3.13) and (3.14) prompt

$$\lim_{j \to \infty} \mathbb{E}^{k(j)} \left[\mathfrak{L}_{k(j)} f(X_s, Y_s) \prod_{l=1}^{m} h_l(X_{s_l}, Y_{s_l}) \right] = \mathbb{E} \left[\mathfrak{L} f(X_s, Y_s) \prod_{l=1}^{m} h_l(X_{s_l}, Y_{s_l}) \right],$$

Finally, we interchange order of integration in (3.12) to arrive at the desired result. $\qquad \square$

In Sections 4 and 5, we frequently resort to the following useful extension of Theorem 4.6. Recall that we introduced $\mathfrak{L}^{Y|X}(x)$ as an operator on $C_b^2(\mathcal{D}^Y)$. The proof is postponed to Appendix B.

Corollary 3.5. *For an arbitrary $z \in \mathcal{D}$, choose a $\mathbb{P}_z \in \mathbf{MP}(\mathfrak{L}, \mathrm{Dom}(\mathfrak{L}))$. Let $f \in C([0, \infty) \times \mathcal{D}^X)$, which is C^1 in its first argument and $f(t, \cdot) \in \mathrm{Dom}(\widetilde{\mathfrak{L}}^X)$ for all $t \geq 0$. Moreover, consider $g \in C^{1,2}([0, \infty) \times \mathcal{D}^Y)$, which is bounded in its second argument. Then, the process*

$$f(t, X_t) g(t, Y_t) - f(0, x) g(0, y) - \int_0^t (\partial_t + \widetilde{\mathfrak{L}}^X + \mathfrak{L}^{Y|X}(x)) f(s, X_s) g(s, Y_s) ds, \quad t \geq 0,$$

is a martingale under $\mathbb{P}_z$.

4. Transform Formula and Existence

The popularity of affine processes largely stems from the fact that the Fourier transform of their marginal distributions is available in semi-explicit form (up to the solution of a system of ODEs). In particular, in the context of affine models many pricing problems in mathematical finance can be solved efficiently by Fourier methods. In this section we show that the characteristic function of MMAPs has a fairly simple form as well, which makes these processes an appealing tool for many modelling tasks. On the theoretical side, our results on the characteristic function of MMAPs permit us to give conditions for the *uniqueness* of the martingale problem associated with the generator $\mathfrak{L}$ from (2.6) and hence for the *existence* of an MMAP.

Consider some $u \in \mathbb{C}^n$ and the function

$$(4.1) \qquad\qquad \mathcal{D}^Y \to \mathbb{C} \,:\, y \mapsto e^{\langle u, y \rangle}.$$

The function (4.1) is an element of $C_b(\mathcal{D}^Y)$ if and only if u belongs to the set

$$\mathcal{U} := \mathbb{C}_-^m \times i\mathbb{R}^{m-n}.$$

Below we discuss expectations of the form $\mathbb{E}_z\left[e^{\langle u, Y_t\rangle}\right]$ for $u \in \mathcal{U}$ and any $z \in \mathcal{D}, t \geq 0$; this includes the characteristic function of Y_t as a special case.

For an affine semigroup (P_t^{aff}), it holds that $P_t^{\text{aff}}e^{\langle u, y\rangle} = \exp(\phi(t,u) + \langle \psi(t,u), y\rangle)$ for $u \in \mathcal{U}$ and continuous ϕ and ψ. The derivatives of $\phi(\cdot, u)$ and $\psi(\cdot, u)$ at $t = 0$ are typically denoted by $F(u)$ and $R(u)$, respectively (their existence is shown in [37]). In what follows we introduce the Markov-modulated analogue of these derivatives. Consider functions $F : \mathcal{D}^X \times \mathcal{U} \to \mathbb{C}$ and $R : \mathcal{U} \to \mathbb{C}^n$ given by

$$
\begin{aligned}
F(x,u) = {} & \langle b(x), u\rangle + \langle a(x)u, u\rangle - c(x) \\
(4.2) \qquad & + \int_{\mathcal{D}^Y \setminus \{0\}} (e^{\langle u, \zeta\rangle} - 1 - \langle u_{\mathcal{J}}, \chi_{\mathcal{J}}(\zeta)\rangle) m(x, d\zeta),
\end{aligned}
$$

$$
(4.3) \qquad R_i(u) = \langle \alpha_i u, u\rangle + \langle \beta_i^+, u\rangle - \gamma_i + \int_{\mathcal{D}^Y \setminus \{0\}} (e^{\langle u, \zeta\rangle} - 1 - \langle u_{\mathcal{J}(i)}, \chi_{\mathcal{J}(i)}(\zeta)\rangle) \mu_i(d\zeta),
$$

$$
(4.4) \qquad R_{\mathcal{J}}(u) = \beta^* u^*.
$$

for $i \in \mathcal{I}$ and where

$$
\begin{aligned}
\beta_i^+ &:= (\beta^\top)_{i\{1,\ldots,n\}} \in \mathbb{R}^n, \quad i \in \mathcal{I}, \\
\beta^* &:= (\beta^\top)_{\mathcal{J}\mathcal{J}} \in \mathbb{R}^{(n-m)\times(n-m)}.
\end{aligned}
$$

Note that R is the same function as in the standard affine case, whereas F is modified (it depends on x). Moreover, F and R clearly separate the x-admissible parameters according to whether they depend on x or not. In the standard affine case, ψ is characterized by a system of ODEs, called generalized Riccati equations, and then $\Phi(t,u) = \exp(\phi(t,u))$ is simply given via the linear ODE $\partial_t \Phi(t,u) = \Phi(t,u)F(u)$. It will turn out that for an MMAP the transform $\mathbb{E}_z\left[e^{\langle u, Y_t\rangle}\right]$ has a similar structure, but the function $\Phi(t,u)$ is replaced by the solution $\varphi(t, x; u)$ of a Cauchy problem that involves the function F and the generator $\mathfrak{L}^X$ of X.

We now introduce the aforementioned system of generalized Riccati equations. Suppose $\psi : [0, \infty) \times \mathcal{U} \to \mathbb{C}^n$ solves

$$
(4.5) \qquad \partial_t \psi(t,u) = R(\psi(t,u)), \quad \psi(0,u) = u.
$$

The above generalized Riccati equations (4.5) for ψ are a well-studied mathematical object and [17, Section 6] prove the existence of a unique solution thereof. Moreover, Equation (4.5) with (4.4) obviously implies that $\psi_{\mathcal{J}}(t,u) = e^{\beta^* t} u^*$. Within our framework, the counterpart of the linear ODE for Φ is the following Cauchy problem, which plays a crucial role throughout this section,

$$
(\text{CP}_{T,u}) \quad \left\{
\begin{array}{l}
\partial_t \varphi(t, x; u) = \mathfrak{L}^X \varphi(t, x; u) + \varphi(t, x; u)F(x, \psi(t, u)), \quad t \in [0, T], \; x \in \mathcal{D}^X, \\
\varphi(0, x; u) = f(x), \quad f \in \text{Dom}(\mathfrak{L}^X),
\end{array}
\right.
$$

where $T > 0$ and $u \in \mathcal{U}$ are to be understood as exogenous parameters. We recall the notion of an important solution concept in the following

Definition 4.1. *A* classical solution *of* $(\text{CP}_{T,u})$ *is a* C^1*-function* $[0,T] \to \text{Dom}(\mathfrak{L}^X) \subset C_\infty(\mathcal{D}^X) : t \mapsto \varphi(t, \cdot; u)$ *satisfying* $(\text{CP}_{T,u})$.

Consider a measure $\mathbb{P}_z \in \mathbf{MP}(\mathfrak{L}, \text{Dom}(\mathfrak{L}))$. In the next result we derive the characteristic function of the one-dimensional distributions of the process Y under $\mathbb{P}_z$.

Theorem 4.2. *Let $z \in \mathcal{D}$, $T \geq t^\star \geq 0$ and $u \in \mathcal{U}$ be arbitrary and consider $\mathbb{P}_z \in \mathbf{MP}(\mathfrak{L}, \mathrm{Dom}(\mathfrak{L}))$. Suppose that for all $f \in \mathrm{Dom}(\mathfrak{L}^X)$ there exists a classical solution of Cauchy problem $(\mathrm{CP}_{T,u})$. Then it holds that*

$$(4.6) \qquad\qquad \mathbb{E}_z\left[e^{\langle u, Y_{t^\star}\rangle}\right] = \varphi^\star(t^\star, x; u)e^{\langle \psi(t^\star, u), y\rangle},$$

where $\psi(\cdot, u) : [0, \infty) \to \mathbb{C}^n$ solves the system of generalized Riccati equations (4.5) and where $\varphi^\star(t, x; u) = \mathbb{E}_z\left[\exp(\int_0^t F(X_s, \psi(t - s, u))ds)\right]$, $t \in [0, T]$.

Comments. The product structure in (4.6) is very convenient for computing Fourier-Laplace transforms. In a first step one computes the function $\psi(t, u)$ by solving the generalized Riccati equations (4.5), analogously as in the standard affine case. In a second step the function φ is either computed using Monte Carlo techniques or, under the added regularity of Proposition 4.4, as a solution of the Cauchy problem (4.9) below. The complexity of this step depends on the form of the generator $\mathfrak{L}^X$ (note that $\mathrm{Dom}(\mathfrak{L}^X) \subset \mathrm{Dom}(\widetilde{\mathfrak{L}}^X)$): if X is a finite state Markov chain, then (4.9) reduces to a set of ODEs; if X is a diffusion process, it is a linear PDE of parabolic type, etc.

Proof of Theorem 4.2. Fix an arbitrary $u \in \mathcal{U}$ and let $\{\rho^k\}_{k\in\mathbb{N}}$ be a sequence of functions $\rho^k \in C_c^\infty(\mathcal{D}^X)$, $0 \leq \rho^k \leq 1$ increasing pointwise to $\mathbb{1}_{\mathcal{D}^X}$. Use $\varphi^k(\cdot; u)$ to denote the solution of $(\mathrm{CP}_{T,u})$ with initial value $\varphi^k(t, x; u) = \rho^k(x)$ and note that there exists a unique solution to (4.5). Accordingly, introduce for each $k \in \mathbb{N}$ the function $g_u^k : [0, T] \times \mathcal{D} \to \mathbb{C}$ by

$$g_u^k(t, x, y) = \varphi^k(t, x; u)e^{\langle\psi(t,u),y\rangle}.$$

For any $u \in \mathcal{U}$ it holds that $\psi(t, u) \in \mathcal{U}$ for all $t \geq 0$ (cf. [17, Proposition 6.4]). Our aim is to show that for any fixed $0 \leq t^\star < T$ the process given by

$$(4.7) \qquad\qquad g_u^k(t^\star - t, X_t, Y_t), \quad t \in [0, T],$$

is a complex-valued martingale. Note that

$$(4.8) \quad \begin{aligned} &\mathfrak{L}^{Y|X}(x)g_u^k(t, x, y) \\ &\qquad = g_u^k(t, x, y)\left(F(x, \psi(t, u)) + \langle R(\psi(t, u)), y\rangle\right), \quad \forall(t, x, y) \in [0, \infty) \times \mathcal{D}. \end{aligned}$$

Using that $[0, T] \ni t \mapsto \varphi^k(t, \cdot; u)$ solves Cauchy problem $(\mathrm{CP}_{T,u})$ and the generalized Riccati equation (4.5) bring us to

$$\begin{aligned} &(\widetilde{\mathfrak{L}}^X + \mathfrak{L}^{Y|X}(x) - \partial_t)g_u^k(t, x, y) \\ &= e^{\langle\psi(t,u),y\rangle}(\widetilde{\mathfrak{L}}^X - \partial_t)\varphi^k(t, x; u) - \langle\partial_t\psi(t, u), y\rangle g_u^k(t, x, y) + \mathfrak{L}^{Y|X}(x)g_u^k(t, x, y) \\ &= -g_u^k(t, x, y)(F(x, \psi(t, u)) + \langle R(\psi(t, u)), y\rangle) + \mathfrak{L}^{Y|X}(x)g_u^k(t, x, y). \end{aligned}$$

Together with Equation (4.8), we consequently get

$$(\partial_t + \widetilde{\mathfrak{L}}^X + \mathfrak{L}^{Y|X}(x))g_u^k(t^\star - t, x, y) = 0, \quad \forall(t, x, y) \in [0, \infty) \times \mathcal{D}.$$

Thus, (4.7) is a martingale by Corollary 3.5.

Next, we resort to the Feynman-Kac Theorem (see e.g. [11, Theorem 17.4.10]) to obtain the representation $\varphi^k(t, x; u) = \mathbb{E}_{(x,y)}\left[\rho^k(X_t)\exp(\int_0^t F(X_s, \psi(t - s, u))ds)\right]$, where the value of the expectation is invariant with respect to y. Using monotone convergence and the conservativeness of (P_t^X), we see that $\lim_{k\to\infty}\varphi^k(t, x; u) = \varphi^\star(t, x; u)$. Consequently, L^1-convergence of $g_u^k(t^\star - t, X_t, Y_t)$ as $k \to \infty$ gives us that $\varphi^\star(t^\star - t, X_t; u)e^{\langle\psi(t^\star - t, u), Y_t\rangle}$, $0 \leq t < T$ is a martingale and taking expectation over (4.7) for $t = t^\star$ leads to

$$\mathbb{E}_z\left[g_u^k(0, X_{t^\star}, Y_{t^\star})\right] = \mathbb{E}_z\left[e^{\langle u, Y_{t^\star}\rangle}\right] = g_u(t^\star, x, y).$$

$\square$

Remark 4.3. From the above proof we see that the assumptions of Theorem 4.2 can in principle be weakened. In fact, the crucial condition to arrive at (4.6) is that $t \mapsto \varphi^\star(t^\star - t, X_t; u)e^{\langle \psi(t^\star - t, u), Y_t \rangle}$ is a martingale, which is guaranteed by the solvability of $(\mathrm{CP}_{T,u})$.

Proposition 4.4. *Let $\varphi^\star(\cdot; u)$ be the function studied in Theorem 4.2 and suppose that*

(i) *for any $x \in \mathcal{D}^X$, $\varphi^\star(\cdot, x; u) \in C^1([0, T])$,*

(ii) *for any $t \in [0, T]$, $F(\cdot, \psi(t, u))$, $\varphi^\star(t, \cdot; u)F(\cdot, \psi(t, u)) \in C_b(\mathcal{D}^X)$.*

Then $\varphi^\star(\cdot; u)$ solves the following Cauchy problem (involving $\widetilde{\mathfrak{L}}^X$ instead of $\mathfrak{L}^X$)

$$(4.9) \qquad \partial_t \varphi^\star(t, x; u) = \widetilde{\mathfrak{L}}^X \varphi^\star(t, x; u) + \varphi^\star(t, x; u)F(x, \psi(t, u)), \quad \varphi^\star(0, \cdot; u) \equiv 1,$$

Remark 4.5. Note that if all x-admissible parameters are constant, then (4.9) reduces to a linear ODE, and we are back in the standard affine case.

Proof of Proposition 4.4. We fix an arbitrary $t \in [0, T)$ and consider some small $s \in (0, T)$. Straightforward computations show that

$$\frac{P_s^X \varphi^\star(t, \cdot; u)(x) - \varphi^\star(t, x; u)}{s}$$
$$= \frac{\varphi^\star(t + s, x; u) - \varphi^\star(t, x; u)}{s}$$
$$+ \frac{1}{s} \mathbb{E}_{(x,y)}\left[e^{\int_0^{t+s} F(X_r, \psi(t+s-r, u))dr} \left(e^{-\int_0^s F(X_r, \psi(t+s-r, u))dr} - 1 \right) \right],$$

where the first term converges by assumption (i) to $\partial_t \varphi^\star(t, x; u)$ as $s \to 0+$. Regarding the second term, we yet again consider the limit as $s \to 0+$, and using dominated convergence along with the fact that X has càdlàg paths said limit is equal to

$$\mathbb{E}_{(x,y)}\left[\lim_{s \to 0+} e^{\int_0^{t+s} F(X_r, \psi(t+s-r, u))dr} \frac{1}{s} \left(e^{-\int_0^s F(X_r, \psi(t+s-r, u))dr} - 1 \right) \right]$$
$$= -\mathbb{E}_{(x,y)}\left[e^{\int_0^t F(X_r, \psi(t-r, u))dr} F(x, \psi(t, u)) \right]$$
$$= -\varphi^\star(t, x; u)F(x, \psi(t, u)),$$

and so we get by condition (ii) that $\varphi^\star(t, \cdot; u) \in \mathrm{Dom}(\widetilde{\mathfrak{L}}^X)$. $\qquad\square$

It is shown in [27, Section 4.4] (in particular Theorem 4.4.2), that if any two solutions for the martingale problem associated with $(\mathfrak{L}, \mathrm{Dom}(\mathfrak{L}))$ generate the same one-dimensional distributions of (X, Y), then the martingale problem is well-posed and (X, Y) has the strong Markov property. In fact, the one-dimensional distributions of (X, Y) can be characterized by similar arguments as in Theorem 4.2, which gives uniqueness for the martingale problem and hence establishes existence of MMAPs. Given their importance for our purposes, we summarize these results in the following

Theorem 4.6. *Suppose that the assumptions of Theorem 4.2 hold. Then, the martingale problem for $(\mathfrak{L}, \mathrm{Dom}(\mathfrak{L}))$ is well-posed and there exists a Markov-modulated affine process (X, Y) corresponding to $(\mathfrak{L}, \mathrm{Dom}(\mathfrak{L}))$. Moreover, (X, Y) has the strong Markov property.*

Proof. Fix arbitrary $u \in i\mathbb{R}^n$, $f \in \mathrm{Dom}(\mathfrak{L}^X)$ and $t \in [0, T]$. Similar to the proof of Theorem 4.2, we then introduce the process

$$\widetilde{g}_u^f(t - s, X_s, Y_s) := \varphi^f(t - s, X_s; u)e^{\langle \psi(t-s, u), Y_s \rangle}, \quad s \in [0, T],$$

where $\varphi^f(\cdot, x; u)$ solves $(\mathrm{CP}_{T,u})$ with $\varphi^f(0, x; u) = f(x)$, for any $x \in \mathcal{D}^X$. The process $\widetilde{g}_u^f(t - s, X_s, Y_s), s \in [0, T]$ is a martingale for any measure solving $\mathbf{MP}(\mathfrak{L}, \mathrm{Dom}(\mathfrak{L}))$. So, for arbitrary $z \in \mathcal{D}$, we pick any two $\mathbb{P}_z^1, \mathbb{P}_z^2 \in \mathbf{MP}(\mathfrak{L}, \mathrm{Dom}(\mathfrak{L}))$, where we denote the corresponding expectations with $\mathbb{E}_z^1$ and $\mathbb{E}_z^2$, respectively. We then have that

$$\mathbb{E}_z^1 \left[f(X_t) e^{\langle u, Y_t \rangle} \right] = \widetilde{g}_u^f(t, x, y) = \mathbb{E}_z^2 \left[f(X_t) e^{\langle u, Y_t \rangle} \right].$$

Since $u, t \geq 0$ and $f \in \mathrm{Dom}(\mathfrak{L}^X)$ were arbitrary, we conclude that the one-dimensional distributions of (X, Y) with fixed starting point z are the same under any solution to $\mathbf{MP}(\mathfrak{L}, \mathrm{Dom}(\mathfrak{L}))$ and with Theorem 4.4.2.a of [27] we arrive at the well-posedness of the martingale problem. The strong Markov-property of (X, Y) under $(\mathbb{P}_z)_{z \in \mathcal{D}} \subset \mathbf{MP}(\mathfrak{L}, \mathrm{Dom}(\mathfrak{L}))$ then follows from Theorem 4.4.2.c of [27]. $\square$

In light of the extensive class of possible models for X together with the weak assumptions on the x-admissible parameters, it is challenging to fully characterize solutions of the Cauchy problem $(\mathrm{CP}_{T,u})$. We are, however, in a position to formulate easy to verify sufficient conditions for the existence of solutions to $(\mathrm{CP}_{T,u})$ (and accordingly for the existence of MMAPs). These conditions depend solely on the regularity of the x-admissible parameters and do not require any further assumptions on $\mathfrak{L}^X$.

Proposition 4.7. *Suppose the parameters $(a, \alpha, b, \beta, c, \gamma, m, \mu)$ underlying $(\mathfrak{L}, \mathrm{Dom}(\mathfrak{L}))$ satisfy the following conditions,*

(i) for any $x \in \mathcal{D}^X$,

$$\int_{\mathcal{D}^Y \setminus Q_0} \|\zeta\| \, m(x, d\zeta) < \infty,$$

where $Q_0 = \{\zeta \in \mathcal{D}^Y \mid |\zeta_i| \leq 1, \, 1 \leq i \leq n\} \setminus \{0\}$,

(ii) $F(\cdot, u) \in C_\infty(\mathcal{D}^X)$ for all $u \in \mathcal{U}$.

Then for all $f \in \mathrm{Dom}(\mathfrak{L}^X)$ there exists a classical solution of Cauchy problem $(\mathrm{CP}_{T,u})$, and thus there exists a Markov-modulated affine process corresponding to $(\mathfrak{L}, \mathrm{Dom}(\mathfrak{L}))$.

Proof. Fix arbitrary $u \in \mathcal{U}$ and $T > 0$. Introduce the vector-valued function $\boldsymbol{G}$ via

$$\boldsymbol{G} : [0, T] \times C_\infty(\mathcal{D}^X) \to C_\infty(\mathcal{D}^X) : (t, \boldsymbol{\varphi}) \mapsto \boldsymbol{\varphi} F(\cdot, \psi(t, u)).$$

The moment condition in (i) ensures that $F(\cdot, u) \in C^1(\mathcal{D}^Y)$ (cf. [17, Lemma 5.3.ii]), so that $\boldsymbol{G}$ is C^1 in its first argument. Further, the function $\boldsymbol{G}$ is continuously Fréchet-differentiable in its second argument and due to condition (ii) the derivative takes values in $C_\infty(\mathcal{D}^X)$. Thus, by Theorem 6.1.5 of [48] there exists for any $f \in \mathrm{Dom}(\mathfrak{L}^X)$ a unique classical solution $[0, T] \ni t \mapsto \varphi(t, \cdot; u)$ to Problem $(\mathrm{CP}_{T,u})$. The result then follows from Theorem 4.6. $\square$

Solving the Cauchy problem is of course substantially simplified in specific applications with narrower model assumptions. For example, if X is a diffusion, then Problem $(\mathrm{CP}_{T,u})$ reduces to a PDE. In such a case, conditions for the existence of a classical solution to the above Cauchy problem are available in various degrees of generality, but are usually similar in character. The coefficients of $\mathfrak{L}^X$ typically need to be locally Lipschitz, grow at most linearly and suffice a local ellipticity assumption. Regarding the perturbation controlled by F, common assumptions translate to the boundedness of $\mathcal{D}^X \times [0, T] \ni (x, t) \mapsto F(x, \psi(t, u))$, and to the condition that $F(x, \psi(t, u))$ is locally Lipschitz (Hölder) in x (t). See e.g. [40, Chapter V] for further details. Some classical references on more general Cauchy problems are [48], [55], [26] or [57].

We now take up Example 2.6 and demonstrate how our results can be used to show existence of certain Markov-modulated affine processes whose coefficients do not obey the standard continuity assumptions frequently found in SDE theory.

Example 4.8. In light of Theorem 4.2 we have that

$$F(x, u) = \mathbb{1}_{\{x \geq 0\}} \frac{u^2}{2} + \mathbb{1}_{\{x < 0\}} \frac{e^{ux} - 1 - u\chi(x)}{x^2}.$$

Recall that in a neighborhood around zero it holds that $\chi(x) = x$, and so with an application of L'Hôpital's rule we get

$$\lim_{x \to 0-} \frac{e^{ux} - 1 - ux}{x^2} = \frac{u^2}{2},$$

for any $u \in \mathcal{U}$, so that F is continuous on $\mathbb{R} \times \mathcal{U}$. Let $f \in C_c^2(\mathcal{D}^X)$ be arbitrary, then $(\mathrm{CP}_{T,u})$ reduces to

$$(\partial_t - \partial_{xx})\varphi(t, x, u) = \varphi(t, x, u)F(x, u), \quad \varphi(0, x, u) = f(x),$$

which is a perturbed heat equation, for which exists a classical solution according to [48, Theorem 8.2.3]. Thus, an application of Theorem 4.6 shows the existence of a Markov process with properties outlined in Example 2.6.

5. Further properties

In this section we study further properties of MMAPs. In Section 5.1 we give conditions ensuring the Feller property of strongly regular Markov-modulated affine semigroups. Sections 5.2 and 5.3 discuss the semimartingale property of MMAPs and the existence of real exponential moments, respectively.

5.1. Feller Property

An alternative approach for studying the existence of (at least strongly regular) MMAPs, is to apply the theory of Hille-Yosida to show that $(\mathfrak{L}, \mathrm{Dom}(\mathfrak{L}))$ can be uniquely extended to a generator of a strongly continuous semigroup. Since the operator $\mathfrak{L}^X$ may be unbounded, this inherently analytical task is challenging. Using Theorem 3.4 and Theorem 4.2 we now show the Feller property of a strongly regular Markov-modulated affine semigroup by probabilistic arguments. Recall that a strongly regular MMAP has continuous x-admissible parameters (note that the Feller property requires that $\mathfrak{L}(\mathrm{Dom}(\mathfrak{L})) \subset C_\infty(\mathcal{D})$). Detailed discussions and results on the generator of Feller processes can be found in e.g. [52], [51], [53], [50, Chapter VII] or [39].

Proposition 5.1. *Suppose the assumptions of Theorem 4.2 hold and that the parameters $(a, \alpha, b, \beta, c, \gamma, m, \mu)$ underlying $(\mathfrak{L}, \mathrm{Dom}(\mathfrak{L}))$ are strongly x-admissible. Then the Markov-modulated affine process (X, Y) corresponding to $(\mathfrak{L}, \mathrm{Dom}(\mathfrak{L}))$ is Feller.*

Further, let $(L, \mathrm{Dom}(L))$ be the infinitesimal generator of the associated Feller semigroup. Then $\mathrm{Dom}(\mathfrak{L})$ is a core of L.

Proof. Theorem 4.6 gives the existence of a sub-Markov semigroup $(P_t)_{t \geq 0}$ corresponding to (X, Y). To show the strong continuity of $(P_t)_{t \geq 0}$, we rely on the classical result [50, Theorem III.2.4], according to which the property that $\lim_{t \to 0+} P_t f(z) = f(z)$, $\forall z \in \mathcal{D}$ and any $f \in C_\infty(\mathcal{D})$ is sufficient for strong continuity. We take up the definition of the set $\overline{\Theta}_0$ from the proof of Lemma 3.1 and pick an arbitrary $\tilde{f} = fh \in \overline{\Theta}_0$. We further fix arbitrary $T > 0$ and $u \in \mathcal{U}$. Similar to the proof of Theorem 4.2, we then introduce the process

$$\widetilde{g}_u^f(t - s, X_s, Y_s) := \varphi^f(t - s, X_s; u)e^{\langle \psi(t-s,u), Y_s \rangle}, \quad s \in [0, T],$$

where $\varphi^f(\cdot, x; u)$ solves $(\mathrm{CP}_{T,u})$ with $\varphi^f(0, x; u) = f(x)$, for any $x \in \mathcal{D}^X$. By a similar argument as in the proof of Theorem 4.2, the process $\widetilde{g}_u^f(t - s, X_s, Y_s)$, $s \in [0, T]$, is a martingale. So,

for any $(x, y) \in \mathcal{D}$, and appropriate $g \in C_c^\infty(\mathbb{R}^{n-m})$, $v \in \mathbb{C}^m_{--}$ (compare with the definition of $\overline{\Theta}_0$), it holds that

$$
\begin{aligned}
P_t \tilde{f}(x, y) &= \mathbb{E}_z \left[f(X_t) \int_{\mathbb{R}^{n-m}} e^{\langle (v, iq), (Y_t^+, Y_t^*) \rangle} g(q) dq \right] \\
&= \int_{\mathbb{R}^{n-m}} \mathbb{E}_z \left[\tilde{g}_{(v, iq)}^f (0, X_t, Y_t) \right] dq \\
&= \int_{\mathbb{R}^{n-m}} \varphi^f(t, x; v, iq) e^{\langle \psi(t, v, iq), y \rangle} g(q) dq,
\end{aligned}
$$

where the second equality follows from Fubini's theorem. Recall the definition of functions F and R in (4.2)–(4.4). Via an application of dominated convergence, we then obtain

$$
\begin{aligned}
\partial_t^+ P_t \tilde{f}(x, y)|_{t=0} &= \int_{\mathbb{R}^{n-m}} \partial_t^+ \left\{ \varphi^f(t, x; v, iq) e^{\langle \psi(t, v, iq), y \rangle} g(q) \right\} |_{t=0} dq \\
&= \int_{\mathbb{R}^{n-m}} \left(\mathfrak{L}^X \varphi^f(0, x; v, iq) + \varphi^f(0, x; v, iq) F(x, \psi(0, v, iq)) \right) e^{\langle \psi(0, v, iq), y \rangle} \\
&\quad + \langle R(\psi(0, v, iq)), y \rangle e^{\langle \psi(0, v, iq), y \rangle} \varphi^f(0, x; v, iq) g(q) dg \\
&= \int_{\mathbb{R}^{n-m}} \left(\mathfrak{L}^X f(x) + F(x, v, iq) + \langle R(v, iq), y \rangle f(x) \right) e^{\langle (v, iq), y \rangle} g(q) dq \\
&= \mathfrak{L} \tilde{f}(x, y),
\end{aligned}
$$

which implies bp-$\lim_{t \to 0+} P_t \tilde{f} = \tilde{f}$. The remaining arguments are identical to the ones from the proofs of [17, Proposition 8.2] and [28, Proposition 6.3] by noting that the linear hull of $\overline{\Theta}_0$ is dense in $C_\infty(\mathcal{D})$, and so we conclude that $(P_t)_{t \geq 0}$ is a Feller semigroup. Finally, the well-posedness of $\mathbf{MP}(\mathfrak{L}, \mathrm{Dom}(\mathfrak{L}))$ (see Theorem 4.6) implies that $\mathrm{Dom}(\mathfrak{L})$ is a core of L, cf. [47]. □

In summary, the continuity of x-admissible parameters and the conditions from Proposition 4.7 are sufficient for (X, Y) to enjoy the Feller property.

5.2. Semimartingale Property

MMAPs can in general explode and get killed, both of which prompt a transition to the cemetery state $\{\Delta\}$. To formalize these concepts we introduce various stopping times; our definitions follow [10]. First, we define the lifetime of Y as

$$
T_\Delta := \inf\{t \geq 0 \mid Y_t = \Delta\}.
$$

Second, we let $\|\Delta\| = \infty$ and we introduce the following sequences of stopping times,

$$
T_k' := \inf\{t \geq 0 \mid \|Y_{t-}\| \geq k \text{ or } \|Y_t\| \geq k\}, \quad k \in \mathbb{N},
$$

$$
T_k := \begin{cases} T_k', & \text{if } T_k' < T_\Delta, \\ \infty, & \text{if } T_k' = T_\Delta, \end{cases} \quad k \in \mathbb{N}.
$$

Moreover, we define the *explosion time* of Y as $T_\infty := \lim_{k \to \infty} T_k$. Then the process Y explodes if and only if $T_\infty < \infty$; killing corresponds to the event $T_\Delta < T_\infty$ and a trajectory of Y belongs to $\mathcal{D}^Y$ for all $t \geq 0$ if and only if $T_\Delta = T_\infty = \infty$. By construction, as $(T_k \wedge k)$ announces T_∞, the explosion time is predictable, which allows us to stop Y at a time before T_∞ and to work on $[0, T_\infty)$ for studying the semimartingale property.

Proposition 5.2. *Let (X, Y) be a Markov-modulated affine process with x-admissible parameters $(a, \alpha, b, \beta, c, \gamma, m, \mu)$, where $c(\cdot) \equiv 0$ and $\gamma = 0$. Then $\mathbb{P}_z(0 < T_\Delta < T_\infty) = 0$ (no killing) and the process Y is a $\mathcal{D}^Y$-valued semimartingale on $[0, T_\infty)$.*

If in addition $(a, \alpha, b, \beta, c, \gamma, m, \mu)$ *are strongly x-admissble, then* Y *admits the* $\{\mathcal{F}_t\}$-*characteristics* (B, C, ν) *with respect to the truncation function* χ. *Here*

$$(5.1) \qquad B_t = \int_0^t \mathbb{1}_{\{t < T_\infty\}} (b(\tilde{X}_s) + \tilde{\beta} Y_s) ds,$$

$$(5.2) \qquad C_t = 2 \int_0^t \mathbb{1}_{\{t < T_\infty\}} \left(a(X_s) + \sum_{i=1}^m \alpha_i Y_s^{+,i} \right) ds,$$

$$(5.3) \qquad \nu(dt, d\xi) = \mathbb{1}_{\{t < T_\infty\}} \left(m(X_t, d\xi) + \sum_{i=1}^m Y_t^{+,i} \mu_i(d\xi) \right) dt,$$

where the function $\tilde{b} : \mathcal{D}^X \to \mathcal{D}^Y$ *is given by* $\tilde{b}(x) = b(x) + \int_{\mathcal{D}^Y \setminus \{0\}} (\chi_{\mathcal{I}}(\xi), 0) m(x, d\xi)$, *and where*

$$\tilde{\beta}_{kl} := \begin{cases} \beta_{kl} + (1 - \delta_{kl}) \int_{\mathcal{D}^Y \setminus \{0\}} \chi_k(\xi) \mu_l(d\xi), & \text{if } l \in \mathcal{I}, \\ \beta_{kl}, & \text{if } l \in \mathcal{J}, \text{ for } 1 \le k \le n. \end{cases}$$

Proof. At first we show that $\mathbb{P}_z(0 < T_\Delta < T_\infty) = 0$. We use a similar argument as in the proof of [10, Lemma 3.1]. Take a sequence $(g_l) \in C_c^\infty(\mathcal{D}^Y)$ with $0 \le g_l \le 1$ and $g_l = 1$ on $D_l := \{y \in \mathcal{D}^Y \mid \|y\| < l\}$. Note that $g_l(y) = 1$ for l large enough and that $\lim_{l \to \infty} g_l(Y_{t \wedge T_k}) = \mathbb{1}_{\{t \wedge T_k < T_\Delta\}}$ $\mathbb{P}_z$-a.s.. For each l we obviously have that $g_l \in \mathrm{Dom}(\mathfrak{L})$, where we tacitly think of $(x, y) \mapsto g_l(y)$ as a constant function in x. Since X is conservative, we have $\widetilde{\mathfrak{L}}^X 1 = 0$ so that the process given by

$$M_t^l := g_l(Y_t) - g_l(y) - \int_0^t \mathfrak{L}^{Y|X}(x) g_l(Y_s) ds, \quad t \ge 0,$$

is a martingale. Take $l > k$ and recall that $c \equiv 0$ and $\gamma = 0$. It then holds that

$$M_{t \wedge T_k}^l = g_l(Y_{t \wedge T_k}) - g_l(y) - \int_0^{t \wedge T_k} \Big(\int_{\mathcal{D}^Y \setminus D_{l-k}} (g_l(Y_s + \zeta) - 1) m(X_s, d\zeta)$$
$$+ \sum_{i=1}^m \int_{\mathcal{D}^Y \setminus D_{l-k}} (g_l(Y_s + \zeta) - 1) Y_s^i \mu_i(d\zeta) \Big) ds.$$

Note that for any $i \in \mathcal{I}$ it holds that

$$(5.4) \qquad \int_0^{t \wedge T_k} \int_{\mathcal{D}^Y \setminus D_{l-k}} |g_l(Y_s + \zeta) - 1| Y_s^i \mu_i(d\zeta) ds \le k \int_0^{t \wedge T_k} \mu(\mathcal{D}^Y \setminus D_{l-k}) ds,$$

which converges almost surely to 0 as $l \to \infty$ by Lebegue's dominated convergence theorem. A similar argument also applies with $m(X_s, \cdot)$. Moreover, for fixed $k \le l$, with inequalities of type (5.4) it is easily seen that almost surely

$$|M_{t \wedge T_k}^l| \le 1 + c_k t,$$

for some constant c_k. Thus, $-M_{t \wedge T_k}^l$ converges in L^1 to $1 - \mathbb{1}_{\{t \wedge T_k < T_\Delta\}} = \mathbb{1}_{\{0 < T_\Delta \le t \wedge T_k\}}$ as $l \to \infty$ for all $t \ge 0$, which shows that $\mathbb{1}_{\{0 < T_\Delta \le \cdot \wedge T_k\}}$ is a martingale. However, as it is non-increasing it needs to be almost surely constant, which implies $\mathbb{P}_z(0 < T_\Delta < T_\infty) = 0$.

Next, for $u \in \mathbb{R}^n$ we introduce the process

$$A(u)_t := i \langle u, B_t \rangle - \frac{1}{2} \langle u, C_t u \rangle + \int \left(e^{i \langle u, y \rangle} - 1 - i \langle u, \chi(y) \rangle \right) \nu([0, t] \times dy), \quad t \in [0, T_\infty),$$

where B, C and ν are as in (5.1)–(5.3). Note that $A(u)_t = \int_0^t F(X_s, u) + \langle R(u), Y_s\rangle ds$, where F and R are specified in (4.2) and (4.3)–(4.4), respectively. The process $e^{i\langle u, Y\rangle} - \int_0^\cdot \mathfrak{L}^{Y|X}(X_s)e^{i\langle u, Y_s\rangle} ds$ is a martingale on $[0, T_\infty)$ by Corollary 3.5. And, for $t \in [0, T_\infty)$

$$e^{i\langle u, Y_t\rangle} - \int_0^t \mathfrak{L}^{Y|X}(X_s)e^{i\langle u, Y_s\rangle} ds = e^{i\langle u, Y_t\rangle} - \int_0^t e^{i\langle u, Y_s\rangle}\left(F(X_s, u) + \langle R(u), Y_s\rangle\right) ds$$

$$= e^{i\langle u, Y_t\rangle} - \int_0^t e^{i\langle u, Y_s\rangle} dA(u)_s,$$

and the result follows from Theorem II.2.42 of [34]. $\qquad\square$

We continue with a few implications of Proposition 5.2. First, take the characteristics (B, C, ν) from Proposition 5.2, set $J_t(\chi) := \sum_{s\leq t}(\Delta Y_s - \chi(\Delta Y_s))$ and denote the integer-valued random measure counting the jumps of Y by $\pi(ds, d\zeta)$. Then, we obtain the following canonical decomposition of Y,

$$(5.5)\qquad Y = Y_0 + Y^c + \int_0^\cdot \int_{\mathcal{D}^Y \setminus \{0\}} \chi(\zeta)\left(\pi(ds, d\zeta) - \nu(ds, d\zeta)\right) + J(\chi) + B,$$

where Y^c denotes the continuous local martingale part of Y, so that $\langle Y^{i,c}, Y^{j,c}\rangle = C^{ij}$ for $i, j \in \mathcal{I} \cup \mathcal{J}$. Second, as a conservative Feller process with càdlàg sample paths, X is a semimartingale (cf. Lemma 3.2 of [53]). Hence under the assumptions of Proposition 5.2, we have that $Z = (X, Y)$ is a semimartingale on $[0, T_\infty)$.

5.3. Exponential moments

In this section we resort to classical semimartingale methods in order to study real exponential moments of Y_t. For this we need a variant of $(\text{CP}_{T,u})$, namely

$$(\text{CP}^\star_{T,u})\quad \begin{cases} \partial_t \widetilde{\varphi}(t, x; u) = \widetilde{\mathfrak{L}}^X \widetilde{\varphi}(t, x; u) + \widetilde{\varphi}(t, x; u)F(x, \widetilde{\psi}(t, u)), & t \in [0, T],\ x \in \mathcal{D}^X, \\ \widetilde{\varphi}(0, \cdot; u) \equiv 1, \end{cases}$$

with $T > 0$ and where F is as in Theorem 4.2, albeit with the modification that $u \in \mathbb{R}^d$, and where $\widetilde{\psi}$ is analogously a modification of ψ, that is $\widetilde{\psi} : [0, T] \times \mathbb{R}^n \to R^n$ solves

$$(5.6)\qquad\qquad\qquad \partial_t \widetilde{\psi}(t, u) = R(\widetilde{\psi}(t, u)), \quad \widetilde{\psi}(0, u) = u.$$

We define a classical solution of $(\text{CP}^\star_{T,u})$ analogously to the previous case of $(\text{CP}_{T,u})$.

Proposition 5.3. *Let $z \in \mathcal{D}$, $T \geq t^\star \geq 0$ and $u \in \mathbb{R}^n$ be arbitrary and consider $\mathbb{P}_z \in \mathbf{MP}(\mathfrak{L}, \text{Dom}(\mathfrak{L}))$. Further, let (X, Y) be a Markov-modulated affine process, where in addition Y is a $\mathcal{D}^Y$-valued semimartingale with $\{\mathcal{F}_t\}$-characteristics (B, C, ν) given by (5.1)–(5.3). Suppose that $\widetilde{\psi}(\cdot, u)$ is C^1 and solves (4.5) and that the function $[0, T] \to C_b(\mathcal{D}^X) : t \mapsto \widetilde{\varphi}(t, \cdot; u)$ is a classical solution of $(\text{CP}^\star_{T,u})$. Suppose the parameters $(a, \alpha, b, \beta, c, \gamma, m, \mu)$ are strongly x-admissible and further satisfy the following conditions,*

(i) $c(\cdot) \equiv 0$ and $\gamma = 0$,

(ii) $\int_{\{|\zeta|>1\}} e^{\langle \psi(t,u),\zeta\rangle} m(x, d\zeta) < \infty,\ \forall (t, x) \in [0, T] \times \mathcal{D}^X$,

(iii) $\int_{\{\zeta_k>1\}} \zeta_k e^{\langle \psi(t,u),\zeta\rangle} \mu_l(d\zeta) < \infty,\ 1 \leq k, l \leq m, \forall t \in [0, T]$.

Then,

$$\mathbb{E}_z\left[e^{\langle u, Y_{t^\star}\rangle}\right] = \widetilde{\varphi}(t^\star, x; u)e^{\langle\widetilde{\psi}(t^\star, u), y\rangle}.$$

Proof. Fix an arbitrary $t^\star \in [0, T]$ and use $\mathcal{F}_{t^\star}^X$ to denote the $\mathbb{P}_z$-completion of $\sigma(X_s : s \in [0, t^\star])$. Conditional on $\mathcal{F}_t^X$, Y_t is distributed as a time-inhomogeneous affine semimartingale evaluated at $t^\star$, the exponential moments of which are handled in [35]. Within our setup, we readily fulfill the conditions 1, 2, 4 and 5 of [35, Theorem 5.1]. Regarding condition 3, we introduce the function $\varphi^0 : [0, T] \times \mathbb{R}^d \times D_{\mathcal{D}^x} \to \mathbb{R}$ via

$$\varphi^0(t, u, f) = \int_0^t F(f(s), \widetilde{\psi}(t - s, u)) ds.$$

Then, an application of [35, Theorem 5.1] gives that

$$\mathbb{E}_z \left[e^{\langle u, Y_{t^\star} \rangle} \mid \mathcal{F}_{t^\star}^X \right] = e^{\varphi^0(t^\star, u, X) + \langle \widetilde{\psi}(t^\star, u), y \rangle}.$$

We need to show that $\mathbb{E}_z \left[e^{\varphi^0(t^\star, u, X)} \right] = \widetilde{\varphi}(t^\star, x; u)$ in order to conclude the proof. To do so, note that by assumption $\widetilde{\varphi}$ solves ($\text{CP}_{T,u}^\star$) and by an application of Corollary 3.5 we get the martingality of

$$M_t^{\widetilde{\varphi}} := \widetilde{\varphi}(t^\star - t, X_t; u) + \int_0^t \widetilde{\varphi}(t^\star - s, X_s; u) F(X_s, \widetilde{\psi}(t^\star - s, u)) ds, \quad t \in [0, T].$$

Integration by parts and collecting terms gives us

$$\widetilde{\varphi}(t^\star - t, X_t; u) \exp \left(\int_0^t F(X_s, \widetilde{\psi}(t^\star - s, u)) ds \right)$$

(5.7)
$$= \int_0^t \exp \left(\int_0^s F(X_r, \widetilde{\psi}(t^\star - r)) dr \right) dM_s^{\widetilde{\varphi}}.$$

The local martingale in (5.7) is in fact a true martingale since the integrand is bounded and since $M^{\widetilde{\varphi}}$ is a square integrable martingale (cf. [49, Theorem IV.11]). So, by rearranging and taking expectations we arrive at the desired result. $\qquad \square$

The general results of the appendix of [17] provide further conditions for the characterization of exponential moments of Y_t. To begin with, let $U \subset \mathbb{C}^n$ be an open neighbourhood of 0. If in addition to the assumptions of Theorem 4.2, $\varphi(t, \cdot)$ and $\psi(t, \cdot)$ have analytic extensions on $U \times \mathcal{D}^X$ and U, respectively, then the exponential moments $\mathbb{E}_{(x,y)} \left[e^{\langle u, Y_t \rangle} \right]$ are finite for all $u \in U \cap \mathbb{R}^n, (x, y) \in \mathcal{D}^X \times \mathcal{D}^Y, 0 < t \leq T$. Moreover, under the same conditions it also holds that

$$\mathbb{E}_{(x,y)} \left[e^{\langle u, Y_t \rangle} \right] = \varphi(t, x; u) e^{\langle \psi(t, u), y \rangle},$$

for all $u \in U$ such that $\text{Re}(u) \in U \cap \mathbb{R}^n, (x, y) \in \mathcal{D}^X \times \mathcal{D}^Y$. As the function ψ, or more precisely the system of generalized Riccati equations defining it, is the same as in the standard affine case, we can outsource the study of suitable extensions for ψ by referencing existing results, most notably the systematic treatment of exponential moments of affine processes in [36]. The study of regularity properties of $u \mapsto \varphi(t, x; u)$ on the other hand is not straightforward as φ is defined via the rather general Cauchy problem ($\text{CP}_{T,u}$).

6. Applications in Mathematical Finance

In this section we discuss applications of MMAPs in mathematical finance. Our goal is to illustrate the wide range of modelling possibilities offered by this class of processes. We begin with an extension of Theorem 4.2 that enables us to solve many pricing problems with Fourier inversion techniques. Moreover, we consider in detail a model for the joint pricing of bonds, equity options and credit derivatives that captures many stylized facts of financial data and we sketch several further applications. Throughout this section we fix some $z \in \mathcal{D}$ and assume that (X, Y) is a Markov-modulated affine semimartingale under $\mathbb{P}_z$.

6.1. Derivative pricing via Fourier inversion techniques

In many financial applications one needs to evaluate expectations of the form

$$(6.1) \qquad Q_t f(z) := \mathbb{E}_z \left[\exp\left(\int_0^t L(Y_s) ds \right) f(X_t, Y_t) \right], \quad t \geq 0, \ f \in B_b(\mathcal{D}),$$

for an affine transformation $L : \mathcal{D}^Y \to \mathbb{R}$, $y \mapsto l + \langle \lambda, y \rangle$, for some $l \in \mathbb{R}$, $\lambda \in \mathbb{R}^n$. This can be done efficiently via Fourier inversion techniques. For this we need to compute for all $(q, u) \in i\mathbb{R}^d \times i\mathbb{R}^n$ the expectation $Q_t f^{(q,u)}(z)$ where $f^{(q,u)}(x, y) = e^{\langle (q,u),(x,y) \rangle}$. As in the standard affine case, this can be achieved using simple modifications of the functions F and R (see (4.2)–(4.4)). Replace $F(x, u)$ by $\widetilde{F}(x, u) = F(x, u) + l$ and $R(u)$ by $\widetilde{R}(u) = R(u) + \lambda$ in the equations of Theorem 4.2, fix some $t > 0$, pick arbitrary $(u, q) \in \mathcal{U} \times i\mathbb{R}^d$ and suppose $\widetilde{\varphi}$ is a classical solution of

$$(\partial_s - \widetilde{\mathfrak{L}}^X)\widetilde{\varphi}(s, x, u) = \widetilde{\varphi}(s, x, u)\widetilde{F}(x, \widetilde{\psi}(s, u)), \quad \widetilde{\varphi}(0, x, u) = e^{\langle q, x \rangle}, \ \forall x \in \mathcal{D}^X,$$

with

$$\partial_s \widetilde{\psi}(s, u) = \widetilde{R}(\widetilde{\psi}(s, u)), \quad \widetilde{\psi}(0, u) = u.$$

Then, a similar argument as in the proof of Theorem 4.2 shows that

$$(6.2) \qquad \mathbb{E}_z \left[\exp\left(\int_0^t L(Y_s) ds \right) e^{\langle (q,u),(X_t, Y_t) \rangle} \right] = \widetilde{\varphi}(t, x, u) e^{\langle \widetilde{\psi}(t,u), y \rangle}.$$

Suppose there exists analytic extensions of $\widetilde{\varphi}(t, x, \cdot)$ and $\widetilde{\psi}(t, \cdot)$ to some $u \in \mathbb{C}^n \backslash \mathcal{U}$, then the transform formula (6.2) also holds for this particular u.

  If $\mathbb{P}_z$ is a risk-neutral pricing measure, then expectations in the form of Equation (6.1) are typically understood as the price of a financial instrument with payoff $f(X_t, Y_t)$ and with risk-free short rate $-L(Y_s)$. Suppose Y_t represents the value of some underlying and we are interested in the price of a derivative with payoff $h(Y_t)$, then it is common to find a representation of h in the form of

$$h(y) = \int_{\mathbb{R}^q} e^{\langle u + iAv, y \rangle} \tilde{h}(v) dv,$$

for some matrix $A \in \mathbb{R}^{n \times q}$ with $q \leq n$, for an appropriate integration kernel $\tilde{h} : \mathbb{R}^q \to \mathbb{C}$, and where u is chosen such that $\mathbb{E}_z \left[e^{\langle u, Y_t \rangle} \right] < \infty$. The integration kernel $\tilde{h}$ can be determined by Fourier-inverting h. For many popular payoff functions, such inversions are explicitly available, see e.g. [29, Section 10.3.1] for univariate examples or [20] (and the references therein) for multivariate payoffs. In such a setup, $Q_t h$ reduces to

$$(6.3) \qquad \int_{\mathbb{R}^q} \widetilde{\varphi}(t, x, u + iAv) e^{\langle \widetilde{\psi}(t, u + iAv), y \rangle} \tilde{h}(v) dv,$$

which is easily computed using numerical integration.

6.2. Joint Pricing of Equity Options and Credit Derivatives

Next we show how Markov-modulated affine diffusions can be used for the joint and dynamically consistent valuation of equity and credit derivatives written on the same underlying firm. More precisely, we consider a firm that may default at some doubly stochastic default time τ (see e.g. [46, Chapter 10.6]) and we introduce a model for the joint dynamics of the risk-free short rate r, of the hazard rate γ of τ, of the pre-default logarithmic stock price p and of the stock price volatility v. In this section we view $\mathbb{P}_z$ as risk-neutral pricing measure. One salient feature of our model is that it allows for negative association between the risk-free short rate and the firm's hazard rate as well as its volatility process. Modelling negative dependence between *positive* diffusions (i.e. processes of 'CIR-type') is not feasible within the standard affine framework, see e.g. [18]. Here we overcome this well-known drawback of affine processes rather naturally via the common dependence of the positive processes on the modulating process X.

The Model. We suppose that X is a solution of the following SDE,

$$(6.4) \qquad\qquad dX_t = -\kappa(X_t - \theta)dt + \sigma\sqrt{X_t(1 - X_t)}dW_t^X,$$

for some standard Brownian motion W^X and with parameters $\theta \in [0, 1], \kappa, \sigma \geq 0$. It follows that X is a Jacobi process on the state space $\mathcal{D}^X = [0, 1]$ with

$$\mathfrak{L}^X f(x) = -\kappa(x - \theta)\frac{\partial f(x)}{\partial x} + \frac{1}{2}\sigma^2(x(1 - x))\frac{\partial^2 f(x)}{\partial x^2}, \quad f \in \mathrm{Dom}(\mathfrak{L}^X) \subset C_c^2(\mathcal{D}^X).$$

The risk-free short rate r and the firm's hazard rate γ are given as weak solution of the following SDEs,

$$(6.5) \qquad\qquad dr_t = \left(\bar{b}_r + b_r X_t + \beta_r r_t\right) dt + \sqrt{2\alpha_r r_t}dW_t^r,$$

$$(6.6) \qquad\qquad d\gamma_t = \left(\bar{b}_\gamma + b_\gamma(1 - X_t) + \beta_\gamma \gamma_t\right) dt + \sqrt{2\alpha_\gamma \gamma_t}dW_t^\gamma,$$

for independent one-dimensional Brownian motions W^r and W^γ and with constants $\alpha_r, \alpha_\gamma > 0$, $\bar{b}_r, b_r, \bar{b}_\gamma, b_\gamma \geq 0$ and $\beta_\gamma, \beta_r \in \mathbb{R}$.

We consider the following extension of the Heston model for the pre-default dynamics of the logarithmic stock price p and its instantaneous variance v

$$(6.7) \qquad\qquad dp_t = \left(r_t + \gamma_t - \frac{1}{2}v_t\right) dt + \sqrt{v_t}dW_t^p,$$

$$(6.8) \qquad\qquad dv_t = \left(\bar{b}_v + b_v(1 - X_t) + \beta_v v_t\right) dt + \sqrt{2\alpha_v v_t}dW_t^v,$$

where W^p and W^v are two Brownian motions with correlation $\rho \in [-1, 1]$. Moreover, we have constants $\alpha_v > 0, \bar{b}_v, b_v \geq 0$ and $\beta_v < 0$. We assume that at τ the stock price jumps to zero, that is we assume that $S_t = \mathbb{1}_{\{\tau > t\}}e^{p_t}$. This gives the following stock price dynamics

$$(6.9) \qquad dS_t = r_t S_t dt + \sqrt{v_t}S_t dW_t^p - S_{t-}dM_t^\tau \quad \text{with } M_t^\tau = \mathbb{1}_{\{\tau \leq t\}} - \int_0^{\tau \wedge t} \gamma_s ds.$$

Note that S is a martingale which justifies the interpretation of $\mathbb{P}_z$ as risk neutral measure. The exogenous process X can be viewed as state of the economy, with values close to 1 (0) representing expansion (recession). Note that the drift of r is increasing in X_t while the drift of γ is decreasing in X_t. The opposite impact of X on the drift of γ and r allows to model negative association between the short rate and the hazard rate. This phenomenon is empirically well documented, cf. [42] or [16]; it might be due to the fact that in times of financial crisis where hazard rates are high, central banks typically lower the reference rates. Moreover, the proposed specification creates natural dependence between the risk-neutral probability of default and the return distribution of p: as the economic environment worsens (X decreases), the risk of default increases, as does the volatility of the stock returns. Stock option pricing with credit risk is for example treated in [7] or [9].

Mathematical aspects. Next we show that the dynamics (6.4), (6.5), (6.7), (6.8) fit into our framework of Markov-modulated affine processes. First note that X is a polynomial process (see [14]) with compact support, and as such it is Feller (cf. Proposition 4.1 of [1]). The coefficients of X satisfy the regularity assumptions of Theorem 8.2.1 in [27], so that $C_c^\infty(\mathcal{D}^X)$ is a core of $(\mathfrak{L}^X, \mathrm{Dom}(\mathfrak{L}^X))$. Overall, Assumption 2.1 is met. Next we let $Y = (Y^1, Y^2, Y^3, Y^4)^\top = (r, \gamma, v, p)^\top$ so that $\mathcal{D}^Y = \mathbb{R}_+^3 \times \mathbb{R}$ and $m = 3, n = 1$. Set

$$\alpha_1 = \begin{pmatrix} \alpha_r & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad \alpha_2 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & \alpha_\gamma & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad \alpha_3 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & \alpha_v & \sqrt{\alpha_v}\rho \\ 0 & 0 & \sqrt{\alpha_v}\rho & 1/2 \end{pmatrix},$$

$$b(x) = \begin{pmatrix} \bar{b}_r + b_r x \\ \bar{b}_\gamma + b_\gamma(1-x) \\ \bar{b}_v + b_v(1-x) \\ 0 \end{pmatrix}, \quad \beta = \begin{pmatrix} \beta_r & 0 & 0 & 0 \\ 0 & \beta_\gamma & 0 & 0 \\ 0 & 0 & \beta_v & 0 \\ 1 & 1 & -1/2 & 0 \end{pmatrix},$$

and let $\alpha = (\alpha_1, \alpha_2, \alpha_3)$. With these definitions the dynamics of Y are governed by the x-admissible parameters $(0, \alpha, b(x), \beta, 0, 0, 0, 0)$.

The Cauchy problem $(\mathrm{CP}_{T,u})$ reduces to the following second order PDE,

$$\partial_t \varphi(t,x,u) = -\kappa(x-\theta)\partial_x \varphi(t,x,u) + \frac{1}{2}\sigma^2(x(1-x))\partial_{xx}\varphi(t,x,u)$$
$$+ \varphi(t,x,u)(\langle b(x), \psi(t,u)\rangle + \langle a(x)\psi(t,u), \psi(t,u)\rangle), \quad u \in \mathcal{U},$$
$$\varphi(0,x,u) = f(x), \quad f \in C^2([0,1]).$$

The existence of a classical solution to the above PDE follows from Proposition 4.7.

Derivative pricing. We work on an enlarged filtration $\mathcal{G}_t = \mathcal{F}_t \vee \sigma(\mathbb{1}_{\{\tau \leq s\}} : 0 \leq s \leq t)$, $t \geq 0$ where $(\mathcal{F}_t)$ is the filtration generated by the processes $Z = (X, Y)$, so that $(\mathcal{G}_t)$ contains information regarding the occurrence of default. Here we discuss the pricing of so-called *survival claims* with payoff $H_T = \mathbb{1}_{\{\tau > T\}}h(Y_T)$; simple examples would be a call option on the stock where $h(Y_T) = (e^{p_T} - K)^+$ or a defaultable bond with zero recovery where $h(Y_T) = 1$. Recall that $\mathbb{P}$ is the risk neutral measure. Hence the price of a survival claim is given by

$$\mathbb{E}_z\left[\mathbb{1}_{\{\tau > T\}}e^{-\int_0^T r_s ds}h(Y_T)\right] = \mathbb{E}_z\left[e^{-\int_0^T (r_s + \gamma_s)ds}h(Y_T)\right],$$

where the equality follows from standard results for stochastic hazard rate models (see e.g. [46, Chapter 10.6]). The right hand side can then be computed via Fourier pricing using pricing can then be done via Fourier pricing and evaluation of relation (6.3). In our model the function $\widetilde{\psi}$ is explicitly available (as in the standard Heston model) as is the transform $\widetilde{h}$ for several popular payoff functions. The one-dimensional PDE for $\widetilde{\varphi}$ can be efficiently solved with numerical methods such as [4]. Alternatively, the theory of BSDEs provides a probabilistic dual perspective of such PDEs, which is studied in [23, Section 4]. More recently, [32] use such a BSDE approach to design deep learning methods to efficiently solve semilinear PDEs. Finally, the integral in (6.3) can be handled with standard Fourier methods as outlined in [8], [41] or [2].

Extensions. Methods for pricing credit derivatives where the payment occurs directly at τ (as in the case of a credit default swap) with a transform formula such as (6.2) are presented in [31]. The above example allows for an immediate extension to a multivariate setting with hazard rates $\gamma^1, \ldots, \gamma^m$, where for each $i \in \mathcal{I}$ the process γ^i follows a model of type (6.5) (albeit with different parameters and with Brownian motions independent of each other). Such a specification provides a useful framework for the analysis of portfolio credit derivatives. Conditional on X, the hazard rates are independent and default dependence amongst the individual firms is generated via their respective loadings on the common factor process X. [31] use a similar setup with X replaced by a finite-state Markov chain for the pricing and analysis of European safe bonds.

6.3. Further Applications

We end this section by sketching two further possible use cases of MMAPs. The first example deals with Markov-modulated Lévy processes and the second one with Markov-modulated Hawkes processes.

Subordination techniques are a popular approach to build tractable multivariate models for financial assets (see e.g. [44] or [45]). However, the flexibility of subordinating Lévy processes

is limited by the fact that the resulting process is typically again of Lévy type. Instead, by using MMAPs it is feasible to construct a larger class of processes, even if the underlying modulated process is conditionally of Lévy type. Additionally, with our class of processes we can precisely target certain aspects of the joint multivariate distribution. As an example, we consider $Y \in \mathbb{R}^n$ as a model for the log prices of n assets and where we aim at jointly controlling the tail behaviour of the n assets solely via their driving process X. We assume that for each $i \in \{1, \ldots, n\}$ the entry Y^i is a pure jump process, and ϱ_i is a random measure selecting the stochastic jump times and sizes of Y^i, that is $\sum_{0 < s \leq t}(Y_s^i - Y_{s-}^i) = \int_0^t \int_{\mathbb{R} \setminus \{0\}} \zeta \varrho_i(d\zeta, ds)$, for $t \geq 0$. The compensator of $\varrho_i(d\zeta, ds)$ is $m_i(X_t, d\zeta)dt$, where X is some exogenous Markov process satisfying Assumption 2.1. We further assume that

$$m^i(x, d\zeta) = \frac{C_i}{|\zeta|^{1+\tilde{Y}_i}} \left(\mathbb{1}_{\{\zeta < 0\}} e^{-G_i(x)|\zeta|} + \mathbb{1}_{\{\zeta > 0\}} e^{-M_i(x)|\zeta|} \right) d\zeta, \quad x \in \mathcal{D}^X,$$

with constants $C_i > 0$, $\tilde{Y}_i < 2$ (note that C_i and $\tilde{Y}_i$ can of course also depend on x, but for the purpose of this example we choose to work with the given simplified setup) and $G_i, M_i \in C_b(\mathcal{D}^X)$ with $G_i(x), M_i(x) \geq 0$ for each x. For fixed x, $m(x, d\zeta)$ is the Lévy kernel of a CGMY process (cf. [6]). Symmetry of the return distribution is controlled by the parameters $G_i(x), M_i(x)$, while for $\tilde{Y}_i < 0$ the process Y^i has finite activity (otherwise it has infinite activity). Suppose that the differences $M_i(x) - G_i(x) > 0$ increase with x for each $i \in \{1, \ldots, n\}$. Then, in such a specification the left tails of the assets returns get heavier as X increases.

Recently, Hawkes process have received a lot of attention for stock price modelling, both from a statistical perspective and for the purpose of option pricing (see [3], [33] [21] or [22]). The framework of MMAPs allows the extension of affine Hawkes processes (i.e. Hawkes processes with exponential excitation kernel) to the situation where the jump intensities depend on an exogenous Markov process. For that, consider self-exciting counting processes $(Y^1, \ldots, Y^k)$ with values in $\mathbb{N}^k$ with respective intensities $(Y^{k+1}, \ldots, Y^m)$, where $m = 2k$. We further consider a Markov process X which is in line with Assumption 2.1. The intensities are given via

$$\begin{pmatrix} dY_t^{k+1} \\ \vdots \\ dY_t^m \end{pmatrix} = \left(\begin{pmatrix} b_{k+1}(X_t) \\ \vdots \\ b_m(X_t) \end{pmatrix} + \beta \begin{pmatrix} Y_t^{k+1} \\ \vdots \\ Y_t^m \end{pmatrix} \right) dt + \tilde{\delta} \begin{pmatrix} dY_t^1 \\ \vdots \\ Y_t^k \end{pmatrix},$$

where $b_{k+1}, \ldots, b_m$ are positive bounded (sufficiently regular) functions on $\mathcal{D}^X$, with $\tilde{\delta} \in \mathbb{R}_+^{k \times k}$, and where $\beta \in \mathbb{R}^{k \times k}$ is a matrix fulfilling the assumptions of its notational equivalent in Definition 2.3. We introduce the auxiliary index set $\mathfrak{J} := (k+1, \ldots, m)$. Subsequently, the generator of (X, Y) is given by

$$\mathfrak{L}^X f(x, y) + \langle (b_{k+1}(x), \ldots, b_m(x)) + \beta y_{\mathfrak{J}}, \nabla_{\mathfrak{J}} f(x, y) \rangle$$
$$+ \sum_{i=1}^k (f(x, y + e_i + \tilde{\delta}_{1,i} e_{k+1} + \cdots + \tilde{\delta}_{k,i} e_m) - f(x, y)) y_{i+k}, \quad f \in C_c^2(\mathcal{D}).$$

We use $\beta^0 \in \mathbb{R}^{m \times m}$ to denote the matrix with entries $\beta_{\mathfrak{J}\mathfrak{J}}^0 = \beta$ and zeroes otherwise. Let $\delta_y(d\zeta)$ denote the Dirac measure on $\mathcal{D}^Y$ centred at y. Then, we set $\mu_j = \delta_{(e_{j-k}, \tilde{\delta}_{\mathcal{I}, (j-k)})}$ for $j \in \mathfrak{J}$, to deduce that the x-admissible parameters are

$$\left(0, 0, (0_k, b_{k+1}(x), \ldots, b_m(x)), \beta^0, 0, 0, 0, \mu \right),$$

where $\mu = (0_k, \mu_{k+1}, \ldots, \mu_m)$ and 0_k denotes a vector of k zeroes.

A. Definitions

Let E be a locally compact and separable space.

Definition A.1. *A family* $(T_t)_{t \geq 0}$ *of linear operators on* $B_b(E)$ *is a* semigroup, *if*

$$T_0 = \mathrm{Id}, \quad T_t T_s f = T_s T_t f = T_{t+s} f \quad \forall f \in B_b(E),\ s,t \geq 0.$$

A semigroup is sub-Markov, *if for all* $t \geq 0$

(A.1) $T_t f \geq 0, \quad \forall f \in B_b(E)$ *with* $f \geq 0,$

(A.2) $T_t f \leq 1, \quad \forall f \in B_b(E)$ *with* $f \leq 1.$

Property (A.1) is typically referred to as positivity preserving and (A.2) is the sub-Markov property.

Definition A.2. *A* Feller semigroup *is a sub-Markov semigroup satisfying the Feller property*

$$T_t f \in C_\infty(E) \quad \forall f \in C_\infty(E),\ t > 0,$$

and which is strongly continuous in $C_\infty(E)$, *i.e.*

$$\lim_{t \to 0} \|T_t f - f\|_\infty = 0 \quad \forall f \in C_\infty(E).$$

Note that in the terminology of [27] a Feller semigroup on $C_\infty(E)$ is a strongly continuous, positivity preserving and conservative contraction semigroup. The above definition A.2 is in line with the one given in e.g. [50].

Definition A.3. *A* C_b-Feller semigroup *is a sub-Markov semigroup satisfying*

$$T_t f \in C_b(E) \quad \forall f \in C_b(E),\ t > 0,$$

and for which $t \mapsto T_t f$ *is continuous in the topology of locally uniform convergence in* $C_b(E)$.

B. Additional Proofs and results

Proof of Proposition 3.5. We begin with the situation, where f and g are independent of the time variable $t \in [0, \infty)$, that is we fix arbitrary $f \in \mathrm{Dom}(\widetilde{\mathfrak{L}}^X)$, $g \in C_b^2(\mathcal{D}^Y)$ and introduce the process

$$M_t^{f,g} := f(X_t)g(Y_t) - f(x)g(y) - \int_0^t (\widetilde{\mathfrak{L}}^X + \mathfrak{L}^{Y|X}(x)) f(X_s)g(Y_s) ds, \quad t \geq 0.$$

Note that by assumption $\widetilde{\mathfrak{L}}^X f \in C_b(\mathcal{D}^X)$. We pick a sequence $\{f_k\} \subset C_c^2(\mathcal{D}^X)$ with $\mathrm{bp\text{-}lim}_{k \to \infty} f_k = f$. By the Feller property of (P_t^X) we have that $\mathfrak{L}^X f_k \in C_\infty(\mathcal{D}^X)$ for each k. Since $C_b(\mathcal{D}^X)$ is dense in $B_b(\mathcal{D}^X)$ with respect to bp-convergence (cf. [27, Proposition 3.4.2]), we know that there exists some $\tilde{f} \in B_b(\mathcal{D}^X)$ with $\mathrm{bp\text{-}lim}_{k \to \infty} \mathfrak{L}^X f_k = \tilde{f}$. Choose arbitrary $t > 0$ and $x \in \mathcal{D}^X$ and consider

$$\frac{P_t^X f_k(x) - f_k(x)}{t} = \frac{1}{t} \int_0^t P_s^X \mathfrak{L}^X f_k(x) ds.$$

Take for the above equation the limit $k \to \infty$ and note that $\widetilde{P}_t^X f \in \mathrm{Dom}(\widetilde{\mathfrak{L}}^X)$ (see [43, Lemma 2.2.3]) to arrive at

$$\frac{1}{t} \int_0^t \widetilde{P}_s^X \widetilde{\mathfrak{L}}^X f(x) ds = \frac{\widetilde{P}^X f(x) - f(x)}{t} = \frac{1}{t} \int_0^t \widetilde{P}^X \tilde{f}(x) ds,$$

which shows that $\mathrm{bp\text{-}lim}_{k \to \infty} \mathfrak{L}^X f_k = \widetilde{\mathfrak{L}}^X f$. Moreover, by Lemma B.1 in Appendix B we can pick a sequence $(g_k) \in C_c^2(\mathcal{D}^Y)$ with $\mathrm{bp\text{-}lim}_{k \to \infty}(g_k, \mathfrak{L}^{Y|X}(x)g_k) = (g, \mathfrak{L}^{Y|X}(x)g)$.

Overall, bp-$\lim_{k\to\infty}(f_k g_k, \mathfrak{L} f_k g_k) = (fg, (\widetilde{\mathfrak{L}}^X + \mathfrak{L}^{Y|X}(x))fg)$, and by dominated convergence (and the associated L^1-convergence), we conclude that $M^{f,g}$ is a martingale. Getting back to the initial (time-dependent) functions f, g, we fix arbitrary $t_2 > t_1 \geq 0$, and we consequently get

$$\mathbb{E}_z\left[f(t_1, X_{t_2})g(t_1, Y_{t_2}) - f(t_1, X_{t_1})g(t_1, Y_{t_1}) \mid \mathcal{F}_{t_1}\right]$$
$$= \mathbb{E}_z\left[\int_{t_1}^{t_2}(\widetilde{\mathfrak{L}}^X + \mathfrak{L}^{Y|X}(x))f(t_1, X_s)g(t_1, Y_s)ds \mid \mathcal{F}_{t_1}\right]$$

Moreover, by the fundamental theorem of calculus, we have

$$\mathbb{E}_z\left[f(t_2, X_{t_2})g(t_2, Y_{t_2}) - f(t_1, X_{t_2})g(t_1, Y_{t_2}) \mid \mathcal{F}_{t_1}\right]$$
$$= \mathbb{E}_z\left[\int_{t_1}^{t_2}\partial_t\big(f(s, X_{t_2})g(s, Y_{t_2})\big)ds \mid \mathcal{F}_{t_1}\right],$$

and so by Lemma 4.3.4 of [27] we arrive at the desired result (set $v = \partial_t fg$ and $w = (\widetilde{\mathfrak{L}}^X + \mathfrak{L}^{Y|X}(x))fg$ to be in line with the notation of that lemma). $\qquad\square$

Lemma B.1. *Let $x \in \mathcal{D}^X$ and $g \in C_b^2(\mathcal{D}^Y)$ be arbitrary. Then, there exists a sequence $(g_k)_{k\in\mathbb{N}}$ with elements in $C_c^2(\mathcal{D}^Y)$ such that bp-$\lim_{k\to\infty}(g_k, \mathfrak{L}^{Y|X}(x)g_k) = (g, \mathfrak{L}^{Y|X}(x)g)$.*

Proof. As in the proof of [17, Proposition 8.2], we choose a function $\rho \in C_c^\infty(\mathbb{R}_+)$ with

$$\rho(r) = \begin{cases} 1, & \text{if } r \leq 1, \\ 0, & \text{if } r > 5, \end{cases}$$

as well as $0 \leq \rho(r) \leq 1$, $|\partial_r \rho(r)| \leq 1$ and $|\partial_r^2 \rho(r)| \leq 1$ for all $r \in \mathbb{R}_+$. We introduce a sequence of functions $g_k \in C_c^2(\mathcal{D}^Y)$, $k \in \mathbb{N}$, via

$$(\text{B.1}) \qquad\qquad g_k(y) := g(y)\rho(\|y\|^2/k).$$

It is easy to see that bp-$\lim_{k\to\infty} g_k = g$. Even more so, for arbitrary $(x, y) \in \mathcal{D}$ it holds that

$$\left|\mathfrak{L}^{Y|X}(x)(g_k - g)(y)\right|$$
$$\leq \sum_{i,l=1}^n \|a_{il}\|_\infty \left\{|\partial_{y_i}\partial_{y_l}g(y)|\big(1 - \rho(\|y\|^2/k)\big)\right.$$
$$\left. + \frac{2|y_i + y_l|}{k}\rho'(\|y\|^2/k)\|(\partial_{y_i} + \partial_{y_l})g\|_\infty + \frac{4|y_l y_i|}{k^2}\rho''(\|y\|^2/k)\|g\|_\infty\right\}$$
$$+ \frac{2}{k}\|g\|_\infty\,\rho'(\|y\|^2/k)\sum_{i=1}^n\|a_{ii}\|_\infty$$
$$+ \sum_{i,l=1}^n|\langle\alpha_{\mathcal{I},il}, y^+\rangle|\left\{|\partial_{y_i}\partial_{y_l}g(y)|\big(1 - \rho(\|y\|^2/k)\big)\right.$$
$$\left. + \frac{2|y_i + y_l|}{k}\rho'(\|y\|^2/k)\|(\partial_{y_i} + \partial_{y_l})g\|_\infty + \frac{4|y_l y_i|}{k^2}\rho''(\|y\|^2/k)\|g\|_\infty\right\}$$
$$+ \frac{2}{k}\|g\|_\infty\,\rho'(\|y\|^2/k)\sum_{i=1}^n|\langle\alpha_{\mathcal{I},il}, y^+\rangle|$$
$$+ \sum_{i=1}^n\|b_i\|_\infty\left\{|\partial_{y_i}g(y)|\big(1 - \rho(\|y\|^2/k)\big) + \frac{2|y_i|}{k}\rho'(\|y\|^2/k)\|g\|_\infty\right\}$$
$$+ \sum_{i=1}^n|(\beta y)_i|\left\{|\partial_{y_i}g(y)|\big(1 - \rho(\|y\|^2/k)\big) + \frac{2|y_i|}{k}\rho'(\|y\|^2/k)\|g\|_\infty\right\}$$

$$+ \|c\|_\infty \, |g(y) - g_k(y)| + \langle \gamma, y^+ \rangle |g(y) - g_k(y)|$$

$$+ \int_{\mathcal{D}^Y \setminus \{0\}} \left| g(y+\zeta)(1 - \rho(\|y+\zeta\|^2/k)) - g(y)(1 - \rho(\|y\|^2/k)) \right.$$

$$- (1 - \rho(\|y\|^2/k))\langle \nabla_{\mathcal{J}} g(y), \chi_{\mathcal{J}}(\zeta)\rangle \big| m(x, d\zeta)$$

$$+ \sum_{j=m+1}^n \frac{2|y_j|}{k} \rho'(\|y\|^2/k) \|g\|_\infty M(x)$$

$$+ \sum_{i=1}^m \int_{\mathcal{D}^Y \setminus \{0\}} \left| g(y+\zeta)(1 - \rho(\|y+\zeta\|^2/k)) - g(y)(1 - \rho(\|y\|^2/k)) \right.$$

$$- (1 - \rho(\|y\|^2/k))\langle \nabla_{\mathcal{J}(i)} g(y), \chi_{\mathcal{J}}(\zeta)\rangle \big| y_i^+ \mu_i(d\zeta)$$

$$+ \sum_{i=1}^m \frac{2 \|y_{\mathcal{J}}\|}{k} \rho'(\|y\|^2/k) \|g\|_\infty y_i^+ \mathcal{M}_i,$$

which converges to 0 as $k \to \infty$ and so we get bp-$\lim_{k \to \infty} \mathfrak{L}^{Y|X}(x)g_k = \mathfrak{L}^{Y|X}(x)g.$ $\square$

Bibliography

[1] Agoitia Hurtado, M. F. D. C., and Schmidt, T. Time-inhomogeneous polynomial processes. *Stochastic Analysis and Applications 38*, 3 (2020), 527–564.

[2] Albrecher, H., Mayer, P., Schoutens, W., and Tistaert, J. The Little Heston Trap. *Wilmott* (2007), 83–92.

[3] Bacry, E., Mastromatteo, I., and Muzy, J.-F. Hawkes processes in finance. *Market Microstructure and Liquidity 01*, 01 (2 2015).

[4] Beckett, G., Mackenzie, J. A., Ramage, A., and Sloan, D. M. On the numerical solution of one-dimensional PDEs using adaptive methods based on equidistribution. *Journal of Computational Physics 167*, 2 (2001), 372–392.

[5] Böttcher, B., Schilling, R., and Wang, J. *Lévy Matters III: Lévy-Type Processes: Construction, Approximation and Sample Path Properties*. Lecture Notes in Mathematics. Springer International Publishing, 2014.

[6] Carr, P., Geman, H., Madan, D. B., and Yor, M. The Fine Structure of Asset Returns: An Empirical Investigation. *Journal of Business 75*, 2 (2002), 305–332.

[7] Carr, P., and Linetsky, V. A jump to default extended CEV model: An application of Bessel processes. *Finance and Stochastics 10*, 3 (2006), 303–330.

[8] Carr, P., and Madan, D. Option valuation using the fast Fourier transform. *Journal of Computational Finance 2*, 4 (1999), 61–73.

[9] Carr, P., and Wu, L. Stock Options and Credit Default Swaps: A Joint Framework for Valuation and Estimation. *Journal of Financial Econometrics* (2009), 1–41.

[10] Cheridito, P., Filipović, D., and Yor, M. Equivalent and absolutely continuous measure changes for jump-diffusion processes. *Annals of Applied Probability 15*, 3 (2005), 1713–1732.

[11] Cohen, S. N., and Elliott, R. J. *Stochastic Calculus and Applications*, 2nd ed. Birkhäuser, 2015.

[12] CUCHIERO, C., FILIPOVIĆ, D., MAYERHOFER, E., AND TEICHMANN, J. Affine processes on positive semidefinite matrices. *Annals of Applied Probability 21*, 2 (2011), 397–463.

[13] CUCHIERO, C., KELLER-RESSEL, M., MAYERHOFER, E., AND TEICHMANN, J. Affine processes on symmetric cones. *Journal of Theoretical Probability 29*, 2 (2016), 359–422.

[14] CUCHIERO, C., KELLER-RESSEL, M., AND TEICHMANN, J. Polynomial processes and their applications to mathematical finance. *Finance and Stochastics 16*, 4 (2012), 711–740.

[15] CUCHIERO, C., AND TEICHMANN, J. Path properties and regularity of affine processes on general state spaces. In *Séminaire de probabilités XLV*, C. Donati-Martin, A. Lejay, and A. Rouault, Eds. Springer, Berlin, 2013, pp. 201–244.

[16] DUFFEE, G. R. The Relation Between Treasury Yields and Corporate Bond Yield Spreads. *The Journal of Finance 53*, 6 (1998), 2225–2241.

[17] DUFFIE, D., FILIPOVIĆ, D., AND SCHACHERMAYER, W. Affine processes and applications in finance. *Annals of Applied Probability 13*, 3 (2003), 984–1053.

[18] DUFFIE, D., AND LIOU, J. Floating-Fixed Credit Spreads. *Financial Analysts Journal 57* (2001), 76–87.

[19] DYNKIN, E. B. *Markov Processes*. Springer, Berlin, Heidelberg, 1965.

[20] EBERLEIN, E., GLAU, K., AND PAPAPANTOLEON, A. Analysis of Fourier transform valuation formulas and applications. *Applied Mathematical Finance 17*, 3 (2010), 211–240.

[21] EL EUCH, O., FUKASAWA, M., AND ROSENBAUM, M. The microstructural foundations of leverage effect and rough volatility. *Finance and Stochastics 22*, 2 (2018), 241–280.

[22] EL EUCH, O., AND ROSENBAUM, M. The characteristic function of rough Heston models. *Mathematical Finance 29*, 1 (2019), 3–38.

[23] EL KAROUI, N., PENG, S., AND QUENEZ, M. C. Backward stochastic differential equations in finance. *Mathematical Finance 7*, 1 (1997), 1–71.

[24] ELLIOTT, R. J., AND OSAKWE, C. J. U. Option pricing for pure jump processes with Markov switching compensators. *Finance and Stochastics 10*, 2 (4 2006), 250–275.

[25] ELLIOTT, R. J., AND SIU, T. K. On Markov-modulated exponential-affine bond price formulae. *Applied Mathematical Finance 16*, 1 (2009), 1–15.

[26] ENGEL, K.-J., AND NAGEL, R. *One-Parameter Semigroups for Linear Evolution Equations*. Springer, New York, 2000.

[27] ETHIER, S., AND KURTZ, T. G. *Markov Processes: Characterization and Convergence*. Wiley, New York, 1986.

[28] FILIPOVIĆ, D. Time-inhomogeneous affine processes. *Stochastic Processes and their Applications 115*, 4 (2005), 639–659.

[29] FILIPOVIĆ, D. *Term-Structure Models. A Graduate Course.* Springer, 2009.

[30] FILIPOVIĆ, D., AND LARSSON, M. Polynomial Jump-Diffusion Models. *Stochastic Systems 10*, 1 (2020), 71–97.

[31] FREY, R., KURT, K., AND DAMIAN, C. How Safe are European Safe Bonds? An Analysis from the Perspective of Modern Credit Risk Models. *Journal of Banking and Finance 119* (2020).

[32] HAN, J., JENTZEN, A., AND WEINAN, E. Solving high-dimensional partial differential equations using deep learning. *Proceedings of the National Academy of Sciences of the United States of America 115*, 34 (8 2018), 8505–8510.

[33] HAWKES, A. G. Hawkes processes and their applications to finance: a review. *Quantitative Finance 18*, 2 (2018), 193–198.

[34] JACOD, J., AND SHIRYAEV, A. S. *Limit Theorems for Stochastic Processes*, 2nd ed. Springer, Berlin, 2003.

[35] KALLSEN, J., AND MUHLE-KARBE, J. Exponentially affine martingales, affine measure changes and exponential moments of affine processes. *Stochastic Processes and their Applications 120*, 2 (2010), 163–181.

[36] KELLER-RESSEL, M., AND MAYERHOFER, E. Exponential moments of affine processes. *Annals of Applied Probability 25*, 2 (2015), 714–752.

[37] KELLER-RESSEL, M., SCHACHERMAYER, W., AND TEICHMANN, J. Affine processes are regular. *Probability Theory and Related Fields 151*, 3-4 (2011), 591–611.

[38] KELLER-RESSEL, M., SCHMIDT, T., AND WARDENGA, R. Affine processes beyond stochastic continuity. *Annals of Applied Probability 29*, 6 (2019), 3387–3437.

[39] KÜHN, F. On martingale problems and Feller processes. *Electronic Journal of Probability 23*, 13 (2018), 1–18.

[40] LADYSCHENSKAJA, O. A., SOLONNIKOV, V. A., AND URALCEVA, N. N. *Linear and Quasilinear Equations of Parabolic Type*. American Mathematical Society, Providence, Rhode Island, 1968.

[41] LEWIS, A. A Simple Option Formula for General Jump-Diffusion and Other Exponential Lévy Processes. *SSRN Electronic Journal* (2001).

[42] LONGSTAFF, F. A., AND SCHWARTZ, E. S. A Simple Approach to Valuing Risky Fixed and Floating Rate Debt. *The Journal of Finance 50*, 3 (1995), 789–819.

[43] LORENZI, L., AND BERTOLDI, M. *Analytical methods for Markov semigroups*. CRC Press, 2006.

[44] LUCIANO, E., AND SCHOUTENS, W. A multivariate jump-driven financial asset model. *Quantitative Finance 6*, 5 (2006), 385–402.

[45] LUCIANO, E., AND SEMERARO, P. Multivariate time changes for Lévy asset models: Characterization and calibration. *Journal of Computational and Applied Mathematics 233*, 8 (2010), 1937–1953.

[46] MCNEIL, A. J., FREY, R., AND EMBRECHTS, P. *Quantitative Risk Management: Concepts, Techniques and Tools*, 2nd ed. Princeton University Press, Princeton, 2015.

[47] OKITALOSHIMA, O., AND VAN CASTEREN, J. A. On the uniqueness of the martingale problem. *International Journal of Mathematics 7*, 6 (1996), 775–810.

[48] PAZY, A. *Semigroups of Linear Operators and Applications to Partial Differential Equations*, 1st ed. Springer, New York, 1983.

[49] PROTTER, P. E. *Stochastic Integration and Differential Equations*, vol. 2.1. Springer Science & Business Media, 2005.

[50] REVUZ, D., AND YOR, M. *Continuous martingales and Brownian motion*, 3rd ed. Springer Science & Business Media, 1999.

[51] SCHILLING, R. L. Conservativeness and Extensions of Feller Semigroups. *Positivity 2* (1998), 239–256.

[52] SCHILLING, R. L. Feller Processes Generated by Pseudo-Differential Operators: On the Hausdorff Dimension of Their Sample Paths. *Journal of Theoretical Probability 11*, 2 (1998), 303–330.

[53] SCHILLING, R. L. Growth and Hölder conditions for the sample paths of Feller processes. *Probability Theory and Related Fields 112* (1998), 565–611.

[54] SCHMIDT, T., TAPPE, S., AND YU, W. Infinite dimensional affine processes. *Stochastic Processes and their Applications 130*, 12 (2020), 7131–7169.

[55] THIEME, H. R. Semiflows generated by Lipschitz perturbations of non-densely defined operators. *Differential and Integral Equations 3*, 6 (1990), 1035–1066.

[56] VAN BEEK, M., MANDJES, M., SPREIJ, P., AND WINANDS, E. Regime switching affine processes with applications to finance. *Finance and Stochastics 24*, 2 (2020), 309–333.

[57] VRABIE, I. *C0-Semigroups and Applications*, first ed. Elsevier, 2003.

[58] WERNER, D. *Funktionalanalysis*, 8th ed. Springer, Berlin Heidelberg, 2018.