MARKOV SWITCHING AFFINE PROCESSES AND APPLICATIONS TO PRICING

Misha van Beek[†], Michel Mandjes[†,‡,§], Peter Spreij[†] and Erik Winands[†]

[†]*Korteweg-de Vries Institute for Mathematics, University of Amsterdam, 1090 GE Amsterdam, The Netherlands*
[‡]*EURANDOM, 5600 MB Eindhoven, The Netherlands*
[§]*CWI, 1090 GB Amsterdam, The Netherlands*
Email: mishavanbeek@uva.nl, M.R.H.Mandjes@uva.nl, spreij@uva.nl, E.M.M.Winands@uva.nl

1. INTRODUCTION

This paper extends derivative pricing based on multivariate affine processes to affine models with Markov switching drift and diffusion coefficients. In the economic and finance literature, models with Markov switching parameters are often said to be regime switching.

In many economic and finance applications, processes fall prey to changes in regime. Regimes are time periods between which the dynamics of these processes are substantially different (Hamilton 1989). E.g. the mean returns, correlations and volatilities of stock prices are different in bull and bear markets, and the mean reversion level of interest rates may be lower in crisis scenarios. Based on this observation, the pricing of derivatives should account for the existence of different regimes.

Furthermore, many financial products benefit from multidimensional analysis. The price of a European call option is better modeled by allowing for stochastic interest rates and stochastic volatility. Also other products require multidimensional analysis directly through their structure. The price of a credit default swap (CDS) is derived from the dynamics of the interest rate and the hazard rate of default of the underlying. When we want to adjust to a price of a derivative for the creditworthiness of its seller, an additional process for the hazard rate of the seller enters into the game. This is known as a credit valuation adjustment (CVA), and together with a similar adjustment for the buyer's creditworthiness, the debit valuation adjustment (DVA), these are common and increasingly important drivers of multivariate analysis (Hull and White 2013).

In this paper, we consider the popular and broad class of multivariate affine processes that is often used to jointly model time series such as interest rates, stochastic volatility, hazard rates and log-asset prices (Duffie et al. 2003). Affine processes include the Vasicek and Cox-Ingersoll-Ross short rate models as special univariate cases. The primary advantage of affine processes over general multivariate processes in general is that the price of many derivatives has a closed form or is implicit in a system of ordinary differential equations (ODEs). ODE solutions are markedly more tractable than the partial differential equations (PDEs) that multivariate processes produce. We generalize multivariate affine processes to include Markov switching drift and diffusion coefficients. Our resulting Markov switching- (MS-)affine process maintains the property of ODE pricing solutions.

There is a rather restraint body of literature on this problem. Elliott and Mamon (2002) consider pricing a bond based on a short rate that follows a univariate Vasicek model with Markov switching mean reversion level. Elliott and Siu (2009) extend this result to bond prices based on a short rate that follows a univariate affine process with Markov switching mean reversion level and (in the Vasicek case) diffusion.

We take a more formal approach and follow the line of argumentation of Filipović (2009, Chapter 10). We derive the characteristic function of the MS-affine process and show that it can be expressed using the solutions of two systems of ODEs. We also prove that these solutions exist and are unique, provided that the parameters of the process are admissible in some sense. The characteristic function is the basis to price a wide variety of payoffs.

Effectively, our main theorem extends all pricing ODEs for affine processes to MS-affine processes. These include CVA and DVA adjustments, CDSs, exchange options, and many more. Moreover, for all these derivatives we may have regime dependent payoffs. The regime dependent payouts are used, for example, when the payoff of a derivative relies on the rating of a counterparty, and for this counterparty we have a rating migration matrix. Each rating (e.g. AAA, AA, etc.) can be seen as a regime in which the dynamics of the processes are different. Another example is when the dynamics of the affine process are different after some policy is introduced, but we are unsure when this policy takes effect. The different regimes would be the different states that the development and implementation of this policy can be in.

This remainder of this paper is outlined as follows. First we define the MS-affine process and the admissibility of its parameters. Then we provide two theorems that can be used for derivative pricing. We conclude with a simple example on how to apply these theorems to a bond price.

2. MODEL AND ANALYSIS

Let W_t be a d-dimensional Brownian motion with filtration $\{\mathcal{G}_t\}$. Let S_t be a continuous time Markov chain with state space $\mathcal{S} = \{1, \ldots, h\}$, filtration $\{\mathcal{H}_t\}$ and generator Q that switches between the regimes in $\mathcal{S}$. W_t and S_t are independent and defined on a filtered probability space $\langle \Omega, \mathcal{F}, \{\mathcal{F}_t\}, \mathbb{P} \rangle$, where $\mathcal{F}_t = \mathcal{G}_t \vee \mathcal{H}_t$.

Definition 2.1 *We call the process X on the canonical state space $\mathcal{X} = \mathbb{R}_+^m \times \mathbb{R}^n$, $m \geq 0$, $n \geq 0$, $m + n = d \geq 1$, MS-affine if*

$$\mathrm{d}X_t = \mu_{S_t}(X_t)\mathrm{d}t + \sigma_{S_t}(X_t)\mathrm{d}W_t, \tag{1}$$

where

$$\sigma_s(x)\sigma_s^\top(x) = a_s + \sum_{i=1}^{d} x_i\alpha_i, \qquad \mu_s(x) = b_s + \sum_{i=1}^{d} x_i\beta_i = b_s + \mathcal{B}x \tag{2}$$

for some $d \times d$*-matrices* a_s *and* α_i*, and* d*-vectors* b_s *and* β_i*, with* $\mathcal{B} = \begin{bmatrix}\beta_1 & \cdots & \beta_d\end{bmatrix}$.

Hence only a_s and b_s are regime dependent, not α_i and β_i.[1]

X_t may stack all sorts of financial variables. For example, if r_t is the short rate, A_t some asset price, V_t the stochastic volatility of the stock price, and h_t the hazard rate of default of the counterparty, then $X_t = (r_t, \ln A_t, V_t, h_t)$ models these processes jointly. For financial applications this model is usually under the risk neutral measure. This implies (among other things) that the drift of $\ln A_t$ is $r_t - \frac{1}{2}V_t$.

For ease of notation, we write $Z_t = e_{S_t} \in \{0,1\}^h$, a vector of zeros with S_t-th entry one. $\mathcal{Z}$ is the state space of Z_t. Then by Elliott (1993),

$$\mathrm{d}Z_t = QZ_t\mathrm{d}t + \mathrm{d}M_t, \tag{3}$$

where M_t is a martingale. Without proof we assume throughout this text that for every $x \in \mathcal{X}$, $z \in \mathcal{Z}$ there exists a unique solution $(X, Z) = (X^x, Z^z)$ of (1) with $X_0 = x$ and $Z_0 = z$.

To ensure that the process does not escape $\mathcal{X}$ we need some admissibility conditions on the parameters in (2). In what follows, we denote $I = \{1, \ldots, m\}$ and $J = \{m+1, \ldots, d\}$. Also, for any sets of indices M and N, and vector v and matrix w, $v_M = [v_i]_{i\in M}$ and $w_{MN} = [w_{ij}]_{i\in M, j\in N}$ are the corresponding sub-vector and sub-matrix.

Definition 2.2 *We call* X *an MS-affine process with admissible parameters if* X *is MS-affine and*

$$\begin{aligned}
a_s, \alpha_i &\ \textit{are symmetric positive semi-definite},\\
a_{sII} &= 0 \textit{ for all } s \in \mathcal{S} \textit{ (and thus } a_{sIJ} = a_{sJI}^\top = 0\textit{)},\\
\alpha_j &= 0 \textit{ for all } j \in J,\\
\alpha_{i,kl} = \alpha_{i,lk} &= 0 \textit{ for } k \in I\backslash\{i\}, \textit{for all } i, l \in \{1, \ldots, d\},\\
b_s &\in \mathcal{X} \textit{ for all } s \in \mathcal{S},\\
\mathcal{B}_{IJ} &= 0,\\
\mathcal{B}_{II} &\ \textit{has nonnegative off-diagonal elements}.
\end{aligned}$$

We now state our main contribution. $\mathrm{diag}(F_s)$ refers to the (block) diagonal matrix from the regime specific matrices $F_1, \ldots, F_h$.

Theorem 2.1 *Let* X *be an MS-affine process with admissible parameters. Let* $u \in \mathrm{i}\mathbb{R}^d$, $t \leq T$, $x \in \mathcal{X}$ *and* $z \in \mathcal{Z}$*. Then there exists unique solutions* $A(t,u) : \mathbb{R}_+ \times \mathrm{i}\mathbb{R}^d \to \mathbb{C}^{d\times d}$ *and* $B(t,u)$:

[1]Taking α_i and β_i regime dependent complicates further analysis and we are not sure whether ODE solutions to the characteristic function are possible in that case.

$\mathbb{R}_+ \times \mathrm{i}\mathbb{R}^d \to \mathbb{C}^d$ *to*

$$\begin{aligned}
\partial_t A(t,u) &= A(t,u)\left(\operatorname{diag}\left(\tfrac{1}{2}B_J(t,u)^\top a_{sJJ}B_J(t,u) + b_s^\top B(t,u)\right) + Q\right), \qquad (4)\\
A(0,u) &= I_h,\\
\partial_t B_i(t,u) &= \tfrac{1}{2}B(t,u)^\top \alpha_i B(t,u) + \beta_i^\top B(t,u), \quad i \in I,\\
\partial_t B_J(t,u) &= \mathcal{B}_{JJ}^\top B_J(t,u),\\
B(0,u) &= u,
\end{aligned}$$

such that the $\mathcal{F}_t$-conditional regime specific characteristic function satisfies

$$\mathbb{E}\left[e^{u^\top X_T} Z_T \middle| \mathcal{F}_t\right] = A(T-t,u)e^{B^\top(T-t,u)X_t} Z_t. \qquad (5)$$

Before proving the above theorem, we state (without proof) the following lemma, which is useful in an MS setting. $\otimes$ denotes the Kronecker product.

Lemma 2.2 *Let $F_{S_t} \in \mathbb{R}^{p\times q}$ be a set of d matrices with Markov switching index, then $(Z_t \otimes I_p)F_{S_t} = \operatorname{diag}(F_{S_t})(Z_t \otimes I_q)$.*

Also, we use the following lemma adapted from Filipović (2009, Lemma 10.1).

Lemma 2.3 *Consider the system of ODEs*

$$\partial_t y(t,y_0) = f(y(t,y_0)), \qquad y(0,y_0) = y_0, \qquad (6)$$

where $f : \mathbb{C}^d \to \mathbb{C}^d$ is a locally Lipschitz continuous function. Then:

1. *For every $y_0 \in \mathbb{C}^d$ there exists a lifetime $t_+(y_0) \in (0,\infty]$ such that there exists a unique solution $y(\cdot,y_0) : [0,t_+(y_0)) \to \mathbb{C}^d$ of (6).*
2. *The domain $\mathcal{D} = \left\{(t,y_0) \in \mathbb{R}_+ \times \mathbb{C}^d \middle| t \le t_+(y_0)\right\}$ is open in $\mathbb{R}_+ \times \mathbb{C}^d$ and maximal in the sense that either $t_+(y_0) = \infty$ or $\lim_{t\uparrow t_+(y_0)} \|y(t,y_0)\| = \infty$, respectively, for all $y_0 \in \mathbb{C}^d$.*

Proof of Theorem 2.1. Define $\Phi_t = A(T-t,u)e^{B(T-t,u)^\top X_t}$. We prove that $\Phi_t Z_t$ is martingale because this implies that $\mathbb{E}\left[e^{u^\top X_T} Z_T \middle| \mathcal{F}_t\right] = \mathbb{E}\left[\Phi_T Z_T | \mathcal{F}_t\right] = \Phi_t Z_t$, and then (5) is true. The dynamics of $\Phi_t Z_t$ follow from Itô's lemma and Lemma 2.2,

$$\begin{aligned}
\mathrm{d}(\Phi_t Z_t) = \mathrm{d}\Phi_t Z_t + \Phi_t \mathrm{d}Z_t &= \Big(\left(\partial_t A(T-t,u)\right) e^{B(T-t,u)^\top X_t}\\
&\quad + \Phi_t \left(\partial_t B(T-t,u)\right)^\top X_t + \Phi_t B(T-t,u)^\top \mu_{S_t}(X_t)\\
&\quad + \tfrac{1}{2}\Phi_t B(T-t,u)^\top \sigma_{S_t}(X_t)\sigma_{S_t}(X_t)^\top B(T-t,u)\Big) Z_t \mathrm{d}t + \Phi_t Q Z_t\\
&\quad + \Phi_t B(T-t,u)^\top \sigma_{S_t}(X_t)\mathrm{d}W_t Z_t + \Phi_t \mathrm{d}M_t\\
&= \Phi_t \left(B(T-t,u)^\top \sigma_{S_t}(X_t)\mathrm{d}W_t Z_t + \mathrm{d}M_t\right).
\end{aligned}$$

Therefore, $\Phi_t Z_t$ is a local martingale. The remaining part of the proof is showing that this local martingale is uniformly bounded, so it is also a martingale.

We know from Filipović (2009, proof of Theorem 10.2) that by admissibility, for any $u \in \mathbb{C}_-^m \times \mathrm{i}\mathbb{R}^n$, $t \in \mathbb{R}_+$ a unique solution $B(t,u) : \mathbb{R}_+ \times \mathbb{C}_-^m \times \mathrm{i}\mathbb{R}^n \to \mathbb{C}_-^m \times \mathrm{i}\mathbb{R}^n$ exists with infinite lifetime, so $\Re\left(B(t,u)^\top x\right) \leq 0$ for all $x \in \mathcal{X}$.

Apply Lemma 2.3 to the vectorization of the ODE of $A(t,u)$ (4), so $y = \mathrm{vec}(A)$, $y_0 = \mathrm{vec}(I_h)$ and f the vectorization of the RHS of (4). f is differentiable by differentiability of $B(t,u)$ and thus locally Lipschitz continuous. Therefore a unique solution for A exists with lifetime $t_+(\mathrm{vec}(I_h)) \in (0,\infty]$. We prove by contradiction that $t_+(\mathrm{vec}(I_h)) = \infty$. Suppose $t_+(\mathrm{vec}(I_h)) < \infty$, then $\lim_{t \uparrow t_+(\mathrm{vec}(I_h))} \|\mathrm{vec}(A(t,u))\| = \infty$. Note that

$$\|\mathrm{vec}(A(t,u))\|^2 = \mathrm{vec}(A(t,u))^* \mathrm{vec}(A(t,u)) = \mathrm{tr}(A(t,u)^* A(t,u)).$$

Define $\lambda = \max_{i=1,\dots,h}\{-\lambda_i, 0\}$, with λ_i the eigenvalues of $Q + Q^\top$, then

$$\begin{aligned}
\partial_t \|\mathrm{vec}(A(t,u))\|^2 &= \mathrm{tr}(\partial_t A(t,u)^* A(t,u) + A(t,u)^* \partial_t A(t,u)) \\
&= \mathrm{tr}\left(\left(Q + Q^\top\right) A(t,u)^* A(t,u)\right) \\
&\quad + 2\,\mathrm{tr}\left(\Re\left(\mathrm{diag}\left(\tfrac{1}{2} B_J(t,u)^\top a_{sJJ} B_J(t,u) + b_s^\top B(t,u)\right)\right) A(t,u)^* A(t,u)\right) \\
&\leq \mathrm{tr}\left(\left(Q + Q^\top\right) A(t,u)^* A(t,u)\right) + \lambda\,\mathrm{tr}\left(A(t,u)^* A(t,u)\right) \\
&\leq \mathrm{tr}\left(Q + Q^\top + \lambda I_h\right) \mathrm{tr}(A(t,u)^* A(t,u)) \\
&= \mathrm{tr}\left(Q + Q^\top + \lambda I_h\right) \|\mathrm{vec}(A(t,u))\|^2.
\end{aligned}$$

For the second equality we have substituted $\partial_t A(t,u)$ with (4). The first inequality follows from $\lambda \geq 0$ and the fact that for all $s \in \mathcal{S}$,

$$\begin{aligned}
&\Re\left(\tfrac{1}{2} B_J(t,u)^\top a_{sJJ} B_J(t,u) + b_s^\top B(t,u)\right) \\
&\quad = \tfrac{1}{2}\Re(B_J(t,u))^\top a_{sJJ} \Re(B_J(t,u)) - \tfrac{1}{2}\Im(B_J(t,u))^\top a_{sJJ} \Im(B_J(t,u)) + b_s^\top \Re(B(t,u)) \leq 0
\end{aligned}$$

by the admissibility restrictions on a_{sJJ} and b_s and the codomain of $B(t,u)$. The second inequality holds because $Q + Q^\top + \lambda I_h$ is positive semi-definite by construction and for any positive definite matrices C and D of the same size it holds that $\mathrm{tr}(CD) \leq \mathrm{tr}(C)\,\mathrm{tr}(D)$. Applying Gronwall's inequality gives $\|\mathrm{vec}(A(t,u))\|^2 \leq h e^{\mathrm{tr}\left(Q + Q^\top + \lambda I_h\right)t}$, for all $t < t_+(\mathrm{vec}(I_h))$. This yields $\lim_{t \uparrow t_+(\mathrm{vec}(I_h))} \|\mathrm{vec}(A(t,u))\| < \infty$, so by contradiction it follows that $t_+(\mathrm{vec}(I_h)) = \infty$; $A(t,u)$ has infinite lifetime for all $u \in \mathbb{C}_-^m \times \mathrm{i}\mathbb{R}^n$.

Combining these results we have that Φ_t and $B(t,u)$ are uniformly bounded for all $t \leq T$, so $\Phi_t Z_t$ is a martingale. ■

Theorem 2.1 is pivotal to derivatives pricing, but cannot be applied directly. Additionally, we need that (5) holds when $u \in \mathbb{R}^d$. Filipović (2009, Theorem 10.3 and Corollary 10.1) proves this for affine processes, and we conjecture that this result extends to MS-affine processes.

3. SIMPLE EXAMPLE

As an example on how to apply the above theorems to derivative pricing, we consider the bond price in a MS-Vasicek short rate model. Take the short rate model $\mathrm{d}r_t = \gamma\left(\mu_{S_t} - r_t\right)\mathrm{d}t + \sigma_{S_t}\mathrm{d}W_t$, where S_t is the continuous time Markov chain that switches between regimes and has generator Q. Introduce the integrator $\mathrm{d}R_t = r_t \mathrm{d}t$, $R_0 = 0$, then $X_t = (r_t, R_t)$ is a MS-affine process, and for $u = (0, -1)$ we have

$$1^\top \mathbb{E}\left[e^{u^\top X_T} Z_T \middle| \mathcal{F}_0\right] = \mathbb{E}\left[e^{-R_T} \middle| \mathcal{F}_0\right] = \mathbb{E}\left[\exp\left(-\int_0^T r_t \mathrm{d}t\right) \middle| \mathcal{F}_0\right].$$

Using Theorem 2.1 we can solve the LHS and thus obtain the price of the bond (the RHS, if it is finite). More examples can be found in Filipović (2009, Chapter 10.3).

References

D. Duffie, D. Filipovic, and W. Schachermayer. Affine processes and applications in finance. *The Annals of Applied Probability*, 13(3):984–1053, 2003.

Robert J. Elliott. New finite-dimensional filters and smoothers for noisily observed Markov chains. *IEEE Transactions on Information Theory*, 39(1):265–271, 1993.

Robert J. Elliott and Rogemar S. Mamon. An interest rate model with a Markovian mean reverting level. *Quantitative Finance*, 2(6):454–458, 2002.

Robert J. Elliott and Tak Kuen Siu. On Markov-modulated exponential-affine bond price formulae. *Applied Mathematical Finance*, 16(1):1–15, 2009.

Damir Filipović. *Term-Structure Models: A Graduate Course*. Springer Finance. Springer, 2009.

James D. Hamilton. A new approach to the economic analysis of nonstationary time series and the business cycle. *Econometrica*, 57(2):357–384, 1989.

John C. Hull and Alan White. Collateral and credit issues in derivatives pricing. Rotman School of Management Working Paper, 2013.