Regime switching affine processes with applications to finance

Misha van Beek[1] · Michel Mandjes[1] ·
Peter Spreij[1,2] · Erik Winands[1]

Abstract We introduce the notion of a regime switching affine process. Informally this is a Markov process that behaves conditionally on each regime as an affine process with specific parameters. To facilitate our analysis, specific restrictions are imposed on these parameters. The regime switches are driven by a Markov chain. We prove that the joint process of the Markov chain and the conditionally affine part is a process with an affine structure on an enlarged state space, conditionally on the starting state of the Markov chain. Like for affine processes, the characteristic function can be expressed in a set of ordinary differential equations that can sometimes be solved analytically. This result unifies several semi-analytical solutions found in the literature for pricing derivatives of specific regime switching processes on smaller state spaces. It also provides a unifying theory that allows us to introduce regime switching to the pricing of many derivatives within the broad class of affine processes. Examples include European options and term structure derivatives with stochastic volatility and default. Essentially, whenever there is a pricing solution based on an affine process, we can extend this to a regime switching affine process without sacrificing the analytical tractability of the affine process.

✉ M. van Beek
mishavanbeek@uva.nl

M.R.H. Mandjes
M.R.H.Mandjes@uva.nl

P. Spreij
spreij@uva.nl

E.M.M. Winands
E.M.M.Winands@uva.nl

[1] Korteweg-de Vries Institute for Mathematics, University of Amsterdam, Amsterdam, The Netherlands

[2] IMAPP, Radboud University, Nijmegen, The Netherlands

1 Introduction

The economic literature has shown that the behaviour of economic quantities varies with the state of the world; see Hamilton [27]. Such a state of the world is usually referred to as a *regime*. The most common example is the distinction between a bull and a bear market, in which stock prices are believed to follow different dynamics. Models incorporating such states are called *regime switching* models.

In a few cases, there are analytically tractable solutions to selected derivatives with regime switching underlying processes. These include the zero-coupon bond price in a Vasiček short rate model with a regime switching mean reversion level and volatility, see Landén [37], or a Cox–Ingersoll–Ross short rate model with regime switching mean reversion level as in Elliott and Siu [19], the price of European options where the underlying stock price process is the exponential of a specific regime switching pure jump process, see Elliott and Osakwe [18], and derivatives pricing with regime switching in a Heston-type stochastic volatility setting; see Elliott et al. [20], Elliott and Lian [16] and Elliott et al. [17].

Apart from these specific cases, there is no overarching theory that allows us to introduce regime switching to more general processes and derivatives, including multivariate extensions. There is also no explanation as to why analytically tractable solutions exist in these and not in other cases. The scope of this paper is to tackle these two points by introducing the class of multivariate *regime switching affine processes*.

This paper is related to some earlier ideas discussed briefly in van Beek et al. [43] on introducing regime switching to affine processes. However, we take a far more general approach here in assuming a generator-based perspective to define the class of regime switching processes, we do not limit ourselves to pure diffusions on the canonical state space, and we extend the results beyond regimes driven by the classical Markov chains.

Regime switching affine processes encapsulate all special cases from the literature mentioned above, but include many more possible extensions. We give a non-exhaustive list of the classes that can now be extended to include regime switching. A set of applications that already includes regime switching follows later.

(1) Multivariate affine term structure models are used throughout finance as the gold standard for yield curve analysis and the valuation of interest rate derivatives such as bills, notes, bonds, strips, caps, floors, swaps and swaptions. See Dai and Singleton [9] and Christensen et al. [4] for examples.

(2) The Heston model is a bivariate affine process that is used ubiquitously in finance for the pricing of European options on equity under the assumption that volatility is stochastic; see Heston [30], Van der Stoep et al. [44] and Fouque and

Saporito [23]. The theory of regime switching affine processes also applies to Heston extensions that include an affine model for rates, such as described in Grzelak and Oosterlee [26].

(3) Next, credit risky securities including corporate bonds are often modelled using an affine model for rates and a CIR-type process for the instantaneous hazard rate of default; see Duffie and Singleton [15], Maboulou and Mashele [39] and White [46]. Together these processes form a multivariate affine process.

(4) Finally, affine processes on generalised state spaces such as the space of positive semidefinite matrices can include stochastic covariances between variables. The processes can be useful in the valuation of options on multiple underlying assets and fixed income products with stochastic correlation between factors and default intensities; see Cuchiero et al. [5]. The Wishart process is an example of such a process that lives on the cone of symmetric positive semidefinite (covariance) matrices. It is used for pricing options with multiple underlying assets in Muhle-Karbe et al. [40], range notes in Chiarella et al. [3] and volatility-equity options in Da Fonseca et al. [8].

There are a myriad of practical applications thinkable with our generalisation of the existing literature. Affine processes are generally multivariate so that we can immediately generalise the literature to multivariate processes with regime switching parameters, as suggested above. One example is in pricing a credit default swap (CDS) with a credit valuation adjustment (CVA). This case has multiple hazard rates for the different involved parties that need to be modelled jointly, possibly together with a short rate. Regime switching affine processes can capture this multivariate aspect, and at the same time introduce regime switching. Giesecke et al. [25] and Frey and Backhaus [24] show that default events exhibit clustering in regimes, supporting this feature. Second, regime switching affine processes enable the pricing of derivatives where the payoff function is regime dependent. This happens when the regimes are the credit rating of some entity, as proposed by Jarrow et al. [31], and when the payoff deviates in the case of a default rating. Third, in the portfolio optimisation space, the mean returns on assets can be modelled as an unobservable Markov chain. Rieder and Bäuerle [41] and Bäuerle and Rieder [2] derive optimal investment strategies in such a setup. Also, Vostrikova and Dong [45] analyse utility maximisation on regime switching Lévy processes. In Sect. 9, we derive the analytical formula for the characteristic function that they consider. Fourth, we can think of specific calibrations of yield curve models as regimes, as in Harms et al. [29]. This allows pricing with future recalibration embedded. A perspective on regime switching different from ours (and the above cited literature) is the Bayesian one in Cuchiero et al. [6], Duembgen and Rogers [12]. These regimes have the interpretation of multiple competing (nested) models, and the probability estimate of each regime represents a Bayesian model averaging weight.

Regime switching affine processes are defined in terms of a continuous-time Markov chain on a specific state space that switches between states, the regimes, and a part with different affine dynamics conditionally on each regime. For example, let the Markov chain switch between the bull and bear market states. Let the short rate follow a Vasiček model reverting to one mean in the bull market, and to another mean in the bear market. Then the joint process of the Markov chain and the short rate is a regime switching affine process.

The literature allows those parameters to vary across regimes that do not multiply with the process. For example, in the drift $\theta(\mu - r_t)$ of the Vasiček model of the short rate r_t, μ may vary across regimes whereas θ may not, because it multiplies with the process (r_t) in the sense that in the expanded affine form $\theta\mu - \theta r_t$, θ and r_t are multiplied. Our definition of regime switching affine processes also allows exactly those parameters to vary across regimes that do not multiply with the process, but in a more general and multivariate setup.

The primary result of this paper is that the characteristic function of our regime switching affine process can be expressed as the solution to ordinary differential equations (ODEs), just like in the purely affine case. Because of the similar (but regime-dependent) exponential affine form of the characteristic function, they are conditionally affine. This result extends the techniques available for the class of affine processes to regime switching affine processes. Hence pricing equations based on affine processes can now be extended to include regime switching on the previously mentioned subset of parameters. This also explains the tractability of the special cases dealt with in the literature, because these can now be seen as specific regime switching affine processes.

The power of this result lies in part in the diversity of affine processes, as explained above. In general, affine processes are use to (jointly) model log-stock prices (with jumps), stochastic volatility, short rates and short rate factors, default intensities, see Duffie et al. [13], and stochastic covariance matrices. Affine processes also include all Lévy processes. Informally, a process X on a state space D is affine if its characteristic function can be written as

$$\mathbb{E}[e^{iu^\top X_t}] = e^{\phi(t,u) + \psi(t,u)^\top X_0}$$

for some complex-valued functions ϕ and ψ. The functions ϕ and ψ solve a system of ODEs. Typical examples of the quantities that are (jointly) enumerated in X are short rates, hazard rates of default, log-stock prices and stochastic volatility. In our regime switching extension, ϕ may depend on the current regime, whereas ψ may not. This restriction directly dictates that the process parameters that determine ψ cannot be different across regimes. Because ϕ and ψ depend on mutually exclusive parameters, all parameters in ϕ may vary across regimes. Naturally, the state space D is not allowed to depend on the regime, to prevent jumping from a regime with a larger state space to a regime with a smaller state space that does not include the current value of the process. For example, if we jump from a negative short rate in a Vasiček model regime with state space $\mathbb{R}$ to a Cox–Ingersoll–Ross short rate model regime with a nonnegative state space $\mathbb{R}_+$, we no longer have a well-defined process.

The result that regime switching affine processes are conditionally affine makes the powerful pricing techniques for affine processes available to regime switching affine processes. The most common technique is *transform pricing*. Using this technique, we can solve complex expectations in terms of ϕ and ψ, such as

$$\mathbb{E}_t\left[\exp\left(\int_t^T (\ell + L^\top X_s)\mathrm{d}s\right) f(X_T)\right],$$

with ℓ scalar and L a vector representing discounting, and f a payoff function on D that satisfies some requirements related to Fourier transformation. For an overview of

examples, see Duffie et al. [14]. In a handful of cases, ϕ and ψ can be solved from the ODEs analytically. If not, then they result from numerically solving the ODEs. Solving such equations is often preferred over numerically solving the partial differential equations (PDEs) that more general processes entail. The ODEs in transform pricing are generally faster to solve than PDEs, and suffer less from numerical instability.

This paper is organised as follows. After the notational setup in Sect. 2 and a brief overview of results on affine processes in Sect. 3, Sect. 4 introduces the (general) regime switching process in terms of its infinitesimal generator. Section 5 defines regime switching affine processes and gives the most important result of the paper, namely a system of ODEs for the characteristic function that is an analogue of the Riccati equations for affine processes. Section 6 extends the validity of this result beyond the standard domain of the characteristic function. Section 7 shows how our analysis in terms of semigroups relates to stochastic differential equations (SDEs). Section 8 discusses discounting and how this affects the ODEs in the previous sections. Finally, Sect. 9 links these techniques to the existing literature on regime switching by viewing a number of special instances from the literature as special cases of the theory developed here.

2 General notation

This section introduces notation and the mathematical context. We consider processes on a general *state space* S. Throughout this paper, we work with various forms of S, either $S = D$, $S = E$ or $S = D \times E$, where D and E are specified as follows. We consider processes on D in the setup of Keller-Ressel et al. [34]. That is, these processes live on the state space $D \subseteq \mathbb{R}^d$ such that the affine hull of D is $\mathbb{R}^d$. We also consider the one-point compactification $D_\infty := D \cup \{\infty\}$ and extend all functions f on D by assuming $f(\infty) = 0$ unless stated otherwise. This compactification is used to describe a point at infinity, i.e., a cemetery where killed or exploded processes go.[1] The state space E can be chosen almost arbitrarily as long as it is locally compact. See Remark 5.2, 1) for a practically relevant setting. Write $\mathcal{C}_b(S)$ for the class of bounded continuous complex-valued functions on S, and $\mathcal{C}_0(S)$ for the subclass of functions that are real-valued and vanishing at infinity; $\mathcal{C}_b(S)$ and $\mathcal{C}_0(S)$ are Banach spaces equipped with the norm $\|f\| = \sup_{x \in S} |f(x)|$.

Central in this paper are time-homogeneous continuous-time Markov processes, denoted by $(X, (\mathbb{P}_x)_{x \in S})$, where $\mathbb{P}_x$ denotes the law of the process starting in $x \in S$ so that $\mathbb{P}_x[X_0 = x] = 1$. This context implies the existence of an associated filtered probability space $(\Omega, \mathcal{F}, \mathbb{F})$, where X is Markov with respect to $\mathbb{F}$. To further aid notation, let (P_t) be the associated semigroup, $P_t f(x) := \mathbb{E}_x[f(X_t)]$ for $f \in \mathcal{C}_b(S)$, with $\mathbb{E}_x$ the expectation under $\mathbb{P}_x$. For general and possibly unbounded functions, we use the same notation, although the expectation need not be defined for all $t \geq 0$.

When we use the term process in the remainder of this paper, we mean a process with the properties defined in this section unless stated otherwise.

---

[1]Intuitively, the killing of a process is a jump to the cemetery state, and explosion means the process wanders off to infinity in finite time. Although such a state may represent default in a bond, there are easier ways to incorporate defaults using discounting. We include the cemetery state nonetheless to remain consistent with the literature.

3 Affine processes

To build the definition of regime switching affine processes, some knowledge of affine processes is necessary. This section gives the relevant setup of affine semigroups; it largely follows Keller-Ressel et al. [34].

We begin with a definition of affine processes. Consider for $u \in \mathcal{U}$ the functions $f_u(x) := \exp(\langle u, x \rangle)$ on D, with $\mathcal{U} \subseteq \mathbb{C}^d$ being the set of complex vectors such that f_u is bounded, i.e., $f_u \in \mathcal{C}_b(D)$. Also introduce $\mathcal{U}_k := \{u \in \mathbb{C}^d : \sup_{x \in D} \Re\langle u, x \rangle \leq k\}$ for $k \in \mathbb{N}$, so that $\bigcup_{k \in \mathbb{N}} \mathcal{U}_k = \mathcal{U}$.

Definition 3.1 A semigroup (P_t) is an *affine semigroup* if it is stochastically continuous (in the sense of Duffie et al. [13, Definition 2.4]) and there exist two functions $\Phi : \mathbb{R}_+ \times \mathcal{U} \to \mathbb{C}$ and $\psi : \mathbb{R}_+ \times \mathcal{U} \to \mathbb{C}^d$ such that for all $t \geq 0$, $x \in D$ and $u \in \mathcal{U}$,

$$P_t f_u(x) = \Phi(t, u) \exp\left(\langle \psi(t, u), x \rangle\right).$$

Remark 3.2 1) Since $i\mathbb{R}^d \subseteq \mathcal{U}$, the definition implies a specification of the form of the characteristic function of X_t with respect to the law $\mathbb{P}_x$.

2) Cuchiero and Teichmann [7] show that every affine process has an affine càdlàg version. **We assume throughout the rest of this paper** that this is the version we are working with.

3) Since there is a one-to-one correspondence between semigroups and processes in the present setup, we call the associated process of affine semigroups affine as well. In the remainder of this paper, we do the same with other properties.

As long as $\Phi(t, u) \neq 0$, we can write $\Phi(t, u) = \exp(\phi(t, u))$; so when we write $P_t f_u(x) = \exp(\phi(t, u) + \langle x, \psi(t, u) \rangle)$, we implicitly assume that $\Phi(t, u) \neq 0$. To explain the deeper relation between Φ and ϕ, introduce for any $u \in \mathcal{U}$ the quantity $\sigma(u) := \inf\{t \geq 0 : \Phi(t, u) = 0\}$ and define $\mathcal{Q}_k := \{(t, u) \in \mathbb{R}_+ \times \mathcal{U}_k : t < \sigma(u)\}$. Then ϕ is a function on $\mathcal{Q} := \bigcup_{k \in \mathbb{N}} \mathcal{Q}_k$. In general, there are multiple choices of ϕ and ψ that describe the same semigroup. **Throughout this paper**, we assume that ϕ and ψ are jointly continuous on $\mathcal{Q}_k$ for $k \in \mathbb{N}$, that $\phi(0, 0) = 0$ and that $\psi(0, 0) = 0$. Keller-Ressel et al. [34, Proposition 2.4 (ii)] prove that these assumptions make ψ and ϕ unique on $\mathcal{Q}$.

An important property of affine semigroups is regularity.

Definition 3.3 An affine semigroup is called *regular* if the derivatives

$$F(u) := \partial_t \Phi(t, u)|_{t=0+}, \qquad R(u) := \partial_t \psi(t, u)|_{t=0+}$$

exist for all $u \in \mathcal{U}$ and are continuous on $\mathcal{U}_k$ for each $k \in \mathbb{N}$.

Note that F is scalar whereas R, whose elements are denoted R^i, is vector-valued. Keller-Ressel et al. [34] prove that every affine semigroup is regular and that these derivatives can be written as

$$F(u) = \frac{1}{2}\langle u, au \rangle + \langle b, u \rangle - c + \int_{\mathbb{R}^d \setminus \{0\}} \big(f_u(\xi) - 1 - \langle u, h(\xi) \rangle \big) m(\mathrm{d}\xi),$$

$$R^i(u) = \frac{1}{2}\langle u, \alpha^i u \rangle + \langle \beta^i, u \rangle - \gamma^i + \int_{\mathbb{R}^d \setminus \{0\}} \big(f_u(\xi) - 1 - \langle h(\xi), u \rangle \big) \mu^i(\mathrm{d}\xi),$$

where $a, \alpha^1, \ldots, \alpha^d$ are $\mathbb{R}^{d \times d}$ matrices, $b, \beta^1, \ldots, \beta^d$ are $\mathbb{R}^d$ vectors, $c, \gamma^1, \ldots, \gamma^d$ are scalars and $m, \mu^1, \ldots, \mu^d$ are signed Borel measures, and $h(x) = x 1_{\{|x| \leq 1\}}$ is a truncation function based on the Euclidean norm.

Fast-forwarding to the application of regime switching, we note that F may be regime dependent, whereas R may not. This implies that only the F-specific parameters a, b, c and $m(\mathrm{d}\xi)$ may vary across regimes. Without this restriction, the characteristic function of such a regime switching affine process loses its conditionally affine structure.

Additionally, for any $u \in \mathcal{U}$, $t, s \geq 0$ with $(t + s, u) \in \mathcal{Q}$ and $(s, \psi(t, u)) \in \mathcal{Q}$, the functions ϕ and ψ satisfy the *flow properties*

$$\Phi(t + s, u) = \Phi(t, u)\Phi\big(s, \psi(t, u)\big), \qquad \Phi(0, u) = 1,$$
$$\psi(t + s, u) = \psi\big(s, \psi(t, u)\big), \qquad \psi(0, u) = u. \tag{3.1}$$

These can be differentiated into a system of ordinary differential equations for $u \in \mathcal{U}$ and $t \in [0, \sigma(u))$, called the *generalised Riccati equations* and given by

$$\partial_t \Phi(t, u) = \Phi(t, u) F\big(\psi(t, u)\big), \qquad \Phi(0, u) = 1,$$
$$\partial_t \psi(t, u) = R\big(\psi(t, u)\big), \qquad \psi(0, u) = u. \tag{3.2}$$

This system is used to find $P_t f_u(x)$, although there are boundary cases, when the solutions to this system are not unique, that require extra care, as described by Keller-Ressel and Mayerhofer [32].

Finally, if the affine semigroup is Feller, then its generator A is given by

$$Af(x) = \frac{1}{2}\mathrm{tr}\big(a(x)\partial_{xx} f(x)\big) + \langle b(x), \partial_x f(x) \rangle + c(x) f(x)$$
$$+ \int_{D \setminus \{0\}} \big(f(x + \xi) - f(x) - \langle h(\xi), \partial_x f(x) \rangle \big) m(x, \mathrm{d}\xi),$$

for all functions $f \in \mathcal{D}(A)$, the domain of A, and with

$$a(x) = a + \alpha^1 x_1 + \cdots + \alpha^d x_d,$$
$$b(x) = b + \beta^1 x_1 + \cdots + \beta^d x_d,$$
$$c(x) = c + \gamma^1 x_1 + \cdots + \gamma^d x_d,$$
$$m(x, \mathrm{d}\xi) = m(\mathrm{d}\xi) + \mu^1(\mathrm{d}\xi) x_1 + \cdots + \mu^d(\mathrm{d}\xi) x_d.$$

4 Regime switching processes

This section introduces *regime switching processes*, which are defined in terms of *generators of Feller semigroups*. Although these regime switching processes are not the focus of the present paper, they are essential to the definition of regime switching affine processes. Informally, regime switching affine processes are regime switching processes defined in terms of generators of *affine* Feller semigroups. We start with some notation and continue with a proof that regime switching processes also satisfy the Feller property. This is important in later sections.

Consider two state spaces D and E as defined in Sect. 2. Consider a function $f : D \times E \to \mathbb{R}$. We define $f^y : D \to \mathbb{R}$ by $f^y(x) = f(x, y)$ for all $(x, y) \in D \times E$. We define f^x similarly.

Definition 4.1 For all $y \in E$, let A^y and B be generators of Feller semigroups on $\mathcal{C}_0(D)$ and $\mathcal{C}_0(E)$, respectively, and also assume that B is bounded. Introduce the linear operator A^* on $\mathcal{C}_0(D \times E)$ through

$$A^* f(x, y) := A^y f^y(x) + B f^x(y).$$

Then A^* is called a *regime switching linear operator*.

Remark 4.2 1) In this definition, we say that the linear operator A^* is *switched* by the *driving* generator B and that E contains the *regimes*. We call the generators A^y the *regime specific* generators and D the *state space of the switched process*. The driving process is denoted by Y, or (Y_t) if the time index is relevant. The switched process is denoted by X, or (X_t) with time index.

2) The restriction to Feller semigroups is weak. For many specific state spaces including the canonical state space, it is known that affine semigroups are Feller, and for the general case it is hypothesised; see Cuchiero and Teichmann [7].

3) The requirement that B is bounded restricts the associated process to pure jumps; see Kolokoltsov [36, Proposition 3.7.1]. This corresponds to the intuition that a process must remain in a state at least for some time to be naturally interpreted as a regime. Examples of pure jump processes are Markov chains and compound Poisson processes. The states that these processes take can be interpreted as regimes because they last for some time.

4) In the above setup, the dynamics of the driving process are independent from the starting point x of the switched process. Indeed, if $f(x, y) = g(y)$ for all $x \in D$, then $A^* f(x, y) = A^y f^y(x) + B f^x(y) = \lim_{t \downarrow 0} \frac{1}{t}(\mathbb{E}_x[f^y] - f^y) + B g(y) = B g(y)$ because f^y is here a constant. Therefore, A^* is bounded on these functions and $P_t^* f(x, y) = \exp(t B) g(y)$, which is independent of x.

For the sake of brevity and readability, **in the remainder of this paper**, we often do not introduce superscript functions f^y and f^x every time we consider an operator on a function with additional variables. So when we write $B f(x, y)$, it is implicit that the operator acts only on the variable y, i.e., this is equivalent to $(B f(x, \cdot))(y)$.

Although it is not directly implied by the definition, A^* is the generator of a Feller semigroup. The following theorem states this and thereby shows that we can think of

A^* as a generator with an associated Feller semigroup and process. The proof relies mostly on results from perturbation theory.

Theorem 4.3 *Every regime switching linear operator A^* on $\mathcal{C}_0(D \times E)$ is the generator of a Feller semigroup.*

Proof Construct the operator A' on $\mathcal{C}_0(D \times E)$ through $A'f(x, y) = A^y f^y(x)$ for all $x \in D$ and $y \in E$. This is possible since $f \in \mathcal{C}_0(D \times E)$ implies that $f^y \in \mathcal{C}_0(D)$ for all $y \in E$.[2] Also extend B in this way to B' on $\mathcal{C}_0(D \times E)$. Clearly, A' and B' are both generators of Feller semigroups and $A^* = A' + B'$. Also, B' is bounded so that its semigroup can be written as $\exp(t B')$ and this implies that $\mathcal{D}(B') = \mathcal{C}_0(D \times E)$.

Denote by R_λ the resolvent on the space of linear operators and choose $\lambda > \|B'\|$. Then λ belongs to the resolvent set of A', called $\rho(A') \supseteq \mathbb{R}_{++}$ (see Ethier and Kurtz [22, Proposition 1.2.1]), where $\mathbb{R}_{++}$ is the set of strictly positive real numbers. Then $\|R_\lambda(A')\| \leq \lambda^{-1}$ yields $\|B' R_\lambda(A')\| < 1$, and thus $(B' R_\lambda(A'))^n \to 0$ as $n \to \infty$. This limit implies the same limit in Cesàro means, and so $B' R_\lambda(A')$ is uniformly ergodic. By Tyran-Kaminska [42, Theorem 1.1], A^* is a closed operator that generates a strongly continuous contraction semigroup. By [22, Theorem 1.2.6], $\mathcal{D}(A^*)$ is dense in $\mathcal{C}_0(D \times E)$ and the range of $\lambda - A^*$ equals $\mathcal{C}_0(D \times E)$. Also, by [22, Theorem 2.2.2], A' and B' satisfy the positive maximum principle. It follows easily that A^* does so, too. Again by [22, Theorem 2.2.2], the closure of A^* is a generator. The result follows from the fact that A^* was already proved to be closed. □

Remark 4.4 The result in Theorem 4.3 may be generalised for definitions that relax the bounded nature of B. Related results exist for strongly continuous semigroups, see e.g. Engel and Nagel [21, Chap. III], but it is unclear whether the same can be done under the more restrictive Feller property. In the present paper, we do not attempt such a generalisation as doing so would move us away from the present, finance-oriented scope (involving regime switching processes) for which B is naturally bounded.

5 Regime switching affine processes and properties

This section introduces the regime switching affine processes that are central to this paper. These are regime switching processes driven by a bounded operator such as a Markov chain, and with affine regime specific generators subject to specific parameter restrictions. For this process, we derive a system of ODEs analogous to the Riccati equations for the affine process.

Definition 5.1 Consider a regime switching generator A^* with regimes E and state space D for the switched process. The generator A^* is *regime switching affine* if

1) the regime specific generators A^y are generators of affine Feller semigroups, and

---

[2]Intuitively, we have constructed the generator of a process that has extra variables Y that do not vary stochastically, but remain in their starting position y that dictates the behaviour of X.

2) the regime specific semigroups have $P_t^y f_u(x) = \Phi^y(y, u) \exp\left(\langle \psi(t, u), x \rangle\right)$ for some functions Φ^y and ψ for all $t \geq 0$, $x \in D$ and $u \in \mathcal{U}$.

Remark 5.2 1) In the practically important special case when B is the generator of a Markov chain on a state space of standard basis vectors, i.e., $E = \{e_1, \ldots, e_p\} \subseteq \mathbb{R}^p$, a $p \times p$ generator matrix Q exists. With $q_{k\ell}$ for $k, \ell = 1, \ldots, p$ denoting the entries of Q, it holds that

$$Bf^x(e_k) = \sum_{\ell=1}^{p} q_{k\ell} f^x(e_\ell), \qquad k = 1, \ldots, p.$$

2) Requiring A^y to generate Feller semigroups is a minor restriction, as explained in Remark 4.2, 2).

An important aspect of the definition of regime switching affine processes is that ψ does not depend on the regime y. We call this independence *regime invariance*, and we use it throughout this paper to ensure parameters or functions do not vary with the regimes.

The associated derivatives of Φ^y and ψ are F^y and R (see Definition 3.3). The function R is regime invariant as a consequence of the regime invariance of ψ. The functional forms of F^y and R (see directly below Definition 3.3) imply that the parameters a^y, b^y, c^y and m^y may be different per regime, whereas α^i, β^i, γ^i and μ^i for $i = 1, \ldots, d$ must be regime invariant.

The primary result of this paper is the functional form of the characteristic function of regime switching affine processes. We consider dual variables $u \in \mathcal{U}$ and $v \in \mathcal{V}$, the set of continuous mappings $v : E \to \mathbb{C}$ such that $y \mapsto \exp\left(\langle v, y \rangle\right)$ is bounded, where we write $\langle v, y \rangle$ for $v(y)$ for homogeneity of notation. When E is a general Banach space, we usually take $v \in E^*$, and when E is finite-dimensional, we take $v \in \mathbb{C}^p$ and $\langle v, y \rangle$ is the usual inner product. In what follows, let $\tilde{X} := (X, Y)$, $\tilde{x} := (x, y)$, $\tilde{\xi} := (\xi, \zeta)$, $\tilde{u} := (u, v)$ and $F^*(u, y) := F^y(u)$, and extend the definition of f_u before Definition 3.1 from $S = D$ to $S = D \times E$ via $f_{\tilde{u}}(\tilde{x}) = \exp\left(\langle u, x \rangle + \langle v, y \rangle\right)$.

Theorem 5.3 *Consider a regime switching affine semigroup (P_t^*) and $\tilde{\mathcal{U}} = \mathcal{U} \times \mathbb{C}^p$. Then there exists a function $\Phi^* : \mathbb{R}_+ \times \tilde{\mathcal{U}} \times E \to \mathbb{C}$ such that for $\tilde{u} \in \tilde{\mathcal{U}}$,*

$$P_t^* f_{\tilde{u}}(\tilde{x}) = \Phi^*(t, \tilde{u}, y) \exp\left(\langle \psi(t, u), x \rangle\right)$$

for all $x \in D$, $y \in E$ and $t \geq 0$, where Φ^* solves the PDE

$$\partial_t \Phi^*(t, \tilde{u}, y) = F^*\left(\psi(t, u), y\right) \Phi^*(t, \tilde{u}, y) + B\Phi^*(t, \tilde{u}, y),$$

$$\Phi^*(0, \tilde{u}, y) = \exp\left(\langle v, y \rangle\right),$$

where $v \in \mathbb{C}^p$ and ψ is as in Definition 5.1.

Proof The proof can be found in Appendix A. $\square$

Remark 5.2, 1) introduced the special case that B generates a Markov chain. Theorem 5.3 simplifies if this is the case, as the following corollary shows. We adopt the notation that diag(a) or diag($a_1, \ldots, a_p$) is the $p \times p$ diagonal matrix with p-vector a on its diagonal. **In the remainder of this paper**, it often holds that $y = e_k$ for some $k = 1, \ldots, p$. In such cases, we simply write F^k instead of F^y, and similarly for other semigroups and functions with superscript y.

Corollary 5.4 *Suppose that B is the generator of a Markov chain with the state space $E = \{e_1, \ldots, e_p\}$ and the generator matrix Q. Then there exists a function $\theta : \mathbb{R}_+ \times \tilde{\mathcal{U}} \to \mathbb{C}^p$ such that*

$$P_t^* f_{\tilde{u}}(\tilde{x}) = \langle \theta(t, \tilde{u}), y \rangle \exp\big(\langle \psi(t, u), x \rangle\big)$$

for all $x \in D$, $y = e_k \in E$ and $t \geq 0$, where θ is the unique solution to the linear ODE

$$\partial_t \theta(t, \tilde{u}) = \text{diag}\Big(F^1\big(\psi(t, u)\big), \ldots, F^p\big(\psi(t, u)\big)\Big)\theta(t, \tilde{u}) + Q\theta(t, \tilde{u}).$$

Proof Due to the nature of the state space, it is always possible to define the vector-valued function θ from Φ^* such that $\langle \theta(t, \tilde{u}), y \rangle = \Phi^*(t, y, \tilde{u})$. Similarly, we can write $F^*(t, y, u) = \langle (F^1(u), \ldots, F^p(u)), y \rangle$. Then the ODE for $\langle \theta(t, \tilde{u}), y \rangle$ is known from Theorem 5.3 as

$$\langle \partial_t \theta(t, \tilde{u}), y \rangle = \partial_t \langle \theta(t, \tilde{u}), y \rangle$$

$$= \Big\langle \Big(F^1\big(\psi(t, u)\big), \ldots, F^p\big(\psi(t, u)\big)\Big), y \Big\rangle \langle \theta(t, \tilde{u}), y \rangle + B\langle \theta(t, \tilde{u}), y \rangle$$

$$= \Big\langle \text{diag}\Big(F^1\big(\psi(t, u)\big), \ldots, F^p\big(\psi(t, u)\big)\Big)\theta(t, \tilde{u}), y \Big\rangle + \langle Q\theta(t, \tilde{u}), y \rangle.$$

Since this holds for every standard basis vector y, the time derivative of θ exists and the ODE follows.

Finally, we need to prove that the ODE has a unique solution. This is true if $F^*(\psi(t, u))$ is continuous in t; see Knobloch and Kappel [35, Satz II.1.1 and Satz II.6.1]. Observe that $\psi(t, u)$ is continuous in t by virtue of its generalised Riccati equation, and has domain $\mathcal{U}$ by the contraction property of Feller semigroups. Moreover, $F(u)$ is continuous on $\mathcal{U}_k$ for all $k \in \mathbb{N}$, hence continuous on $\mathcal{U}$, and so the composition $F^*(\psi(t, u))$ is continuous in t. □

The remainder of this paper focuses on the setup as in Corollary 5.4. We refer to this setup as a regime switching affine process *driven by a Markov chain*. One might intuitively expect that this case gives analytically tractable results because X and Y can be viewed as jointly affine on a state space $D \times E$. However, it turns out that this is not possible as Markov chains are not affine processes to begin with.

The question arises whether the regime invariance restrictions of ψ (or, equivalently, R or α^i, β^i, γ^i and μ^i) in Theorem 5.3 can be relaxed. It appears that this is not straightforward and perhaps not possible. We conjecture that the restriction here is that $\partial_t P_t^* f_{\tilde{u}}(\tilde{x})|_{t=0+} = (G(\tilde{u}) + \langle H(\tilde{u}), \tilde{x} \rangle) f_{\tilde{u}}(\tilde{x})$ for some functions G and H, as is the case for affine processes; see Duffie et al. [13, Remark 2.8, where $G(u) = F(u)$

and $H(u) = R(u)$]. It is easy to show that this equation holds for regime switching affine processes driven by a Markov chain by taking derivatives, to get

$$\partial_t P_t^* f_{\tilde{u}}(\tilde{x})|_{t=0+} = \langle \partial_t^+ \theta(t, \tilde{u})|_{t=0}, y \rangle \exp(\langle u, x \rangle)$$
$$+ \langle (\exp(v_1), \ldots, \exp(v_p)), y \rangle \langle R(u), x \rangle \exp(\langle u, x \rangle)$$
$$= \langle \operatorname{diag}(\exp(-v_1), \ldots, \exp(-v_p)) \partial_t^+ \theta(t, \tilde{u})|_{t=0}, y \rangle f_{\tilde{u}}(\tilde{x})$$
$$+ \langle R(u), x \rangle f_{\tilde{u}}(\tilde{x}),$$

and by setting $H(\tilde{u}) = (R(u), \operatorname{diag}(\exp(-v_1), \ldots, \exp(-v_p)) \partial_t^+ \theta(t, \tilde{u})|_{t=0})$ and $G(\tilde{u}) = 0$. In the second equality, the above derivation depends on the property that $\exp(\langle v, y \rangle) = \langle (\exp(v_1), \ldots, \exp(v_p)), y \rangle$, which is only possible because the elements of E can be written as $y = e_k$ for some k, regardless of the form of the ODE for θ. This step also clearly breaks down when R is regime dependent.

6 Enlarging the domain of the characteristic function

So far, we have generally required that $u \in \mathcal{U}$. In financial applications of affine processes, option prices are usually expressed as functions of $P_t f_u$, where (P_t) is an affine semigroup and $u \in \mathbb{C}^d \supseteq \mathcal{U}$. For most applications, $\mathcal{U}$ is too small a set to work with, and this carries over to the application of regime switching affine processes. Even for relatively standard univariate diffusion processes, specific parameters lead to explosions in the exponential moments, i.e., $P_t^y f_u(x)$ does not exist. However, for any $u \in \mathbb{C}$, Φ^y and ψ exist up to some time T. This time naturally depends on u. This section extends the validity of Theorem 5.3 to u beyond $\mathcal{U}$, yet up to a time $T \leq \infty$.

In the world of affine processes, Duffie et al. [13] provide a toolset to determine if a solution $\Phi(t, u)$, $\psi(t, u)$ to the generalised Riccati equations up to some $T > 0$ for $u \in \mathbb{C}^d$ satisfies Definition 3.1. The steps usually involve checking that F and R^i, $i = 1, \ldots, d$, are analytic on some open domain, concluding from this that Φ and ψ must therefore also be analytic, and finally using regularity properties of characteristic functions to expand the domain on which the solutions of the generalised Riccati equations describe the exponential moments of the affine process. Our case is roughly equivalent so that we can extend the most important results for affine processes to regime switching affine processes. We start with the analytic nature of the Riccati solutions. Lemma 6.1 can be viewed as the regime switching affine generalisation of [13, Lemma 6.5 ii] for the setup in Corollary 5.4, i.e., driven by a Markov chain. Its proof closely follows [13].

Lemma 6.1 *Consider a regime switching affine semigroup (P_t^*) driven by a Markov chain. Suppose that all F^k and R are analytic on some open set U in $\mathbb{C}^d$. Let $T \leq \infty$ be such that a local U-valued solution ψ to the generalised Riccati equations exists on $(0, T)$ for all $u \in U$. Then:*

(i) *For all $v \in \mathbb{C}^p$, there also exists a solution θ to the ODE in Theorem 5.3 on $(0, T)$ that is unique given ψ.*

(ii) *ψ and θ have unique analytic extensions on $(0, T) \times U$ and $(0, T) \times U \times \mathbb{C}^p$, respectively.*

Proof Consider first the existence of θ. That ψ has a unique analytic extension on $(0, T) \times U$ follows directly from Dieudonné [11, Theorem 10.8.2]. The function ψ is U-valued; so since F^k is analytic on U, the vector-valued composition $(F^1(\psi(t, u)), \ldots, F^1(\psi(t, u)))$ is analytic on $(0, T)$ for every $u \in U$ and therefore continuous. Using the linearity of the ODE in Theorem 5.3, it follows from Knobloch and Kappel [35, Satz II.1.1 and Satz II.6.1] that this ODE has a unique solution χ^u on $(0, T)$ for every starting point $w \in \mathbb{C}^p$, not only those that can be written as an exponent of v. Defining $\theta(t, \tilde{u}) = \chi^u(t, w)$ for $w \in (\mathbb{C} \backslash \{0\})^p$, where $w_k = \exp(v_k)$, gives the existence of a unique solution θ given ψ.

To prove the final statement of the lemma, consider the composite system

$$\partial_t \begin{pmatrix} \psi(t, u) \\ \theta(t, u, w) \end{pmatrix} = G\big(\psi(t, u), \chi(t, u, w)\big), \qquad \psi(0, u) = u, \chi(0, u, w) = w,$$

$$G(u, w) = \begin{pmatrix} R(u) \\ (\text{diag}(F^*(u)) + Q)w \end{pmatrix}.$$

Then G is analytic on the open domain $U \times \mathbb{C}^p$, and there exists a $U \times \mathbb{C}^p$-valued local solution on $(0, T)$ for every $(u, w) \in U \times \mathbb{C}^p$. Using Dieudonné [11, Theorem 10.8.2] again, ψ and χ have a unique analytic extension on $(0, t) \times U$ and $(0, t) \times U \times \mathbb{C}^p$, respectively. The result for θ follows by changing the domain as above. $\qquad\square$

With the analytic nature of ψ and θ established, Theorem 6.2 below provides the necessary conditions to conclude that the ODE solutions describe the exponential moments of regime switching affine processes. This theorem is the regime switching affine generalisation of Duffie et al. [13, Theorem 2.16 ii].

Theorem 6.2 *Consider a regime switching affine semigroup (P_t^*) driven by a Markov chain. Let $t \geq 0$ and let U be an open convex neighbourhood of 0 in $\mathbb{C}^d$. Suppose that $\psi(t, \cdot)$ and $\theta(t, \cdot)$ have analytic extensions on U and $U \times \mathbb{C}^p$, respectively. Then*:

(i) *The exponential moments $P_t^* f_{\tilde{u}}(\tilde{x})$ are finite for all $u \in U \cap \mathbb{R}^d$, $v \in \mathbb{R}^p$, $x \in D$ and $y \in E$.*

(ii) *$P_t^* f_{\tilde{u}}(\tilde{x}) = \langle \theta(t, \tilde{u}), y \rangle \exp(\langle \psi(t, u), x \rangle)$ holds for all $u \in U$ such that $\Re(u) \in U \cap \mathbb{R}^d$, $v \in \mathbb{C}^p$, $x \in D$ and $y \in E$.*

Proof The proof of [13, Theorem 2.16 ii] is sufficient here as it relies in no way on the form of the characteristic function or the state space, only on its analytic nature which is satisfied by assumption. The only difference is that we do not require the process to be conservative. This is allowed because the proof works for any bounded measure (ν in [13, Appendix A]), not only probability measures. For a killed process, we have $f_u(\infty) = 0$ by assumption; so for the purpose of computing exponential moments, we might as well adjust that the measure value is zero on jumps to this point at infinity. This means that whenever a process includes killing, the exponential moments deflate due to strict subadditivity of the measure rather than the definition $f_u(\infty) = 0$, with the exact same effect. $\qquad\square$

A deeper analysis of analytical extensions in the context of affine processes can be found in Keller-Ressel and Mayerhofer [32].

7 SDE characterisation

We have given a definition of regime switching affine processes driven by a Markov chain in terms of the generator. For many applications, SDEs are more intuitive. Also, practitioners more often think in terms of SDEs than semigroups. This section gives an SDE formulation using the existing link for affine processes. For completeness, we state the result for affine processes without regime switching first.

For any Markov process, consider the sequence $\tau_n = \inf\{t \geq 0 : \|X_t - X_0\| > n\}$ and define the explosion time $\tau_\infty := \lim_{n\to\infty} \tau_n$. Recall the definitions of a, b and m from the end of Sect. 3. The following proposition is adapted from Keller-Ressel et al. [33, Corollary 3.11].

Proposition 7.1 *Let X be an affine process and suppose that the killing terms vanish, i.e., $c = 0$ and $\gamma = 0$. Then under every $\mathbb{P}_x$, $x \in D$, the process X is a D-valued semimartingale on $[\![0, \tau_\infty[\![$ with canonical semimartingale representation*

$$dX_t = b(X_t)dt + \sqrt{a(X_t)}dW_t$$
$$+ \int_{\mathbb{R}^d} h(\xi)\big(\nu_X(\omega; dt, d\xi) - m(X_{t-}, d\xi)\big) + \int_{\mathbb{R}^d} \big(\xi - h(\xi)\big)\nu_X(\omega; dt, d\xi),$$

where $\sqrt{\cdot}$ indicates the Cholesky decomposition, (W_t) is a d-dimensional Brownian motion, ν_X is the random measure associated with the jumps of X, h is a truncation function and a, b and m are as defined in Sect. 3.

The following proposition is the regime switching counterpart to Proposition 7.1 for the Markov chain case. Recall that we can write $Y = e_K$ for $K \in \{0, \ldots, p\}$ in this particular case.

Proposition 7.2 *Let $\tilde{X} = (X, Y) = (X, e_K)$ be a regime switching affine process and suppose that the killing terms vanish, i.e., $c^k = 0$ for $k = 1, \ldots, p$ and $\gamma = 0$. Then under every $\mathbb{P}_{(x,y)}$, $x \in D$, $y \in E$, the process (X, Y) is a D-valued semimartingale on $[\![0, \tau_\infty[\![$ with canonical semimartingale representation*

$$dX_t = b^{K_t}(X_t)dt + \sqrt{a^{K_t}(X_t)}dW_t$$
$$+ \int_{\mathbb{R}^d} h(\xi)\big(\nu_X(\omega; dt, d\xi) - m^{K_t}(X_t, d\xi)\big) + \int_{\mathbb{R}^d} \big(\xi - h(\xi)\big)\nu_X(\omega; dt, d\xi),$$

$$dY_t = \sum_{k=1}^p Y_{k,t} \sum_{\ell \neq k} (e_\ell - e_k)dN_t^{q_{k\ell}},$$

where a^k, b^k and m^k are the regime specific parameters corresponding to A^k and $(N_t^{q_{k\ell}})$ are independent Poisson processes with intensity $q_{k\ell}$ that count the transitions from e_k to e_ℓ up to time t.

Proof We prove the statement by letting (X, Y) follow the canonical representation in the statement, and showing that its semigroup has the correct parameters.

Let $u \in \mathcal{U}$ be such that ψ and θ exist for all $t \geq 0$ by Theorem 5.3. Based on this, let $T \geq 0$ and define $M_t^X := \exp{(\langle \psi(T-t,u), X_t \rangle)}$ and $M_t^Y := \langle \theta(T-t,\tilde{u}), Y_t \rangle$. Then let $M_t := M_t^X M_t^Y$. With the ansatz that M is a martingale, we find

$$P_T^* f_{\tilde{u}}(x, y) = \mathbb{E}_{(x,y)}[\exp{(\langle u, X_T \rangle + \langle v, Y_T \rangle)}]$$
$$= \mathbb{E}_{(x,y)}\left[\left\langle \left(\exp{(v_1)}, \ldots, \exp{(v_p)}\right), Y_T \right\rangle \exp{(\langle u, X_T \rangle)}\right]$$
$$= \mathbb{E}_{(x,y)}\left[\langle \theta(0, \tilde{u}), Y_T \rangle \exp{\left(\langle \psi(0, u), X_T \rangle\right)}\right]$$
$$= \mathbb{E}_{(x,y)}[M_T]$$
$$= M_0$$
$$= \langle \theta(T, \tilde{u}), y \rangle \exp{\left(\langle \psi(T, u), x \rangle\right)}.$$

This means that if M is a martingale, then the semigroup (P_t^*) of (X, Y) has the correct form. We now show that M is indeed a martingale.

By Itô's lemma and the definitions of F^k and R, we have that

$$dM_t^X = M_{t-}^X \bigg(-\langle \partial_T \psi(T-t, u), X_{t-} \rangle + \langle \psi(T-t, u), b^{K_{t-}}(X_{t-}) \rangle$$
$$+ \frac{1}{2}\langle a^{K_{t-}}(X_{t-})\psi(T-t, u), \psi(T-t, u) \rangle$$
$$+ \int_{\mathbb{R}^d} \left(\exp{\left(\langle \psi(T-t, u), \xi \rangle\right)} - 1 - \langle \psi(T-t, u), h(\xi) \rangle \right)$$
$$\times m^{K_{t-}}(X_{t-}, d\xi) \bigg) dt + d\tilde{M}_t^X$$
$$= M_{t-}^X F^{K_{t-}}\left(\psi(T-t, u)\right) dt + d\tilde{M}_t^X,$$

where $\tilde{M}^X$ stands for the local martingale part of M^X. By Elliott and Osakwe [18, (7)], we can write $dY_t = Q^\top Y_{t-} dt + d\tilde{Y}_t$, where $\tilde{Y}$ is a p-dimensional martingale. So with the martingale $\tilde{M}_t^Y = \int_0^t \langle \theta(T-s, \tilde{u}), d\tilde{Y}_s \rangle$, we get

$$dM_t^Y = \left(-\langle \partial_T \theta(T-t, \tilde{u}), Y_{t-} \rangle + \langle \theta(T-t, \tilde{u}), Q^\top Y_{t-} \rangle \right) dt + d\tilde{M}_t^Y$$
$$= \left(-\left\langle \left(\text{diag}\left(F^*(\psi(T-t, u)) \right) + Q \right) \theta(T-t, \tilde{u}), Y_{t-} \right\rangle \right.$$
$$\left. + \langle \theta(T-t, \tilde{u}), Q^\top Y_{t-} \rangle \right) dt + d\tilde{M}_t^Y$$
$$= -\left\langle \text{diag}\left(F^*(\psi(T-t, u)) \right) \theta(T-t, \tilde{u}), Y_{t-} \right\rangle dt + d\tilde{M}_t^Y$$
$$= -M_t^Y F^{K_{t-}}\left(\psi(T-t, u)\right) dt + d\tilde{M}_t^Y,$$

and by independence of the jump measures of Y and X,

$$
\begin{aligned}
\mathrm{d}M_t &= M_{t-}^X \mathrm{d}M_t^Y + M_{t-}^Y \mathrm{d}M_t^X \\
&= -M_{t-}^X M_t^Y F^{K_{t-}} \big(\psi(T-t,u)\big)\mathrm{d}t - M_{t-}^X \mathrm{d}\tilde{M}_t^Y \\
&\quad + M_{t-}^Y M_{t-}^X F^{K_{t-}} \big(\psi(T-t,u)\big)\mathrm{d}t + M_{t-}^Y \mathrm{d}\tilde{M}_t^X \\
&= -M_{t-}^X \mathrm{d}\tilde{M}_t^Y + M_{t-}^Y \mathrm{d}\tilde{M}_t^X.
\end{aligned}
$$

Hence M is a local martingale. But M^X and Y are uniformly bounded, and so M is a true martingale. $\qquad\square$

8 Discounting

Derivatives pricing often requires some form of discounting, either with an interest rate or a hazard rate of default, or both; see Lando [38]. This section explains discounting, and that discounting for regime switching affine processes driven by a Markov chain boils down to the same small modification of the generalised Riccati equations as discounting for affine processes.

Consider an affine process X with semigroup (P_t). Discounting means that instead of $P_t f(x)$, we are interested in

$$
Q_t f(x) := \mathbb{E}_x\left[\exp\left(\int_0^t L(X_s)\mathrm{d}s \right) f(X_t) \right],
$$

where $L : D \to \mathbb{R}$ given by

$$
L(x) := \ell + \lambda^1 x_1 + \cdots + \lambda^d x_d
$$

is the discounting rate. This is usually a short rate or a hazard rate, or a combination. For (Q_t), $Q_t f_u(x)$ still has the exponential affine form from Definition 3.1. Naturally, unless $\ell = \lambda^1 = \cdots = \lambda^d = 0$, Φ and ψ are not the same as for (P_t). However, it is well known that a small adaptation of F and R solves this problem. Effectively $F(u)$ becomes $F(u) + \ell$ and $R^i(u)$ becomes $R^i(u) + \lambda^i$ for all $i = 1, \ldots, d$.

For a regime switching affine process (X, Y), the natural equivalent is to see it as a conditionally affine process and use the exact same notation to obtain

$$
Q_t^* f(x, y) := \mathbb{E}_{(x,y)}\left[\exp\left(\int_0^t L(X_s, Y_s)\mathrm{d}s \right) f(X_t, Y_t) \right],
$$

where $L : D \times E \to \mathbb{R}$ is given by

$$
L(x, y) := \ell + \lambda^1 x_1 + \cdots + \lambda^d x_d + \lambda^{d+1} y_1 + \cdots + \lambda^{d+p} y_p.
$$

The same modifications apply to F^k and R, giving $F^k(u) + \ell + \lambda^{d+k}$ and $R^i(u) + \lambda^i$.

Alternatively, we can assume without loss of generality that $\ell = 0$ since the constant ℓ can be added to all the $\lambda^{d+1}, \ldots, \lambda^{d+p}$, keeping L the same. Then, using a slightly different notation, we can write

$$Q_t^* f(x, y) = \mathbb{E}_{(x,y)} \left[\exp \left(\int_0^t \langle L^*(X_s), Y_s \rangle \mathrm{d}s \right) f(X_t, Y_t) \right], \qquad (8.1)$$

where $L^* : D \times E \to \mathbb{R}^p$ is now given by

$$L_k^*(x) := \ell^k + \lambda^1 x_1 + \cdots + \lambda^d x_d. \qquad (8.2)$$

The modifications to F^k and R are now $F^k(u) + \ell^k$ for $k = 1, \ldots, p$ and $R^i(u) + \lambda^i$ for $i = 1, \ldots, d$.

9 Applications

This section explains why the most important cases of regime switching processes from the literature are in fact regime switching affine processes. First, we consider the regime switching short rate models in Landén [37] and Elliott and Siu [19], and show how the price of a zero-coupon bond is expressed in both models. Second, we derive the characteristic function of the regime switching univariate Lévy processes analysed by Elliott and Osakwe [18]. Third, we derive the characteristic function of a regime switching Heston model of the type in Elliott and Lian [16], Elliott et al. [17, 20]. In the final example, we explore a credit derivative with different interest rate and credit regimes. For this application, we do not base the model directly on the literature. The section concludes with a few high-level notes on hedging in the context of regime switching affine processes.

9.1 Riskless bond

Landén [37] and Elliott and Siu [19] consider the Vasiček and Cox–Ingersoll–Ross short rate models (with translated notation such that X_t is the short rate and K_t is the regime index as before)

$$\mathrm{d}X_t = (b^{K_t} + \beta X_t)\mathrm{d}t + \sqrt{a^{K_t}}\mathrm{d}W_t$$

and

$$\mathrm{d}X_t = (b^{K_t} + \beta X_t)\mathrm{d}t + \sqrt{\alpha X_t}\mathrm{d}W_t.$$

Using Proposition 7.2, it is easy to see how these are special cases of regime switching affine processes. The authors find that the price of a zero-coupon bond at t with maturity $T > t$ in these short rate models is

$$P(t, T, x, k) := \mathbb{E} \left[\exp \left(-\int_t^T X_s \mathrm{d}s \right) \bigg| X_t = x, K_t = k \right] = \mathrm{e}^{A(t,T,k) - B(t,T)x}.$$

In our notation in (8.1), this is

$$P(t, T, x, k) = \mathbb{E}_{(x, e_k)} \left[\exp \left(\int_0^{T-t} \langle L^*(X_s), Y_s \rangle \mathrm{d}s \right) \right] = Q^*_{T-t} f_0(x, e_k),$$

where $L^*_k(x) = \ell^k + \lambda^1 x = -x$ is the discount rate as in (8.2) and f_0 is simply $f_{\tilde{u}}$ with $\tilde{u} = 0$. Note that due to $\ell^k = 0$, the discounting formula is regime invariant and the regime only influences the price indirectly through the dynamics of the short rate (r_t). The authors use a system of differential equations that is a specific case of the ODE in Theorem 5.3, and use $A(t, T, r, k) = \ln \theta_k(T - t, 0)$.

9.2 European option price

In the next example, we look at option pricing in a Lévy setting. Elliott and Osakwe [18] choose a regime switching univariate Lévy process with the Lévy kernel of the variance gamma process. Their arguments for this specific process are empirical, but theoretically, their derivation checks out for any other univariate Lévy kernel. Translated to our setup, they choose $\alpha^i = 0$, $\beta^i = 0$, $\gamma^i = 0$, $\mu^i = 0$ for $i = 1$ since there is only one dimension, and $a^k = 0$, b^k free, $c^k = 0$ and $m^k(\mathrm{d}\xi)\mathrm{d}t$ are the compensators of variance gamma processes for $k = 1, \ldots, p$.

With these parameters, it follows that $R^i(u) = 0$ so that $\psi(t, u) = u$. The function $F^k(u)$ is the regime specific log-characteristic function

$$F^k(u) = ub^k + \int_{\mathbb{R}\setminus\{0\}} \left(e^{u\xi} - 1 - uh(\xi) \right) m^k(\mathrm{d}\xi).$$

Using Theorem 5.3, we get the characteristic function

$$\mathbb{E}[\exp(uX_t)|X_0 = 0, K_0 = k] = P^*_t f_{(u,0)}(0, e_k) = \theta_k(t, \tilde{u}),$$

where $\tilde{u} = (u, 0)$, $u \in i\mathbb{R}$ and θ is $p \times 1$ and the solution to

$$\partial_t \theta(t, \tilde{u}) = \left(\mathrm{diag}\left(F^*(u) \right) + Q \right) \theta(t, \tilde{u}), \qquad \theta(0, \tilde{u}) = 1,$$

where 1 is a vector of ones, and with straightforward closed-form solution

$$\theta(t, \tilde{u}) = \exp \left(\left(\mathrm{diag}\left(F^*(u) \right) + Q \right) t \right) 1.$$

This result is obtained in an entirely different fashion by Elliott and Osakwe [18, Proposition 2]. By Lemma 6.1 and Theorem 6.2, we may expand u beyond $i\mathbb{R}$ to a convex open neighbourhood of 0 in $\mathbb{C}$, as long as F is analytic on this neighbourhood. Note how in our setup, we can trivially generalise to any multivariate process (X_t), as in the setup of Vostrikova and Dong [45]. Variants of this result have been established in various contexts related to Markov additive processes; see e.g. Asmussen [1, Proposition XI.2.2].

9.3 Stochastic volatility

Next, we cast the regime switching Heston model of Elliott et al. [17] in the regime switching affine process format. The interpretation of volatility in terms of regimes is not new, but has become more important in the low volatility period of 2017; see Hamilton and Susmel [28]. Let $\ln S_t$ be the log-stock price, V_t the (instantaneous) variance and K_t the regime index. The regime switching Heston model has

$$d \ln S_t = \left(r - \frac{1}{2} V_t \right) dt + \sqrt{V_t} dW_t^1,$$

$$dV_t = \kappa (\theta_{K_t} - V_t) dt + \nu \sqrt{V_t} \left(\rho dW_t^1 + \sqrt{1 - \rho^2} dW_t^2 \right).$$

When $X_t = (\ln S_t, V_t)$, it is easy to see this in the light of Proposition 7.2 as

$$\begin{pmatrix} d \ln S_t \\ dV_t \end{pmatrix} = \left(\begin{pmatrix} r \\ \kappa \theta_{K_t} \end{pmatrix} + \begin{pmatrix} 0 & -\frac{1}{2} \\ 0 & \kappa \end{pmatrix} \begin{pmatrix} \ln S_t \\ V_t \end{pmatrix} \right) dt$$

$$+ \sqrt{\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \ln S_t + \begin{pmatrix} 1 & \nu\rho \\ \nu\rho & \nu^2 \end{pmatrix} V_t} \begin{pmatrix} dW_t^1 \\ dW_t^2 \end{pmatrix}.$$

From this formulation, we can read off all parameters required for R^i and F^k as

$$a^k = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}, \qquad \alpha^1 = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}, \qquad \alpha^2 = \begin{pmatrix} 1 & \nu\rho \\ \nu\rho & \nu^2 \end{pmatrix},$$

$$b^k = \begin{pmatrix} r \\ \kappa \theta_k \end{pmatrix}, \qquad \beta^1 = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \qquad \beta^2 = \begin{pmatrix} -\frac{1}{2} \\ \kappa \end{pmatrix}.$$

The killing and jump parameters c^k, γ^i, m^k and μ^i are all zero, and we note that a^k is zero and thus regime invariant. From here on, it is straightforward to numerically solve the Riccati equations for $\psi(t, u)$ and $\theta(t, \tilde{u})$ using (3.2) and Theorem 5.3 to get the characteristic function. With this, we can price European options through standard Fourier transformation as in Elliott et al. [17], or we can price the variance swap in Elliott and Lian [16], Elliott et al. [20].

9.4 Defaultable bond

Lastly, we consider an example from credit risk, namely the price of a defaultable zero-coupon bond of a bank, with no recovery value. As far as we know, this example does not correspond to a model found in the literature. This simplified case can be extended to include recovery value and can be used to price a credit default swap, possibly with counterparty credit risk, using the principles in Duffie et al. [14]. We consider a situation in which in search of yield, banks tend to take on more risky loans in a low interest rate environment. We model this by constructing two regimes indexed by a process (K_t), one stressed regime with a relatively low mean reversion level for the interest rate (r_t) with a corresponding high mean reversion level for the

bank's default hazard rate (h_t), and one vice versa. Both processes are correlated CIR processes under the risk neutral measure,

$$\mathrm{d}r_t = \theta_r(\mu^r_{K_t} - r_t)\mathrm{d}t + \sigma_r\sqrt{r_t}\mathrm{d}W^1_t,$$

$$\mathrm{d}h_t = \theta_h(\mu^h_{K_t} - h_t)\mathrm{d}t + \sigma_h\sqrt{h_t}\big(\rho^2\mathrm{d}W^1_t + \sqrt{1-\rho^2}\mathrm{d}W^2_t\big),$$

where $\mu^r_1 < \mu^r_2$ and $\mu^h_1 > \mu^h_2$. It is well known from Lando [38, Proposition 3.1] that by the Markov property of the involved processes, the price of such a bond is given by

$$P(t, T, r, \lambda, k) = 1_{\{\tau > t\}}\, \mathbb{E}\left[\exp\left(-\int_t^T (r_s + h_s)\mathrm{d}s\right)\Big| r_t = r, h_t = h, K_t = k\right],$$

where τ is the default time corresponding to the intensity process h and T is the maturity. To translate this to our setup, write $X = (r, h)$. It is straightforward to derive the parameters of the regime switching process for this case as it closely resembles the Heston model in the previous example. Using (8.1), we get

$$P(t, T, r, h, k) = 1_{\{\tau > t\}}\, \mathbb{E}_{(x, e_k)}\left[\exp\left(\int_0^{T-t} \langle L^*(X_s), Y_s\rangle \mathrm{d}s\right)\right] = Q^*_{T-t} f_0(x, e_k),$$

where $L^*_k(x) = \ell^k + \lambda^1 x_1 + \lambda^2 x_2 = -r - h$ is the discount rate as in (8.2). Note that due to $\ell^k = 0$, we have regime invariance in the discounting, meaning that $\langle L^*(X_t), Y_t\rangle = -r_t - h_t$. This pricing problem can be solved numerically using the ODE in Theorem 5.3.

9.5 Notes on hedging

In the Black–Scholes world (and several generalisations to other diffusion processes), hedging a contingent claim is well understood and standard. After switching to an equivalent martingale measure, we can use a unique self-financing strategy to replicate the payoff of the contingent claim almost surely. With jumps involved, market incompleteness generally (but not always) arises, i.e., the contingent claim cannot be exactly replicated. It is beyond the scope of this paper to go into detail, but worthwhile to highlight the implications when using regime switching processes for hedging purposes.

In such applications, the set of generators A^y may cause hedging errors due to incompleteness, for example because these are the generators of processes that include certain jump components. These hedging errors may be regime specific, i.e., in some regimes it may be harder to hedge than in others. But the perturbation by the generator B may lead to an additional source of hedging error that is caused by non-tradable regime changes. This is clearest in the case when we try to hedge a European option in a regime switching Black–Scholes world where only the volatility is different in each regime. In the trivial case that the regime is observed and constant (i.e., each state is absorbing), the contingent claim can be replicated as in the Black–Scholes case. If the regime is observed and can be traded, then the risk of jumping to another

regime is priced, and the claim can be replicated again with appropriate modifications to the absorbing case. However, if the regime cannot be traded, then in general a hedging error emerges. In this specific Black–Scholes case, the hedging error can be fully attributed to the regime switching nature of the process, since the generators A^y correspond to different Black–Scholes diffusions. The hedging error can be minimised using for instance mean–variance hedging and locally risk minimising trading strategies. Di Masi et al. [10] show explicitly how this can be done for the regime switching volatility case described above. It is conceivable that their approach can be mimicked for our switching regime processes, but as said above, this is beyond the scope of the present paper.

Appendix A: Proof of Theorem 5.3

The proof consists of three parts. We first prove the form of $P_t^* f_{\tilde{u}}(\tilde{x})$ through the existence of θ. Second, given that θ exists, we find the ODE it satisfies. Finally, we prove that this ODE always has a unique solution.

To prove the existence of θ, consider A' and B' as in the proof of Theorem 4.3. Let (P_t') be the semigroup that corresponds to A'. This semigroup exists since A' is the generator of a Feller semigroup. From the construction of A', it is immediate that $P_t' f(x, y) = P_t^y f^y(x)$. Intuitively, this shows that (P_t') is the semigroup of a regime switching affine process when every regime is absorbing.

Define $P_t^{(0)} := P_t'$ and recursively, for all $n = 1, 2, \dots$ and $t \geq 0$,

$$P_t^{(n)} := \int_0^t P_{t-s}' B' P_s^{(n-1)} \mathrm{d}s.$$

Then by Engel and Nagel [21, Theorem III.1.10], we have

$$P_t^* = \sum_{n=1}^{\infty} P_t^{(n)}. \tag{A.1}$$

We prove the theorem by showing that $P_t^{(n)} f_{\tilde{u}}(\tilde{x})$ obeys a specific factorisation that gives $P_t^* f_{\tilde{u}}(\tilde{x})$ the desired form as stated in the theorem.

First, define the functions g_n recursively through

$$g_0(t, y, \tilde{u}) = \Phi^y(t, u) \exp(\langle v, y \rangle),$$

$$g_n(t, y, \tilde{u}) = \int_0^t \Phi^y(t - s, \psi(s, u)) B' g_{n-1}(s, y, \tilde{u}) ds.$$

We prove by induction for $n \geq 0$ that

$$P_t^{(n)} f_{\tilde{u}}(\tilde{x}) = \exp(\langle \psi(t, u), x \rangle) g_n(t, y, \tilde{u}). \tag{A.2}$$

This holds for $n = 0$ because

$$P_t^{(0)} f_{\tilde{u}}(\tilde{x}) = P_t' f_{\tilde{u}}(\tilde{x}) = f_v(y) P_t^y f_u(x) = \exp(\langle v, y \rangle) \Phi^y(t, u) \exp(\langle \psi(t, u), x \rangle).$$

Suppose (A.2) holds for $n - 1$. Then it must also hold for n since by using (3.1) in the fifth equality below, we get

$$
\begin{aligned}
P_t^{(n)} f_{\tilde{u}}(\tilde{x}) &= \int_0^t P_{t-s}' B' P_s^{(n-1)} f_{\tilde{u}}(\tilde{x}) ds \\
&= \int_0^t P_{t-s}' B' \exp(\langle \psi(s, u), x \rangle) g_{n-1}(s, y, \tilde{u}) ds \\
&= \int_0^t \left(P_{t-s}' \exp(\langle \psi(s, u), x \rangle) \right) B' g_{n-1}(s, y, \tilde{u}) ds \\
&= \int_0^t \Phi^y(t - s, \psi(s, u)) \exp\left(\langle \psi(t - s, \psi(s, u)), x \rangle \right) B' g_{n-1}(s, y, \tilde{u}) ds \\
&= \int_0^t \Phi^y(t - s, \psi(s, u)) \exp(\langle \psi(t, u), x \rangle) B' g_{n-1}(s, y, \tilde{u}) ds \\
&= \exp(\langle \psi(t, u), x \rangle) \int_0^t \Phi^y(t - s, \psi(s, u)) B' g_{n-1}(s, y, \tilde{u}) ds \\
&= \exp(\langle \psi(t, u), x \rangle) g_n(t, y, \tilde{u}).
\end{aligned}
$$

At this point, it follows from (A.1) and (A.2) that

$$P_t^* f_{\tilde{u}}(\tilde{x}) = \Phi^*(t, \tilde{u}, y) \exp(\langle \psi(t, u), x \rangle) \tag{A.3}$$

for all $x \in D$, $y \in E$ and $t \geq 0$ and some function Φ^*. The finiteness of Φ^* follows from basic properties of the complex norm and the fact that Feller semigroups are contractive, so that

$$|\Phi^*(t, \tilde{u}, y)| = \frac{|P_t^* f_{\tilde{u}}(\tilde{x})|}{|\exp(\langle \psi(t, u), x \rangle)|} \leq \frac{|f_{\tilde{u}}(\tilde{x})|}{|\exp(\langle \psi(t, u), x \rangle)|} < \infty.$$

We proceed with the derivation of the ODE for Φ. Taking the time derivative on the left-hand side of (A.3) gives

$$\partial_t P_t^* f_{\tilde{u}}(\tilde{x}) = A^* \big(P_t^* f_{\tilde{u}}(\tilde{x}) \big)$$
$$= A^* \Phi^*(t, \tilde{u}, y) \exp \big(\langle \psi(t, u), x \rangle \big)$$
$$= \Phi^*(t, \tilde{u}, y) A^y \exp \big(\langle \psi(t, u), x \rangle \big) + \exp \big(\langle \psi(t, u), x \rangle \big) B \Phi^*(t, \tilde{u}, y)$$
$$= \Phi^*(t, \tilde{u}, y) \big(\partial_t \phi^y(t, u) + \langle \partial_t \psi(t, u), x \rangle \big) \exp \big(\langle \psi(t, u), x \rangle \big)$$
$$+ \exp \big(\langle \psi(t, u), x \rangle \big) B \Phi^*(t, \tilde{u}, y).$$

Similarly, taking the derivative on the right-hand side gives

$$\partial_t \Phi^*(t, \tilde{u}, y) \exp \big(\langle \psi(t, u), x \rangle \big)$$
$$= \exp \big(\langle \psi(t, u), x \rangle \big) \partial_t \Phi^*(t, \tilde{u}, y) + \Phi^*(t, \tilde{u}, y) \partial_t \exp \big(\langle \psi(t, u), x \rangle \big)$$
$$= \exp \big(\langle \psi(t, u), x \rangle \big) \partial_t \Phi^*(t, \tilde{u}, y) + \Phi^*(t, \tilde{u}, y) \langle \partial_t \psi(t, u), x \rangle \exp \big(\langle \psi(t, u), x \rangle \big).$$

Equating both sides and multiplying by $\exp \big(- \langle \psi(t, u), x \rangle \big)$ yields

$$\partial_t \Phi^*(t, \tilde{u}, y) = -\Phi^*(t, \tilde{u}, y) \langle \partial_t \psi(t, u), x \rangle$$
$$+ \Phi^*(t, \tilde{u}, y) \big(\partial_t \phi^y(t, u) + \langle \partial_t \psi(t, u), x \rangle \big) + B \Phi^*(t, \tilde{u}, y)$$
$$= \Phi^*(t, \tilde{u}, y) \partial_t \phi^y(t, u) + B \Phi^*(t, \tilde{u}, y)$$
$$= F^y \big(\psi(t, u) \big) \Phi^*(t, \tilde{u}, y) + B \Phi^*(t, \tilde{u}, y),$$

which is what we wanted to prove. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

References

1. Asmussen, S.: Applied Probability and Queues. Springer, Berlin (2003)
2. Bäuerle, N., Rieder, U.: Portfolio optimization with jumps and unobservable intensity process. Math. Finance **17**, 205–224 (2007)
3. Chiarella, C., Da Fonseca, J., Grasselli, M.: Pricing range notes within Wishart affine models. Insur. Math. Econ. **58**, 193–203 (2014)
4. Christensen, J.H.E., Diebold, F.X., Rudebusch, G.D.: The affine arbitrage-free class of Nelson–Siegel term structure models. J. Econom. **164**, 4–20 (2011)
5. Cuchiero, C., Filipović, D., Mayerhofer, E., Teichmann, J.: Affine processes on positive semidefinite matrices. Ann. Appl. Probab. **21**, 397–463 (2011)
6. Cuchiero, C., Klein, I., Teichmann, J.: A fundamental theorem of asset pricing for continuous time large financial markets in a two filtration setting. Preprint (2017). arXiv:1705.02087
7. Cuchiero, C., Teichmann, J.: Path properties and regularity of affine processes on general state spaces. In: Donati-Martin, C., et al. (eds.) Séminaire de Probabilités XLV. Lecture Notes in Mathematics, vol. 2078, pp. 201–244. Springer, Berlin (2013)
8. Da Fonseca, J., Gnoatto, A., Grasselli, M.: Analytic pricing of volatility-equity options within Wishart-based stochastic volatility models. Oper. Res. Lett. **43**, 601–607 (2015)
9. Dai, Q., Singleton, K.J.: Specification analysis of affine term structure models. J. Finance **55**, 1943–1978 (2000)
10. Di Masi, G.B., Kabanov, Y.M., Runggaldier, W.J.: Mean-variance hedging of options on stocks with Markov volatilities. Theory Probab. Appl. **39**, 172–182 (1995)
11. Dieudonné, J.: Foundations of Modern Analysis. Academic Press, San Diego (1969)
12. Duembgen, M., Rogers, L.C.G.: Estimate nothing. Quant. Finance **14**, 2065–2072 (2014)
13. Duffie, D., Filipović, D., Schachermayer, W.: Affine processes and applications in finance. Ann. Appl. Probab. **13**, 984–1053 (2003)

14. Duffie, D., Pan, J., Singleton, K.: Transform analysis and asset pricing for affine jump-diffusions. Econometrica **68**, 1343–1376 (2000)

15. Duffie, D., Singleton, K.J.: Modeling term structures of defaultable bonds. Rev. Financ. Stud. **12**, 687–720 (1999)

16. Elliott, R.J., Lian, G.-H.: Pricing variance and volatility swaps in a stochastic volatility model with regime switching: discrete observations case. Quant. Finance **13**, 687–698 (2013)

17. Elliott, R.J., Nishide, K., Osakwe, C.-J.U.: Heston-type stochastic volatility with a Markov switching regime. J. Futures Mark. **36**, 902–919 (2016)

18. Elliott, R.J., Osakwe, C.-J.U.: Option pricing for pure jump processes with Markov switching compensators. Finance Stoch. **10**, 250–275 (2006)

19. Elliott, R.J., Siu, T.K.: On Markov-modulated exponential-affine bond price formulae. Appl. Math. Finance **16**, 1–15 (2009)

20. Elliott, R.J., Siu, T.K., Chan, L.: Pricing volatility swaps under Heston's stochastic volatility model with regime switching. Appl. Math. Finance **14**, 41–62 (2007)

21. Engel, K.-J., Nagel, R.: A Short Course on Operator Semigroups. Springer, Berlin (2006)

22. Ethier, S.N., Kurtz, T.G.: Markov Processes: Characterization and Convergence. Wiley, New York (2009)

23. Fouque, J.-P., Saporito, Y.F.: Heston stochastic vol-of-vol model for joint calibration of VIX and S&P 500 options. Quant. Finance **18**, 1003–1016 (2018)

24. Frey, R., Backhaus, J.: Portfolio credit risk models with interacting default intensities: a Markovian approach. Preprint, University of Leipzig (2004). Available online at https://pdfs.semanticscholar.org/cbf5/e49076569f6762465C88ddaee544a87a4C4C.pdf

25. Giesecke, K., Longstaff, F.A., Schaefer, S., Strebulaev, I.: Corporate bond default risk: a 150-year perspective. J. Financ. Econ. **102**, 233–250 (2011)

26. Grzelak, L.A., Oosterlee, C.W.: On the Heston model with stochastic interest rates. SIAM J. Financ. Math. **2**, 255–286 (2011)

27. Hamilton, J.D.: A new approach to the economic analysis of nonstationary time series and the business cycle. Econometrica **57**, 357–384 (1989)

28. Hamilton, J.D., Susmel, R.: Autoregressive conditional heteroskedasticity and changes in regime. J. Econom. **64**, 307–333 (1994)

29. Harms, P., Stefanovits, D., Teichmann, J., Wüthrich, M.V.: Consistent recalibration of yield curve models. Math. Finance **28**, 757–799 (2018)

30. Heston, S.L.: A closed-form solution for options with stochastic volatility with applications to bond and currency options. Rev. Financ. Stud. **6**, 327–343 (1993)

31. Jarrow, R.A., Lando, D., Turnbull, S.M.: A Markov model for the term structure of credit risk spreads. Rev. Financ. Stud. **10**, 481–523 (1997)

32. Keller-Ressel, M., Mayerhofer, E.: Exponential moments of affine processes. Ann. Appl. Probab. **25**, 714–752 (2015)

33. Keller-Ressel, M., Schachermayer, W., Teichmann, J.: Affine processes are regular. Probab. Theory Relat. Fields **151**, 591–611 (2011)

34. Keller-Ressel, M., Schachermayer, W., Teichmann, J.: Regularity of affine processes on general state spaces. Electron. J. Probab. **18**(43), 1–17 (2013)

35. Knobloch, H.W., Kappel, F.: Gewöhnliche Differentialgleichungen (2013). Vieweg+Teubner Verlag

36. Kolokoltsov, V.N.: Markov Processes, Semigroups and Generators. De Gruyter, Berlin (2011)

37. Landén, C.: Bond pricing in a hidden Markov model of the short rate. Finance Stoch. **4**, 371–389 (2000)

38. Lando, D.: On Cox processes and credit risky securities. Rev. Deriv. Res. **2**, 99–120 (1998)

39. Maboulou, A.P.B., Mashele, H.P.: Credit derivative valuation and parameter estimation for multi-factor affine CIR-type hazard rate model. J. Math. Finance **5**, 273–285 (2015)

40. Muhle-Karbe, J., Pfaffel, O., Stelzer, R.: Option pricing in multivariate stochastic volatility models of OU type. SIAM J. Financ. Math. **3**, 66–94 (2012)

41. Rieder, U., Bäuerle, N.: Portfolio optimization with unobservable Markov-modulated drift process. J. Appl. Probab. **42**, 362–378 (2005)

42. Tyran-Kaminska, M.: Ergodic theorems and perturbations of contraction semigroups. Stud. Math. **195**, 147–155 (2009)

43. van Beek, M., Mandjes, M., Spreij, P., Winands, E.: Markov switching affine processes and applications to pricing. In: Vanmaele, M., et al. (eds.) Actuarial and Financial Mathematics Conference,

Interplay Between Finance and Insurance, pp. 97–102. Koninklijke Vlaamse Academie van België voor Wetenschappen en Kunsten, Brussels (2014)

44. van der Stoep, A.W., Grzelak, L.A., Oosterlee, C.W.: The Heston stochastic-local volatility model: efficient Monte Carlo simulation. Int. J. Theor. Appl. Finance **17**, 1450045 (2014)

45. Vostrikova, L., Dong, Y.: Utility maximization for Lévy switching models. Preprint (2018). arXiv:1807.08982

46. White, A.: Pricing credit default swap subject to counterparty risk and collateralization (2018). arXiv:1803.07843