A Didactic Note on Affine Stochastic Volatility Models

Jan KALLSEN*

HVB-Stiftungsinstitut für Finanzmathematik, Zentrum Mathematik
TU München, Boltzmannstraße 3, D-85747 Garching bei München, Germany.
kallsen@ma.tum.de

Summary. Many stochastic volatility (SV) models in the literature are based on an affine structure, which makes them handy for analytical calculations. The underlying general class of affine Markov processes has been characterized completely and investigated thoroughly by Duffie, Filipovic, and Schachermayer (2003). In this note, we take a look at this set of processes and, in particular, affine SV models from the point of view of semimartingales and time changes. In the course of doing so, we explain the intuition behind semimartingale characteristics.

Key words: semimartingale characteristics, affine process, time change, stochastic volatility

Mathematics Subject Classification (2000): 60G99, 91B70

1 Introduction

Semimartingale calculus is by now a standard tool which is covered in many textbooks. However, this holds true to a lesser extent for the notion of semimartingale characteristics – despite of its practical use in many applications. A first goal of this note is to convince readers (who are not already convinced) that semimartingale characteristics are a very natural and intuitive concept.

We do so in Section 2 by taking ordinary calculus as a starting point and by restricting attention to the important special case of absolutely continuous characteristics. We argue that differential characteristics and certain martingale problems can be viewed as natural counterparts or extensions of derivatives and ordinary differential equations (ODE's). In this sense, affine processes are the solutions to particularly simple martingale problems, which extend affine ODE's to the stochastic case. They are considered in Section 3.

*This paper has been inspired by fruitful discussions with Arnd Pauwels.

Affine processes have been characterized completely and investigated thoroughly in an extremely useful and impressive paper by Duffie et al. ([7], henceforth DFS). They work predominantly in the context of Markov processes and their generators. But in a semimartingale setting, their results yield an explicit solution to the affine martingale problem.

Next to interest rate theory and credit risk, stochastic volatility (SV) models constitute one of the main areas in finance where the power of the affine structure has been exploited. In Section 4 we review a number of affine SV models under the perspective of semimartingale characteristics.

Unexplained notation is typically used as in [12]. Superscripts refer generally to coordinates of a vector or vector-valued process rather than powers. The few exceptions as e.g. $e^x, \sigma^2, v_t^{1/\alpha}$ should be obvious from the context. The notion of a *Lévy process* $X = (X_t)_{t\in\mathbb{R}_+}$ is applied slightly ambigiously. In the presence of a given filtration $\mathbf{F} = (\mathcal{F}_t)_{t\in\mathbb{R}_+}$, X is supposed to denote a Lévy process relative to this filtration (PIIS in the language of [12]), otherwise an *intrinsic* Lévy process in the sense of [19], i.e. a PIIS relative to its own natural filtration.

2 Differential semimartingale calculus

In this section we want to provide non-experts in the field with an intuitive feeling for semimartingale characteristics. It is not the aim to explain the mathematics behind this concept in detail. This is done exemplarily in the standard reference [12] (henceforth JS) or in [11], [23].

We hope that the reader does not feel offended by the following digression on $\mathbb{R}^d$-valued deterministic functions $X = (X_t)_{t\in\mathbb{R}_+}$ of time. Specifically, linear functions $X_t = bt$ are distinguished by constant growth. They are completely characterized by a single vector $b \in \mathbb{R}^d$. Many arbitrary functions behave "locally" as linear ones. This local behavior is expressed in terms of the derivative $\frac{d}{dt}X_t$ of X at time $t \in \mathbb{R}_+$. Of course, linear functions are up to the starting value X_0 the only ones with constant derivative. In many applications, functions occur as solutions to ODE's rather than explicitly, i.e. their derivative is expressed implicitly as

$$\frac{d}{dt}X_t = f(X_t), \quad X_0 = x_0. \tag{2.1}$$

In simple cases, the solution to the initial value problem (2.1) can be found in a closed form, e.g., if f is a linear or, more generally, an affine function. Linear ODE's are solved by exponential functions.

We now want to extend the above concepts to a probabilistic setting. Firstly note that stochastic processes $(X_t)_{t\in\mathbb{R}_+}$ are nothing else but random functions of time. A natural interpretation of constant growth in stochastic terms is *stationary, independent increments*. Therefore, the *Lévy pocesses (processes with stationary, independent increments)* can be viewed as random counterparts of linear functions. This is also reflected by the importance of Lévy processes in applications. The slope b of a linear function is paralleled by the *Lévy–Khintchine triplet* (b, c, F) of a Lévy process, where the vector $b \in \mathbb{R}^d$ stands for a linear drift as in the deterministic case, the symmetric non-negative $d \times d$ matrix c denotes the covariance matrix of the Brownian motion part of the process, and the Lévy measure F on $\mathbb{R}^d$ reflects the intensity of jumps of different sizes. By virtue of the Lévy–Khintchine formula, this triplet characterizes the distribution of a Lévy process X uniquely. Indeed, we have $Ee^{i\lambda^\top X_t} = e^{t\psi(i\lambda)}$, where the *Lévy exponent* ψ is given by

$$\psi(u) = u^\top b + \frac{1}{2}u^\top cu + \int (e^{u^\top x} - 1 - u^\top h(x))F(dx) \tag{2.2}$$

and $h : \mathbb{R}^d \to \mathbb{R}^d$ denotes a fixed truncation function as, e.g., $h(x) = x1_{\{|x|\leq 1\}}$. If h is replaced with another truncation function $\widetilde{h}$, only the drift coefficient b changes according to

$$b(\widetilde{h}) = b(h) + \int (\widetilde{h}(x) - h(x))F(dx). \tag{2.3}$$

It may seem less obvious how to extend derivatives and initial value problems to the stochastic case. A classical approach is provided within the theory of Markov processes. *Infinitesimal generators* describe the local behaviour of a Markov process X in terms of the current value X_t, which means that they naturally generalize ODE's. In this note, however, we focus instead on *semimartingale characteristics* and *martingale problems* as an alternative tool. Although the general theory behind Markov processes and semimartingales looks quite different in the first place, there exist close relationships between the corresponding concepts (cf. [11], [8]).

Finally, one can use *stochastic differential equations (SDE's)* to describe a process in terms of its local behavior. Even though there is a natural connection between martingale problems and SDE's, "linear" martingale problems do not correspond to linear SDE's as we shall see below.

The characteristics of a $\mathbb{R}^d$-valued semimartingale X can be defined in several equivalent ways. In the following definition they occur in an equation which resembles (2.2).

Definition 1. *Suppose that B is a predictable $\mathbb{R}^d$-valued process, C a predictable process whose values are non-negative symmetric $d \times d$ matrices, both with components of finite variation, and ν a predictable random measure on $\mathbb{R}_+ \times \mathbb{R}^d$ (i.e. a family $(\nu(\omega; \cdot))_{\omega\in\Omega}$ of measures on $\mathbb{R}_+ \times \mathbb{R}^d$ with a certain predictability property, cf. JS for details). Then (B, C, ν) is called* characteristics *of X if and only if $e^{i\lambda^\top X} - \int_0^\cdot e^{i\lambda^\top X_{t-}} d\Psi_t(i\lambda)$ is a local martingale for any $\lambda \in \mathbb{R}^d$, where*

$$\Psi_t(u) := u^\top B_t + \frac{1}{2}u^\top C_t u + \int_{[0,t]\times\mathbb{R}^d} (e^{u^\top x} - 1 - u^\top h(x))\nu(d(s, x)).$$

It can be shown that any semimartingale has unique characteristics up to a P-null set. This integral version of the characteristics can alternatively be written in differential form. More specifically, there exist an increasing predictable process A, predictable processes b, c, and a transition kernel F from $(\Omega \times \mathbb{R}_+, \mathcal{P})$ into $(\mathbb{R}^d, \mathcal{B}^d)$ such that

$$B_t = \int_0^t b_s dA_s, \quad C_t = \int_0^t c_s dA_s, \quad \nu([0,t] \times G) = \int_0^t F_s(G) dA_s, \quad G \in \mathcal{B}^d.$$

This decomposition is, of course, not unique. However, in most applications the characteristics (B, C, ν) are actually absolutely continuous, which means that one may choose $A_t = t$. In this case we call the triplet (b, c, F) *differential characteristics* of X. It is unique up to some $P(d\omega) \otimes dt$-null set.

Definition 2. *Suppose that b is a predictable $\mathbb{R}^d$-valued process, c a predictable process whose values are non-negative symmetric $d \times d$ matrices, and F a transition kernel from $(\Omega \times \mathbb{R}_+, \mathcal{P})$ to $(\mathbb{R}^d, \mathcal{B}^d)$ such that $F_.(\{0\}) = 0$ and $\int (1 \wedge |x|^2) F_.(dx) < \infty$. We call the triplet (b, c, F)* differential characteristics *of X if $e^{i\lambda^\top X} - \int_0^\cdot e^{i\lambda^\top X_{t-}} \psi_t(i\lambda) dt$ is a local martingale for any $\lambda \in \mathbb{R}^d$, where*

$$\psi_t(u) := u^\top b_t + \frac{1}{2} u^\top c_t u + \int_{\mathbb{R}^d} (e^{u^\top x} - 1 - u^\top h(x)) F_t(dx)$$

denotes the Lévy exponent of $(b, c, F)(\omega, t)$. For want of a handy notation in the literature, we write $\partial X := (b, c, F)$ in this case.

From an intuitive viewpoint one can interpret the differential characteristics as a local Lévy–Khintchine triplet. Very loosely speaking, a semimartingale with differential characteristics (b, c, F) resembles locally after t a Lévy process with triplet $(b, c, F)(\omega, t)$. Since this local behaviour may depend on the history up to t, the differential characteristics may be random albeit predictable. In this sense, the connection between Lévy processes and differential characteristics parallels the one between linear functions and derivatives of deterministic functions. In fact, b equals the ordinary derivative if X has absolutely continuous paths (and $c = 0$, $F = 0$ in this case). As is well-known, X is a Lévy process if and only if the differential characteristics are deterministic and constant (cf. JS, II.4.19):

Proposition 1 (Lévy process). *A $\mathbb{R}^d$-valued semimartingale X, $X_0 = 0$, is a Lévy process if and only if it has a version (b, c, F) of the differential characteristic which does not depend on (ω, t). In this case, (b, c, F) equals the Lévy-Khintchine triplet.*

As for the ordinary derivative, a number of rules allows to calculate the differential characteristics comfortably by using Lévy processes as building blocks.

Proposition 2 (Stochastic integration). *Let X be a $\mathbb{R}^d$-valued semimartingale and H a $\mathbb{R}^{n\times d}$-valued predictable process with $H^{j\cdot} \in L(X)$, $j = 1,\dots,n$ (i.e. integrable with respect to X). If $\partial X = (b,c,F)$, then the differential characteristics of the $\mathbb{R}^n$-valued integral process*

$$H \bullet X := (H^{j\cdot} \bullet X)_{j=1,\dots,n}$$

equals $\partial(H \bullet X) = (\widetilde{b}, \widetilde{c}, \widetilde{F})$, where

$$\begin{aligned} \widetilde{b}_t &= H_t b_t + \int (\widetilde{h}(H_t x) - H_t h(x)) F_t(dx), \\ \widetilde{c}_t &= H_t c_t H_t^\top, \\ \widetilde{F}_t(G) &= \int 1_G(H_t x) F_t(dx), \quad G \in \mathcal{B}^n. \end{aligned}$$

Here, $\widetilde{h} : \mathbb{R}^n \to \mathbb{R}^n$ denotes the truncation function which is used on $\mathbb{R}^n$.

Variants of Proposition 2 are stated in JS, IX.5.3 or [17], Lemma 3. The effect of C^2-functions on the characteristics follows directly from Itô's formula (cf. [9], Corollary A.6):

Proposition 3 (C^2-function). *Let X be a $\mathbb{R}^d$-valued semimartingale with differential characteristics $\partial X = (b,c,F)$. Suppose that $f : U \to \mathbb{R}^n$ is twice continuously differentiable on some open subset $U \subset \mathbb{R}^d$ such that X, X_- are U-valued. Then the $\mathbb{R}^n$-valued semimartingale $f(X)$ has differential characteristics $\partial(f(X)) = (\widetilde{b}, \widetilde{c}, \widetilde{F})$, where*

$$\begin{aligned} \widetilde{b}_t^i &= \sum_{k=1}^d \partial_k f^i(X_{t-}) b_t^k + \frac{1}{2} \sum_{k,l=1}^d \partial_{kl} f^i(X_{t-}) c_t^{kl} \\ &\quad + \int \left(\widetilde{h}^i \left(f(X_{t-} + x) - f(X_{t-}) \right) - \sum_{k=1}^d \partial_k f^i(X_{t-}) h^k(x) \right) F_t(dx), \\ \widetilde{c}_t^{ij} &= \sum_{k,l=1}^d \partial_k f^i(X_{t-}) c_t^{kl} \partial_l f^j(X_{t-}), \\ \widetilde{F}_t(G) &= \int 1_G \left(f(X_{t-} + x) - f(X_{t-}) \right) F_t(dx), \quad G \in \mathcal{B}^n. \end{aligned}$$

Here, ∂_k etc. denote partial derivatives and $\widetilde{h}$ again the truncation function on $\mathbb{R}^n$.

A Girsanov-type theorem due to Jacod and Mémin studies the behaviour of the characteristics under absolutely continuous changes of the probability measure (cf. JS, III.3.24). We state here the following version.

Proposition 4 (Change of the probability measure). *Let X be a $\mathbb{R}^d$-valued semimartingale with differential characteristics $\partial X = (b, c, F)$. Suppose that $\widetilde{P} \overset{\text{loc}}{\ll} P$ with the density process*

$$Z = \mathcal{E}(H \bullet X^c + W * (\mu^X - \nu^X)) \tag{2.4}$$

for some $H \in L(X^c)$, $W \in G_{\text{loc}}(\mu^X)$, where X^c denotes the continuous martingale part of X and μ^X, ν^X the random measure of jumps of X and its compensator (cf. JS for details). Then the differential characteristics $(\widetilde{b}, \widetilde{c}, \widetilde{F})$ of X relative to $\widetilde{P}$ are given by

$$\begin{aligned} \widetilde{b}_t &= b_t + H_t^\top c_t + \int W(t,x)h(x)F_t(dx), \\ \widetilde{c}_t &= c_t, \\ \widetilde{F}_t(G) &= \int 1_G(x)(1 + W(t,x))F_t(dx), \quad G \in \mathcal{B}^n. \end{aligned}$$

In applications, the density process can typically be stated in the form (2.4). Alternatively, one may use a version of Proposition 4 where $(\widetilde{b}, \widetilde{c}, \widetilde{F})$ is expressed in terms of the joint characteristics of (X, Z) (cf. [15], Lemma 5.1).

Finally, we consider the effect of absolutely continuous time changes (cf. [17], Lemma 5 and [11], Chapter 10 for details). They play an important role in SV models as we shall see in Section 4.

Proposition 5 (Absolutely continuous time change). *Let X be a $\mathbb{R}^d$-valued semimartingale with differential characteristics $\partial X = (b, c, F)$. Suppose that $(T_\theta)_{\theta \in \mathbb{R}_+}$ is a finite, absolutely continuous time change (i.e. T_θ is a finite stopping time for any θ and $T_\theta = \int_0^\theta \dot{T}_\rho d\rho$ with non-negative derivative $\dot{T}_\rho$).*

Then the time-changed process $(\widetilde{X}_\theta)_{\theta \in \mathbb{R}_+} := ((X \circ T)_\theta)_{\theta \in \mathbb{R}_+} := (X_{T_\theta})_{\theta \in \mathbb{R}_+}$ is a semimartingale relative to the time-changed filtration

$$(\widetilde{\mathcal{F}}_\theta)_{\theta \in \mathbb{R}_+} := (\mathcal{F}_{T_\theta})_{\theta \in \mathbb{R}_+}$$

with differential characteristics $\partial \widetilde{X} = (\widetilde{b}, \widetilde{c}, \widetilde{F})$ given by

$$\begin{aligned} \widetilde{b}_\theta &= b_{T_\theta} \dot{T}_\theta, \\ \widetilde{c}_\theta &= c_{T_\theta} \dot{T}_\theta, \\ \widetilde{F}_\theta(G) &= F_{T_\theta}(G) \dot{T}_\theta, \quad G \in \mathcal{B}^n. \end{aligned}$$

Let us now turn to the stochastic counterpart of the initial value problem (2.1), where the local dynamics of X are expressed in terms of X itself. This can be interpreted as a special case of a martingale problem in the sense of JS, III.2.4 and III.2.18.

Definition 3. *Suppose that P_0 is a distribution on $\mathbb{R}^d$ and functions $\beta : \mathbb{R}^d \times \mathbb{R}_+ \to \mathbb{R}^d$, $\gamma : \mathbb{R}^d \times \mathbb{R}_+ \to \mathbb{R}^{d \times d}$, $\varphi : \mathbb{R}^d \times \mathbb{R}_+ \times \mathcal{B}^d \to \mathbb{R}_+$ are given.*

We call $(\Omega, \mathcal{F}, \mathbf{F}, P, X)$ solution *to the martingale problem related to* P_0 *and* (β, γ, φ) *if* X *is a semimartingale on* $(\Omega, \mathcal{F}, \mathbf{F}, P)$ *such that* $\mathcal{L}(X_0) = P_0$ *and* $\partial X = (b, c, F)$ *with*

$$\begin{aligned} b_t(\omega) &= \beta(X_{t-}(\omega), t), \\ c_t(\omega) &= \gamma(X_{t-}(\omega), t), \\ F_t(\omega, G) &= \varphi(X_{t-}(\omega), t, G). \end{aligned} \tag{2.5}$$

More in line with the common language of martingale problems, one may also call the distribution P^X of X *solution* to the martingale problem. In any case, *uniqueness* refers only to the law P^X because solution processes on different probability spaces cannot be reasonably compared otherwise.

Since ODE's are particular cases of this kind of martingale problems, one cannot expect that unique solutions generally exist, let alone to solve them (cf. JS, III.2c and [11] in this respect). In this note we will only consider particularly simple martingale problems, namely linear and affine ones.

3 Affine processes

Parallel to affine ODE's, we assume that the differential characteristics (2.5) are affine functions of X_{t-} in the following sense:

$$\begin{aligned} \beta((x^1, \ldots, x^d), t) &= \beta_0 + \sum_{j=1}^{d} x^j \beta_j, \\ \gamma((x^1, \ldots, x^d), t) &= \gamma_0 + \sum_{j=1}^{d} x^j \gamma_j, \\ \varphi((x^1, \ldots, x^d), t, G) &= \varphi_0(G) + \sum_{j=1}^{d} x^j \varphi_j(G), \end{aligned} \tag{3.1}$$

where $(\beta_j, \gamma_j, \varphi_j)$, $j = 0, \ldots, d$ are given Lévy–Khintchine triplets on $\mathbb{R}^d$. As in the deterministic case, it is possible not only to prove existence of a unique solution but also to solve the affine martingale problem related to (3.1) in a sense explicitly. This has been done by DFS. More precisely, they characterize affine *Markov* processes and their laws. However, applied to the present setup one obtains the statement below on affine martingale problems (cf. Theorem 3.1).

It is obvious that the $d+1$ Lévy–Khintchine triplets $(\beta_j, \gamma_j, \varphi_j)$ cannot be chosen arbitrarily. It has to be ensured that the local covariance matrix c and the local jump measure F in the differential characteristics $\partial X = (b, c, F)$ of the solution remain positive even if some of the components X^j turn negative. This leads to a number of conditions:

Definition 4. *Let $m, n \in \mathbb{N}$ with $m + n = d$. Lévy–Khintchine triplets $(\beta_j, \gamma_j, \varphi_j)$, $j = 0, \ldots, d$ are called* admissible *if the following conditions hold:*

$$\left.\begin{array}{r} \beta_j^k - \int h^k(x)\varphi_j(dx) \geq 0 \\ \varphi_j((\mathbb{R}_+^m \times \mathbb{R}^n)^C) = 0 \\ \int h^k(x)\varphi_j(dx) < \infty \end{array}\right\} \text{if } 0 \leq j \leq m, \quad 1 \leq k \leq m, \quad k \neq j;$$

$$\gamma_j^{kl} = 0 \qquad \text{if } 0 \leq j \leq m, \quad 1 \leq k, l \leq m \quad \text{unless } k = l = j;$$

$$\beta_j^k = 0 \qquad \text{if } j \geq m+1, \quad 1 \leq k \leq m;$$

$$\left.\begin{array}{r} \gamma_j = 0 \\ \varphi_j = 0 \end{array}\right\} \text{if } j \geq m+1.$$

A deep result of DFS shows that the martingale problem related to (3.1) has a unique solution for essentially any admissible choice of triplets:

Theorem 3.1. *Let $(\beta_j, \gamma_j, \varphi_j)$, $j = 0, \ldots, d$, be admissible Lévy–Khintchine triplets and denote by ψ_j the corresponding Lévy exponents in the sense of (2.2). Suppose in addition that*

$$\int_{\{|x|\geq 1\}} |x|^k \varphi_j(dx) < \infty, \quad 1 \leq j, k \leq m. \tag{3.2}$$

Then the martingale problem related to (β, γ, φ) as in (3.1) and any initial distribution P_0 on $\mathbb{R}_+^m \times \mathbb{R}^n$ has a solution $(\Omega, \mathcal{F}, \mathbf{F}, P, X)$, where X is $\mathbb{R}_+^m \times \mathbb{R}^n$-valued. Its distribution is uniquely characterized by its conditional characteristic function

$$E\left(e^{i\lambda^\top X_{s+t}}\middle|\mathcal{F}_s\right) = \exp\left(\Psi^0(t, i\lambda) + \Psi^{(1,\ldots,d)}(t, i\lambda)^\top X_s\right), \quad \lambda \in \mathbb{R}^d, \tag{3.3}$$

where the mappings $\Psi^{(1,\ldots,d)} = (\Psi^1, \ldots, \Psi^d) : \mathbb{R}_+ \times (\mathbb{C}_-^m \times i\mathbb{R}^n) \to (\mathbb{C}_-^m \times i\mathbb{R}^n)$ and $\Psi^0 : \mathbb{R}_+ \times (\mathbb{C}_-^m \times i\mathbb{R}^n) \to \mathbb{C}$ solve the following system of generalized Riccati equations:

$$\Psi^0(0, u) = 0, \quad \Psi^{(1,\ldots,d)}(0, u) = u,$$

$$\frac{d}{dt}\Psi^j(t, u) = -\psi_j(\Psi^{(1,\ldots,d)}(t, u)), \quad j = 0, \ldots, d \tag{3.4}$$

(and $\mathbb{C}_-^m := \{z \in \mathbb{C}^m : \mathrm{Re}(z^j) \leq 0, j = 1, \ldots, m\}$).

Proof. Up to two details, the assertion follows directly from DFS, Theorems 2.7, 2.12 and Lemma 9.2. Equation (3.3) is derived in DFS under the additional assumptions that the initial distribution is of degenerate form $P_0 = \epsilon_x$ for $x \in \mathbb{R}_+^m \times \mathbb{R}^n$ and that the filtration $\mathbf{F}$ is generated by X. Hence, it suffices to reduce the general statement to this case.

Let $(\mathbb{D}^d, \mathcal{D}^d, \mathbf{D}^d)$ be the Skorohod path space of $\mathbb{R}^d$-valued càdlàg functions on $\mathbb{R}_+$ endowed with its natural filtration (cf. JS, Chapter VI). Denote by Y the canonical process, i.e. $Y_t(\alpha) = \alpha(t)$ for $\alpha \in \mathbb{D}^d$.

Fix $s \in \mathbb{R}_+, \omega \in \Omega$. From the characterization in Definition 1 (more precisely, from the slightly more general formulation in JS, II.2.42, because we do not know in the first place that Y is a semimartingale) it follows that Y has differential characteristics of the form (2.5) and (3.1) relative to the probability measure $\widetilde{P}_{s,\omega} := P^{(X_{s+t})_{t\in\mathbb{R}_+}|\mathcal{F}_s}(\omega,\cdot)$ on $(\mathbb{D}^d, \mathcal{D}^d)$ (except for some P-null set of ω's). Therefore, Y solves the affine martingale problem corresponding to (3.1) and it has degenerate initial distribution $\widetilde{P}^{Y_0}_{s,\omega} = \epsilon_{X_s(\omega)}$. Theorem 2.12 in DFS yields that

$$\begin{aligned} E\left(e^{i\lambda^\top X_{s+t}}\middle|\mathcal{F}_s\right)(\omega) &= \widetilde{E}_{s,\omega} e^{i\lambda^\top Y_t} \\ &= \widetilde{E}_{s,\omega}\left(\widetilde{E}_{s,\omega}\left(e^{i\lambda^\top Y_t}\middle|\mathcal{D}_0\right)\right) \\ &= \widetilde{E}_{s,\omega}\exp\left(\Psi^0(t,i\lambda) + \Psi^{(1,\dots,d)}(t,i\lambda)^\top Y_0\right), \\ &= \exp\left(\Psi^0(t,i\lambda) + \Psi^{(1,\dots,d)}(t,i\lambda)^\top X_s(\omega)\right) \end{aligned}$$

for P-almost all $\omega \in \Omega$. □

Remarks.

1. The restriction $X^1, \dots, X^m \geq 0$ has to be naturally imposed because otherwise $\gamma(X_{t-}, t)$, $\varphi(X_{t-}, t, G)$ in (3.1) may turn negative which does not make sense. The remaining n components $X^{m+1}, \dots, X^d$, on the other hand, affect the characteristics of X only through the drift rate β_j. Due to the conditions $\gamma_j = 0$, $\varphi_j = 0$, $j \geq m+1$, parts of the ODE system (3.4) are reduced actually to simple integrals and linear equations which can be solved in closed form (cf. (2.13)–(2.15) in DFS and Corollary 3.2 below for a special case).
2. Condition (3.2) guaranties that the solution process does not explode in finite time and hence is a semimartingale on $\mathbb{R}_+$ in the usual sense. It can be relaxed by a weaker necessary and sufficient condition (cf. DFS, Proposition 9.1).
3. By introducing the zeroth component $X^0_t = 1$, it is easy to see that an affine process in $\mathbb{R}^m_+ \times \mathbb{R}^n \subset \mathbb{R}^d$ can be interpreted as a process with *linear* characteristics in $\mathbb{R}^{1+m}_+ \times \mathbb{R}^n \subset \mathbb{R}^{1+d}$. Since the solution to linear ODE's are exponential functions, one could be tempted to call the solutions to such linear martingale problems "stochastic exponentials." However, this notion usually refers to the solutions to linear SDE's and the latter typically do not have linear characteristics. For example, Propositions 1 and 2 yield that the differential characteristics of the geometric Wiener process $X_t = 1 + \int_0^t X_s dW_s$ are of the form $\partial X = (0, X^2, 0)$. Hence they are quadratic rather than linear in X.

4. Observe that the solution depends on the involved triplets only through their Lévy exponents, which is agreeable for concrete models where the latter are known in closed form.

For such applications as, e.g., estimation purposes it is useful to dispose of a closed form expression of the finite-dimensional marginals. It follows by induction from Theorem 3.1.

Corollary 3.1. *The joint characteristic function of* $X_{t_1}, \ldots, X_{t_\nu}$ *is given by*

$$E\exp\left(i\sum_{k=1}^{\nu}\lambda^{k\cdot}X_{t_k}\right)$$
$$=\hat{P}_0\left(\Psi_\nu(t_1-t_0,\ldots,t_\nu-t_{\nu-1};i\lambda^{1\cdot},\ldots,i\lambda^{\nu\cdot})\right)\exp\left(\sum_{k=1}^{\nu}\Psi^0(t_k-t_{k-1},i\lambda^{k\cdot})\right),$$

for any $0=t_0\leq t_1\leq\cdots\leq t_\nu$ *and any* $\lambda\in\mathbb{R}^{\nu\times d}$, *where* $\hat{P}_0(u):=\int e^{ux}P_0(dx)$ *and* Ψ_ν *is defined recursively via*

$$\Psi_1(\tau_1;u_1):=\Psi^{(1,\ldots,d)}(\tau_1,u_1)$$

and

$$\Psi_k(\tau_1,\ldots,\tau_k;u_1,\ldots,u_k)$$
$$:=\Psi_{k-1}\left(\tau_1,\ldots,\tau_{k-1};u_1,\ldots,u_{k-2},u_{k-1}+\Psi^{(1,\ldots,d)}(\tau_k,u_k)\right).$$

Since an affine process is characterized by at most $d+1$ Lévy–Khintchine triplets, one may wonder whether it can in fact be expressed pathwise in terms of $d+1$ Lévy processes with the corresponding triplets. We give a partial answer to this question.

Theorem 3.2 (Time change representation of affine processes). *Let* X *be an affine process as in Theorem 3.1. On a possibly enlarged probability space, there exist intrinsic* $\mathbb{R}^d$*-valued Lévy processes* $L^{(j)}$ *with triplets* $(\beta_j,\gamma_j,\varphi_j)$, $j=0,\ldots,d$, *such that*

$$X_t=X_0+L_t^{(0)}+\sum_{j=1}^{d}L^{(j)}_{\Theta_t^j}\,, \tag{3.5}$$

where

$$\Theta_t^j=\int_0^t X^j_{s-}ds. \tag{3.6}$$

Proof. By an enlargement of the probability space $(\Omega,\mathcal{F},P)$ we refer, specifically, to a space of the form $(\Omega\times\mathbb{D}^{d'},\mathcal{F}\otimes\mathcal{D}^{d'},P')$ such that $P'(A\times\mathbb{D}^{d'})=P(A)$ for $A\in\mathcal{F}$. Here $\mathbb{D}^{d'}$ denotes as before the space of $\mathbb{R}^{d'}$-valued càdlàg functions. The process X is identified with the process X' on the enlarged space which is given by $X'_t(\omega,\alpha):=X_t(\omega)$ for $(\omega,\alpha)\in\Omega\times\mathbb{D}^{d'}$.

Step 1: Firstly, we choose triplets $(\widetilde{\beta}_j, \widetilde{\gamma}_j, \widetilde{\varphi}_j)$, $j = 0, \ldots, (d+2)d$, on $\mathbb{R}^{(d+2)d}$ as follows. For $j = 0, \ldots, d$, we define $(\widetilde{\beta}_j, \widetilde{\gamma}_j, \widetilde{\varphi}_j)$ as the Lévy–Khintchine triplet of the $\mathbb{R}^{(d+2)d}$-valued Lévy process $(V, U^0, \ldots, U^d)$ given by

$$U^k := \begin{cases} V & \text{if } k = j \\ 0 \in \mathbb{R}^d & \text{if } k \neq j, \end{cases}$$

where V denotes a $\mathbb{R}^d$-valued Lévy process with triplet $(\beta_j, \gamma_j, \varphi_j)$. For $j > d$, we set $(\widetilde{\beta}_j, \widetilde{\gamma}_j, \widetilde{\varphi}_j) = (0, 0, 0)$. One verifies easily that the new triplets $(\widetilde{\beta}_j, \widetilde{\gamma}_j, \widetilde{\varphi}_j), j = 0, \ldots, (d+2)d$ are admissible (with $\widetilde{d} := (d+2)d$, $\widetilde{m} := m$, $\widetilde{n} := \widetilde{d} - m$). By Theorem 3.1 (resp. DFS) there is an $\mathbb{R}^{(d+2)d}$-valued affine process $(\widetilde{X}, \widetilde{Y}^0, \ldots, \widetilde{Y}^d)$ corresponding to the initial distribution $\widetilde{P}_0 = P_0 \otimes \bigotimes_{j=0}^d \epsilon_0$ and the triplets $(\widetilde{\beta}_j, \widetilde{\gamma}_j, \widetilde{\varphi}_j)$; namely, the canonical process on the path space $(\mathbb{D}^{(d+2)d}, \mathcal{D}^{(d+2)d}, \mathbf{D}^{(d+2)d})$ relative to some law Q on that space.

Step 2: By applying Proposition 3 to the mapping $f(x, y^0, \ldots, y^d) = x$, we observe that the characteristics of the first d components $\widetilde{X}$ coincide with those of the original $\mathbb{R}^d$-valued affine process X. Since P_0 is the distribution of both X_0 and $\widetilde{X}_0$, we have that $P^X = Q^{\widetilde{X}}$, i.e. the laws of X and $\widetilde{X}$ coincide as well.

Step 3: On the product space $(\Omega', \mathcal{F}') := (\Omega \times \mathbb{D}^{(d+1)d}, \mathcal{F} \otimes \mathcal{D}^{(d+1)d})$ define a probability measure

$$P'(d\omega \times dy) := P(d\omega) Q^{(\widetilde{Y}^0, \ldots, \widetilde{Y}^d) | \widetilde{X} = X(\omega)}(dy)$$

and a $\mathbb{R}^{(d+2)d}$-valued process $(X', Y^0, \ldots, Y^d)$ with

$$(X', Y^0, \ldots, Y^d)_t(\omega, y) := (X_t(\omega), y(t)).$$

Its distribution $P'^{(X', Y^0, \ldots, Y^d)}$ equals Q by Step 2. If the filtration $\mathbf{F}'$ on $(\Omega', \mathcal{F}')$ is chosen to be generated by $(X', Y^0, \ldots, Y^d)$, then this process is affine in the sense of Theorem 3.1 corresponding to the triplets $(\widetilde{\beta}_j, \widetilde{\gamma}_j, \widetilde{\varphi}_j)$. As suggested before Step 1, we identify X' on the enlarged space with X on the original space.

Step 4: Applying Proposition 3 to the mapping

$$f(x, y^0, \ldots, y^d) = x - \sum_{j=0}^d y^j$$

yields that $X - \sum_{j=0}^d Y^j$ has differential characteristics $(0, 0, 0)$, which implies that it is constant, i.e.

$$X = X_0 + \sum_{j=0}^d Y^j.$$

Step 5: Finally, applying Proposition 3 to $f(x, y^0, \dots, y^d) = y^j$ yields that Y^j has differential characteristics

$$\partial Y^j = (X_-^j \beta_j, X_-^j \gamma_j, X_-^j \varphi_j) \tag{3.7}$$

for $j = 1, \dots, d$ and $\partial Y^0 = (\beta_0, \gamma_0, \varphi_0)$. In particular, $L^{(0)} := Y^0$ is a Lévy process.

Step 6: Let $j \in \{m+1, \dots, d\}$. Since $\gamma_j = 0$, $\varphi_j = 0$, we have that

$$Y_t^j = \beta_j \int_0^t X_{s-}^j ds = L_{\Theta_t^j}^{(j)}$$

for the deterministic Lévy process $L_\theta^{(j)} := \beta_j \theta$ and the (not necessarily increasing) "time change" (3.6).

Step 7: Now, let $j \in \{1, \dots, m\}$. For $\theta \in \mathbb{R}_+$ define

$$T_\theta^j := \inf\{t \in \mathbb{R}_+ : \ \Theta_t^j > \theta\}.$$

Since $\Theta^j = (\Theta_t^j)_{t \in \mathbb{R}_+}$ is adapted, we have that its inverse $T^j = (T_\theta^j)_{\theta \in \mathbb{R}_+}$ is a time change in the sense of [11], §10.1a.

For $H := 1_{\{X_-^j = 0\}}$ we have $\partial(H \bullet Y^j) = (0, 0, 0)$ by Proposition 2, which implies that $H \bullet Y^j = 0$. For fixed $\omega' \in \Omega'$ consider $u < v$ with $\Theta_u^j = \Theta_v^j$. Then $(u, v] \subset \{t \in \mathbb{R}_+ : X_{t-}^j(\omega') = 0\}$, which implies that

$$Y_v^j - Y_u^j = H \bullet Y_v^j - H \bullet Y_u^j = 0.$$

In view of [11], (10.14), it follows that Y^j is T^j-adapted.

Define the time-changed process $L^{(j)} := Y^j \circ T^j$ (in the sense of [11], (10.6) if $T_\theta^j = \infty$ for finite θ, i.e. if $\Theta_\infty^j < \infty$). The integral characteristics of $L^{(j)}$ relative to the corresponding time-changed filtration equal $(\widetilde{B}, \widetilde{C}, \widetilde{\nu})$ with

$$\widetilde{B}_\theta = B_{T_\theta^j}, \quad \widetilde{C}_\theta = C_{T_\theta^j}, \quad \widetilde{\nu}([0, \theta] \times \cdot) = \nu([0, T_\theta^j] \times \cdot), \tag{3.8}$$

where (B, C, ν) denote the integral characteristics of Y^j. This is stated in [16], Lemma 5, for the case $\Theta_\infty^j = \infty$. In the general case $L^{(j)}$ may only be a semimartingale on $[\![0, \Theta_\infty^j[\![$ in the sense of [11], (5.4). Then (3.8) holds on this stochastic interval as can be deduced from [11], (10.17), (10.27).

Consequently,

$$\widetilde{B}_\theta = B_{T_\theta^j} = \beta_j \int_0^{T_\theta^j} X_{s-}^j ds = \beta_j (\Theta^j \circ T^j)_\theta = \beta_j \theta$$

and accordingly for $\widetilde{C}, \widetilde{\nu}$ if $\theta < \Theta_\infty^j$. This means that $L^{(j)}$ is a "Lévy process on $[\![0, \Theta_\infty^j[\![$" in the sense that its characteristics on $[\![0, \Theta_\infty^j[\![$ equal those of a Lévy process with triplet $(\beta_j, \gamma_j, \varphi_j)$.

Step 8: By "glueing" $(L^{(j)}_\theta)_{\theta\in[0,\Theta^j_\infty)}$ together with another Lévy process on $[\![\Theta^j_\infty,\infty[\![$ having the same triplet, we extend $L^{(j)}$ to the whole $\mathbb{R}_+$. This can be done along the lines of [11], (10.32) and §10.2b after an enlargement of the probability space.

Since Y^j is T^j-adapted (cf. Step 7), we have $Y^j_t = L^{(j)}_{\Theta^j_t}$ for any $t\in\mathbb{R}_+$. The assertion follows now from Step 4. □

The previous result is not entirely satisfactory in some aspects. E.g., it is not shown that X is a measurable function of $L^{(j)}, j = 0,\dots,d$, i.e., loosely speaking, that X is a *strong* solution of the time change equations (3.5)-(3.6).

For the purposes of the subsequent section, let us state a simple special case of Theorem 3.1. We suppose that $m = n = 1$, where the second component X^2 will denote a logarithmic asset price in the affine SV models considered below. We assume that it has no mean-reverting term. Secondly, we suppose that the "volatility" process X^1 is of the Ornstein–Uhlenbeck type. This means that the Riccati-type equation (3.4) is an affine ODE, which can be solved explicitly.

Corollary 3.2. *In the case $m = n = 1$ suppose that $(\beta_j,\gamma_j,\varphi_j)$, $j = 0,1,2$, are Lévy–Khintchine triplets such that*

$$\begin{aligned}
\beta^1_0 - \textstyle\int h^1(x)\varphi_0(dx) &\geq 0, \\
\gamma^{kl}_0 &= 0 \qquad \text{unless } k=l=2,\\
\varphi_0((\mathbb{R}_+\times\mathbb{R})^C) &= 0,\\
\textstyle\int h^1(x)\varphi_0(dx) &< \infty,\\[1em]
\gamma^{kl}_1 &= 0 \qquad \text{unless } k=l=2,\\
\varphi_1((\{0\}\times\mathbb{R})^C) &= 0,\\[1em]
(\beta_2,\gamma_2,\varphi_2) &= (0,0,0).
\end{aligned}$$

Then the martingale problem related to (β,γ,φ) as in (3.1) and any initial distribution P_0 on $\mathbb{R}_+\times\mathbb{R}$ has a solution $(\Omega,\mathcal{F},\mathbf{F},P,X)$, where X is $\mathbb{R}_+\times\mathbb{R}$-valued. Its distribution is uniquely characterized by its conditional characteristic function

$$E\left(e^{i\lambda^1X^1_{s+t}+i\lambda^2X^2_{s+t}}\middle|\mathcal{F}_s\right) = \exp\left(\Psi^0(t,i\lambda^1,i\lambda^2)+\Psi^1(t,i\lambda^1,i\lambda^2)X^1_s+i\lambda^2X^2_s\right),$$

where $\Psi^j:\mathbb{R}_+\times(\mathbb{C}_-\times i\mathbb{R})\to\mathbb{C}$, $j=0,1$, are given by

$$\begin{aligned}
\Psi^1(t,u^1,u^2) &= e^{\beta^1_1t}u^1 - \frac{1-e^{\beta^1_1t}}{\beta^1_1}\psi_1(0,u^2),\\
\Psi^0(t,u^1,u^2) &= \int_0^t \psi_0(\Psi^1(s,u^1,u^2),u^2)ds.
\end{aligned}$$

4 Affine stochastic volatility models

In the empirical literature, a number of so-called stylized facts has been reported repeatedly, namely semi-heavy tails in the return distribution, volatility clustering, and a negative correlation between changes in volatility and asset prices (*leverage effect*). These features are reflected in the SV models that have been suggested. At the same time, it seems desirable to work in settings which are analytically tractable. Here, affine models play an important role. The fact that the characteristic function is known in closed or semi-closed form opens the door to derivative pricing, calibration, hedging, and estimation.

If the model is set up under the risk-neutral measure, European option prices can be computed by Laplace transform methods. This approach relies on the fact that the characteristic function or Laplace transform can be interpreted as a set of prices of complex-valued contingent claims. A large class of arbitrary payoffs can be represented explicitly as a linear combination or, more precisely, integral of such "simple" claims (cf. e.g. [4], [20]). As far as estimation is concerned, the knowledge of the joint characteristic function can be exploited for generalized moment estimators (cf. [13] and [26] for an overview).

Typically, (broad-sense) stochastic volatility models fall into two groups. Either market activity is expressed in terms of the time-varying *size* or magnitude of price movements, or alternatively, by their *speed* or arrival rate. The models of the first group are often stated in terms of an equation

$$dX_t = \sigma_t dL_t, \tag{4.1}$$

possibly modified by an additional drift term. Here, X denotes the logarithm of an asset price and L a Lévy process as, e.g., Brownian motion. In this equation, the SV process σ affects the size of relative price moves.

Models of the second kind arise from time-changed Lévy processes

$$X_t = X_0 + L_{V_t}. \tag{4.2}$$

Again, L denotes a Lévy process and X the logarithm of the asset price. Here, the time change $V_t = \int_0^t v_s ds$ affects the speed of price moves. Often V_t is interpreted as business time. Measured on this operational time scale, log prices evolve homogeneously but due to randomly changing trading activity v_t, this is not true relative to calender time.

If the Lévy process L is a Wiener process and if L, σ, respectively L, v, are independent, then the two approaches lead essentially to the same models. This fact is due to the self-similarity of Brownian motion and it is reflected by Propositions 2 and 5, where the choice $v_t = \sigma_t^2$ leads to the same differential characteristics of X in either case. Again due to self-similarity, the correspondence between (4.1) and (4.2) remains true for α-stable Lévy motions L. In this case, $v_t = \sigma_t^\alpha$ leads to the same characteristics (cf. also [17] in this respect). For general Lévy processes, however, (4.1) and (4.2) lead to quite different models because the change of measure in Proposition 2 does not lead to a multiple of F as in Proposition 5. Except for α-stable Lévy motions L, models of type (4.1), in general, do not lead to affine processes. Typically, the distribution of X is not known in closed form.

Another important distinction refers to the sources of randomness that drive the Lévy process L and the volatility process σ resp. v in (4.1) and (4.2). In the simplest case, these two are supposed to be independent. This, however, excludes the above-mentioned leverage effect, i.e. it does not allow for negative correlation between volatility and asset price changes. Whereas such a correlation can be incorporated easily in models of type (4.1), this is less obvious in (4.2) because L and v live on different time scales (business vs. calender time).

The other extreme would be to use a common source of randomness for both L and σ or L and v, respectively. This can be interpreted in the sense that changes in volatility are caused by changes in asset prices. This spirit underlies the ARCH-type models in the econometric literature. An interesting and natural continuous-time extension of GARCH(1,1) has recently been suggested in [18]. But since ARCH models are based on rescaling the innovations in the sense of (4.1), they do not lead to an affine structure. Nevertheless, the idea to use a common driver for volatility and price moves can be carried out in the context of affine processes as well (cf. Subsection 4.6).

We will now discuss a number of well-known affine SV models from the point of view of characteristics. For a more exhaustive coverage of the literature, see DFS and [5]. We express the characteristics of the affine processes in terms of triplets (3.1). By straightforward insertion one can derive closed-form expressions for the corresponding Lévy exponents $\psi_j, j = 0, \dots, d$, in terms of the Lévy exponents of the involved Lévy processes and the additional parameters in the corresponding model.

In all the examples, it is implicitly assumed that the filtration is generated by the affine process under consideration (cf. the last remark of Subsection 4.8 in this context). Moreover, we assume generally that the identity $h(x) = x$ is used as "truncation" function because this simplifies some of the expressions considerably. This choice implies that the corresponding Lévy measures have first moments in the tails. The general formulation without such moment assumptions can be derived immediately from (2.3).

4.1 Stein and Stein (1991)

Slightly generalized, the model in [24] is of the form

$$
\begin{aligned}
dX_t &= (\mu + \delta\sigma_t^2)dt + \sigma_t dW_t, \\
d\sigma_t &= (\kappa - \lambda\sigma_t)dt + \alpha dZ_t
\end{aligned}
\tag{4.3}
$$

with constants $\kappa \geq 0, \mu, \delta, \lambda, \alpha$ and Wiener processes W, Z having constant correlation ρ. As can be seen from straightforward application of Propositions

1-3, neither (σ, X) nor (σ^2, X) have affine characteristics in the sense of (3.1) unless the parameters are chosen in a very specific way (e.g. $\kappa = 0$). However, the $\mathbb{R}^3$-valued process (σ, σ^2, X) is "almost" the solution to an affine martingale problem related with (3.1), namely, for $(\beta_j, \gamma_j, \varphi_j)$, $j = 0, \ldots, 3$, given by

$$\begin{aligned}
(\beta_0, \gamma_0, \varphi_0) &= \left(\begin{pmatrix} \kappa \\ \alpha^2 \\ \mu \end{pmatrix}, \begin{pmatrix} \alpha^2 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, 0 \right), \\
(\beta_1, \gamma_1, \varphi_1) &= \left(\begin{pmatrix} -\lambda \\ 2\kappa \\ 0 \end{pmatrix}, \begin{pmatrix} 0 & 2\alpha^2 & \alpha\rho \\ 2\alpha^2 & 0 & 0 \\ \alpha\rho & 0 & 0 \end{pmatrix}, 0 \right), \\
(\beta_2, \gamma_2, \varphi_2) &= \left(\begin{pmatrix} 0 \\ -2\lambda \\ \delta \end{pmatrix}, \begin{pmatrix} 0 & 0 & 0 \\ 0 & 4\alpha^2 & 2\alpha\rho \\ 0 & 2\alpha\rho & 1 \end{pmatrix}, 0 \right), \\
(\beta_3, \gamma_3, \varphi_3) &= (0, 0, 0).
\end{aligned}$$

Since γ_1 is not non-negative definite, $(\beta_1, \gamma_1, \varphi_1)$ is not a Lévy–Khintchine triplet in the usual sense and hence Theorem 3.1 cannot be applied. Nevertheless, the Riccati-type equation (3.4) leads to the correct characteristic function in this case (see, e.g., the derivation in [22]). The process (σ, σ^2, X) is closely related to the non-degenerate example in DFS, Subsection 12.2 of an affine Markov process with a non-standard maximal domain.

4.2 Heston (1993)

If κ is chosen to be 0 in the Ornstein-Uhlenbeck equation (4.3), then the Stein and Stein model reduces to a special case of the model in [10]:

$$\begin{aligned}
dX_t &= (\mu + \delta v_t)dt + \sqrt{v_t}dW_t, \\
dv_t &= (\kappa - \lambda v_t)dt + \sigma\sqrt{v_t}dZ_t.
\end{aligned} \tag{4.4}$$

Here, $\kappa \geq 0, \mu, \delta, \lambda, \sigma$ denote constants and W, Z Wiener processes with constant correlation ρ. Calculation of the characteristics yields that (v, X) is an affine process with triplets $(\beta_j, \gamma_j, \varphi_j)$, $j = 0, 1, 2$, in (3.1) given by

$$\begin{aligned}
(\beta_0, \gamma_0, \varphi_0) &= \left(\begin{pmatrix} \kappa \\ \mu \end{pmatrix}, 0, 0 \right), \\
(\beta_1, \gamma_1, \varphi_1) &= \left(\begin{pmatrix} -\lambda \\ \delta \end{pmatrix}, \begin{pmatrix} \sigma^2 & \sigma\rho \\ \sigma\rho & 1 \end{pmatrix}, 0 \right), \\
(\beta_2, \gamma_2, \varphi_2) &= (0, 0, 0).
\end{aligned}$$

4.3 Barndorff-Nielsen and Shephard (2001)

In the article [1] (henceforth BNS) it is considered a model of the form

$$\begin{aligned} dX_t &= (\mu + \delta v_{t-})dt + \sqrt{v_{t-}}dW_t + \rho dZ_t, \\ dv_t &= -\lambda v_{t-}dt + dZ_t. \end{aligned} \tag{4.5}$$

Here, $\mu, \delta, \rho, \lambda$ denote constants, W a Wiener processes, and Z a subordinator (i.e. an increasing Lévy process) with Lévy–Khintchine triplet $(b^Z, 0, F^Z)$. Compared to the Heston model, the square-root process (4.4) is replaced with a Lévy-driven Ornstein–Uhlenbeck (OU) process. Since W and Z are necessarily independent, leverage is introduced by the ρdZ_t term. Again, Propositions 1 and 2 yield that (v, X) is an affine process with triplets $(\beta_j, \gamma_j, \varphi_j)$, $j = 0, 1, 2$, in (3.1) of the form

$$\beta_0 = \begin{pmatrix} b^Z \\ \mu + \rho b^Z \end{pmatrix}, \quad \gamma_0 = 0, \quad \varphi_0(G) = \int 1_G(y, \rho y)F^Z(dy), \quad G \in \mathcal{B}^2,$$

$$\begin{aligned} (\beta_1, \gamma_1, \varphi_1) &= \left(\begin{pmatrix} -\lambda \\ \delta \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}, 0 \right), \\ (\beta_2, \gamma_2, \varphi_2) &= (0, 0, 0). \end{aligned}$$

Due to the simple structure of the characteristics, we are in the situation of Corollary 3.2.

BNS consider also a slightly extended version of the above model. They argue that the autocorrelation pattern of volatility is not appropriately matched by a single OU process. As a way out they suggest a linear combination of independent OU processes, i.e. a model of the form

$$\begin{aligned} dX_t &= (\mu + \delta v_{t-})dt + \sqrt{v_{t-}}dW_t + \sum_{k=1}^{\nu} \rho_k dZ_t^k, \\ v_t &= \sum_{k=1}^{\nu} \alpha_k v_t^{(k)}, \\ dv_t^{(k)} &= -\lambda_k v_{t-}^{(k)} dt + dZ_t^k, \end{aligned}$$

with constants $\alpha_1, \dots, \alpha_\nu \geq 0$, $\mu, \delta, \rho_1, \dots, \rho_\nu, \lambda_1, \dots, \lambda_\nu$, a Wiener processes W, and a $\mathbb{R}^\nu$-valued Lévy process Z with triplet $(b^Z, 0, F^Z)$ whose components are independent subordinators. $(v^{(1)}, \dots, v^{(\nu)}, v, X)$ is a $\mathbb{R}^{\nu+2}$-valued affine process whose triplets $(\beta_j, \gamma_j, \varphi_j)$, $j = 0, \dots, \nu + 2$ are of the form

$$\beta_0 = \begin{pmatrix} b^{Z^1} \\ \vdots \\ b^{Z^\nu} \\ \sum_k \alpha_k b^{Z^k} \\ \mu + \sum_k \rho_k b^{Z^k} \end{pmatrix}, \quad \gamma_0 = 0,$$

$$\varphi_0(G) = \int 1_G(y^1, \dots, y^\nu, \textstyle\sum_{k=1}^{\nu} \alpha_k y^k, \sum_{k=1}^{\nu} \rho_k y^k)F^Z(dy), \quad G \in \mathcal{B}^{\nu+2},$$

$$(\beta_k, \gamma_k, \varphi_k) = \left((0, \dots, 0, -\lambda_k, 0, \dots, 0, -\alpha_k\lambda_k, 0)^\top, 0, 0\right), \ k = 1, \dots, \nu,$$

$$(\beta_{\nu+1}, \gamma_{\nu+1}, \varphi_{\nu+1}) = \left(\begin{pmatrix} 0 \\ \vdots \\ 0 \\ \delta \end{pmatrix}, \begin{pmatrix} 0 & \dots & 0 & 0 \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \cdots & 0 & 0 \\ 0 & \cdots & 0 & 1 \end{pmatrix}, 0 \right),$$

$$(\beta_{\nu+2}, \gamma_{\nu+2}, \varphi_{\nu+2}) = (0, 0, 0).$$

In order to preserve this affine structure, the subordinators $Z^1, \dots, Z^\nu$ do not have to be independent. The other extreme case $Z^1 = \dots = Z^\nu$, leads to the realm of continuous-time ARMA processes proposed in [2].

4.4 Carr, Geman, Madan, Yor (2003)

The paper [3] (henceforth CGMY) generalizes both the Heston and the BNS model by allowing for jumps in the asset price. As noted at the beginning of this section, one must consider time changes in order to preserve the affine structure unless the driver of the asset price changes is a stable Lévy motion (cf. Subsection 4.5).

The analogue of the Heston model is

$$\begin{aligned} X_t &= X_0 + \mu t + L_{V_t} + \rho(v_t - v_0), \\ dV_t &= v_t dt, \\ dv_t &= (\kappa - \lambda v_t)dt + \sigma\sqrt{v_t}dZ_t, \end{aligned} \tag{4.6}$$

where $\kappa \geq 0, \mu, \rho, \lambda, \sigma$ are constants, L denotes a Lévy process with triplet (b^L, c^L, F^L) and Z an independent Wiener process. Again, (v, X) is an affine process whose triplets $(\beta_j, \gamma_j, \varphi_j)$, $j = 0, 1, 2$ meet the equations

$$(\beta_0, \gamma_0, \varphi_0) = \left(\begin{pmatrix} \kappa \\ \mu + \rho\kappa \end{pmatrix}, 0, 0 \right),$$

$$\beta_1 = \begin{pmatrix} -\lambda \\ b^L - \rho\lambda \end{pmatrix}, \ \gamma_1 = \begin{pmatrix} \sigma^2 & \sigma^2\rho \\ \sigma^2\rho & \sigma^2\rho^2 + c^L \end{pmatrix}, \ \varphi_1(G) = \int 1_G(0, x) F^L(dx),$$

$$(\beta_2, \gamma_2, \varphi_2) = (0, 0, 0).$$

Observe that we recover the characteristics of the Heston model – up to a rescaling of the volatility process v – if L is chosen to be a Brownian motion with drift.

PROOF. It remains to be shown that the differential characteristics of (v, X) are as claimed above. Note that ∂v and $\partial(L \circ V)$ are obtained from Propositions 2 and 5, respectively. For any $\mathbb{R}^2$-valued semimartingale Y with $\partial Y = (b, c, F)$

we have $\partial Y^1 = (b^1, c^{11}, F^1)$ with $F^1(G) := F((G \setminus \{0\}) \times \mathbb{R})$ and likewise for Y^2, e.g., by Proposition 3.

Since v does not jump, this yields $F_t(G) = \int 1_G(0,x) F_t^{L\circ V}(dx)$, $G \in \mathcal{B}$, for the joint Lévy measure F of $(v, L \circ V)$. Consequently, $\partial(v, L \circ V) =: (b, c, F)$ is completely determined if we know c^{12} $(= c^{21})$. Since L is independent of Z and hence of v, it follows from some technical arguments that $\langle v, L \circ V \rangle = 0$, which implies that $c^{12} = 0$ by JS, II.2.6. Applying Proposition 3 to the mapping $f(y,x) = (y, x + \rho y)$ yields $\partial(v, X)$ in the case $\mu = 0$. The modification $\mu \neq 0$ just affects the drift coefficient of X. □

In order to generalize the BNS model, the square-root process (4.6) is replaced with a Lévy-driven OU process:

$$\begin{aligned} X_t &= X_0 + \mu t + L_{V_t} + \rho Z_t, \\ dV_t &= v_{t-} dt, \\ dv_t &= -\lambda v_{t-} dt + dZ_t. \end{aligned} \tag{4.7}$$

Here, μ, ρ, λ are constants and L, Z denote independent Lévy processes with triplets (b^L, c^L, F^L) and $(b^Z, 0, F^Z)$, respectively, and Z is supposed to be increasing. The triplets $(\beta_j, \gamma_j, \varphi_j)$, $j = 0, 1, 2$, of the affine process (v, X) are given by

$$\beta_0 = \begin{pmatrix} b^Z \\ \mu + \rho b^Z \end{pmatrix}, \quad \gamma_0 = 0, \quad \varphi_0(G) = \int 1_G(y, \rho y) F^Z(dy), \quad G \in \mathcal{B}^2,$$

$$\beta_1 = \begin{pmatrix} -\lambda \\ b^L \end{pmatrix}, \quad \gamma_1 = \begin{pmatrix} 0 & 0 \\ 0 & c^L \end{pmatrix}, \quad \varphi_1(G) = \int 1_G(0, x) F^L(dy), \quad G \in \mathcal{B}^2,$$

$$(\beta_2, \gamma_2, \varphi_2) = (0, 0, 0).$$

For a Brownian motion with drift L, we recover the dynamics of the BNS model (4.5). As in that case, Corollary 3.2 can be applied.

Proof. The differential characteristics of (v, X) are derived similarly as above. Again, ∂v and $\partial(L \circ V)$ are obtained from Propositions 2 and 5, respectively. If we write $\partial(v, L \circ V) =: (b, c, F)$, then $c^{11} = 0$ and hence also $c^{12} = c^{21} = 0$. The marginal of the instantaneous Lévy measure F_t are given by the corresponding Lévy measures of v and $L \circ V$, respectively. Since L is independent of Z, we have that v and $L \circ V$ never jump at the same time (up to some P-null set). Consequently, F is concentrated on the coordinate axes, which implies that $F(G) = \int 1_G(y, 0) F^v(dy) + \int 1_G(0, x) F^{L\circ V}(dx)$. As above, Proposition 3 yields the characteristics of $(v, \widetilde{X})$ for $\widetilde{X} := L_{V_t} + \rho v_t$. Since $dX_t = d\widetilde{X}_t + (\mu + \lambda v_t)dt$, a correction of the drift yields $\partial(v, X)$. □

4.5 Carr and Wu (2003)

The study [5] considers a modification of the Heston model where the Wiener process W is replaced by an α-stable Lévy motion L with $\alpha \in (1,2)$ and Lévy–Khintchine triplet $(0,0,F^L)$:

$$\begin{aligned} dX_t &= \mu dt + v_t^{1/\alpha} dL_t, \\ dv_t &= (\kappa - \lambda v_t)dt + \sigma\sqrt{v_t} dZ_t. \end{aligned}$$

The self-similarity of L is reflected by the fact that $\int 1_G(c^{1/\alpha}x)F^L(dx) = cF^L(G)$ for $c > 0$, $G \in \mathcal{B}$. An application of Propositions 1 and 2 shows that (v,X) is an affine process with triplets $(\beta_j, \gamma_j, \varphi_j)$, $j = 0,1,2$, of the form

$$(\beta_0, \gamma_0, \varphi_0) = \left(\begin{pmatrix} \kappa \\ \mu \end{pmatrix}, 0, 0 \right),$$

$$\beta_1 = \begin{pmatrix} -\lambda \\ 0 \end{pmatrix}, \quad \gamma_1 = \begin{pmatrix} \sigma^2 & 0 \\ 0 & 0 \end{pmatrix}, \quad \varphi_1(G) = \int 1_G(0,x)F^L(dy), \quad G \in \mathcal{B}^2,$$

$$(\beta_2, \gamma_2, \varphi_2) = (0,0,0).$$

4.6 Carr and Wu (2004) and affine ARCH-like models

In the paper [6] it is considered a number of models, two of which could be written in the form

$$X_t = X_0 + \mu t + L_{V_t}, \tag{4.8}$$

$$dV_t = v_{t-} dt, \tag{4.9}$$

$$v_t = v_0 + \kappa t + Z_{V_t} \tag{4.10}$$

with constants $\kappa \geq 0$, μ and a Lévy process (Z,L) in $\mathbb{R}^2$ with triplet $(b^{(Z,L)}, c^{(Z,L)}, F^{(Z,L)})$, where Z has only non-negative jumps and finite expected value $E(Z_1)$.

Note that the above equation $v_t = v_0 + \kappa t + Z_{\int_0^t v_{s-} ds}$ is implicit. It may not be evident in the first place that a unique solution to this time change equation exists. On the other hand, the affine martingale problem corresponding to triplets $(\beta_j, \gamma_j, \varphi_j)$, $j = 0,1,2$, of the form

$$\begin{aligned} (\beta_0, \gamma_0, \varphi_0) &= \left(\begin{pmatrix} \kappa \\ \mu \end{pmatrix}, 0, 0 \right), \\ (\beta_1, \gamma_1, \varphi_1) &= \left(b^{(Z,L)}, c^{(Z,L)}, F^{(Z,L)} \right), \\ (\beta_2, \gamma_2, \varphi_2) &= (0,0,0) \end{aligned}$$

has a unique solution by Theorem 3.1. In view of Theorem 3.2, the solution process (v,X) can be expressed in the form (4.8)–(4.10) for some Lévy process (Z,L) with triplet $(b^{(Z,L)}, c^{(Z,L)}, F^{(Z,L)})$.

The paper [6] discusses two particular cases of the above setup, namely a joint compound Poisson process with drift (Z, L) and, alternatively, the completely dependent case $Z_t = -\lambda t - \sigma L_t$ with constants λ, σ and some totally skewed α-stable Lévy motion L, where $\alpha \in (1, 2]$. The latter model has an ARCH-like structure in the sense that the same source of randomness L drives both the volatility and the asset price process. This extends to a more general situation where L is an arbitrary Lévy process and $\Delta Z_t = f(\Delta L_t)$ for some deterministic function $f : \mathbb{R} \to \mathbb{R}_+$ as e.g. $f(x) = x^2$. If L or f are asymmetric, such models allow for leverage. A drawback of this setup is that it is not of the simple structure in Corollary 3.2. Non-trivial ODE's may have to be solved in order to obtain the characteristic function.

4.7 A model with flexible leverage

Any affine SV model can be defined directly in terms of the involved Lévy–Khintchine triplets, sometimes in the simple form of Corollary 3.2. Since this leads automatically to handy formulas for characteristic functions as well as differential characteristics, there is in principle no need for a stochastic differential equation or the like. Still, concrete equations of the above type may be useful in order to reduce generality and to give more insight in the structure of a model.

Observe that the dependence structure between changes in asset prices and volatility in (4.7) is quite restrictive in the sense that any rise ΔZ_t in volatility results in a perfectly correlated move $\rho \Delta Z_t$ of the asset. This cannot be relaxed easily by considering dependent Lévy processes $L,\ Z$ because these two live on different time scales. In this subsection, we suggest a class of models in the spirit of (4.7), which is more flexible as far as the leverage effect is concerned. Nevertheless, we retain the simple structure of Corollary 3.2, where no Riccati-type equations have to be solved.

$$\begin{aligned} X_t &= X_0 + L_{V_t} + Y_t, \\ dV_t &= v_{t-} dt, \\ dv_t &= -\lambda v_{t-} dt + dZ_t. \end{aligned} \tag{4.11}$$

Here, λ is a constant and L a Lévy process with triplet (b^L, c^L, F^L), which is assumed to be independent of another Lévy process (Z, Y) in $\mathbb{R}^2$ with triplet $(b^{(Z,Y)}, c^{(Z,Y)}, F^{(Z,Y)})$ and Z is supposed to be a subordinator. As before, (v, X) is an affine process with triplets $(\beta_j, \gamma_j, \varphi_j)$, $j = 0, 1, 2$, given by

$$(\beta_0, \gamma_0, \varphi_0) = \left(b^{(Z,Y)}, c^{(Z,Y)}, F^{(Z,Y)}\right), \tag{4.12}$$

$$\beta_1 = \begin{pmatrix} -\lambda \\ b^L \end{pmatrix}, \quad \gamma_1 = \begin{pmatrix} 0 & 0 \\ 0 & c^L \end{pmatrix}, \quad \varphi_1(G) = \int 1_G(0, x) F^L(dx), \quad G \in \mathcal{B}^2,$$

$$(\beta_2, \gamma_2, \varphi_2) = (0, 0, 0)\,.$$

PROOF. This follows similarly as in Subsection 4.4. In a first step, one derives $\partial(v, Y)$ and $\partial(L \circ V)$ from Propositions 2 and 5. Since these two processes have zero covariation and never jump together, this leads to the joint characteristics $\partial(v, Y, L \circ V)$ in the same way as for (4.7). Applying Proposition 3 yields $\partial(v, X)$. □

The model (4.11) remains vague about how to choose the dependence structure between Z and Y. Therefore, we consider the following more concrete special case of the above setup:

$$\begin{aligned} X_t &= X_0 + \mu t + L_{V_t} + U_{Z_t}, \\ dV_t &= v_{t-}dt, \\ dv_t &= (\kappa - \lambda v_{t-})dt + dZ_t, \end{aligned}$$

where $\kappa \geq 0, \lambda$ are constants and L, U, Z three independent Lévy processes. The triplet of L is denoted by (b^L, c^L, F^L) and Z is supposed to be a subordinator which equals the sum of its jumps, i.e. with triplet $(b^Z, 0, F^Z)$ where $b^Z = \int zF^Z(dz)$. The triplets in (3.1) of the affine process (v, X) are of the form

$$\beta_0 = \begin{pmatrix} \kappa + b^Z \\ \mu + b^Z E(U_1) \end{pmatrix}, \quad \gamma_0 = 0, \quad \varphi_0(G) = \int 1_G(z, x) P^{U_z}(dx) F^Z(dz),$$

$$\beta_1 = \begin{pmatrix} -\lambda \\ b^L \end{pmatrix}, \quad \gamma_1 = \begin{pmatrix} 0 & 0 \\ 0 & c^L \end{pmatrix}, \quad \varphi_1(G) = \int 1_G(0, x) \varphi^L(dx), \quad G \in \mathcal{B}^2,$$

$$(\beta_2, \gamma_2, \varphi_2) = (0, 0, 0),$$

where P^{U_θ} denotes the law of U_θ for $\theta \in \mathbb{R}_+$. Since the structure of the corresponding Lévy exponent ψ_0 is less obvious in this case, we express it explicitly in terms of the Lévy exponents ψ^L, ψ^U, ψ^Z of L, Z, U, respectively.

$$\begin{aligned} \psi_0(u^1, u^2) &= \kappa u^1 + \mu u^2 + \psi^Z\left(u^1 + \psi^U(u^2)\right), \\ \psi_1(u^1, u^2) &= -\lambda u^1 + \psi^L(u^2) \end{aligned}$$

PROOF. To determine the triplets (4.12), it remains to derive the joint characteristics of $(\widetilde{Z}, \widetilde{Y})_t := (\kappa t + Z_t, \mu t + U_{Z_t})$. Note that $(\widetilde{Z}_t - \kappa t, \widetilde{Y}_t - \mu t) = \widetilde{U} \circ Z$ for the $\mathbb{R}^2$-valued Lévy process $\widetilde{U}_\theta = (\theta, U_\theta)$. Here, Proposition 5 cannot be applied because the time change Z is not continuous. But [21], Theorem 30.1, yields that $\widetilde{U} \circ Z$ is a Lévy process with triplet $(b^{\widetilde{U} \circ Z}, 0, F^{\widetilde{U} \circ Z})$, where

$$b^{\widetilde{U} \circ Z} = \begin{pmatrix} b^Z \\ b^Z E(U_1) \end{pmatrix}, \quad F^{\widetilde{U} \circ Z}(G) = \int 1_G(z, x) P^{U_z}(dx) F^Z(dz), \quad G \in \mathcal{B}^2.$$

Since

$$\begin{aligned} E\exp\left(u^1\widetilde{Z}_t+u^2\widetilde{Y}_t\right) &= E\left(E\left(\exp\left(u^1(\kappa t+Z_t)+u^2(\mu t+U_{Z_t})\right)\big| Z\right)\right) \\ &= \exp(u^1\kappa t+u^2\mu t)E\exp\left(u^1Z_t+Z_t\psi^U(u^2)\right) \\ &= \exp\left(t\left(\kappa u^1+\mu u^2+\psi^Z\left(u^1+\psi^U(u^2)\right)\right)\right), \end{aligned}$$

the Lévy exponent of $(\widetilde{Z},\widetilde{Y})$ is given by

$$\psi^{(\widetilde{Z},\widetilde{Y})}(u^1,u^2)=\kappa u^1+\mu u^2+\psi^Z(u^1+\psi^U(u^2)).$$

The Lévy exponents ψ_0,ψ_1 follow now directly from (4.12). □

To be more specific, assume that $U_\theta=\rho+\sigma W_\theta$ for some Wiener process W and constants ρ,σ, in which case $\psi^U(u)=\rho u+\frac{\sigma^2}{2}u$. This means that, conditionally on an upward move Δv of the "volatility" process v, the log asset price X exhibits a Gaussian jump with mean $\rho\Delta v$ and variance $\sigma^2\Delta v$. For $\sigma=0$ we are back in the setup of (4.7). For L,Z one may e.g. choose any of the tried and tested processes in CGMY.

4.8 Further remarks

Ordinary versus stochastic exponential

In the literature, positive asset prices are modelled typically either as ordinary or as stochastic exponential, i.e.

$$S_t=S_0e^{X_t}=S_0\mathcal{E}(\widetilde{X})_t.$$

Above, we considered the first representation in terms of X or, more precisely, $X+\log(S_0)$. In [16], the process $\widetilde{X}$ is called the *exponential transform* of X. One can compute $\widetilde{X}$ from X and vice versa quite easily. It is well-known that X is a Lévy process if and only if $\widetilde{X}$ is a Lévy process. A similar statement holds for the affine SV models above, where the differential characteristics of (v,X) (resp. $(v^{(1)},\dots,v^{(\nu)},v,X)$ in the BNS case) do not depend on X_t. By applying Propositions 3 and 2 one observes in a straightforward manner that $(v,\widetilde{X})$ (resp. $(v^{(1)},\dots,v^{(\nu)},v,\widetilde{X})$) is affine as well. However, for purposes of estimation or option pricing it is often more convenient to work with X rather than $\widetilde{X}$.

Statistical versus risk-neutral modelling

Statistical estimators based on historical data yield parameters of the model under the physical probability measure P. By contrast, option pricing and calibration refers to expectations relative to some risk-neutral measure $\widetilde{P}$. For both purposes, affine models offer considerable computational advantages. Therefore one may wonder whether a given measure change preserves the affine structure. This can be checked quite easily with the help of Proposition 4. E.g., if X is a $\mathbb{R}^d$-valued affine process and if, in (2.4), $H_t(\omega)$ is constant and $W(\omega, t, x)$ depends only on x, then the affine structure carries over to $\widetilde{P}$. Only the triplets in (3.1) have to be adapted accordingly.

Martingale property of the asset price

Suppose that the process X in the examples above denotes the logarithm of a discounted asset price. If the model is set up under some "risk-neutral" probability measure, one would like e^X to be a martingale or at least a local martingale. The latter property can be directly read from the characteristics. If $\partial X = (b, c, F)$, then e^X is a local martingale if and only if $Ee^{X_0} < \infty$ and

$$b_t + \frac{c_t}{2} + \int (e^x - 1 - h(x))F_t(dx) = 0, \quad t \in \mathbb{R}_+, \tag{4.13}$$

(cf. [16], Theorems 2.19, 2.18). In the context of the affine SV processes (v, X) in the previous examples (i.e., in particular, with $\psi_2 = 0$), Expression (4.13) equals

$$\psi_0(0,1) + \psi_1(0,1)v_t.$$

Since v_t is random, both $\psi_0(0,1)$ and $\psi_1(0,1)$ typically have to be 0 in order for e^X to be a local martingale.

It is a more delicate to decide whether e^X is a true martingale. This holds automatically if X is a Lévy process (cf. [14], Lemma 4.4). In the affine case a sufficient condition can be derived from DFS.

Proposition 1. *Let (v, X) be an affine SV process as in the previous examples (and hence the conditions in Theorem 3.1 hold). Suppose that $(0,1) \in U$ and $(0,0) \in U$ for an open convex set $U \subset \mathbb{C}^2$ such that, for any $u \in U$,*

1. *$\psi_j(\mathrm{Re}(u)) < \infty, \quad j = 0, 1,$*
2. *there exists an U-valued solution $\Psi^{(1,2)}(\cdot, u)$ on $\mathbb{R}_+$ to the initial value problem (3.4).*

If $Ee^{X_0} < \infty$ and $\psi_0(0,1) = 0 = \psi_1(0,1)$, then e^X is a martingale.

Proof. From Lemmas 5.3, 6.5 and Theorem 2.16 in DFS it follows that (3.3) holds also for $\lambda = (0, -i)$, i.e.

$$E(e^{X_{s+t}}|\mathcal{F}_s) = \exp\left(\Psi^0(t,0,1) + \Psi^1(t,0,1)^\top (v_s, X_s)\right) = e^{X_s},$$

which yields the assertion. □

The previous result carries over to the BNS case $(v^{(1)}, \dots, v^{(\nu)}, v, X)$ or to more general affine situations. The point is to verify that exponential moments can actually be calculated from (3.3).

Observability of the volatility process

In the examples above we assumed implicitly that the affine process under consideration as, e.g., (v, X) is adapted to the given filtration. In practice, however, only the logarithmic asset price X but not the volatility process v can be observed directly. Therefore, the canonical filtration of X would be a natural choice. Fortunately, v is typically adapted to the latter if X is driven by an infinite activity process. The intuitive reason is that one can recover v in an almost sure fashion from X by observing the quadratic variation of the continuous martingale part or by counting the jumps in the purely discontinuous case (cf. e.g. [25], Theorem 1). This holds even in models with leverage as e.g. (4.7) if the Lévy measure of L has considerably more mass near the origin than the one of Z.

References

1. Barndorff-Nielsen O., Shephard N.: Non-Gaussian Ornstein-Uhlenbeck-based models and some of their uses in financial economics. *Journal of the Royal Statistical Society, Series B* **63**, 167–241 (2001)
2. Brockwell P.: Representations of continuous-time ARMA processes. *Journal of Applied Probability* **41A**, 375–382 (2004)
3. Carr P., Geman H., Madan D., Yor M.: Stochastic volatility for Lévy processes. *Mathematical Finance* **13**, 345–382 (2003)
4. Carr P., Madan D.: Option valuation using the fast Fourier transform. *The Journal of Computational Finance* **2**, 61–73 (1999)
5. Carr P., Wu L.: The finite moment log stable process and option pricing. *The Journal of Finance* **58**, 753–777 (2003)
6. Carr P., Wu L.: Time-changed Lévy processes and option pricing. *Journal of Financial Economics* **71**, 113–141 (2004)
7. Duffie D., Filipovic D., Schachermayer, W.: Affine processes and applications in finance. *The Annals of Applied Probability* **13** 984–1053 (2003)
8. Ethier S., Kurtz T.: *Markov Processes. Characterization and Convergence.* New York: Wiley 1986
9. Goll T., Kallsen J.: Optimal portfolios for logarithmic utility. *Stochastic Processes and their Applications* **89**, 31–48 (2000)
10. Heston S.: A closed-form solution for options with stochastic volatilities with applications to bond and currency options. *The Review of Financial Studies* **6**, 327–343 (1993)
11. Jacod J.: *Calcul Stochastique et Problèmes de Martingales*, volume 714 of *Lecture Notes in Mathematics.* Berlin: Springer 1979
12. Jacod J., Shiryaev A.: *Limit Theorems for Stochastic Processes.* Berlin: Springer, second edition 2003.
13. Jiang G., Knight J.: Estimation of continuous-time processes via the empirical characteristic function. *Journal of Business & Economic Statistics*, **20**, 198–212 (2002)
14. Kallsen J.: Optimal portfolios for exponential Lévy processes. *Mathematical Methods of Operations Research* **51**, 357–374 (2000)

15. Kallsen J.: σ-localization and σ-martingales. *Theory of Probability and Its Applications* **48**, 152–163 (2004)
16. Kallsen J., Shiryaev A.: The cumulant process and Esscher's change of measure. *Finance & Stochastics* **6**, 397–428 (2002)
17. Kallsen J., Shiryaev A.: Time change representation of stochastic integrals. *Theory of Probability and Its Applications* **46**, 522–528 (2002)
18. Klüppelberg C., Lindner A., Maller R.: A continuous-time GARCH process driven by a Lévy process: Stationarity and second order behaviour. *Journal of Applied Probability* **41**, 601–622 (2004)
19. Protter, P.: *Stochastic Integration and Differential Equations.* Berlin: Springer, second edition 2004
20. Raible, S.: *Lévy Processes in Finance: Theory, Numerics, and Empirical Facts.* Dissertation Universität Freiburg 2000
21. Sato, K.: *Lévy Processes and Infinitely Divisible Distributions.* Cambridge: Cambridge University Press 1999
22. Schöbel R., Zhu, J.: Stochastic volatility with an Ornstein-Uhlenbeck process: An extension. *European Finance Review*, **3** 23–46 (1999)
23. Shiryaev, A.: *Essentials of Stochastic Finance.* Singapore: World Scientific 1999
24. Stein E., Stein J.: Stock price distributions with stochastic volatility: An analytic approach. *The Review of Financial Studies* **4**, 727–752 (1991)
25. Winkel M.: The recovery problem for time-changed Lévy processes. Preprint 2001.
26. Yu J.: Empirical characteristic function estimation and its applications. *Econometric Reviews* **23** 93–123 (2004)