MOMENT EXPLOSIONS AND LONG-TERM BEHAVIOR OF AFFINE STOCHASTIC VOLATILITY MODELS

MARTIN KELLER-RESSEL

ABSTRACT. We consider a class of asset pricing models, where the risk-neutral joint process of log-price and its stochastic variance is an affine process in the sense of Duffie, Filipovic, and Schachermayer [2003]. First we obtain conditions for the price process to be conservative and a martingale. Then we present some results on the long-term behavior of the model, including an expression for the invariant distribution of the stochastic variance process. We study moment explosions of the price process, and provide explicit expressions for the time at which a moment of given order becomes infinite. We discuss applications of these results, in particular to the asymptotics of the implied volatility smile, and conclude with some calculations for the Heston model, a model of Bates and the Barndorff-Nielsen-Shephard model.

1. INTRODUCTION

Duffie, Pan, and Singleton [2000] introduced the notion of an affine jump-diffusion, which is a jump-diffusion process, whose drift vector, instantaneous covariance matrix and arrival rate of jumps all depend in an affine way on the state vector. Duffie, Pan, and Singleton remark that models built on affine processes provide a balanced tradeoff between analytical tractability and complexity, making them an attractive choice for applications in mathematical finance. In particular they mention applications to the pricing of options in stochastic volatility models and note that the models of Heston [1993], Bates [1996, 2000] , and Bakshi, Cao, and Chen [1997] fall into the affine class. To these, we could add the more recent models of Barndorff-Nielsen and Shephard [2001] and Carr and Wu [2004], which are also of affine type.

Duffie, Filipovic, and Schachermayer [2003] subsequently extended the class of affine jump-diffusions, defining an *affine process* as a time-homogenous Markov process, whose characteristic function is the exponential of an affine function of the state vector. It turns out that this class coincides for a large part with the class of affine jump-diffusions, but also allows for infinite activity of jumps and for killing or explosions of the process. Duffie, Filipovic, and Schachermayer aim to give a rigorous mathematical foundation to the theory of affine processes, covering many aspects, such as the characterization of an affine process in terms of the 'admissible parameters' (comparable to the *characteristic triplet* of a Lévy process) and properties of the ordinary differential equations ('generalized Riccati equations') that are implied by the process.

In this article we study stochastic volatility models, comprised of a log-price process $(X_t)_{t\geq 0}$ and a stochastic variance process $(V_t)_{t\geq 0}$, such that the joint process $(X_t, V_t)_{t\geq 0}$ is an *affine process.* We will show that many properties of such a model, including its long-term behavior and moment explosions, can be analyzed by studying differential equations of the generalized Riccati type. Our results on the long-term behavior are formulated as asymptotic results for the cumulant generating function of the stock price, as time goes to infinity. Asymptotics of this type have been used by Lewis [2000] to obtain large-time-to-maturity results for the implied volatility smile of stochastic volatility models via a saddlepoint expansion. The issue of moment explosions in stochastic volatility models has recently received much attention, due to the articles of Andersen and Piterbarg [2007] and Lions and Musiela [2007]. Moment explosions are intimately connected to large-strike asymptotics of the implied volatility smile via results of Lee [2004], that have later been expanded by Benaim and Friz [2006].

In the first part of the paper we introduce our main assumption, that the joint process $(X_t, V_t)_{t\geq 0}$ is affine, and recapitulate the main results of Duffie et al. [2003]. We derive necessary and sufficient conditions for conservativeness of the process and for the martingale property of the discounted price process $S_t = \exp(X_t)$. At the end of Section 2 we add two assumptions, and give a precise definition of the class of affine stochastic volatility models, which constitutes the main subject of this article. In Section 3 we derive our central results on long-term properties of an affine stochastic volatility model, providing conditions for the existence of an invariant distribution of the stochastic variance process, and characterizing this distribution in terms of its cumulant generating function. We also give results on the long-term properties of the price process, showing that as time tends to infinity, the marginal distributions of the price process approach those of an exponential-Lévy process. The characteristic exponent of this Lévy process can be derived directly from the specification of the affine stochastic volatility model. Both results are obtained by applying qualitative ODE theory to the generalized Riccati equations introduced in the first part.
In Section 4 we study moment explosions of the price process, and show that an explicit representation for the time of moment explosion can be given – not only for the primary model, but also for the model in the stationary variance regime. In Section 5 we outline applications of our results to the asymptotics of implied volatilities and of implied forward volatilities. We briefly discuss the results of Lee [2004] and point out the connection between the stationary variance regime and the pricing of forward-start options, when the time until the start of the contract is large. We conclude in Section 6 with explicit calculations for several models to which our results apply, such as the Heston model, a Heston model with added jumps, a model of Bates, and the Barndorff-Nielsen-Shephard model.

2. Affine Stochastic Volatility Models

2.1. **Definition and the generalized Riccati equations.** We consider an asset-pricing model of the following kind: The interest rate r is non-negative and constant, and the asset price $(S_t)_{t\geq 0}$ is given by

$$S_t = \exp(rt + X_t) \quad t \geq 0 \,,$$

such that $(X_t)_{t \geq 0}$ is the discounted log-price process starting at $X_0 \in \mathbb{R}$ a.s. The discounted price process is simply $\exp(X_t)$, such that we will assume in the remainder that $r = 0$, and that $(S_t)_{t \geq 0}$ is already discounted. Denote by $(V_t)_{t \geq 0}$ another process, starting at $V_0 > 0$ a.s., which can be interpreted as stochastic variance process of $(X_t)_{t \geq 0}$, but may also control the arrival rate of jumps. The following assumptions are made on the joint process $(X_t, V_t)_{t \geq 0}$:

A1: $(X_t, V_t)_{t \geq 0}$ is a stochastically continuous, time-homogeneous Markov process.

A2: The cumulant generating function $\Phi_t(u, w)$ of (X_t, V_t) is of a particular affine form: We assume that there exist functions $\phi(t, u, w)$ and $\psi(t, u, w)$ such that

$$\Phi_t(u, w) := \log \mathbb{E}\left[\exp(uX_t + wV_t)|\, X_0, V_0\right] = \phi(t, u, w) + V_0\psi(t, u, w) + X_0 u$$

for all $(t, u, w) \in \mathbb{R}_{\geqslant 0} \times \mathbb{C}^2$, where the expectation exists.

By convention, the logarithm above denotes the principal branch of the complex logarithm. Assumptions A1 and A2 make $(X_t, V_t)_{t \geq 0}$ an affine process in the sense of Duffie et al. [2003]. The term $X_0 u$ in the cumulant generating function $\Phi_t(u, w)$ corresponds to a reasonable homogeneity assumption on the model: If the starting value X_0 of the price process is shifted by x, also X_t is simply shifted by x for any $t \geq 0$. Note that Assumption A2 also implies that the variance process $(V_t)_{t \geq 0}$ is a Markov process in its own right. We do not yet make the assumption that $(S_t)_{t \geq 0}$ is conservative (i.e. without explosions or killing) or even a martingale. Instead it will be our first goal in Section 2.2 to obtain necessary and sufficient conditions for these properties.

Applying the law of iterated expectations to $\Phi_t(u, w)$ yields the following 'flow-equations' for ϕ and ψ: (see also Duffie et al. [2003, Eq. (3.8)–(3.9)])

$$(2.1) \qquad \begin{aligned} \phi(t + s, u, w) &= \phi(t, u, w) + \phi(s, u, \psi(t, u, w)), \\ \psi(t + s, u, w) &= \psi(s, u, \psi(t, u, w)), \end{aligned}$$

for all $t, s \geq 0$. The following result will be crucial:

Theorem 2.1. *Suppose that $|\phi(\tau, u, \eta)| < \infty$ and $|\psi(\tau, u, \eta)| < \infty$ for some $(\tau, u, \eta) \in \mathbb{R}_{\geqslant 0} \times \mathbb{C}^2$. Then, for all $t \in [0, \tau]$ and $w \in \mathbb{C}$ with $\mathrm{Re}\, w \leq \mathrm{Re}\, \eta$*

$$|\phi(t, u, w)| < \infty, \qquad |\phi(t, u, w)| < \infty,$$

and the derivatives

$$(2.2) \qquad F(u, w) := \left.\frac{\partial}{\partial t}\phi(t, u, w)\right|_{t=0+}, \qquad R(u, w) := \left.\frac{\partial}{\partial t}\psi(t, u, w)\right|_{t=0+}$$

exist. Moreover, for $t \in [0, \tau)$, ϕ and ψ satisfy the generalized Riccati equations

$$(2.3\text{a}) \qquad \partial_t\phi(t, u, w) = F(u, \psi(t, u, w)), \quad \phi(0, u, w) = 0$$

$$(2.3\text{b}) \qquad \partial_t\psi(t, u, w) = R(u, \psi(t, u, w)), \quad \psi(0, u, w) = w\,.$$

The above theorem is 'essentially' proven in Duffie et al. [2003], but under slightly different conditions[1]. Note that the differential equations (2.3) follow immediately from the flow equations (2.1) by taking the derivative with respect to s, and evaluating at $s = 0$. They are called *generalized Riccati Equations* since they degenerate

---

[1]Duffie et al. assume differentiability of ϕ and ψ with respect to t ('regularity') a priori, while in our case we can deduce it directly from Assumption A2. A proof is given in the appendix.

into (classical) Riccati equations with quadratic functions F and R, if $(X_t, V_t)_{t\geq 0}$ is a pure diffusion process.

Note that the first Riccati equation is just an integral in disguise, and ϕ may be written explicitly as

$$(2.4) \qquad \phi(t, u, w) = \int_0^t F(u, \psi(s, u, w)) \, ds \ .$$

Also the solution ψ of the second Riccati equation can be represented at least implicitly in the following way: Suppose that $\psi(t, u, w)$ is a non-stationary local solution on $[0, \delta)$ of (2.3b). Then $R(u, \psi(t, u, w)) \neq 0$ for all $t \in [0, \delta)$, and $\psi(t, u, w)$ is a strictly monotone function of t; dividing both sides of (2.3b) by $R(u, \psi(t, u, w))$, integrating from 0 to $t < \delta$, and substituting $\eta = \psi(s, u, w)$ yields

$$(2.5) \qquad \int_w^{\psi(t,u,w)} \frac{d\eta}{R(u,\eta)} ds = t \ .$$

Another important result that can be found in Duffie et al. [2003] states that F and R must be of Lévy-Khintchine form, i.e.

$$(2.6\text{a}) \qquad F(u, w) = (u, w) \cdot \frac{a}{2} \cdot \begin{pmatrix} u \\ w \end{pmatrix} + b \cdot \begin{pmatrix} u \\ w \end{pmatrix} - c$$

$$+ \int_{D\backslash\{0\}} \left(e^{xu+yw} - 1 - \omega_F(x, y) \cdot \begin{pmatrix} u \\ w \end{pmatrix} \right) m(dx, dy) \ ,$$

$$(2.6\text{b}) \qquad R(u, w) = (u, w) \cdot \frac{\alpha}{2} \cdot \begin{pmatrix} u \\ w \end{pmatrix} + \beta \cdot \begin{pmatrix} u \\ w \end{pmatrix} - \gamma$$

$$+ \int_{D\backslash\{0\}} \left(e^{xu+yw} - 1 - \omega_R(x, y) \cdot \begin{pmatrix} u \\ w \end{pmatrix} \right) \mu(dx, dy)$$

where $D = \mathbb{R} \times \mathbb{R}_{\geq 0}$, and ω_F, ω_R are suitable truncation functions, which we fix by defining

$$\omega_F(x, y) = \begin{pmatrix} \frac{x}{1+x^2} \\ 0 \end{pmatrix} \qquad \text{and} \qquad \omega_R(x, y) = \begin{pmatrix} \frac{x}{1+x^2} \\ \frac{y}{1+y^2} \end{pmatrix} \ .$$

Moreover the parameters $(a, \alpha, b, \beta, c, \gamma, m, \mu)$ satisfy the following admissibility conditions:

- a, α are positive semi-definite 2×2-matrices, and $a_{12} = a_{21} = a_{22} = 0$.
- $b \in D$ and $\beta \in \mathbb{R}^2$.
- $c, \gamma \in \mathbb{R}_{\geq 0}$
- m and μ are Lévy measures on D, and $\int_{D\backslash\{0\}} \left((x^2 + y) \wedge 1 \right) m(dx, dy) < \infty$.

The affine form of the cumulant generating function, the generalized Riccati equations and finally the Lévy-Khintchine decomposition (2.6) lead to the following interpretation of F and R: F characterizes the state-independent dynamic of the process (X_t, V_t) while R characterizes its state-dependent dynamic. Both F and R decompose into a diffusion part, a drift part, a jump part and an instantaneous killing rate. Hence $a + \alpha V_t$ can be regarded as instantaneous covariance matrix of $(X_t, V_t)_{t\geq 0}$, $b + V_t\beta$ as the instantaneous drift, $m(dx, dy) + V_t\mu(dx, dy)$ as instantaneous arrival rate of jumps with jump heights in $(dx \times dy)$, and finally $c + \gamma V_t$ as the instantaneous killing rate.

The following Lemma establishes some important properties of F and R as functions of *real-valued* arguments. A proof is given in the appendix.

Lemma 2.2. *(a) F and R are proper closed convex functions on $\mathbb{R}^2$.*
(b) F and R are analytic in the interior of their effective domain.
(c) Let U be a one-dimensional affine subspace of $\mathbb{R}^2$. Then $F|_U$ is either a strictly convex or an affine function. The same holds for $R|_U$.
(d) If $(u, w) \in \operatorname{dom} F$, then also $(u, \eta) \in \operatorname{dom} F$ for all $\eta \leq w$. The same holds for R.

Remark 2.3. As usual in convex analysis, we regard F and R as functions defined on all of $\mathbb{R}^2$, that may attain values in $\mathbb{R} \cup \{+\infty\}$. The set $\{(u, w) : F(u, w) < \infty\}$ is called effective domain of F, and denoted by $\operatorname{dom} F$.

We define a function $\chi(u)$, that will appear in several conditions throughout this article. Corollary 3.5 gives an interpretation of χ as a rate of convergence for the asymptotic behavior of the cumulant generating function of $(X_t)_{t \geq 0}$.

Definition 2.4. For each $u \in \mathbb{R}$ where $R(u, 0) < \infty$, define $\chi(u)$ as

$$\chi(u) := \left. \frac{\partial R}{\partial w}(u, w) \right|_{w=0} .$$

$\chi(u)$ is well-defined at least as a limit as $w \uparrow 0$, possibly taking the value $+\infty$; it can be written explicitly as

$$\chi(u) = \alpha_{12}u + \beta_1 + \int_{D \setminus \{0\}} y \left(e^{xu} - \frac{1}{1 + y^2} \right) \mu(dx, dy) .$$

Note that also $\chi(u)$ is a convex function.

2.2. **Explosions and the martingale property.** We are interested in conditions under which $S_t = \exp(X_t)$ is conservative and a martingale. If such conditions are satisfied, $(S_t)_{t \geq 0}$ may serve as the price process under the risk-neutral measure in an arbitrage-free asset pricing model. The following theorem gives sufficient and necessary conditions:

Theorem 2.5. *Suppose (X_t, V_t) satisfies Assumptions A1 and A2. Then the following holds:*
(a) $(S_t)_{t \geq 0}$ is conservative if and only if $F(0, 0) = R(0, 0) = 0$ and

$$(2.7) \qquad \int_{0-} \frac{d\eta}{R(0, \eta)} = -\infty ;$$

(b) $(S_t)_{t \geq 0}$ is a martingale if and only if it is conservative, $F(1, 0) = R(1, 0) = 0$ and

$$(2.8) \qquad \int_{0-} \frac{d\eta}{R(1, \eta)} = -\infty .$$

Remark 2.6. The notation $\int_{0-}$ denotes an integral over an arbitrarily small left neighborhood of 0.

By (2.6) the condition $F(0, 0) = R(0, 0) = 0$ is equivalent to $c = \gamma = 0$, i.e. obviously the killing rate has to be zero for the process to be conservative. As will be seen in the proof, the integral conditions (2.7) and (2.8) are related to a

uniqueness condition for non-Lipschitz ODEs, which has been discovered by Osgood [1898].

The following Corollary gives easy-to-check sufficient conditions:

Corollary 2.7. *Suppose (X_t, V_t) satisfies Assumptions A1 and A2.*

(a) If $F(0,0) = R(0,0) = 0$ and $\chi(0) < \infty$ then $(S_t)_{t\geq 0}$ is conservative.

(b) If $(S_t)_{t\geq 0}$ is conservative, $F(1,0) = R(1,0) = 0$ and $\chi(1) < \infty$, then $(S_t)_{t\geq 0}$ is a martingale.

Proof. For a proof of 2.5a we refer to [Filipović, 2001, Th. 4.11]. Statement 2.5b can be shown in a similar way:

Since (X_t, V_t) is Markovian, we have for all $0 \leq s \leq t$, that

$$\mathbb{E}[S_t | \mathcal{F}_s] = S_s \exp\left(\phi(t-s, 1, 0) + V_s \psi(t-s, 1, 0)\right) .$$

We have assumed that $V_0 > 0$ a.s., such that $(S_t)_{t\geq 0}$ is a martingale if and only if $(X_t)_{t\geq 0}$ is conservative and $\psi(t, 1, 0) = \phi(t, 1, 0) \equiv 0$ for all $t \in \mathbb{R}_{\geqslant 0}$.

We show Corollary 2.7 and the first implication of 2.5b: Suppose that $(S_t)_{t\geq 0}$ is conservative and that $F(1,0) = R(1,0) = 0$. By Theorem 2.1 $\psi(t, 1, w)$ solves the differential equation

$$(2.9) \qquad \frac{\partial}{\partial t}\psi(t, 1, w) = R(1, \psi(t, 1, w)), \qquad \psi(0, 1, w) = w$$

for all $w \leq 0$. Since $R(1,0) = 0$ it is clear that $\widetilde{\psi}(t, 1, 0) \equiv 0$ satisfies this ODE for the initial value $w = 0$. To deduce that $\widetilde{\psi}(t, 1, 0) = \psi(t, 1, 0)$ however, we need to know whether the solution is unique. Since $R(1, w)$ is continuously differentiable for $w < 0$, it satisfies a Lipschitz condition on $(-\infty, 0)$. If $\chi(1) < \infty$, the Lipschitz condition can be extended to $(-\infty, 0]$, and $\psi(t, 1, 0) \equiv 0$ is the unique solution. Without the assumption that $\chi(1) < \infty$, we substitute Lipschitz' condition by Osgood's condition[2] (2.8): Suppose that (2.8) holds, and there exists a non-zero solution $\widetilde{\psi}$ such that $\widetilde{\psi}(t_1, 1, 0) < 0$ for some $t_1 > 0$. Then for all $t < t_1$ such that ψ remains non-zero on $[t, t_1]$ we have (similarly to (2.5)) that

$$(2.10) \qquad \int_{\widetilde{\psi}(t_1, 1, 0)}^{\widetilde{\psi}(t, 1, 0)} \frac{d\eta}{R(1, \eta)} = t - t_1 .$$

Assume that $t_0 \geq 0$ is the first point left of t_1 such that $\widetilde{\psi}(t_0, 1, w) = 0$. Letting $t \downarrow t_0$, the left side of (2.10) tends to $-\infty$, whereas the right side remains bounded, leading to a contradiction. We conclude that $\psi(t, 1, 0) \equiv 0$ is the unique solution of (2.9). Finally equation (2.4) together with $F(1,0) = 0$ yields that also $\phi(t, 1, 0) \equiv 0$ for all $t \in \mathbb{R}_{\geqslant 0}$ and we have shown that $(S_t)_{t\geq 0}$ is a martingale.

For the other direction of 2.5b note that $(S_t)_{t\geq 0}$ being a martingale implies that $\phi = \psi \equiv 0$ solve the generalized Riccati equations and thus that $F(1,0) = R(1,0) = 0$. It remains to show (2.8). Assume that (2.8) does not hold. Then, for each $t > 0$, (2.10) with $t_1 = 0$ implicitly defines a solution $\widetilde{\psi}(t, 1, 0)$ of the generalized Riccati equation (2.9), satisfying $\widetilde{\psi}(t, 1, 0) < 0$ for all $t > 0$. By uniqueness of the solution $\psi(t, 1, w)$ for $w < 0$ and the flow property (2.1), we have $\widetilde{\psi}(t + s, 1, 0) = \psi(t, 1, \widetilde{\psi}(s, 1, 0))$ for t, s small enough. Letting $s \downarrow 0$ we obtain $\psi(t, 1, 0) = \widetilde{\psi}(t, 1, 0) < 0$, which is a contradiction to $\psi \equiv 0$. $\qquad\square$

---

[2]See Osgood [1898]

We add now two assumptions to A1 and A2 and complete our definition of an affine stochastic volatility model:

A3: The discounted price process $S_t = e^{X_t}$ is a martingale.
A4: $R(u, 0) \neq 0$ for some $u \in \mathbb{R}$.

Assumption A4 excludes models where the distribution of $(X_t)_{t \geq 0}$ does not depend at all on the volatility state V_0. In such a case we can not speak of a true stochastic volatility model, and it will be beneficial to avoid these degenerate cases. We are now ready to give our definition of an affine stochastic volatility model:

Definition 2.8. The process $(X_t, V_t)_{t \geq 0}$ is called an affine stochastic volatility model, if it satisfies assumptions A1 – A4.

A simple consequence of this definition, that will often be used is the following:

Lemma 2.9. *Let $(X_t, V_t)_{t \geq 0}$ be an affine stochastic volatility model. Then $R(u, 0)$ is a strictly convex function, satisfying $R(0, 0) = R(1, 0) = 0$.*

Proof. From assumption A3 and Theorem 2.5 it follows that $R(0, 0) = R(1, 0) = 0$. Lemma 2.2 implies that $R(u, 0)$ is either strictly convex or an affine function. Assume it is affine. Then $R(u, 0) = 0$ for all $u \in \mathbb{R}$. This contradicts A4, such that we conclude that $R(u, 0)$ is a strictly convex function. □

3. Long-term asymptotics

In this section we study the behavior of an affine stochastic volatility model as $t \to \infty$. We focus first on the stochastic variance process $(V_t)_{t \geq 0}$. Under mild assumptions this process will converge in law to its invariant distribution:

3.1. **Stationarity of the variance process.**

Proposition 3.1. *Suppose that A1 and A2 hold, that $\chi(0) < 0$ and the Lévy measure m satisfies the logarithmic moment condition*

$$\int_{y > 1} (\log y) \, m(dx, dy) < \infty.$$

Then $(V_t)_{t \geq 0}$ converges in law to its unique invariant distribution L, which has the cumulant generating function

$$(3.1) \qquad l(w) = \int_w^0 \frac{F(0, \eta)}{R(0, \eta)} \, d\eta \qquad (w \leq 0) .$$

Keller-Ressel and Steiner [2008] show that under the given conditions the process $(V_t)_{t \geq 0}$ converges in law to a limit distribution L, whose cumulant generating function can be represented by (3.1). A short argument at the end of this paragraph shows that the limit distribution is also the unique invariant distribution of $(V_t)_{t \geq 0}$. First we make the following definition: Given some affine stochastic volatility model $(X_t, V_t)_{t \geq 0}$, we introduce the process $(\widetilde{X}_t, \widetilde{V}_t)_{t \geq 0}$, defined as the Markov process with the same transition probabilities as $(X_t, \bar{V}_t)_{t \geq 0}$, but started with $X_0 = 0$ and V_0 distributed according to L. We will refer to $(\widetilde{X}_t, \widetilde{V}_t)_{t \geq 0}$ as the stochastic volatility model $(X_t, V_t)_{t \geq 0}$ 'in the stationary variance regime'. We also define the associated price process $\widetilde{S}_t := \exp(rt + \widetilde{X}_t)$. As we discuss in Section 5 the process $(\widetilde{X}_t, \widetilde{V}_t)_{t \geq 0}$ can be related to the pricing of forward-starting options,

when the time until the start of the contract is large.

The cumulant generating function of $(\widetilde{X}_t, \widetilde{V}_t)$ is given by

(3.2)
$$\log \mathbb{E}[e^{u\widetilde{X}_t + w\widetilde{V}_t}] = \log \mathbb{E}\left[\exp\left(\phi(t,u,w) + \widetilde{V}_0 \psi(t,u,w)\right)\right] = \phi(t,u,w) + l(\psi(t,u,w)).$$

We verify now that L is indeed an invariant distribution of $(V_t)_{t\geq 0}$:

$$(3.3) \quad \mathbb{E}\left[\exp\left(w\widetilde{V}_t\right)\right] = \exp\left(\phi(t,0,w) + l(\psi(t,0,w))\right) =$$

$$= \exp\left(\int_0^t F(0,\psi(s,0,w))ds + \int_{\psi(t,0,w)}^0 \frac{F(0,\eta)}{R(0,\eta)}\,d\eta\right) =$$

$$= \exp\left(\int_w^{\psi(t,0,w)} \frac{F(0,\eta)}{R(0,\eta)}\,d\eta + \int_{\psi(t,0,w)}^0 \frac{F(0,\eta)}{R(0,\eta)}\,d\eta\right) = \exp(l(w)),$$

where we have used that under the conditions of the Proposition above, $\psi(t,0,w)$ is a strictly monotone function converging to 0 as $t \to \infty$. (cf. Keller-Ressel and Steiner [2008]). To see that L is unique, assume that there exists another invariant distribution L', and let $(V_t')_{t\geq 0}$ be the variance process started with V_0' distributed according to L'. Again we use that $\phi(t,u,w) \to l(w)$ and $\psi(t,0,w) \to 0$ as $t \to \infty$ (see Keller-Ressel and Steiner [2008]), and get that

$$\lim_{t\to\infty} \mathbb{E}\left[\exp(wV_t')\right] = \mathbb{E}\left[\lim_{t\to\infty} \exp\left(\phi(t,0,w) + V_0'\psi(t,0,w)\right)\right] = \mathbb{E}[\exp(l(w))] = e^{l(w)},$$

for all $w \leq 0$, in contradiction to the invariance of L'.

3.2. **Long-term behavior of the log-price process.** We have seen that $(V_t)_{t\geq 0}$ converges to a limit distribution, but we do not expect the same for the log-price process $(X_t)_{t\geq 0}$. Nevertheless, it can be shown that the rescaled cumulant generating function $\frac{1}{t} \log \mathbb{E}\left[e^{X_t u}\right]$ converges under suitable conditions to a limit $h(u)$, that is again the cumulant generating function of some infinitely divisible random variable. This result can be interpreted such, that for large t the marginal distributions of $(X_t)_{t\geq 0}$ are 'close' to the marginal distributions of a Lévy process with characteristic exponent $h(u)$. Furthermore, $h(u)$ can be directly obtained from the functions F and R, without knowledge of the explicit forms of ϕ and ψ. We start with a preparatory Lemma:

Lemma 3.2. *Let $(X_t, V_t)_{t\geq 0}$ be an affine stochastic volatility model and suppose that $\chi(0) < 0$ and $\chi(1) < 0$. Then there exist a maximal interval I and a unique function $w \in C(I) \cap C^1(I^\circ)$, such that*

$$R(u, w(u)) = 0 \qquad \text{for all} \quad u \in I$$

and $w(0) = w(1) = 0$.
Moreover it holds that $[0,1] \subseteq I$, $w(u) < 0$ for all $u \in (0,1)$; $w(u) > 0$ for all $u \in I \setminus [0,1]$; and

$$(3.4) \qquad\qquad \frac{\partial R}{\partial w}(u, w(u)) < 0$$

for all $u \in I^\circ$.

We show Lemma 3.2 together with the next result, which makes the connection to the qualitative properties of the generalized Riccati equations.

Lemma 3.3. *(a) For each $u \in I^\circ$, $w(u)$ is an asymptotically stable equilibrium point of the generalized Riccati equation* (2.3b).
(b) For $u \in I^\circ$, there exists at most one other equilibrium point $\widetilde{w}(u) \neq w(u)$, and if it exists, it is necessarily unstable and satisfies $\widetilde{w}(u) > \max(0, w(u))$.
(c) For $u \in \mathbb{R} \setminus I$, no equilibrium point exists.

Proof. Define $L = \{(u, w) : R(u, w) \le 0\}$. As the level set of the closed convex function R, it is a closed and convex set. For all $u \in \mathbb{R}$, define $w(u) = \inf \{w : (u, w) \in L\}$, and $I = \{u \in \mathbb{R} : w(u) < \infty\}$. Clearly $w(u)$ is a continuous convex function, and I a subinterval of $\mathbb{R}$. We will now show that $w(u)$ and I satisfy all properties stated in Lemma 3.2. By assumption A3 and Theorem 2.5, $R(0, 0) = R(1, 0) = 0$; together with Lemma 2.2 it follows that the set $[0, 1] \times (-\infty, 0]$ is contained in $\operatorname{dom} R$. Since $R(u, 0)$ is by Lemma 2.9 strictly convex, and also $\chi(u)$ is convex, we deduce that $R(u, 0) < 0$ and $\frac{\partial R}{\partial w}(u, 0) = \chi(u) < 0$ for all $u \in (0, 1)$. In addition $R(u, w)$, as a function of w, is either affine or strictly convex, such that there exists a unique point $w(u)$, where $R(u, w(u)) = 0$, and necessarily $\frac{\partial R}{\partial w}(u, w(u)) < 0$. It is clear that for $u \in (0, 1)$ $w(u)$ coincides with the function defined above, and that $w(u) < 0$. At $u = 0$ we have that $R(0, 0) = 0$ and $\chi(0) < 0$, implying that $w(0) = 0$. A symmetrical argument at $u = 1$ shows that $w(1) = 0$, and thus that $[0, 1] \subseteq I$.
We show next that $w(u) \in C^1(I^\circ)$: Define $u_+ = \sup I$, and $w_+ = \lim_{u \uparrow u_+} w(u)$; u_-, w_- are defined symmetrically at the left boundary of I. Note that $u_\pm$ and $w_\pm$ can take infinite values. Define the open set

$$K := \{(\lambda u_- + (1 - \lambda)u_+, w) : \lambda \in (0, 1), w < \lambda w_- + (1 - \lambda)w_+\} \ .$$

Lemma 2.2 implies that K is contained in the interior of $\operatorname{dom} R$. On the other hand, the graph of w, restricted to I°, i.e. the set $\{(u, w(u)) : u \in I^\circ\}$, is clearly contained in K. Since R is by Lemma 2.2 an analytic function in the interior of its effective domain, the implicit function theorem implies that $w(u) \in C^1(I^\circ)$. In addition it follows that $\frac{\partial R}{\partial w}(u, w(u)) \neq 0$ for all $u \in I^\circ$, such that the assertion $\frac{\partial R}{\partial w}(u, w(u)) < 0$, which we have shown above for $u \in (0, 1)$, can be extended to all of I°. The claim that $w(u) > 0$ for $u \in I \setminus [0, 1]$ can easily be derived from the convexity of $w(u)$, and the fact that $w(u) < 0$ inside $(0, 1)$ and $w(0) = w(1) = 0$.

We have now proved most part of Lemma 3.2 (except for the uniqueness), and turn towards Lemma 3.3: Since $R(u, w(u)) = 0$ and $\frac{\partial R}{\partial w}(u, w(u)) < 0$ for all $u \in I^\circ$, $w(u)$ must be an asymptotically stable equilibrium point of the generalized Riccati equation 2.3b, showing 3.3a. Assume now that for some $u \in I^\circ$ there exists a point $\widetilde{w}(u) \neq w(u)$ such that $R(u, \widetilde{w}(u)) = 0$. By Lemma 2.2, $R(u, w)$ is, as a function of w, either strictly convex or affine. If it is affine, it has a unique root, and $\widetilde{w}(u)$ cannot exist. If it is strictly convex, there can exist a single point $\widetilde{w}(u)$ other than $w(u)$, such that $R(u, \widetilde{w}(u)) = 0$. Necessarily $\widetilde{w}(u) > w(u)$ and $\frac{\partial R}{\partial w}(u, \widetilde{w}(u)) > 0$. This shows that $\widetilde{w}(u)$ is an unstable equilibrium point of the generalized Riccati equation for ψ. In addition $\widetilde{w}(u) > w(u)$, and in particular the fact that $\widetilde{w}(0) > 0$ and $\widetilde{w}(1) > 0$ shows the uniqueness of $w(u)$ in the sense of Lemma 3.2. To see that $\widetilde{w}(u) > \max(0, w(u))$, note that we only have to show that $\widetilde{w}(u) > 0$, whenever $w(u) < 0$. This is the case only for $u \in (0, 1)$. Assume that $\widetilde{w}(u) \le 0$ for $u \in (0, 1)$. Then the convexity of R and $\frac{\partial R}{\partial w}(u, \widetilde{w}(u)) > 0$ would imply that $R(u, 0) \ge 0$ for some $u \in (0, 1)$. This is impossible by Lemma 2.9, and we have shown 3.3b. Finally

3.3c follows directly from the definition of $w(u)$ as $w(u) = \inf\{w : (u, w) \in L\}$ and I as the effective domain of $w(u)$. $\square$

We are now ready to show our main result on the long-term properties of the log-price process $(X_t)_{t \geq 0}$.

Theorem 3.4. *Let $(X_t, V_t)_{t \geq 0}$ be an affine stochastic volatility model and suppose that $\chi(0) < 0$ and $\chi(1) < 0$. Let $w(u)$ be given by Lemma 3.2 and define*

$$h(u) = F(u, w(u)), \qquad J = \{u \in I : F(u, w(u)) < \infty\} \ .$$

Then $[0, 1] \subseteq J \subseteq I$; $w(u)$ and $h(u)$ are cumulant generating functions of infinitely divisible random variables and

$$(3.5a) \qquad \qquad \lim_{t \to \infty} \psi(t, u, 0) = w(u) \quad \text{for all} \quad u \in I \ ;$$

$$(3.5b) \qquad \qquad \lim_{t \to \infty} \frac{1}{t} \phi(t, u, 0) = h(u) \quad \text{for all} \quad u \in J \ .$$

Corollary 3.5. *Under the conditions of Theorem 3.4, the following holds:*

$$(3.6a) \qquad \qquad \sup_{u \in [0,1]} |\psi(t, u, 0) - w(u)| \leq C \exp(-\mathfrak{X} \cdot T) \ ;$$

$$(3.6b) \qquad \qquad \sup_{u \in [0,1]} \left| \frac{1}{t} \phi(t, u, 0) - h(u) \right| \leq \Omega C \exp(-\mathfrak{X} \cdot T) \ ;$$

for some constant C, and with

$$\mathfrak{X} = \inf_{u \in [0,1]} |\chi(u)| \quad \text{and} \quad \Omega = \sup_{u \in [0,1]} \left. \frac{\partial}{\partial w} F(u, w) \right|_{w=0}$$

Proof. Let $u \in [0, 1]$. By Lemma 3.2 $(u, w(u)) \in [0, 1] \times (-\infty, 0]$. By Theorem 2.5 $F(0, 0) = F(1, 0) = 0$, such that Lemma 2.2 guarantees that $[0, 1] \times (-\infty, 0] \subseteq \text{dom } F$. It follows that $[0, 1] \subseteq J$. Define

$$z(t, u) = \psi(t, u, 0) - w(u) \ .$$

Inserting into the generalized Riccati equation 2.3b,

$$\frac{\partial}{\partial t} z(t, u) = R(u, \psi(t, u, 0)) = R(u, \psi(t, u, 0)) - R(u, w(u)), \quad \text{and} \quad z(0, u) = w(u) \ .$$

If $\psi(t, u, 0) \leq 0$ we can bound the right hand side by

$$R(u, \psi(t, u, 0)) - R(u, w(u)) \leq z(t, u) \frac{\partial R}{\partial w}(u, 0) = z(t, u)\chi(u),$$

using convexity of R. By Gronwall's inequality

$$z(t, u) \leq |w(u)| \exp(\chi(u)t) \ .$$

Since χ is convex, $\chi(0) < 0$ and $\chi(1) < 0$, we have shown (3.6a). The estimate

$$|\phi(t, u) - h(u)| =$$

$$= \left| \frac{1}{t} \int_0^t (F(u, \phi(s, u)) - F(u, w(u))) \ ds \right| \leq \left| \frac{\partial F}{\partial w}(u, 0) \right| \cdot |\psi(t, u) - w(u)|$$

yields (3.6b) and we have shown Corollary 3.5.

Let now $u \in I^\circ \setminus [0, 1]$. Combining Lemma 2.9 and Lemma 3.3 we have that $R(u, w) > 0$ for all $w \in [0, w(u))$, and $R(u, w(u)) = 0$. It follows that the initial value $\psi(0, u, 0) = 0$ is in the basin of attraction of the stable equilibrium point $w(u)$

and thus that $\psi(t, u, 0)$ is strictly increasing and converging to $w(u)$. An additional argument may be needed at the boundary of I: Let $u_+ = \sup I$ and assume that $u_+ \in I$ (i.e. I is right-closed). Since $(u_+, w) \in \operatorname{dom} R$ for all $w \leq w(u_+)$, we can define $\frac{\partial R}{\partial w}(u_+, w(u_+))$ at least as a limit for $w \uparrow w(u_+)$. By Lemma 3.2 either $\frac{\partial R}{\partial w}(u_+, w(u_+)) < 0$ or $\frac{\partial R}{\partial w}(u_+, w(u_+)) = 0$. In the first case we can argue as in the interior of I that $w(u_+)$ is an asymptotically stable equilibrium point. In the second case we use once more that by Lemma 2.2 $R(u_+, w)$ is, as a function of w, either strictly convex or affine. If it is affine, it must be equal to 0, and thus $R(u_+, 0) = 0$, in contradiction to Lemma 2.9. Hence it is strictly convex, and attains its minimum at $w(u_+)$. This implies that $R(u_+, w) > 0$ for all $w \in [0, w(u_+))$ and we conclude that $\psi(t, u_+, 0)$ converges to $w(u_+)$.[3] For $u_- = \inf I$, a symmetrical argument applies.

Assertion (3.5b) follows immediately from the representation (2.4), and

$$\lim_{t \to \infty} \frac{1}{t} \phi(t, u, 0) = \lim_{t \to \infty} \frac{1}{t} \int_0^t F(u, \psi(s, u, 0)) \, ds = F(u, w(u))$$

for all $u \in J$.

We have shown that the sequence of infinitely divisible cumulant generating functions $\psi(t, u, 0)$ converges on I to a function $w(u)$ that is continuous in a right neighborhood of 0. This is sufficient to imply that $w(u)$ is again the cumulant generating function of an infinitely divisible random variable (See Feller [1971, VIII.1, Example (e)] for the convergence part, and Sato [1999, Lemma 7.8] for the infinite divisibility.). The same argument can be applied to ϕ and $h(u)$, and we have shown Theorem 3.4. □

4. Moment explosions

In this section we continue to study the time evolution of moments $\mathbb{E}[S_t^u] = \mathbb{E}[e^{X_t u}]$ of the price process in an affine stochastic volatility model. We are interested in the phenomenon that in a stochastic volatility model, moments of the price process can explode (become infinite) in finite time. For stochastic volatility models of the CEV-type – a class including the Heston model, but no models with jumps – moment explosions have been studied by Andersen and Piterbarg [2007] and Lions and Musiela [2007]. In the context of option pricing, an interesting result of Lee [2004] connects the existence of moments of the stock price process to the steepness of the smile for deep in-the-money or out-of-the-money options. Our first result shows that in an affine stochastic volatility model a simple explicit expression for the time of moment explosion can be given:

4.1. **Moment explosions.** By definition, the u-th moment of S_t, i.e. $\mathbb{E}[S_t^u]$ is given by $S_0^u \exp\left(\phi(t, u, 0) + V_0 \psi(t, u, 0)\right)$. We define the **time of moment explosion** for the moment of order u by

$$T_*(u) = \sup \{t : \mathbb{E}[S_t^u] < \infty\} \ .$$

It is obvious from the Markov property that $\mathbb{E}[S_t^u]$ is finite for all $t < T_*(u)$ and infinite for all $t > T_*(u)$. As in the previous section, the main result follows from a qualitative analysis of the generalized Riccati equations (2.3).

---

[3]Even though $\psi(t, u_+, 0)$ converges to $w(u_+)$, note that $w(u_+)$ is not a stable equilibrium point in the usual sense. This is due to the fact that solutions from a *right*-neighborhood $N \cap (w(u_+), \infty)$ will *diverge* from $w(u_+)$ to $+\infty$.

Theorem 4.1. *Suppose the conditions of Theorem 3.4 hold. Define* $J = \{u \in I : F(u, w(u)) < \infty\}$,

$$f_+(u) := \sup\{w \geq 0 : F(u, w) < \infty\} \; ,$$
$$r_+(u) := \sup\{w \geq 0 : R(u, w) < \infty\} \; ,$$

and suppose that $F(u, 0) < \infty$, $R(u, 0) < \infty$ *and* $\chi(u) < \infty$.
(a) If $u \in J$, *then*

$$T_*(u) = +\infty \; .$$

(b) If $u \in \mathbb{R} \setminus J$, *then*

$$T_*(u) = \int_0^{\min(f_+(u), r_+(u))} \frac{d\eta}{R(u, \eta)} \; .$$

If $F(u, 0) = \infty$, $R(u, 0) = \infty$ *or* $\chi(u) = \infty$ *then*
(c)

$$T_*(u) = 0 \; .$$

Proof. Suppose that $u \in J$. Then Theorem 3.4 implies that both $\psi(t, u, 0)$ and $\phi(t, u, 0)$ are finite for all $t \geq 0$. This proves (a). Let now $u \in \mathbb{R} \setminus J$, $F(u, 0) < \infty$, $R(u, 0) < \infty$ and $\chi(u) < \infty$. To prove (b) we start by analyzing the maximal lifetime of solutions to the generalized Riccati equation

$$(4.1) \qquad \frac{\partial}{\partial t}\psi(t, u, 0) = R(u, \psi(t, u, 0)), \quad \psi(0, u, 0) = 0 \; .$$

Define $M = [0, r_+(u))$ and note that $R(u, .) \in C(M)$. Since $u \notin [0, 1]$, Lemma 2.9 implies that $R(u, 0) > 0$. It is clear, that at least a local solution $\psi(t, u, 0)$ to the ODE exists, which satisfies $0 \leq \psi(t, u, 0) \leq r_+(u)$ and is an increasing function of t as long as it can be continued. Using a standard extension theorem (e.g. Hartman [1982, Lem. I.3.1]) the local solution $\psi(t, u, 0)$ has a maximal extension to an interval $[0, T(u))$, such that one of the following holds:

(i) $T(u) = \infty$, or
(ii) $T(u) < \infty$ and $\psi(t, u, 0)$ comes arbitrarily close to the boundary of M, i.e.

$$\limsup_{t \to T(u)} \psi(t, u, 0) = r_+(u) \; .$$

Consider case (i). Since ψ is increasing, its limit for $t \to \infty$ exists, but can be infinite. Suppose $\lim_{t \to \infty} \psi(t) = \alpha < \infty$. Then α must be a stationary point, i.e. $R(u, \alpha) = 0$, but this is impossible by Lemma 3.3. The case that $\alpha = \infty$ is only possible if $r_+(u) = \infty$, such that in this case $\lim_{t \to T(u)} \psi(t, u, 0) = r_+(u)$. Consider case (ii). Since ψ is increasing the limes superior can be replaced by an ordinary limit and we get $\lim_{t \to T(u)} \psi(t, u, 0) = r_+(u)$ as before.

Let now T_n be a sequence such that $T_n \uparrow T(u)$. By (2.5) it holds that

$$(4.2) \qquad \int_0^{\psi(T_n, u, 0)} \frac{d\eta}{R(u, \eta)} ds = T_n \; .$$

Letting $n \to \infty$ we obtain that $T(u) = \int_0^{r_+(u)} \frac{d\eta}{R(u, \eta)} ds$.

We can write the time of moment explosion $T_*(u)$ as the maximum joint lifetime of $\phi(t, u, 0)$ and $\psi(t, u, 0)$, i.e. $T_*(u) = \sup\{t \geq 0 : \phi(t, u, 0) < \infty \wedge \psi(t, u, 0) < \infty\}$. By the integral representation (2.4) it is clear that if $f_+(u) \geq r_+(u)$, $\phi(t, u, 0)$ is finite whenever $\psi(t, u, 0)$ is finite and $T_*(u) = T(u)$. If $f_+(u) < r_+(u)$ then $\psi(T_*(u), u, 0) = f_+(u)$. Inserting into the representation (4.2) yields (b).

For assertion (c), let $F(u, 0) = \infty$, $R(u, 0) = \infty$, or $\chi(u) = \infty$. In the first case, $\phi(t, u, 0)$ does not exist beyond $t = 0$. In the other cases no local solution to the generalized Riccati equation (4.1) exists, such that $\psi(t, u, 0)$ explodes immediately. □

4.2. Moment explosions in the stationary variance regime.

In Section 3.1 we have introduced $(\widetilde{X}_t, \widetilde{V}_t)_{t \geq 0}$ as the model in the stationary variance regime. The moment explosions of this process can be analyzed in a similar manner as above. We define the time of moment explosion in the stationary variance regime by

$$T_*^S(u) := \sup\left\{T \geq 0 : \mathbb{E}[\widetilde{S}_T^u] < \infty\right\} ;$$

the superscript 'S' stands for 'stationary'.

The analogue to Theorem 4.1 is the following result:

Theorem 4.2. *Suppose the conditions of Theorem 3.4 hold. Define $f_+(u), r_+(u)$ as in Theorem 4.1, and in addition*

$$l_+ := \sup\{w > 0 : l(w) < \infty\} .$$

Suppose that $F(u, 0) < \infty$, $R(u, 0) < \infty$ and $\chi(0) < \infty$.

(a) If $u \in J$ and $w(u) \leq l_+$, then

$$T_*^S(u) = +\infty .$$

(b) If $u \in \mathbb{R} \setminus J$ or $w(u) > l_+$, then

$$T_*^S(u) = \int_0^{\min(f_+(u), r_+(u), l_+)} \frac{d\eta}{R(u, \eta)} .$$

If $F(u, 0) = \infty$, $R(u, 0) = \infty$ or $\chi(0) = \infty$, then

(c)

$$T_*^S(u) = 0 .$$

Corollary 4.3. *Under the conditions of Theorem 4.2,*

$$T_*^S(u) \leq T_*(u), \qquad \text{for all } u \in \mathbb{R}$$

Proof. By equation (3.2), the moment $\mathbb{E}[\widetilde{S}_t^u]$ is given by

$$\mathbb{E}[\widetilde{S}_t^u] = \exp\left(\phi(t, u, 0) + l(\psi(t, u, 0))\right) .$$

This expression is finite, if $\phi(t, u, 0)$ and $\psi(t, u, 0)$ are finite, *and* if $\psi(t, u, 0) < l_+$. It is infinite if $\phi(t, u, 0)$ or $\psi(t, u, 0)$ are infinite, or if $\psi(t, u, 0) > l_+$. The rest of the proof can be carried out as for Theorem 4.1. Note, that now even for $u \in J$, the moment can explode, if l_+ is reached by $\psi(t, u, 0)$ before the stationary point $w(u)$. Corollary 4.3 follows easily by comparing the range of integration and the conditions for case (a) and (b) between Theorem 4.1 and Theorem 4.2. □

5. Applications

5.1. **Smile behavior at extreme strikes.** In the preceding section, we have kept u fixed, and looked at the first time $T_*(u)$ that the moment $\mathbb{E}[S_t^u]$ becomes infinite. It will now be more convenient to reverse the roles of T and u, and for a given time t to define the **upper critical moment** by

$$u_+(t) = \sup\{u \geq 1 : \mathbb{E}[S_t^u] < \infty\} = \sup\{u \geq 1 : T_*(u) < t\} \ ,$$

and the **lower critical moment** by

$$u_-(t) = \inf\{u \leq 0 : \mathbb{E}[S_t^u] < \infty\} = \inf\{u \leq 0 : T_*(u) < t\} \ .$$

It is seen that $u_-(T)$ and $u_+(T)$ can be defined as the generalized inverse of $T_*(u)$ on $(-\infty, 0]$ and $[1, \infty)$ respectively. In addition it is easily derived from Jensen's inequality, that

$$\mathbb{E}[S_t^u] < \infty \qquad \text{for all } u \in (u_-(t), u_+(t)), \quad \text{and}$$
$$\mathbb{E}[S_t^u] = \infty \qquad \text{for all } u \in \mathbb{R} \setminus [u_-(t), u_+(t)] \ .$$

The results of Lee [2004] relate the explosion of moments to the 'wing behavior' of the implied volatility smile, i.e. the shape of the smile for strikes that are deep in-the-money or out-of-the-money. To give a precise statement, let ξ be the log-moneyness, which for a European option with time-to-maturity T and strike K is given by $\xi = \log\left(\frac{K}{e^{rT}S_0}\right)$.

Proposition 5.1 (Lee's moment formula)**.** *Let $V(T, \xi)$ be the implied Black-Scholes-Variance of a European call with time-to-maturity T and log-moneyness ξ. Then*

$$\limsup_{\xi \to -\infty} \frac{V(T, \xi)}{|\xi|} = \frac{\varsigma(-u_-(T))}{T}$$

and

$$\limsup_{\xi \to \infty} \frac{V(T, \xi)}{|\xi|} = \frac{\varsigma(u_+(T) - 1)}{T}$$

where $\varsigma(x) = 2 - 4\left(\sqrt{x^2 + x} - x\right)$ and $u_\pm(T)$ are the critical moment functions.

The function ς is strictly decreasing on $\mathbb{R}_{\geqslant 0}$, mapping 0 to 2, and ∞ to 0. Thus for fixed time-to-maturity T, the steepness of the smile is decreasing with $|u_\pm(T)|$. A finite critical moment $u_\pm(T)$ implies asymptotically linear behavior of $V(T, \xi)$ in ξ, and an infinite critical moment implies sublinear behavior of $V(T, \xi)$. It is also evident that $u_-(T)$ determines the 'left' side of the volatility smile, also known as small-strike, in-the-money-call or out-of-the-money-put side; $u_+(T)$ determines the 'right' side, or large-strike, out-of-the-money-call, in-the-money-put side. Finally we mention that Lee's result has been extended and strengthened by Benaim and Friz [2006] from a 'lim sup' to a genuine limit under conditions related to regular variation of the underlying distribution function.

5.2. **Forward-smile behavior.** The forward smile is derived from the prices of forward-start options. For a forward-start call option – all options we consider are European – a start date τ, a strike date $T + \tau$ and a moneyness ratio M are agreed upon today (at time $t = 0$). The option then yields at time $T + \tau$ a payoff of $\left(\frac{S_{T+\tau}}{S_\tau} - M\right)_+$, i.e. the relative return over the time period from τ to $\tau + T$,

reduced by M and floored at 0. Under the pricing measure the value of such an option at $t = 0$ is given by

$$(5.1) \qquad e^{-r(T+\tau)}\mathbb{E}\left[\left(\frac{S_{T+\tau}}{S_\tau} - M\right)_+\right] = e^{-\tau r}\mathbb{E}\left[\left(e^{X_{T+\tau}-X_\tau} - e^\xi\right)_+\right] ,$$

where we define the log-moneyness ξ of a forward-start option as $\xi = \log M + rT$. Forward-start options are not just interesting in their own right, but are used as building blocks of more complex derivatives, such as Cliquet options (see Gatheral [2006, Chapter 10]).

Analogously to plain vanilla options, we can define the **implied forward volatility** $\sigma(\tau, T, \xi)$, by comparing the forward option price to the price of an option with identical payoff in the Black-Scholes model. Note that the implied forward volatility depends also on τ, the starting time of the contract. For $\tau = 0$, the implied volatility of a plain vanilla option is retrieved. More interesting is the behavior for $\tau > 0$. Intuitively, we expect the implied volatility (and the option price) to increase with τ in a stochastic volatility model, since the uncertainty of the variance V_τ at the starting date of the option has to be priced in. In an affine stochastic volatility model, it will be seen that under mild conditions, the implied forward volatilities $\sigma(\tau, T, \xi)$ actually converge to a limit as $\tau \to \infty$. Not surprisingly, this behavior is related to the convergence of $(V_t)_{t\geq 0}$ to its invariant distribution. In the limit $\tau \to \infty$, the pricing of a forward-start option is equivalent to the pricing of a plain vanilla option *in the stationary variance regime* (cf. Section 3.1).

Proposition 5.2. *Let $(X_t, V_t)_{t\geq 0}$ be an affine stochastic volatility model, satisfying the conditions of Proposition 3.1. Let $\sigma(\tau, T, \xi)$ be the implied forward volatility in this model. Then*

$$\lim_{\tau\to\infty} \sigma(\tau, T, \xi) = \widetilde{\sigma}(T, \xi) ,$$

where $\widetilde{\sigma}(T, \xi)$ is the implied volatility of a European call with payoff $\left(e^{\widetilde{X}_T} - e^\xi\right)_+$, and $\widetilde{X}_T$ is the log-price process of the model in the stationary variance regime.

Proof. We can write the price of a forward-start call as

$$C(\tau, T, \xi) = e^{-\tau r}\mathbb{E}\left[\left(e^{X_{T+\tau}-X_\tau} - e^\xi\right)_+\right] = e^{-r\tau}\mathbb{E}\left[\mathbb{E}^{(0, V_\tau)}\left[\left(e^{X_T} - e^\xi\right)_+\right]\right] .$$

Denote by $C^{\mathrm{BS}}(T, \xi, \sigma)$ the (plain vanilla) call price in a Black-Scholes model with volatility σ and the normalization $S_0 = 1$. It is easy to see that the price of a forward-start option in the Black-Scholes model is just the discounted plain vanilla price, i.e. $C^{\mathrm{BS}}(\tau, T, \xi, \sigma) = e^{-r\tau}C^{\mathrm{BS}}(T, \xi, \sigma)$. By definition, the implied forward volatility of the call $C(\tau, T, \xi)$ satisfies

$$C^{\mathrm{BS}}(T, \xi, \sigma(\tau, T, \xi)) = e^{r\tau}C(\tau, T, \xi) = \mathbb{E}\left[\mathbb{E}^{(0, V_\tau)}\left[\left(e^{X_T} - e^\xi\right)_+\right]\right] .$$

Taking the limit $\tau \to \infty$ on both sides we obtain

$$C^{\mathrm{BS}}(T, \xi, \lim_{\tau\to\infty} \sigma(\tau, T, \xi)) = \mathbb{E}\left[\lim_{\tau\to\infty}\mathbb{E}^{(0, V_\tau)}\left[\left(e^{X_T} - e^\xi\right)_+\right]\right] = \mathbb{E}\left[\left(e^{\widetilde{X}_T} - e^\xi\right)_+\right] ,$$

using dominated convergence. It is well known that the above equation allows a unique solution in terms of the Black-Scholes implied volatility, and we get $\widetilde{\sigma}(T, \xi) = \lim_{\tau\to\infty} \sigma(\tau, T, \xi)$. $\square$

Combining Lee's moment formula with our results on moment explosions under the stationary variance regime (Theorem 4.2), asymptotics of $\widetilde{\sigma}(T,\xi)$ for $\xi \to \pm\infty$ can be derived.

6. Examples

6.1. **The Heston model with and without jumps.** In the model of Heston [1993], the log-price $(X_t)_{t\geq 0}$ and the corresponding variance process $(V_t)_{t\geq 0}$ are given under the risk-neutral measure by the SDE

$$dX_t = -\frac{V_t}{2}\,dt + \sqrt{V_t}\,dW_t^1$$

$$dV_t = -\lambda(V_t - \theta)\,dt + \zeta\sqrt{V_t}\,dW_t^2$$

where W_t^1, W_t^2 are Brownian motions with correlation parameter ρ, and $\zeta, \lambda, \theta > 0$. In affine form, the model is written as

(6.1a) $$F(u,w) = \lambda\theta w$$

(6.1b) $$R(u,w) = \frac{1}{2}(u^2 - u) + \frac{\zeta^2}{2}w^2 - \lambda w + uw\rho\zeta\;.$$

It is easily calculated that χ is given by $\chi(u) = \rho\zeta u - \lambda$. We will first analyze the long term behavior of $(X_t)_{t\geq 0}$, with the help of Theorem 3.4. To satisfy the condition $\chi(1) < 0$ we need $\lambda > \zeta\rho$. Note that this condition is always satisfied if $\rho \leq 0$, the case that is typical for applications. Solving a quadratic equation we find that

$$w(u) = \frac{(\lambda - u\rho\zeta) - \sqrt{(\lambda - u\rho\zeta)^2 - \zeta^2(u^2 - u)}}{\zeta^2}\;, \quad \text{and} \quad h(u) = \lambda\theta w(u)\;.$$

Denoting the term under the square root by $\Delta(u)$, we see that $w(u)$ and $h(u)$ are both defined on $J = I = \{u : \Delta(u) \geq 0\}$. Since R is a second order polynomial in the Heston model, the equilibrium points of the generalized Riccati equation for ψ form an ellipse in the (u,w)-plane, and $w(u)$ is given by its lower part – see Figure 1 for an illustration. Interestingly, $w(u)$, and also $h(u)$, are cumulant generating functions of a Normal Inverse Gaussian distribution (cf. Barndorff-Nielsen [1997, Eq. (2.4)]). Thus, for large t, the price process of the Heston model is, in terms of its marginal distributions, close to a Normal-Inverse-Gaussian exponential-Lévy model.

Next we consider moment explosions in the Heston model. As mentioned above, moment explosions in the Heston model (and other models) have already been studied by Andersen and Piterbarg [2007]. Nevertheless this will provide a first test of Theorem 4.1: In the case of the Heston model it is easily determined from (6.1) that $f_+(u) = r_+(u) = \infty$. Calculating the integral in case (b) of Theorem 4.1, we obtain

(6.2) $$T_*(u) = \begin{cases} +\infty & \Delta(u) \geq 0 \\ \frac{2}{\sqrt{-\Delta(u)}}\left(\arctan\frac{\sqrt{-\Delta(u)}}{\chi(u)} + \pi\mathbf{1}_{\{\chi(u)<0\}}\right) & \Delta(u) < 0\;. \end{cases}$$

In Figure 2 a plot of this function for typical parameter values is shown. Note that Andersen and Piterbarg [2007] distinguish an additional case where $\chi(u) < 0$, but $\Delta(u) > 0$. A little calculation shows that this can only happen if $\chi(1) \geq 0$, a case
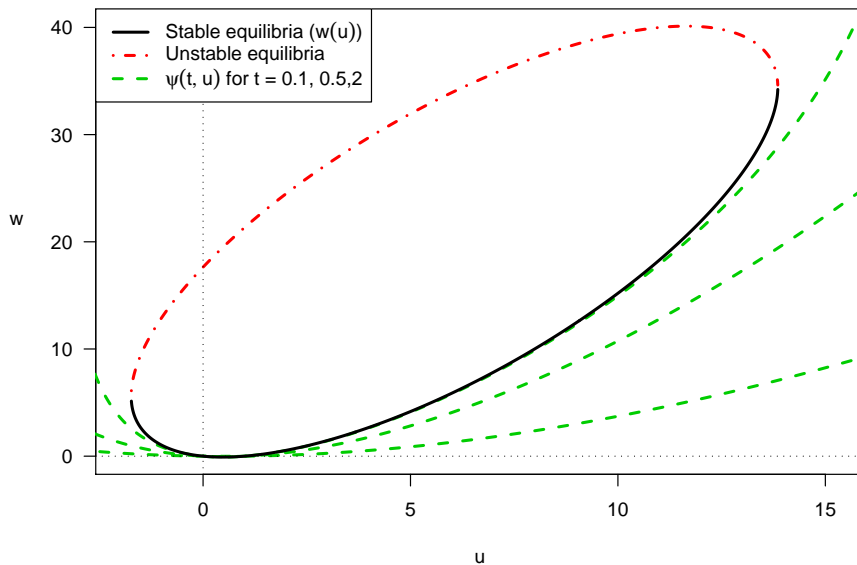
FIGURE 1. This plot shows the stable and unstable equilibria of the generalized Riccati equation of a Heston model with parameters $\rho = -0.7165$, $\zeta = 0.3877$, $\lambda = 1.3253$ and $\theta = 0.0354$ (taken from Gatheral [2006, Table 3.2]). It can be seen how the solutions $\psi(t, u)$ converge to the stable equilibrium points, which form the lower boundary of an ellipse in the (u, w)-plane.

that is precluded by our assumptions in Theorem 3.4, and never occurs when $\rho \leq 0$.

We will now study the effect of adding jumps to the Heston model. The simplest case is the addition of an independent jump component with constant activity: Let $(J_t)_{t \geq 0}$ be a pure-jump Lévy process, independent of $(W_t^{1,2})_{t \geq 0}$ and define the Heston-with-jumps model by

$$dX_t = \left(\delta - \frac{V_t}{2} \right) dt + \sqrt{V_t}\, dW_t^1 + dJ_t$$

$$dV_t = -\lambda(V_t - \theta)\, dt + \zeta \sqrt{V_t}\, dW_t^2 \ .$$

The drift δ is determined by the martingale condition for $(S_t)_{t \geq 0}$. To make the example simpler, we assume that $(J_t)_{t \geq 0}$ jumps only downwards. This is equivalent to saying that the Lévy measure $m(dx)$ of $(J_t)_{t \geq 0}$ is supported on $(-\infty, 0)$. The affine form of the model is

(6.3) $$F(u, w) = \lambda \theta w + \widetilde{\kappa}(u)$$

(6.4) $$R(u, w) = \frac{1}{2}(u^2 - u) + \frac{\zeta^2}{2} w^2 - \lambda w + uw\rho\zeta \ ,$$

where $\widetilde{\kappa}(u)$ is the compensated cumulant generating function of the jump part, i.e.

$$\widetilde{\kappa}(u) = \int_{(-\infty,0)} (e^{xu} - 1)\, m(dx) - u \int_{(-\infty,0)} (e^x - 1)\, m(dx)\ .$$

Let $\kappa_- < 0$ be the number such that $\widetilde{\kappa}(u)$ is finite on (κ_-, ∞) and infinite outside. For example, if the absolute jump heights are exponentially distributed with an expected jump size of $1/\alpha$, then $\kappa_- = -\alpha$.

To analyze the explosion times of this model, note that R, and thus $\chi(u)$, $w(u)$, I and $r_+(u)$ have not changed compared to the Heston model. As long as $u > \kappa_-$, the explosion time $T_*(u)$ is the same as in the Heston model. However, if $u \leq \kappa_-$, $F(u,0) = \infty$ and by Theorem 4.1, $T_*(u) = 0$. Thus, the addition of jumps to the Heston model has the effect of truncating the explosion time to zero, whenever $u \leq \kappa_-$.

From the viewpoint of the critical moment functions, $u_+(t)$ does not change compared to the Heston model, but $u_-(t)$ does; in the model with jumps it is given by

$$u_-^{\mathrm{Jump}}(t) = u_-^{\mathrm{Heston}}(t) \vee \kappa_-\ .$$

Since u_- is increasing with t, it makes sense to define a cutoff time $T_\sharp$ by

$$T_\sharp = \sup \left\{ t \geq 0 : u_-^{\mathrm{Heston}}(t) = \kappa_- \right\} = T_*(\kappa_-)\ ,$$

such that

$$u_-^{\mathrm{Heston}}(t) < u_-^{\mathrm{Jump}}(t), \quad \text{if} \quad t < T_\sharp$$
$$u_-^{\mathrm{Heston}}(t) = u_-^{\mathrm{Jump}}(t), \quad \text{if} \quad t \geq T_\sharp\ .$$

In Figure 2 a comparison of the critical moment functions in the Heston model with and without jumps can be seen. By Lee's moment formula, the critical moment $u_-(t)$ moving closer to 0 will cause the left side of the implied volatility smile to become steeper. Thus the net effect of adding the jump component $(J_t)_{t \geq 0}$ to the Heston model, is a steepening of the left side of the smile for maturities smaller than $T_\sharp$. For times larger than $T_\sharp$, the asymptotic behavior of the smile (in the sense of Lee's formula) is *exactly* the same as in the Heston model without jumps. This corresponds well to the frequently made observation (see e.g. Gatheral [2006, Chapter 5]) that a Heston model with jumps can be fitted well by first fitting a (jump-free) Heston model to long maturities, and then calibrating only the additional parameters to the full smile. In fact Gatheral proposes (on heuristical grounds) the concept of a 'critical time' T, after which the influence of an independent jump component on the implied volatility smile can be neglected. The analysis of the Heston model with jumps is of course easily extended to the case that $(J_t)_{t \geq 0}$ is not one-sided. In that case the effects discussed above will be seen to affect also the right side of the implied volatility smile.

6.2. **A model of Bates.** We consider now the model given by

$$dX_t = \left(\delta - \frac{V_t}{2} \right) dt + \sqrt{V_t}\, dW_t^1 + \int_D x\, \widetilde{N}(V_t, dt, dx)$$
$$dV_t = -\lambda(V_t - \theta)\, dt + \zeta \sqrt{V_t}\, dW_t^2\ .$$

where as before $\lambda, \theta, \zeta > 0$ and the Brownian motions are correlated with correlation ρ. The jump component is given by $\widetilde{N}(V_t, dt, dx) = N(V_t, dt, dx) - n(V_t, dt, dx)$, where $N(V_t, dt, dx)$ is a Poisson random measure, and its compensator $n(V_t, dt, dx)$

is of the *state-dependent* form $V_t \mu(dx)dt$, with $\mu(dx)$ the Lévy measure given in (2.6). A model of this kind has been proposed by Bates [2000] to explain the time-variation of jump-risk implicit in observed option prices. Bates also proposes a second variance factor, which we omit in this example, in order to remain in the scope of Definition 2.8. It would however not be difficult to extend our approach to the two-factor Bates model, since the two proposed variance-factors are mutually independent, causing the corresponding generalized Riccati equations to decouple. Since it is affine, the above model can be characterized in terms of the functions F and R:

$$(6.5) \qquad F(u,w) = \lambda\theta w$$

$$(6.6) \qquad R(u,w) = \frac{1}{2}(u^2 - u) + \frac{\zeta^2}{2}w^2 - \lambda w + uw\rho\zeta + \widetilde{\kappa}(u) \ .$$

where $\widetilde{\kappa}(u)$ is the compensated cumulant generating function of the Lévy measure μ. As in the Heston model we can obtain $w(u)$ and $h(u)$ explicitly, and get

$$h(u) = \frac{-\chi(u) - \sqrt{\Delta(u)}}{\zeta^2} \ , \quad \text{and} \quad h(u) = \lambda\theta w(u) \ ,$$

where $\chi(u) = \rho\zeta u - \lambda$ and $\Delta(u) = \chi(u)^2 - \zeta^2(u^2 - u + 2\widetilde{\kappa}(u))$. Both $w(u)$ and $h(u)$ are defined on $I = J = \{u : \Delta(u) \geq 0\}$. The time of moment explosion can again be calculated explicitly, and is given by

$$(6.7) \qquad T_*(u) = \begin{cases} +\infty & \Delta(u) > 0 \\ \frac{2}{\sqrt{-\Delta(u)}}\left(\arctan\frac{\sqrt{-\Delta(u)}}{\chi(u)} + \pi\mathbf{1}_{\{\chi(u) < 0\}}\right) & -\infty < \Delta(u) < 0 \\ 0 & \Delta(u) = -\infty \ . \end{cases}$$

6.3. **The Barndorff-Nielsen-Shephard model.** The Barndorff-Nielsen-Shephard (BNS) model was introduced by Barndorff-Nielsen and Shephard [2001] as a model for asset pricing. In SDE form it is given in the risk-neutral case by

$$dX_t = (\delta - \frac{1}{2}V_t)dt + \sqrt{V_t}\,dW_t + \rho\,dJ_{\lambda t}$$
$$dV_t = -\lambda V_t\,dt + dJ_{\lambda t}$$

where $\lambda > 0$, $\rho < 0$ and $(J_t)_{t\geq 0}$ is a Lévy subordinator, i.e. a pure jump Lévy process that increases a.s. The drift δ is determined by the martingale condition for $(S_t)_{t\geq 0}$. The time-scaling $J_{\lambda t}$ is introduced by Barndorff-Nielsen and Shephard to make the invariant distribution of the variance process independent of λ. The distinctive features of the BNS model are that the variance process has no diffusion component, i.e. moves purely by jumps and that the negative correlation between variance and price movements is achieved by simultaneous jumps in $(V_t)_{t\geq 0}$ and $(X_t)_{t\geq 0}$. The BNS model is an affine stochastic volatility model, and F and R are given by

$$(6.8) \qquad F(u,w) = \lambda\kappa(w + \rho u) - u\lambda\kappa(\rho)$$

$$(6.9) \qquad R(u,w) = \frac{1}{2}(u^2 - u) - \lambda w$$

where $\kappa(u)$ is the cumulant generating function of $(J_t)_{t\geq 0}$.

We simply have $\chi(u) = -\lambda$ and $w(u)$ from Lemma 3.2 is given by

$$w(u) = \frac{1}{2\lambda}(u^2 - u) \ .$$

It follows that

$$h(u) = \lambda\kappa\left(\frac{u^2}{2\lambda} + u\left(\rho - \frac{1}{2\lambda}\right)\right) - u\lambda\kappa(\rho) \ .$$

This expression can be interpreted as cumulant generating function of a Brownian motion with variance $\frac{1}{\lambda}$ and drift $\rho - \frac{1}{2\lambda}$, subordinated by the Lévy process $J_{\lambda t}$ and then mean-corrected to satisfy the martingale condition.

To analyze moment explosions in the BNS model, let $\kappa_+ := \sup\{u > 0 : \kappa(u) < \infty\}$. It is easy to see that f_+ is given by $f_+ = \max(\kappa_+ - \rho u, 0)$. Since $r_+ = \infty$, we have that the explosion time for the moment of order u is given by

$$T_*(u) = \int_0^{f_+} \frac{d\eta}{R(u,\eta)} = -\frac{1}{\lambda}\log\left(1 - \frac{2\lambda(\max(\kappa_+ - \rho u, 0))}{u(u-1)}\right) \ .$$

The critical moment functions $u_\pm(T)$ can be obtained explicitly by solving a quadratic equation, and are given by

$$u_\pm(t) = \frac{1}{2} - \frac{\rho\lambda}{1 - e^{-\lambda t}} \pm \sqrt{\frac{1}{4} + \frac{(2\kappa_+ - \rho)\lambda}{1 - e^{-\lambda t}} + \frac{\rho^2\lambda^2}{(1 - e^{-\lambda t})^2}} \ .$$

The large-strike asymptotics for the implied volatility smile in the sense of Lee can be explicitly calculated by inserting $u_\pm$ into Proposition 5.1.

6.4. **The Heston model in the stationary variance regime.** In the Heston model the limit distribution of the variance process $(V_t)_{t\geq 0}$ is a Gamma distribution with parameters $(-\frac{2\lambda\theta}{\zeta^2}, \frac{2\lambda}{\zeta^2})$. This is well-known, but can also be obtained by applying Proposition 3.1. The cumulant generating function $l(w)$ is thus given by

$$l(w) = -\frac{2\lambda\theta}{\zeta^2}\log\left(1 - \frac{\zeta^2}{2\lambda}w\right) \ ,$$

defined on $(-\infty, \frac{2\lambda}{\zeta^2})$, such that $l_+ = \frac{2\lambda}{\zeta^2}$. As before we have that $\chi(u) = \rho\zeta u - \lambda$, and we assume that $\chi(1) < 0$. In addition we define $\chi^+(u) = \rho\zeta u + \lambda$. By Theorem 4.2, the explosion time in the stationary regime is given by

$$(6.10) \quad T_*^S(u) = \int_0^{2\lambda/\zeta^2} \frac{d\eta}{R(u,\eta)} =$$

$$= \begin{cases} \infty & \sqrt{\Delta(u)} > -\chi^+(u), \\ \frac{1}{\sqrt{\Delta}}\log\left|\frac{\chi^+\chi + 2\lambda\sqrt{\Delta} - \Delta}{\chi^+\chi - 2\lambda\sqrt{\Delta} - \Delta}\right| & 0 < \sqrt{\Delta(u)} < -\chi^+(u), \\ \frac{2}{\sqrt{-\Delta}}\arctan\left(\frac{2\lambda\sqrt{-\Delta}}{\chi^+\chi - \Delta} + \pi\mathbf{1}_{\{\chi^+\chi < \Delta\}}\right) & \Delta(u) < 0 \ . \end{cases}$$

In Figure 2 $T_*^S(u)$ is plotted together with $T_*(u)$ for the Heston model.

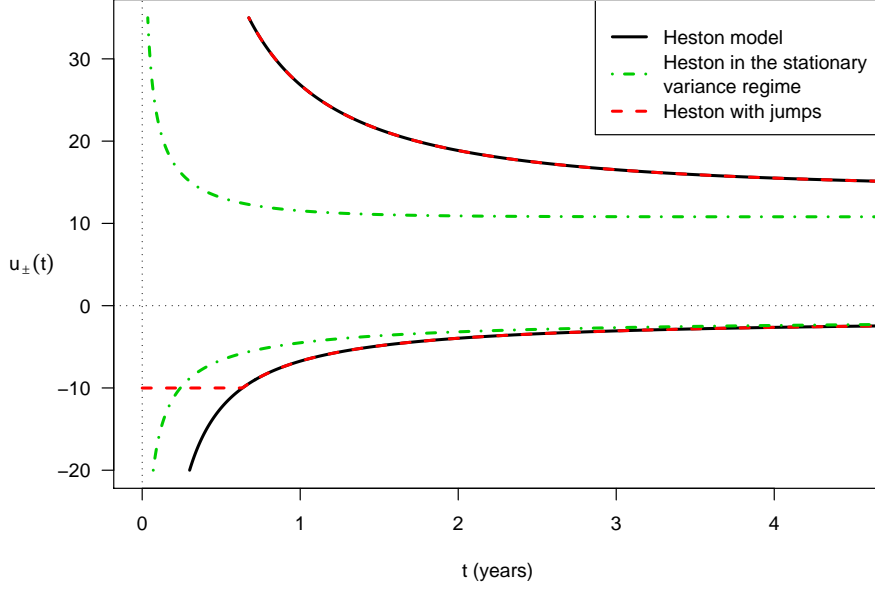FIGURE 2. This plot shows the critical moment functions $u_\pm(t)$ for a Heston model with the same parameters as in Figure 1. Also shown are $u_\pm^S(t)$ for the model in the stationary variance regime, and $u_\pm^{\mathrm{Jump}}(t)$ for the Heston model with an independent jump component, whose negative jump heights are exponentially distributed with mean $\alpha = -0.1$. Note that $u_\pm^{\mathrm{Jump}}(t)$ coincides with $u_\pm(t)$ everywhere except in the lower left corner of the plot.

6.5. **The BNS model in the stationary variance regime.** In the BNS model, the cumulant generating function of the limit distribution L of the variance process is given by Proposition 3.1 by

$$l(w) = \int_0^w \frac{\kappa(\eta)}{\eta} d\eta \ ,$$

provided the log-moment condition $\int_{y>1} (\log y)\, \mu(dy) < \infty$ holds for the Lévy measure of $(J_t)_{t\geq 0}$. The above integral is finite as long as $w \in (-\infty, \kappa_+)$, and infinite outside. Thus $l_+ = \kappa_+$. In Section 6.3 we obtained that $f_+(u) = \kappa_+ - \rho u$, such that the time of moment explosion under stationary variance is given by

$$T_*^S(u) = \int_0^{min(f_+(u), l_+)} \frac{d\eta}{R(u, \eta)} = -\frac{1}{\lambda} \log\left(1 - \frac{2\lambda k(u)}{u(u-1)}\right) \ ,$$

where $k(u) = \kappa_+$ for $u \geq 1$ and $k(u) = \max(\kappa_+ - \rho u, 0)$ for $u \leq 0$. Again, this expression can be inverted to give the critical moment functions in the stationary

variance case. By definition $\rho \leq 0$, such that we obtain

$$u_-^S(T) = \frac{1}{2} - \frac{\rho\lambda}{1 - e^{-\lambda T}} - \sqrt{\frac{1}{4} + \frac{(2\kappa_+ - \rho)\lambda}{1 - e^{-\lambda T}} + \frac{\rho^2\lambda^2}{(1 - e^{-\lambda T})^2}}$$

$$u_+^S(T) = \frac{1}{2} + \sqrt{\frac{1}{4} + \frac{2\kappa_+ \lambda}{1 - e^{-\lambda T}}} \ .$$

APPENDIX A. ADDITIONAL PROOFS

Proof of Theorem 2.1. Let $t \leq \tau$. By the flow equation we can write

$$\phi(\tau, u, \eta) = \phi(t, u, \eta) + \phi(\tau - t, u, \psi(t, u, \eta))$$
$$\psi(\tau, u, \eta) = \psi(\tau - t, u, \psi(t, u, \eta)) \ .$$

Since the left sides are finite by assumption, it follows that also $\phi(t, u, \eta)$ and $\psi(t, u, \eta)$ are. V_t is non-negative, such that

$$|\mathbb{E}\left[\exp\left(uX_t + wV_t\right)\right]| \leq |\mathbb{E}\left[\exp\left(uX_t + \eta V_t\right)\right]| \ ,$$

whenever $\operatorname{Re} w \leq \operatorname{Re} \eta$. Thus $\phi(t, u, w)$ and $\psi(t, u, w)$ exist for all $w \in \mathbb{C}$ with $\operatorname{Re} w \leq \operatorname{Re} \eta$. As a particular case we can conclude that $\phi(t, u, w)$ and $\psi(t, u, w)$ exist for all (u, w) in $\mathcal{U} := \left\{(u, w) \in \mathbb{C}^2 : \operatorname{Re} u = 0, \operatorname{Re} w \leq 0\right\}$.

We also define $\mathcal{U}^\circ := \left\{(u, w) \in \mathbb{C}^2 : \operatorname{Re} u = 0, \operatorname{Re} w < 0\right\}$, and show next that $\phi(t, u, w)$ and $\psi(t, u, w)$ are (right-)differentiable at $t = 0$ for all $(u, w) \in \mathcal{U}^\circ$. The key idea of our proof is originally due to Montgomery and Zippin [1955], and has also been presented in Filipović and Teichmann [2003] and Dawson and Li [2006]. First note that the identity

$$\mathbb{E}\left[wV_t e^{uX_t + wV_t}\right] = \left(\frac{\partial}{\partial w}\phi(t, u, w) + V_0\frac{\partial}{\partial w}\psi(t, u, w)\right) \exp\left(\phi(t, u, w) + V_0\psi(t, u, w) + X_0 u\right)$$

shows that $\frac{\partial}{\partial w}\phi(t, u, w)$ and $\frac{\partial}{\partial w}\psi(t, u, w)$ exist, and are continuous for all $t \leq \tau$ and $(u, w) \in \mathcal{U}^\circ$. By Taylor expansion it holds that

$$\int_0^s \psi(r, u, \psi(t, u, w)) \, dr - \int_0^s \psi(r, u, w) \, dr = \int_0^s \frac{\partial}{\partial w}\psi(r, u, w) \, dr \, (\psi(t, u, w) - w)$$

$$\text{(A.1)} \hspace{6cm} + o\big(|\psi(t, u, w) - w|\big) \ .$$

On the other side, using the flow property, we calculate

$$\int_0^s \psi(r, u, \psi(t, u, w)) \, dr - \int_0^s \psi(r, u, w) \, dr = \int_0^s \psi(r + t, u, w) \, dr - \int_0^s \psi(r, u, w) \, dr =$$

$$\text{(A.2)}$$

$$= \int_t^{s+t} \psi(r, u, w) \, dr - \int_0^s \psi(r, u, w) \, dr = \int_0^t \psi(r + s, u, w) \, dr - \int_0^t \psi(r, u, w) \, dr \ .$$

Denoting the last expression by $I(s, t)$, and putting (A.1) and (A.2) together, we obtain

$$\lim_{t \to 0} \frac{\left|\frac{1}{s}I(s, t)\right|}{|\psi(t, u, w) - w|} = \left|\frac{1}{s}\int_0^s \frac{\partial}{\partial w}\psi(t, u, w) \, dr\right| \ .$$

Thus, writing $M_s = \frac{1}{s}\int_0^s \frac{\partial}{\partial w}\psi(t, u, w) \, dr$, we have

$$\lim_{t \to 0} \frac{1}{t}|\psi(t, u, w) - w| = \left|\lim_{t \to 0} \frac{I(s, t)}{st}\right| \cdot |M_s|^{-1} = \left|\frac{\psi(s, u, w) - w}{s}\right| |M_s|^{-1} \ .$$

But M_s is a continuous function of s, and $\lim_{s\to 0} M_s = \frac{\partial}{\partial w}\psi(0, u, w) = 1$, such that for s small enough $M_s \neq 0$. We conclude that the left hand side is finite, and using (A.1) we obtain that

$$\lim_{t\to 0}\frac{\psi(t, u, w) - w}{t} = \left(\frac{\psi(s, u, w) - w}{s}\right)\cdot\left(\frac{1}{s}\int_0^s\frac{\partial}{\partial w}\psi(r, u, w)\, dr\right)^{-1}.$$

The finiteness of the right hand side implies the existence of the limit on the left. In addition the right hand side is continuous for $(u, w) \in \mathcal{U}^\circ$, showing that also the left hand side is. A similar calculation for $\phi(t, u, w)$ shows that

$$\lim_{t\to 0}\frac{\phi(t, u, w)}{t} = \frac{\phi(s, u, w)}{s} - \lim_{t\to 0}\left(\frac{\psi(t, u, w) - w}{t}\right)\cdot\left(\frac{1}{s}\int_0^s\frac{\partial}{\partial w}\phi(r, u, w)\, dr\right),$$

allowing the same conclusions for $\phi(t, u, w)$. We have thus shown that the time-derivatives of $\phi(t, u, w)$ and $\psi(t, u, w)$ at $t = 0$ exist, and are continuous in $\mathcal{U}^\circ$. Combining Duffie et al. [2003, Proposition 7.2] and Duffie et al. [2003, Proposition 6.4] the differentiability can be extended from $\mathcal{U}^\circ$ to $\mathcal{U}$, and we have shown that $(X_t, V_t)_{t\geq 0}$ is a *regular* affine process. The rest of Theorem 2.1 follows now as in Duffie et al. [2003, Theorem 2.7] $\qquad\qquad\square$

Proof of Lemma 2.2. We prove the assertions of Lemma 2.2 for F; they follow analogously for R. By the Lévy-Khintchine representation (2.6), $F(u, w) + c$ is the cumulant generating functions of some infinitely divisible random variables, say X. Writing $z = (u, w) \in \mathbb{R}^2$, and using Hölder's inequality it holds for any $\lambda \in [0, 1]$ that

$$(A.3)\quad F(\lambda z_1 + (1 - \lambda)z_2) = \log\mathbb{E}\left[e^{\lambda\langle z_1, X\rangle}e^{(1-\lambda)\langle z_2, X\rangle}\right] - c \leq$$

$$\leq \lambda\log\mathbb{E}\left[e^{\langle z_1, X\rangle}\right] + (1 - \lambda)\mathbb{E}\left[e^{\langle z_2, X\rangle}\right] - c = \lambda F(z_1) + (1 - \lambda)F(z_2),$$

showing convexity of F. In addition equality in (A.3) holds if and only if $ke^{\langle z_1, X\rangle} = e^{\langle z_2, X\rangle}$ a.s. for some $k > 0$. This in turn is equivalent to $\langle z_1 - z_2, X\rangle$ being constant a.s. Choosing now z_1 and $z_2 \neq z_1$ from some one-dimensional affine subspace $U = \{p + \langle q, x\rangle : x \in \mathbb{R}\}$ of $\mathbb{R}^2$, we see that either $\langle q, X\rangle$ is constant a.s. in which case $F|_U$ is affine, or it is not constant, in which case strict inequality holds in (A.3) for all $z_1, z_2 \in U$, showing (c).

Let $L_\alpha = \{z : F(z) \leq \alpha\}$ be a level set of F, and $z_n \in L_\alpha$ a sequence converging to z. Then by Fatou's Lemma

$$\log\mathbb{E}[e^{\langle z, X\rangle}] - c \leq \liminf_{n\to\infty}\log\mathbb{E}[e^{\langle z_n, X\rangle}] - c \leq \alpha,$$

showing that $z \in L_\alpha$ and thus that F is a closed convex function. Finally F is proper, because $F(0, 0) = c > -\infty$, showing (a).

Next we show analyticity: Consider the random variables $X_n := X\mathbf{1}_{\{|X|\leq n\}}$. Since they are bounded, their Laplace transforms, and hence also their cumulant generating functions are entire functions on $\mathbb{C}^2$, and thus analytic on $\mathbb{R}^2$. As a uniform limit of analytic functions $F(u, w)$ is analytic in the interior of $\operatorname{dom} F$, showing (b). Assertion (d) follows directly from Theorem 2.1. $\qquad\qquad\square$

References

Leif B. G. Andersen and Vladimir V. Piterbarg. Moment explosions in stochastic volatility models. *Finance and Stochastics*, 11:29–50, 2007.

G. Bakshi, C. Cao, and Z. Chen. Empirical performance of alternative option pricing models. *Journal of Finance*, 52:2003–2049, 1997.

Ole E. Barndorff-Nielsen. Processes of normal inverse Gaussian type. *Finance and Stochastics*, 2:41 – 68, 1997.

Ole E. Barndorff-Nielsen and Neil Shephard. Non-Gaussian Ornstein–Uhlenbeck-based models and some of their uses in financial economics. *Journal of the Royal Statistical Society B*, 63:167–241, 2001.

David S. Bates. Post-'87 crash fears in the S&P 500 futures option market. *Journal of Econometrics*, 94:181–238, 2000.

David S. Bates. Jump and stochastic volatility: exchange rate processes implicit in Deutsche Mark options. *The Review of Financial Studies*, 9:69–107, 1996.

Shalom Benaim and Peter Friz. Regular variation and smile asymptotics. arXiv:math/0603146v2, 2006.

Peter Carr and Liuren Wu. Time-changed Lévy processes and option pricing. *Journal of Financial Economics*, 71(1):113–141, 2004.

D. A. Dawson and Zenghu Li. Skew convolution semigroups and affine Markov processes. *The Annals of Probability*, 34(3):1103–1142, 2006.

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

Darrell Duffie, Jun Pan, and Kenneth Singleton. Transform analysis and asset pricing for affine jump-diffusions. *Econometrica*, 68(6):1343 – 1376, 2000.

William Feller. *An introduction to probability theory and its applications.* John Wiley & Sons, 2nd edition, 1971.

Damir Filipović. A general characterization of one factor affine term structure models. *Finance and Stochastics*, 5:389–412, 2001.

Damir Filipović and Josef Teichmann. Regularity of finite-dimensional realizations for evolution equations. *Journal of Functional Analyis*, 197:433–446, 2003.

Jim Gatheral. *The Volatilty Surface.* Wiley Finance, 2006.

Philip Hartman. *Ordinary Differential Equations.* Birkhäuser, 1982.

S. Heston. A closed-form solution of options with stochastic volatility with applications to bond and currency options. *The Review of Financial Studies*, 6:327–343, 1993.

Martin Keller-Ressel and Thomas Steiner. Yield curve shapes and the asymptotic short rate distribution in affine one-factor models. *Finance and Stochastics*, 12: 149 – 172, 2008.

Roger Lee. The moment formula for implied volatility at extreme strikes. *Mathematical Finance*, 14(3):469–480, 2004.

Alan L. Lewis. *Option Valuation under Stochastic Volatility.* Finance Press, 2000.

P. Lions and M. Musiela. Correlations and bounds for stochastic volatility models. *Annales de l'Institut Henri Poincaré*, 24:1–16, 2007.

Deane Montgomery and Leo Zippin. *Topological Transformation Groups.* Interscience Publishers, Inc., 1955.

W. F. Osgood. Beweis der Existenz einer Lösung der Differentialgleichung $\frac{dy}{dx} = f(x, y)$ ohne Hinzunahme der Cauchy-Lipschitz'schen Bedingung. *Monatshefte für Mathematik und Physik*, 9:331–345, 1898.

Ken-Iti Sato. *Lévy processes and infinitely divisible distributions.* Cambridge University Press, 1999.

Vienna University of Technology, Wiedner Hauptstrasse 8–10, A-1040 Wien, Austria
E-mail address: mkeller@fam.tuwien.ac.at