Affine forward variance models

Jim Gatheral, Baruch College, CUNY,
jim.gatheral@baruch.cuny.edu,

Martin Keller-Ressel, TU Dresden,
Martin.Keller-Ressel@tu-dresden.de

January 19, 2018

Abstract

We introduce the class of affine forward variance (AFV) models of which both the conventional Heston model and the rough Heston model are special cases. We show that AFV models can be characterized by the affine form of their cumulant generating function, which can be obtained as solution of a convolution Riccati equation. We further introduce the class of affine forward order flow intensity (AFI) models, which are structurally similar to AFV models, but driven by jump processes, and which include Hawkes-type models. We show that the cumulant generating function of an AFI model satisfies a generalized convolution Riccati equation and that a high-frequency limit of AFI models converges in distribution to the AFV model.

Contents

1 Introduction

The class of affine processes introduced in [DFS03], consists of all continuous-time Markov processes taking values in $\mathbb{R}^m_{\geqslant 0} \times \mathbb{R}^n$, whose log-characteristic function depends in an affine way on the initial state vector of the process. Affine processes have proved particularly convenient for financial modeling, typically giving rise to models with tractable formulae for the values of financial claims; the perennially popular Heston model [Hes93] is just one (and perhaps the most famous) example of such a model.

In this paper, we introduce the class of *affine forward variance (AFV) models* of which classical Markovian affine stochastic volatility models turn out to be a special case. By writing our model in forward variance form, we are able to provide a unique characterization of a much wider class of affine stochastic volatility models, which includes non-Markovian models, such as the rough Heston model of [EER16]. It will also become evident that the class of AFV models is closely related to the affine Volterra processes introduced in the influential paper [JLP17].

Inspired by the original derivation [EER16] of the rough Heston model as a limit of simple pure jump models of order flow, we further introduce the class of *affine forward order flow intensity* (AFI) models. These model are structurally similar to affine forward variance models and generalize the simple order flow model of [EER16], by allowing arbitrary order size distributions and more general decay of the self-excitation of order flow. We define a high-frequency limit in which such models give rise to continuous affine forward variance models. In so doing, we generalize and simplify previous such derivations.

Our paper proceeds as follows. In Section 2, we introduce the class of affine forward variance models and show that a forward variance model has an affine cumulant generating function (CGF) if and only if it can be written in a very specific form. We further show that the CGF can be obtained as the unique global solution of a convolution Riccati equation closely related to the Volterra-Riccati equations of [JLP17]. In Section 3, we introduce the class of AFI models, showing that the CGF of such models solves a *generalized* convolution Riccati equation. In Section 4, we show that AFI models become AFV models in a high-frequency limit, where order arrivals are extremely frequent and order sizes extremely small.

2 Affine forward variance models

2.1 A characterization of affine forward variance models

On a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with right-continuous filtration $(\mathcal{F}_t)_{t\geqslant 0}$, following [BG12], we consider a forward variance model of the form

$$dS_t = S_t \alpha(V_t) \left(\rho dW_t + \sqrt{1-\rho^2} dW_t^{\perp}\right) \tag{2.1a}$$

$$d\xi_t(T) = \eta_t(T;\omega) dW_t, \quad t \in (0,T), \tag{2.1b}$$

where $W, W^\perp$ are independent Brownian motions, the $\mathbb{R}_{\geqslant 0}$-valued stochastic process $\eta_t(T;\omega)$ is progressively measurable for all $T > 0$ and ξ is linked to the instantaneous variance V by

$$\xi_t(T) = \mathbb{E}\left[V_T \middle| \mathcal{F}_t\right]. \tag{2.2}$$

Due to the right-continuity of $(\mathcal{F}_t)$, we can recover V_t from $\xi_t(T)$ as $V_t = \lim_{T\downarrow t} \xi_t(T)$. The initial conditions of the forward variance model (2.1) are the initial stock price S_0 and the initial forward variance $(\xi_0(T)_{T>0})$. The first equation can be written in terms of the logarithm $X_t = \log S_t$ as

$$dX_t = -\frac{\alpha(V_t)^2}{2}dt + \alpha(V_t)\left(\rho dW_t + \sqrt{1-\rho^2}dW_t^\perp\right).$$

We assume that $\mathbb{P}$ is a martingale measure for S, such that $\mathbb{E}[S_t] = \mathbb{E}\left[e^{X_t}\right] = S_0$. In particular, this implies by Jensen's inequality that the moments $\mathbb{E}[S_t^u]$ are finite for all $u \in [0,1]$.

Remark 2.1. As noted earlier in [BG12], all conventional finite-dimensional Markovian stochastic volatility models may be cast as forward variance models.

Definition 2.2. We say that the process (X, ξ) has an *affine cumulant generating function* determined by $g(t;u)$, if its conditional cumulant generating function is of the form

$$\log \mathbb{E}\left[e^{u(X_T - X_t)}\middle| \mathcal{F}_t\right] = \int_t^T g(T-s;u)\xi_t(s)ds. \tag{2.3}$$

for all $u \in [0,1]$, $0 \leqslant t \leqslant T$ and $g(.;u)$ is $\mathbb{R}_{\leqslant 0}$-valued and continuous on $[0,T]$ for all $T > 0$ and $u \in [0,1]$.

Remark 2.3. Alternatively, we could consider (2.3) with imaginary parameter $u = iz$ for $z \in \mathbb{R}$, i.e. an *affine log-characteristic function* as in [EER16]. However, it will turn out that restricting to real parameters greatly simplifies the mathematical treatment.

Convolution integrals, as in the exponent of (2.3), will appear frequently in the following calculations and so it will prove useful to introduce the convolution operation $\star$. For functions with multiple arguments or subscripts, we use the convention that convolution acts on the first argument, excluding subscripts. Other arguments or subscripts are passed on to the result. With this convention (2.3) can be written succinctly as

$$\mathbb{E}\left[e^{u(X_T - X_t)}\middle| \mathcal{F}_t\right] = \exp\Big((g \star \xi)_t(T;u)\Big). \tag{2.4}$$

We impose the following technical assumptions on the forward variance model:

Assumption 2.4. (a) For $dt \otimes d\mathbb{P}$-almost all (t,ω) it holds that $T \mapsto \eta_t(T;\omega)$ is right-continuous, decreasing and non-zero.

(b) For any $T > 0$ the integrability condition

$$\int_0^T \left(\int_0^T \eta_s(s+r;\omega)^2 ds \right)^{1/2} dr < \infty \tag{2.5}$$

holds for almost all $\omega \in \Omega$.

Note that if η decomposes as $\eta_t(T;\omega) = Y_t(\omega)\kappa(T-t)$ with Y_t non-negative, then (a) is equivalent to κ being decreasing and the integrability condition (b) is equivalent to

$$\int_0^T Y_s(\omega)^2 ds < \infty, \ a.s. \quad \text{and} \quad \int_0^T \kappa(r)dr < \infty,$$

In particular, condition (b) is weak enough to accommodate kernels κ with integrable singularity at 0.

Theorem 2.5. *Under Assumption 2.4 the model (X, ξ) has an affine CGF if and only if*

$$\alpha(V_t) = a\sqrt{V_t} \tag{2.6a}$$

$$\eta_t(T;\omega) = \sqrt{V_t(\omega)}\kappa(T-t) \tag{2.6b}$$

for some $a \geqslant 0$ and a deterministic, non-negative decreasing kernel κ, which satisfies $\int_0^T \kappa(r)dr < \infty$ for all $T > 0$.
Moreover, $g(.;u) : \mathbb{R}_{\geqslant 0} \to \mathbb{R}_{\leqslant 0}$ in (2.3) is the unique global solution of the convolution Riccati equation

$$g(t;u) = R_V\Big(u, \int_0^t \kappa(t-s)g(s;u)ds\Big) = R_V\Big(u, (\kappa \star g)(t;u)\Big), \quad t \geqslant 0 \tag{2.7}$$

where

$$R_V(u,w) = \frac{a^2}{2}(u^2 - u) + a\rho uw + \frac{1}{2}w^2. \tag{2.8}$$

Remark 2.6. Alternatively, the CGF $g(t;u)$ can be written as

$$g(t;u) = R_V(u, f(t,u)),$$

where $f(t,u)$ is the unique global solution of the non-linear Volterra equation

$$f(t;u) = \int_0^t \kappa(t-s)R_V(u, f(t,s))ds. \tag{2.9}$$

See Appendix 5 for further discussion of non-linear Volterra equations and for the equivalence of equations (2.7) and (2.9).

Remark 2.7. Note that the instantaneous variance process $V_t = \lim_{T\downarrow t} \xi_t(T)$ of an AFV model can be written as

$$V_t = \xi_0(t) + \int_0^t \kappa(t-s)\sqrt{V_s} dW_s,$$

and is therefore an affine Volterra process in the sense of [JLP17]. It has been shown in [JLP17] that under certain conditions, such processes possess an affine log-characteristic function and that this function satisfies the nonlinear Volterra equation (2.9). Our result imposes weaker assumptions on the kernel κ and adds the converse direction which characterizes (2.6) as necessary *and* sufficient. Note that our formulation in forward variance form avoids working directly with the (non-semimartingale) process V and therefore simplifies many of the arguments.

2.2 Two examples: Heston and rough Heston models

Example 2.8 (The Heston model). The Heston model [Hes93] is given by

$$dS_t = S_t\sqrt{V_t}\left(\rho dW_t + \sqrt{1-\rho^2} dW_t^{\perp}\right) \tag{2.10a}$$

$$dV_t = -\lambda(V_t - \theta)dt + \zeta\sqrt{V_t} dW_t. \tag{2.10b}$$

A simple calculation shows that

$$\xi_t(T) = \mathbb{E}\left[V_T \middle| \mathcal{F}_t\right] = \theta\left(1 - e^{-\lambda(T-t)}\right) + e^{-\lambda(T-t)} V_t.$$

Hence,

$$d\xi_t(T) = \zeta e^{-\lambda(T-t)}\sqrt{V_t} dW_t$$

and it follows that the Heston model can be written as an affine forward variance model with

$$\alpha(V_t) = \sqrt{V_t}, \qquad \kappa(x) = \zeta e^{-\lambda x}$$

and initial forward variance

$$\xi_0(T) = \theta\left(1 - e^{-\lambda T}\right) + e^{-\lambda T} V_0 = V_0 + (\theta - V_0)\lambda \int_0^T \kappa(s) ds.$$

To obtain the Riccati ODEs for the Heston model in the usual form (see e.g. [KR11]), let $\psi(.;u)$ be a C^1-function such that

$$g(t;u) = \frac{\partial}{\partial t}\psi(t;u) + \lambda\psi(t;u), \quad \text{and} \quad \psi(0;u) = 0.$$

By partial integration we obtain

$$(\kappa \star g)(t;u) = \zeta \int_0^t e^{-\lambda(t-s)} g(s;u) = \zeta\psi(t;u).$$

Inserting into the convolution Riccati equation (2.7) yields

$$\frac{\partial}{\partial t}\psi(t;u) = \frac{1}{2}(u^2 - u) + (\zeta\rho u - \lambda)\psi(t;u) + \frac{\zeta^2}{2}\psi(t;u)^2, \qquad \psi(t;0) = 0,$$

in accordance with [KR11]. Furthermore, it is straightforward to show that

$$\int_t^T g(t-s;u)\xi_t(s)ds = \phi(t,u) + V_t\psi(t;u),$$

with $\phi(t,u) = \lambda\theta\int_0^t \psi(s,u)ds$. ◇

Example 2.9 (The rough Heston model). In the rough Heston model, introduced in [EER16], (2.10b) is replaced by

$$V_t = V_0 + \frac{1}{\Gamma(\alpha)}\int_0^t (t-s)^{\alpha-1}\lambda(\theta - V_s)ds + \frac{\zeta}{\Gamma(\alpha)}\int_0^t (t-s)^{\alpha-1}\sqrt{V_s}dW_s \quad (2.11)$$

where $\alpha \in (1/2, 1)$ is related to the 'roughness' of the paths of V. In [ER17] it is shown that the forward variance in the rough Heston model satisfies

$$d\xi_t(T) = \kappa(T-s)\sqrt{V_t}dW_t,$$

with the kernel

$$\kappa(x) = \zeta x^{\alpha-1}E_{\alpha,\alpha}(-\lambda x^{\alpha})$$

and where $E_{\alpha,\beta}(x)$ denotes the generalized Mittag-Leffler function (cf. [EMOT81], [Pod98, Sec. 1.2]). Thus, the rough Heston model is an affine forward variance model in the sense of Theorem 2.5. The initial forward variance is given by (cf. [ER17, Prop. 3.1])

$$\xi_0(T) = V_0 + (\theta - V_0)\lambda\int_0^T \kappa(s)ds.$$

To obtain the *fractional* Riccati equation (cf. [EER16, Eq. (24)]) for the rough Heston model set

$$\psi(t;u) = \frac{1}{\zeta}(\kappa \star g)(t;u) = \frac{1}{\Gamma(\alpha)}\int_0^t (t-s)^{\alpha-1}E_{\alpha,\alpha}(-\lambda(t-s))g(s;u)ds.$$

By [EER16, Lem. A.2] $\psi(t;u)$ satisfies

$$D^{\alpha}\psi(t;u) + \lambda\psi(t;u) = g(s;u)$$

where D^{α} denotes the Riemann-Liouville fractional derivative of order α. Inserting into the convolution Riccati equation (2.7) yields

$$D^{\alpha}\psi(t;u) = \frac{1}{2}(u^2 - u) + (\zeta\rho u - \lambda)\psi(t;u) + \frac{\zeta^2}{2}\psi(t;u)^2, \qquad \psi(t;0) = 0,$$

in accordance with [EER16, Eq (24)]. Denote by $I^\alpha f = \frac{1}{\Gamma(\alpha)} \int_0^\infty (t-s)^{\alpha-1} f(s)ds$ the Riemann-Liouville fractional integral of order α and write $\mathbf{1}$ for the function of constant value one. The exponent in (2.3) can be transformed as follows:

$$\begin{aligned}
\int_0^\cdot g(t-s;u)\xi_0(s)ds &= V_0(g \star \mathbf{1}) + (\theta - V_0)\frac{\lambda}{\zeta}(g \star \kappa \star \mathbf{1}) = \\
&= V_0\left((g - \lambda\,\psi) \star \mathbf{1}\right) + \theta\lambda\,(\psi \star \mathbf{1}) = \\
&= V_0(D^\alpha \psi) \star \mathbf{1} + \theta\lambda \int_0^\cdot \psi(s;u) = \\
&= V_0 I^{1-\alpha}\psi + \theta\lambda \int_0^\cdot \psi(s;u),
\end{aligned}$$

which is the same as [EER16, Eq (23)]. ◇

2.3 Proof of the characterization result

To prepare for the proof of Theorem 2.5, we introduce the following notation: Given a continuous function $g : \mathbb{R}_{\geqslant 0} \times [0,1] \to \mathbb{R}, (t,u) \mapsto g(t;u)$, we set

$$G_t = (g \star \xi)_t(T;u) = \int_t^T g(T-s;u)\xi_t(s)ds, \tag{2.12}$$

$$M_t = \exp\left(uX_t + G_t\right). \tag{2.13}$$

If (X,ξ) has an affine CGF determined by $g(t;u)$ then it follows from (2.3) that M is a martingale. Conversely, if M is a martingale, then (2.3) follows by taking conditional expectations. Hence, the affine property of (X,ξ) can be characterized in terms of the martingale property of M. In order to apply Itô's formula to M we represent G as an Itô process. The calculation is analogous to the drift computation in the Heath-Jarrow-Morton-model (cf. [Fil09, Ch. 6]) and uses the stochastic Fubini theorem to interchange stochastic integral and Lebesgue integral.

Lemma 2.10. *Let G be given as in* (2.12) *and let Assumption 2.4 hold. Then G can be written in Itô process form as*

$$G_t = \int_0^T g(T-s;u)\xi_0(s)ds - \int_0^t g(T-s;u)V_s ds + \int_0^t v_s(T;u,\omega)dW_s,$$

where

$$v_t(T,u,\omega) = (g \star \eta)_t(T,u,\omega) = \int_t^T g(T-r;u)\eta_t(r;\omega)dr. \tag{2.14}$$

Proof. Following [Fil09, p. 94] closely, we compute

$$\begin{aligned}
G_t &= \int_t^T g(T-s;u)\xi_t(s)ds = \\
&= \int_t^T g(T-s;u)\xi_0(s)ds + \int_t^T \int_0^t g(T-s;u)\eta_r(s;\omega)dW_r\, ds \overset{stoch.Fub.}{=} \\
&= \int_t^T g(T-s;u)\xi_0(s)ds + \int_0^t \int_t^T g(T-s;u)\eta_r(s;\omega)ds\, dW_r = \\
&= \int_0^T g(T-s;u)\xi_0(s)ds + \int_0^t \int_r^T g(T-s;u)\eta_r(s;\omega)ds\, dW_r - \\
&\quad - \int_0^t g(T-s;u)\xi_0(s)ds - \int_0^t \int_r^t g(T-s;u)\eta_r(s;\omega)ds\, dW_r \overset{stoch.Fub.}{=} \\
&= \int_0^T g(T-s;u)\xi_0(s)ds + \int_0^t \int_r^T g(T-s;u)\eta_r(s;\omega)ds\, dW_r - \\
&\quad - \int_0^t g(T-s;u) \underbrace{\left(\xi_0(s)ds + \int_0^s \eta_r(s;\omega)dW_r\right)}_{=V_s} ds.
\end{aligned}$$

To justify the application of the stochastic Fubini theorem, we use the condition given in [Ver12, Thm. 2.2]: For $r, x \in [0,T]$ set

$$\psi(r,x;\omega) = g(T-(r+x);u)\eta_r(r+x;\omega),$$

where we extend g by zero whenever $r+x > T$. By [Ver12, Thm. 2.2] a sufficient condition for the exchange of integrals is given by

$$\int_0^T \left(\int_0^T |\psi(r,x;\omega)|^2 dr\right)^{1/2} dx < \infty, \ a.s.$$

Since $g(t;u)$ is continuous on $[0,T]$ for each $u \in [0,1]$ this integrability condition depends only on $\eta_r(r+x;\omega)$ and is implied by Assumption 2.4. □

We are now prepared to prove Theorem 2.5.

Proof of Theorem 2.5. Fix $(T,u) \in (0,\infty) \times \mathbb{R}$. We apply Itô's formula to $M_t = \exp(uX_t + G_t)$ and obtain, using Lemma 2.10,

$$\begin{aligned}
\frac{dM_t}{M_t} &= u\, dX_t + dG_t + \frac{u^2}{2} d[X,X]_t + u\, d[X,G]_t + \frac{1}{2} d[G,G]_t = \qquad (2.15)\\
&= \text{loc.mg.}+ \\
&+ \left\{\frac{1}{2}(u^2-u)\alpha(V_t)^2 - g(T-t;u)V_t + u\rho\alpha(V_t)v_t(T;\omega) + \frac{1}{2}v_t(T;\omega)^2\right\} dt,
\end{aligned}$$

where 'loc.mg.' stands for the local martingale part, which we need not compute explicitly. If (X, ξ) has an affine CGF determined by $g(t; u)$ then M is a local martingale and all dt-terms must vanish. This implies that

$$\frac{1}{2}(u^2 - u)\alpha(V_t)^2 - g(T-t; u)V_t + u\rho\alpha(V_t)v_t(T;\omega) + \frac{1}{2}v_t(T;\omega)^2 = 0 \quad (2.16)$$

for all $u \in [0, 1]$ and for $dt \otimes d\mathbb{P}$-almost all $(t, \omega) \in [0, T] \times \Omega$. Evaluating the equation at three different $u_1, u_2, u_3 \in \mathbb{R}$ and arranging in matrix-vector-form, we obtain

$$\underbrace{\frac{1}{2}\begin{pmatrix} u_1^2 - u_1 & 2u_1\rho & 1 \\ u_2^2 - u_2 & 2u_2\rho & 1 \\ u_3^2 - u_3 & 2u_3\rho & 1 \end{pmatrix}}_{:=M(u_1,u_2,u_3)} \cdot \begin{pmatrix} \alpha(V_t)^2 \\ \alpha(V_t)v_t(T;\omega) \\ v_t(T,\omega)^2 \end{pmatrix} = \begin{pmatrix} g(T-t, u_1) \\ g(T-t, u_2) \\ g(T-t, u_3) \end{pmatrix} V_t.$$

Clearly, unless $\rho = 0$, we can find some $(u_1, u_2, u_3) \in \mathbb{R}^3$, such that $M(u_1, u_2, u_3)$ is invertible. Thus, we obtain

$$\begin{pmatrix} \alpha(V_t)^2 \\ \alpha(V_t)v_t(T;\omega) \\ v_t(T;\omega)^2 \end{pmatrix} = M(u_1, u_2, u_3)^{-1} \begin{pmatrix} g(T-t, u_1) \\ g(T-t, u_2) \\ g(T-t, u_3) \end{pmatrix} V_t,$$

showing that all elements of the vector on the left side must in fact be *linear* functions of V_t and *time-homogeneous*, in the sense that their time-dependency can be reduced to dependency on the difference $T - t$. In the case $\rho = 0$, this argument can be easily adapted by omiting the middle row and column of each vector and matrix involved.

In all cases, we deduce that $\alpha(V_t)$ is proportional to $\sqrt{V_t}$ and that $v_t(T;\omega) = \sqrt{V_t(\omega)}\widetilde{v}(T-t)$ for some deterministic $\widetilde{v}(r)$. But $v_t(T;\omega)$ is the convolution of g and η in the sense of (2.14), such that also $\eta_t(T;\omega)$ must decompose as

$$\eta_t(T;\omega) = \sqrt{V_t}\,\kappa(T-t)$$

for some deterministic $\kappa(r)$. We write

$$(\kappa \star g)(T; u) = \int_0^T g(T-s; u)\kappa(s)ds$$

and, after canceling the common linear factor V_t and setting $\tau = T - t$, (2.16) becomes

$$\frac{1}{2}(u^2 - u) - g(\tau; u) + u\rho \cdot (\kappa \star g)(\tau; u) + \frac{1}{2}(\kappa \star g)(\tau; u)^2 = 0,$$

which is the convolution Riccati equation (2.7). The integrability condition on κ is implied by Assumption 2.4.

We now turn to the claim that the convolution Riccati equation (2.7) has a unique global solution. To this end, we set $H_u(w) = R_V(u, w), u \in (0, 1)$ and show that H_u satisfies the conditions of Corollary A.7. In particular, it is easy to check that for all $u \in (0, 1)$,

- $H_u(w)$ is a finite, strictly convex function on $(-\infty, 0]$ and satisfies $H_u(0) < 0$;
- $H_u(w)$ has a single root $H_u(w_*(u)) = 0$ in $(-\infty, 0]$.

Thus, we conclude from Corollary A.7 the existence of a unique global solution $g(t;u)$ to (2.7) for all $u \in (0,1)$. Moreover, $g(t;u) \leqslant 0$ for all $(t,u) \in \mathbb{R}_{\geqslant 0} \times (0,1)$ by estimate (A.11). We can add the boundary cases $u \in \{0,1\}$, observing that they yield the constant global solution $g(t;u) \equiv 0$, which must be unique by [GLS90, Thm. 13.1.2]. We can now easily complete the proof and show the converse direction of the theorem, by reversing the above arguments: Assume that (2.6) holds true and let G and M be defined as in (2.12) and (2.13), with g the solution of the Riccati equation. Applying Itô's formula as above, we see that in (2.15) all dt-terms vanish and conclude that $M_t = \exp(uX_t + G_t)$ is a local martingale. From $g(t,u) \leqslant 0$ it follows that M is bounded and hence a true martingale. Thus

$$\mathbb{E}\left[e^{uX_T} \middle| \mathcal{F}_t \right] = \mathbb{E}\left[M_T \middle| \mathcal{F}_t \right] = M_t = \exp\left(uX_t + \int_t^T g(T-s;u)\xi_t(s)ds \right), \tag{2.17}$$

for all $u \in [0,1]$ showing (2.3). □

3 Affine forward order flow intensity models

We now introduce a class of models for market order flow, which are structurally similar to the forward variance models. These models consist of a log-price X and a *forward intensity process* $\xi_t(T)$, which models the expectation (at time t) of the future intensity of order flow (at time T). The forward intensity $\xi_t(T)$ has a role similar to the forward variance, and we call the resulting model a *affine forward order flow intensity* (AFI) model.[1] The AFI model is driven purely by the arrival of market orders, which are represented by two independent pure-jump semimartingales J_t^+, J_t^- of finite activity and with intensity λ_{t-}. Both processes jump only upwards and represent the arrival of buy and sell orders respectively. For simplicity, we assume that the distribution of buy and sell orders is the same and given by a probability measure $\zeta(dx)$ on $\mathbb{R}_{\geqslant 0}$. We assume that $\int_0^\infty e^x \zeta(dx) < \infty$; in particular, also the first moment $\int_0^\infty x\zeta(dx)$ exists. In addition, we assume that the order flow processes are self-exciting, in the sense that each arriving order positively impacts the intensity process. This impact can be asymmetric, i.e. the degree of self-excitement may be different for buy-

[1]The strong empirical correlation between order volume (as a proxy for intensity) and return variance is well-documented in the literature (see e.g. [GRT92]). Therefore the parallels between AFV and AFI models should not come as a complete surprise.

and sell-orders. Together this leads to the specification of the AFI model as

$$dX_t = -\lambda_{t-} m_X dt + dJ_t^+ - dJ_t^-, \tag{3.1a}$$

$$d\xi_t(T) = \kappa(T-t)\left(\frac{1}{\gamma^+} d\widetilde{J}_t^+ + \frac{1}{\gamma^-} d\widetilde{J}_t^-\right). \tag{3.1b}$$

where $\kappa : \mathbb{R}_{\geqslant 0} \to \mathbb{R}_{\geqslant 0}$ is an integrable, decreasing non-zero function ('kernel'), $\gamma^\pm$ are positive constants, m_X is determined by the martingale condition on $S = e^X$ and $\widetilde{J}_t^\pm$ denote the *compensated* order flow processes, i.e. $\widetilde{J}_t^\pm := J_t^\pm - m_\zeta \int_0^t \lambda_{s-} ds$, where

$$m_\zeta = (\gamma_+ + \gamma_-) \int_0^\infty x \, \zeta(dx).$$

Finally, $\xi_t(u)$ is linked to λ_t by

$$\xi_t(u) = \mathbb{E}\left[\lambda_u | \mathcal{F}_t\right]. \tag{3.2}$$

Since $(\mathcal{F}_t)_{t\geqslant 0}$ is right-continuous, we have $\lambda_t = \lim_{u\downarrow t} \xi_t(u)$. Setting

$$J_t^X = J_t^+ - J_t^-, \qquad \widetilde{J}_t^\lambda = \frac{1}{\gamma^+}\widetilde{J}_t^+ + \frac{1}{\gamma^-}\widetilde{J}_t^- \tag{3.3}$$

we can rewrite (3.1) as

$$dX_t = -\lambda_{t-} m_X dt + dJ_t^X,$$
$$d\xi_t(T) = \kappa(T-t) d\widetilde{J}_t^\lambda.$$

Note that compensated jump processes are local martingales, such that also $\xi_t(u)$ is (at least locally) a martingale, which is consistent with (3.2).

We now discuss the jump processes and the compensators of their random jump measures in more detail. Recall that we have assumed the same order size distribution $\zeta(dx)$ for both buy and sell orders. Hence the random jumps of $J^\pm$ are compensated by

$$d\nu_t^\pm(dx) = \lambda_{t-}\zeta(\pm dx)dt,$$

where x represents jump size. While J^+ and J^- are independent, it is important to note that J_t^X and J_t^λ are not. Instead, they move by simultaneous jumps. Thus, the predictable compensator of the jump measure of (J^X, J^λ) is given by

$$d\nu_t^{(X,\lambda)}(dx, dy) = \lambda_{t-}\chi(dx, dy)dt, \qquad \text{where}$$
$$\chi(dx, dy) = \Big(\mathbf{1}_{\{x\geqslant 0\}}\mathbf{1}_{\{x=\gamma_+ y\}}\zeta(dx) + \mathbf{1}_{\{x\leqslant 0\}}\mathbf{1}_{\{x=-\gamma_- y\}}\zeta(-dx)\Big).$$

Note that the measure of joint jump heights $\chi(dx, dy)$ is concentrated on the line segments $x = \gamma_+ y$, $(x \geqslant 0)$ and $x = -\gamma_- y$, $(x \leqslant 0)$ due to the simultaneity of jumps. In addition, we define

$$\psi(u) = \int_0^\infty (e^{ux} - 1)\, \zeta(dx) \tag{3.4}$$

and calculate

$$\int_{\mathbb{R}\times\mathbb{R}_{\geqslant 0}} \left(e^{ux+wy}-1\right)\chi(dx,dy) = \psi\big(u+w\gamma^{+}\big)+\psi\big(-u+w\gamma^{-}\big).$$

Applying Itô's formula for jump processes to e^X it is easy to see that the martingale condition implies that

$$m_X = \psi(1)+\psi(-1).$$

The following theorem is the analogue of Theorem 2.5 and shows the structural similarity between affine forward variance models and AFI models.

Theorem 3.1. *The AFI model* (3.1) *has an affine CGF in the sense of Definition 2.2. Moreover,* $g(.;u):\mathbb{R}_{\geqslant 0}\to\mathbb{R}_{\leqslant 0}$ *in* (2.3) *is the unique global solution of the generalized convolution Riccati equation*

$$g(t;u) = R_\lambda\Big(u, \int_0^t \kappa(t-s)g(s;u)ds\Big) = R_\lambda\Big(u,(\kappa\star g)(t;u)\Big), \tag{3.5}$$

where

$$R_\lambda(u,w) = \psi\big(u+w\gamma^{+}\big)+\psi\big(-u+w\gamma^{-}\big)-um_X-wm_\zeta, \tag{3.6}$$

with ψ *as in* (3.4).

Proof. Essentially, we proceed as in the second part of the proof of Theorem 2.5. Let G be defined as in (2.12) and set $M_t=\exp(uX_t+G_t)$. Applying the same argument as in the proof of Lemma 2.10, but replacing Brownian motion by the pure-jump-martingale $\widetilde{J}^X$ we obtain

$$G_t = \int_0^T g(T-s;u)\xi_0(s)ds - \int_0^t g(T-s;u)\lambda_{s-}ds + \int_0^t v_s(T;u)d\widetilde{J}^X_s,$$

where

$$v_t(T;u) = \int_t^T \kappa(r-t)g(T-r;u)dr.$$

Applying the Itô-formula with jumps to M we obtain

$$M_t = M_0 + \int_0^t M_{s-}\left(udX_t+dG_t\right) + \sum_{0\leqslant s\leqslant t} M_{s-}\left(e^{u\Delta X_s+\Delta G_s}-1-u\Delta X_s-\Delta G_s\right)$$

and compensating the jumps yields

$$\begin{aligned}\frac{dM_t}{M_t} &= \text{loc. mg.} - \lambda_{t-}(um_X+v_t(T;u)m_\zeta)dt - g(T-t;u)\lambda_{t-}dt+ \\ &+\lambda_{t-}\int_{\mathbb{R}\times\mathbb{R}_{\geqslant 0}}\Big(e^{ux+yv_t(T;u)}-1\Big)\chi(dx,dy)dt,\end{aligned} \tag{3.7}$$

where 'loc. mg.' denotes a local martingale part that we need not compute explicitly. We see that the dt-terms vanish, if

$$g(\tau; u) = R_\lambda\Big(u, \int_0^\tau \kappa(\tau - s) g(s; u) ds\Big),$$

i.e. if the generalized convolution Riccati equation (3.5) has a solution for $0 \leqslant \tau \leqslant T - t$.

To show that there exists a unique global solution of (3.5), we set $H_u(w) = R_\lambda(u, w), u \in (0, 1)$ and show that H_u satisfies the conditions of Corollary A.7. In particular, for all $u \in (0, 1)$,

- $H_u(w)$ is a finite, strictly convex function on $(-\infty, 0]$ and satisfies $H_u(0) < 0$;
- $H_u(w)$ has a single root $H_u(w_*(u)) = 0$ in $(-\infty, 0]$.

Indeed, note that strict convexity is inherited from ψ, cf. (3.4). In addition, convexity of the exponential function implies, for $u \in (0, 1)$, that

$$e^{ux} = e^{u \cdot x + (1-u) \cdot 0} \leqslant u e^x + (1 - u) e^0 < u e^x + 1,$$

and hence that

$$H_u(0) = \int_0^\infty \left(e^{ux} - 1 - u e^x\right) \zeta(dx) + \int_0^\infty \left(e^{-ux} - 1 - u e^{-x}\right) \zeta(dx) < 0.$$

Finally, the existence of the root $w_*(u)$ follows from the fact that

$$\lim_{w \to -\infty} \int_0^\infty \left(e^{(\pm u + \gamma^\pm w)x} - 1 - \gamma^\pm w x\right) \zeta(dx) = +\infty,$$

which implies that $\lim_{w \to -\infty} H_u(w) = +\infty$. In summary, H_u satisfies all conditions of Cor. A.7 and we conclude the existence of a unique global solution $g(t; u)$ of the Riccati equation for all $u \in (0, 1)$. Moreover, $g(t; u) \leqslant 0$ for all $(t, u) \in \mathbb{R}_{\geqslant 0} \times (0, 1)$ by estimate (A.11). We can add the boundary cases $u \in \{0, 1\}$, observing that they yield the constant global solution $g(t; u) \equiv 0$, which must be unique by [GLS90, Thm. 13.1.2]. From (3.7) we conclude that $M_t = \exp(uX_t + G_t)$ is a local martingale. From $g(t, u) \leqslant 0$ it follows that M is bounded and hence a true martingale. By the same argument as in (2.17) this shows the affine CGF (2.3). □

Example 3.2 (The bivariate Hawkes process of [EER16]). Consider (3.1a), driven by a bivariate Hawkes process (J^+, J^-) with unit jump size (i.e., $\zeta(dx) = \delta_1(dx)$), common kernel ϕ, and common intensity λ_t, given by

$$\lambda_t = \mu + \int_0^t \phi(t - s) \left(\frac{1}{\gamma^+} dJ_s^+ + \frac{1}{\gamma^-} dJ_s^+\right),$$

as in [EER16, Sec. 2]. Assume that $\frac{1}{\gamma^+} + \frac{1}{\gamma^-} = 1$ and that ϕ satisfies the stability condition $\int_0^\infty \phi(s)ds < 1$.[2] Then the kernel ϕ has an 'inverse kernel' κ (cf. [BMM15, Def. 2]), also called 'resolvent' of ϕ (cf. [GLS90, Ch. 2], [JLP17, Sec. 2]), which is defined as solution of the functional equation

$$\kappa \star \phi = \kappa - \phi. \tag{3.8}$$

In terms of this inverse kernel κ, the Hawkes intensity λ has the martingale representation (cf. [BMM15, Eq. (45)])

$$\lambda_t = \mu + \mu \int_0^t \kappa(t-u)du + \int_0^t \kappa(t-u)d\tilde{J}^\lambda_u,$$

with $\tilde{J}^\lambda$ as in (3.3). Taking conditional expectations and using the martingale property of $\tilde{J}^\lambda$ yields

$$\mathbb{E}\left[\lambda_T \middle| \mathcal{F}_t\right] = \mu + \mu \int_0^T \kappa(T-u)du + \int_0^t \kappa(T-u)d\tilde{J}^\lambda_u,$$

and hence

$$d\xi_t(T) = d\mathbb{E}\left[\lambda_T \middle| \mathcal{F}_t\right] = \kappa(T-t)d\tilde{J}^\lambda_t,$$

which shows that the model can be cast as AFI model with kernel κ. Denoting by $\widehat{\phi}, \widehat{\kappa}$ the Laplace transforms of ϕ, κ, the resolvent equation (3.8) becomes

$$\widehat{\kappa} \cdot \widehat{\phi} = \widehat{\kappa} - \widehat{\phi}, \tag{3.9}$$

after taking Laplace transforms. For concrete specifications of ϕ this allows to find the corresponding κ. Consider, for example

$$\phi(x) = \zeta e^{-(\lambda+\zeta)x} \qquad \text{with Laplace tf.} \qquad \widehat{\phi}(z) = \frac{\zeta}{1+\lambda+\zeta}. \tag{3.10}$$

Tor this ϕ we obtain from (3.9) the 'inverse kernel'

$$\kappa(x) = \zeta e^{-\lambda x} \qquad \text{with Laplace tf.} \qquad \widehat{\kappa}(z) = \frac{\zeta}{1+\lambda},$$

i.e., the kernel of the Heston model in forward variance form; see Example 2.8. Furthermore, the Hawkes kernel

$$\phi(x) = \zeta x^{\alpha-1} E_{\alpha,\alpha}(-(\lambda+\zeta)x^\alpha) \tag{3.11}$$

has Laplace transform $\widehat{\phi}(z) = \zeta/(z^\alpha + \lambda + \zeta)$ (cf. [HMS11, Eq. (7.5)]). Thus its resolvent is

$$\kappa(x) = \zeta x^{\alpha-1} E_{\alpha,\alpha}(-\lambda x^\alpha), \tag{3.12}$$

the kernel of the *rough* Heston model in forward variance form; see Example 2.9.

[2] The stability condition ensures stationarity of the process (J^+, J^-), cf. [BMM15].

4 High-frequency limit of the AFI model

We proceed to show that the AFV model is the high-frequency limit of the AFI model. This limit is closely related to the limits of 'nearly unstable' Hawkes processes considered in [JR15, JR16, EER16], see Example 4.3 below.

4.1 A first convergence result

We introduce a small parameter ϵ and rescale(3.1) as

$$dX_t^\epsilon = -\lambda_{t-}^\epsilon m_X dt + dJ_t^{\epsilon,+} - dJ_t^{\epsilon,-}, \tag{4.1a}$$

$$d\xi_t^\epsilon(T) = \kappa^\epsilon(T-t)\left(\frac{1}{\gamma^+}d\widetilde{J}_t^{\epsilon,+} + \frac{1}{\gamma^-}d\widetilde{J}^{\epsilon,-}\right), \tag{4.1b}$$

where $J^{\epsilon,\pm}$ are independent pure jump semimartingales with (conditional) intensity

$$\lambda_{t-}^\epsilon = \tfrac{1}{\epsilon}\lambda_{t-}, \qquad \xi_t(T) = \mathbb{E}\left[\lambda_T^\epsilon \middle| \mathcal{F}_t\right] = \tfrac{1}{\epsilon}\xi_t(T),$$

and jump height distribution

$$\zeta^{\epsilon,\pm}(dx) = \zeta(\pm dx/\sqrt{\epsilon}).$$

Thus, as $\epsilon \downarrow 0$, the frequency of jumps increases proportional to $1/\epsilon$, while the size of jumps shrinks proportional to $\sqrt{\epsilon}$. The kernel is scaled as

$$\kappa^\epsilon(x) = \frac{1}{\epsilon}\kappa(x),$$

and the initial conditions of (4.1) are given by $X_0^\epsilon = X_0$ and $\xi_0^\epsilon(T) = \frac{1}{\epsilon}\xi_0(T)$. Under the given scaling, the quantities from (3.4) and below transform as

$$\begin{aligned}\psi^\epsilon(u) &= \psi(\sqrt{\epsilon}u)\\ m_X^\epsilon &= \psi(\sqrt{\epsilon}) + \psi(-\sqrt{\epsilon})\\ m_\zeta^\epsilon &= \sqrt{\epsilon}m_\zeta\end{aligned}$$

and we write

$$R^\epsilon(u,w) = \psi^\epsilon\big(u + w\gamma^{\epsilon,+}\big) + \psi^\epsilon\big(-u + w\gamma^{\epsilon,-}\big) - um_X^\epsilon - wm_\zeta^\epsilon.$$

Lemma 4.1. *Given $\gamma^\pm > 0$ and the jump height distribution $\zeta(dx)$, set*

$$|\gamma| = \sqrt{(\gamma^+)^2 + (\gamma^-)^2} \qquad c = |\gamma|\sqrt{\int_0^\infty x^2\zeta(dx)}$$

$$a = \sqrt{2\int_0^\infty x^2\zeta(dx)}, \qquad \rho = \frac{\gamma^+ - \gamma^-}{\sqrt{2}|\gamma|}.$$

Then

$$\lim_{\epsilon\to 0}\frac{1}{\epsilon}R^\epsilon(u,w) = \frac{a^2}{2}(u^2-u) + ac\rho uw + \frac{c^2}{2}w^2 = R_V(u,cw)$$

with $R_V(u,w)$ as in (2.8). Moreover, also the partial derivatives with respect to u and w converge, i.e.

$$\lim_{\epsilon\to 0}\frac{1}{\epsilon}\frac{\partial R^\epsilon}{\partial u}(u,w) = \frac{\partial R_V}{\partial u}(u,cw) = \frac{a^2}{2}(2u-1)+ac\rho w$$
$$\lim_{\epsilon\to 0}\frac{1}{\epsilon}\frac{\partial R^\epsilon}{\partial w}(u,w) = \frac{\partial R_V}{\partial w}(u,cw) = cw + a\rho u.$$

Proof. We can write

$$R^\epsilon(u,w) = \int_0^\infty b^\epsilon(x;u,w)\zeta(dx) \tag{4.2}$$

where

$$b^\epsilon(x;u,w) = -u(e^{\sqrt{\epsilon}x}+e^{-\sqrt{\epsilon}x}) - w(\gamma^+ + \gamma^-)\sqrt{\epsilon}x + \\ + \exp\Big((u+w\gamma^+)\sqrt{\epsilon}x\Big) - 1 + \exp\Big((-u+w\gamma^-)\sqrt{\epsilon}x\Big) - 1.$$

Expanding in powers of $\sqrt{\epsilon}x$ yields

$$b^\epsilon(x;u,w) = \epsilon x^2\left((u^2-u) + uw(\gamma^+ - \gamma^-) + \frac{w^2}{2}\left((\gamma^+)^2) + (\gamma^-)^2\right)\right) + \mathcal{O}(\epsilon^{3/2}x^3).$$

Hence,

$$\lim_{\epsilon\to 0}\frac{1}{\epsilon}R^\epsilon(u,w) = \int_0^\infty x^2\zeta(dx)\cdot\left(u^2 - u + \sqrt{2}|\gamma|\rho uw + \frac{w^2}{2}|\gamma|^2\right) = R_V(u,cw),$$

where exchanging limit and integral is justified by dominated convergence and the integrability condition $\int_0^\infty e^x\zeta dx < \infty$.

To show the convergence of partial derivatives, we take partial derivatives in (4.2) to obtain

$$\frac{\partial R^\epsilon}{\partial u}(u,w) = \int_0^\infty \frac{\partial b^\epsilon}{\partial u}(x;u,w)\zeta(dx).$$

Since R^ϵ is convex, its difference quotients converges monotonically, and monotone convergence can be used to exchange derivative and integral. Expanding $\frac{\partial b^\epsilon}{\partial u}(x;u,w)$ in powers of $\sqrt{\epsilon}x$, a direct calculation yields the desired limit. The proof for the $\frac{\partial}{\partial w}$-derivative is analogous. □

Combining Lemma 4.1 with Theorem 2.5 and 3.1 yields a first distributional convergence result.

Proposition 4.2. *Let $(X^\epsilon, \xi^\epsilon)$ be the rescaled AFI model (4.1). Define a, ρ, c as in Lemma 4.1 and set $\kappa_V(x) = c\kappa(x)$. Then, for any $t \geqslant 0$,*

$$X_t^\epsilon \xrightarrow{\epsilon\to 0} X_t \qquad \textit{in distribution,} \tag{4.3}$$

where (X,ξ) is a forward variance model with parameters a and ρ, and kernel κ_V.

Proof. By Theorem 3.1, $g^\epsilon(t;u)$ in the CGF (2.3) of X^ϵ is the unique global solution of the generalized convolution Riccati equation (3.5) and hence satisfies

$$\tfrac{1}{\epsilon}g^\epsilon(t;u) = \tfrac{1}{\epsilon}R^\epsilon\Big(u, \kappa^\epsilon \star g^\epsilon(t;u)\Big) = \tfrac{1}{\epsilon}R^\epsilon\Big(u, \kappa \star \left(\tfrac{1}{\epsilon}g^\epsilon\right)(t;u)\Big). \tag{4.4}$$

Note that $\frac{1}{\epsilon}R^\epsilon(u,w)$ is jointly continuous in all variables, and by Lemma 4.1 converges to $R_V(u,cw)$ as $\epsilon \to 0$. By Corollary A.7 (4.4) can be transformed into a non-linear Volterra equation of type (A.6), whose solution depends jointly continuous on (t,ϵ,u) by [GLS90, Thm. 13.1.1]. We conclude that $\frac{1}{\epsilon}g^\epsilon(t;u)$ converges, uniformly for (t,u) in compacts, to $g(t;u)$ as $\epsilon \to 0$, where $g(t;u)$ is the unique solution (cf. Theorem 2.5) of

$$g(t;u) = R_V\Big(u, c\kappa \star g(t;u)\Big) = R_V\Big(u, \kappa_V \star g(t;u)\Big).$$

Using Theorems 2.5 and 3.1, we conclude that

$$\mathbb{E}\left[e^{uX_t^\epsilon}\right] = \exp\left(uX_0 + \int_0^t g^\epsilon(t-s;u)\xi_0^\epsilon(s)ds\right) \to$$
$$\exp\left(uX_0 + \int_0^t g(t-s;u)\xi_0(s)ds\right) = \mathbb{E}\left[e^{uX_t}\right],$$

as $\epsilon \to 0$, i.e., the moment generating function of X_t^ϵ converges to the moment generating function of X_t on $u \in [0,1]$. By [Bil86, Prob. 30.4], convergence of moment generating functions on a (non-empty) interval implies the convergence in distribution in (4.12). ☐

The following example shows that the scaling in (4.1) is related to the 'nearly unstable' limit of Hawkes models in [EER16].

Example 4.3 (Nearly unstable limit of bivariate Hawkes processes). We continue Example 3.2 and consider the bivariate Hawkes process from [EER16] with Mittag-Leffler kernel (3.11). Introduce a small parameter ϵ and scale the kernel as

$$\phi_\epsilon(x) = \tfrac{\zeta}{\epsilon}x^{\alpha-1}E_{\alpha,\alpha}(-(\lambda + \tfrac{\zeta}{\epsilon})x^\alpha).$$

In terms of its Laplace transform, this scaling becomes $\widehat{\phi}_\epsilon(z) = \zeta/(\epsilon(z^\alpha+\lambda)+\zeta)$. In particular, we have

$$\int_0^\infty \phi_\epsilon(x)dx = \widehat{\phi}(0) = \frac{\zeta}{\epsilon\lambda+\zeta} \to 1,$$

i.e. as $\epsilon \to 0$ the stability condition of the Hawkes process approaches the critical value 1 (hence 'nearly unstable'). From (3.9), the Laplace transform of the resolvent kernel $\kappa_\epsilon(x)$ can be determined as

$$\widehat{\kappa}_\epsilon(z) = \frac{\widehat{\phi}_\epsilon(z)}{1-\widehat{\phi}_\epsilon(z)} = \frac{\zeta/\epsilon}{z^\alpha+\lambda}$$

and thus the resolvent kernel is given by

$$\kappa_\epsilon(x) = \tfrac{\zeta}{\epsilon} x^{\alpha-1} E_{\alpha,\alpha}(-\lambda x^\alpha) = \tfrac{1}{\epsilon}\kappa(x).$$

Together with square-root scaling of the jump size we are exactly in the setting of (4.1) and conclude from Proposition 4.2 that the (univariate) marginal distributions of X converge to those of the corresponding AFV model, i.e. the rough Heston model (cf. Example 2.9). Theorem 4.9 below strengthens this result to convergence of all finite-dimensional marginal distributions. $\diamond$

4.2 The joint moment generating function

In this subsection, we derive results on the joint moment generating function of log-price and forward variance and of the finite-dimensional marginal distributions of X.

Assumption 4.4. We assume that (X, ξ) is either an AFV model (2.1) or an AFI model (3.1), and we write $R(u, w)$ for the corresponding function $R_V(u, w)$ or $R_\lambda(u, w)$. In addition we denote, for any $u \in (0, 1)$, by $w_*(u)$ the unique root where

$$R(u, w_*(u)) = 0, \quad \text{and} \quad w_*(u) < 0.$$

Note that the function $R(u, w)$ has already been studied in the context of affine stochastic volatility models in [KR11, Lem. 3.2ff]. In particular, we note that $R(u, w)$ and $w_*(u)$ are convex functions for $u \in [0, 1], w \leqslant 0$ and continuously differentiable on the interior of their domain.

Proposition 4.5. *Let (X, ξ) be an AFV or an AFI model and let $R(u, w)$ and $w_*(u)$ be defined as in Assumption 4.4. Let $\Delta > 0$, $T' = T + \Delta$, and let h be a piecewise continuous $\mathbb{R}_{\leqslant 0}$-valued function on $[0, \Delta]$, such that $w_*(u) < \int_0^\Delta \kappa(\Delta - s)h(s)ds$. Then*

$$\mathbb{E}\left[\exp\left(uX_T + \int_T^{T'} h(T'-s)\xi_T(s)ds\right)\middle|\mathcal{F}_t\right] = \\ = \exp\left(uX_t + \int_t^{T'} g(T'-s; u, h)\xi_t(s)ds\right), \quad (4.5)$$

where $g(.; u, h) : \mathbb{R}_{\geqslant 0} \to \mathbb{R}_{\leqslant 0}$ is the unique solution of the (generalized) convolution Riccati equation

$$g(t; u, h) = R\Big(u, \int_0^t \kappa(t-s)g(s; u, h)ds\Big), \quad t \in [\Delta, \infty) \quad (4.6)$$

with initial condition

$$g(t; u, h) = h(t), \qquad t \in [0, \Delta). \quad (4.7)$$

Remark 4.6. Note that the expression (4.5) for the joint moment generating function does not correspond to the exponential-affine transform formula (4.6) of [JLP17]. Specifically, h constant in (4.5) would give the joint moment generation function of X_T and the forward variance swap $\int_T^{T'} \xi_T(s)\, ds$. In contrast, f constant in (4.6) of [JLP17] would give the the joint moment generation function of X_T and quadratic variation $\int_0^T V_s\, ds$.

Proof. The existence of a unique, $\mathbb{R}_{\leqslant 0}$-valued solution to (4.6) with initial condition (4.7) follows from an application of Corollary A.8 with $H_u(w) = R(u, w)$. In the proofs of Theorem 2.5 and Theorem 3.1, we have already established that H_u satisfies the necessary conditions to apply the corollary. Next, we define $G_t^\Delta = \int_t^{T'} g(T' - s; u, h)\xi_t(s)ds$ and specialize to the forward variance case. By Lemma 2.10, it holds that

$$dG_t^\Delta = -g(T' - t; u, h)V_t dt + \left(\int_t^{T'} g(T' - r; u, h)\kappa(r - t)dr\right) V_t dW_t.$$

Applying Itô's formula to $M_t^\Delta = \exp(uX_t + G_t^\Delta)$, as in the proof of Theorem 2.5, we see that M_t^Δ is a local martingale on $[0, T]$ if

$$g(T' - t; u, h) = R\left(u, \int_t^{T'} g(T' - r; u, h)\kappa(r - t)dr\right).$$

Setting $\tau = T' - t \in [\Delta, T']$ this is exactly (4.6). We conclude that M_t^Δ is a local martingale on $[0, T]$, and – being bounded – even a true martingale. Using the initial condition (4.7), we observe that

$$\mathbb{E}\left[\exp\left(uX_T + \int_T^{T'} h(T' - s)\xi_T(s)ds\right)\middle|\mathcal{F}_t\right] = \mathbb{E}\left[M_T^\Delta\middle|\mathcal{F}_t\right] =$$
$$= M_t = \exp\left(uX_t + \int_t^{T'} g(T' - s; u, h)\xi_t(s)ds\right),$$

showing (2.3). The proof in the AFI case is analogous with the following modifications: W_t has to be substituted by the pure-jump martingale $\tilde{J}_t^X$ and V_t by the intensity λ_{t-}. Itô's formula for jump processes can then be applied as in the proof of Theorem 3.1. □

Proposition 4.7. *Let (X, ξ) be an AFV or an AFI model and let $R(u, w)$ and $w_*(u)$ be defined as in Assumption 4.4. Let $t_0 \leqslant t_1 \leqslant \cdots t_n = T$ and $u = (u_0, \ldots, u_{n-1}) \in (0, 1)^n$ be such that $w_*(u_0) \leqslant w_*(u_2) \leqslant \cdots w_*(u_{n-1})$. Then, for all $k \in \{0, \ldots, n-1\}$,*

$$\mathbb{E}\left[\exp\left(u_k(X_{t_{k+1}} - X_{t_k}) + \cdots + u_{n-1}(X_T - X_{t_{n-1}})\right)\middle|\mathcal{F}_{t_k}\right] =$$
$$= \exp\left(\int_{t_k}^T g_k(T - s; u)\xi_{t_k}(s)ds\right), \quad (4.8)$$

where the functions g_k are recursively defined as the solutions of the convolution Riccati equations

$$g_k(t;u) = R\Big(u_k, \int_0^t \kappa(t-s) g_k(s;u)ds\Big), \quad t \in [T - t_{k+1}, T - t_k) \tag{4.9}$$

with initial conditions

$$g_k(t;u) = g_{k+1}(t;u), \qquad t \in [0, T - t_{k+1}). \tag{4.10}$$

Remark 4.8. Note that for $k = n-1$ equation (4.9) becomes a (generalized) convolution Riccati equation *without* initial condition and (4.10) becomes void (i.e. a condition over an empty set).

Proof. We show the result by backward induction on k: For $k = n-1$ the proposition is equivalent to Theorem 2.5, when (X,ξ) is an AFV model, and to Theorem 3.1, when (X,ξ) is an AFI model. Setting $\Delta_k := T - t_k$, we obtain from (A.15) in Corollary A.8 that

$$w_*(u_{n-1}) < \int_0^{\Delta_{n-1}} \kappa(\Delta_{n-1} - s) g_{n-1}(s;u)ds. \tag{4.11}$$

For the induction step assume that (4.8) has been shown for a certain k and that (4.11) holds with $n-1$ replaced by k. Writing

$$Z_{k-1} := \exp\big(u_{k-1}(X_{t_k} - X_{t_{k-1}}) + \cdots + u_n(X_T - X_{t_{n-1}})\big)$$

and applying the tower law of conditional expectations, we have

$$\begin{aligned} \mathbb{E}\left[Z_{k-1} \middle| \mathcal{F}_{t_{k-1}}\right] &= \mathbb{E}\left[\exp\big(u_{k-1}(X_{t_k} - X_{t_{k-1}})\big) \cdot \mathbb{E}\left[Z_k \middle| \mathcal{F}_{t_k}\right] \middle| \mathcal{F}_{t_{k-1}}\right] = \\ &= \mathbb{E}\left[\exp\left(u_{k-1}(X_{t_k} - X_{t_{k-1}}) + \int_{t_k}^T g_k(T-s;u)\xi_{t_k}(s)ds \right) \middle| \mathcal{F}_{t_{k-1}} \right]. \end{aligned}$$

Since

$$w_*(u_{k-1}) \leqslant w_*(u_k) < \int_0^{\Delta_k} \kappa(\Delta_k - s) g_k(s;u)ds$$

we may apply Proposition 4.5 with Δ_k and obtain (4.8) with g_{k-1} as solution of (4.9) with initial condition (4.10). Finally, (4.11) holds with $n-1$ replaced by $k-1$, using the estimate (A.15) from Corollary A.8. □

4.3 Convergence of finite-dimensional marginal distributions

Theorem 4.9. *Let $(X^\epsilon, \xi^\epsilon)$ be the rescaled AFI model (4.1), define a, ρ, c as in Lemma 4.1 and set $\kappa_V(x) = c\kappa(x)$. Then, for any $n \in \mathbb{N}$ and $0 = t_0 \leqslant t_1 \leqslant \cdots t_n = T$,*

$$\big(X^\epsilon_{t_0}, \ldots, X^\epsilon_{t_n}\big) \xrightarrow{\epsilon \to 0} (X_{t_0}, \ldots, X_{t_n}) \qquad \textit{in distribution}, \tag{4.12}$$

where (X,ξ) is the AFV model with parameters a, ρ, c and kernel κ_V.

Proof. By Lemma 4.1 $\frac{1}{\epsilon}R^\epsilon(u,w)$ converges to $R_V(u,cw)$ and the same holds true for the partial derivatives with respect to u and w. Therefore, by the implicit function theorem, also $w_*^\epsilon(u)$ and $\frac{\partial}{\partial u}w_*^\epsilon(u)$ converge to $\frac{1}{c}w_*(u)$ and $\frac{1}{c}w_*'(u)$ as $\epsilon \to 0$ for all $u \in (0,1)$. Moreover, since the w_*^ϵ are convex functions of u, the convergence is uniform on compacts (cf. [Roc70, Thm. 10.8]). The limit $\frac{1}{c}w_*(u)$ can be calculated explicitly and is given by

$$\tfrac{1}{c}w_*(u) = \tfrac{a}{c}\left(-\rho u + \sqrt{\rho^2 u^2 + (u-u^2)}\right).$$

It is easy to see that w_* is decreasing on $(0,u_*)$ and increasing on $(u_*,1)$, where

$$u_* := \begin{cases} \frac{1}{2}\frac{1-|\rho|}{1-\rho^2} & \text{if } \rho \in (0,1) \\ \frac{1}{4} & \text{if } |\rho| = 1. \end{cases}$$

We conclude that there is $N \in \mathbb{N}$ and a closed interval $I \subset (0,u_*)$ with non-empty interior, such that $u \mapsto w_*(u)$ and $u \mapsto w_*^\epsilon(u)$ are decreasing on I for all $\epsilon \leqslant 1/N$. Introduce the set

$$D := \{u \in I^n : u_0 \geqslant u_2 \geqslant \cdots \geqslant u_{n-1}\} \subset (0,1)^n$$

and note that also D is closed with non-empty interior. In addition, $w_*^\epsilon(u_0) \leqslant w_*^\epsilon(u_2) \leqslant \cdots w_*^\epsilon(u_{n-1})$ for all $u = (u_0,\ldots,u_{n-1}) \in D$ and $\epsilon \leqslant 1/N$, and the same holds for w_*. From Proposition 4.7 we conclude that the joint moment generating function of the increments $(X_{t_1}^\epsilon - X_{t_0}^\epsilon, X_{t_2}^\epsilon - X_{t_1}^\epsilon, \ldots, X_{t_n}^\epsilon - X_{t_{n-1}}^\epsilon)$ is of the form

$$\begin{aligned} Z^\epsilon(u) &:= \mathbb{E}\left[\exp\left(u_0(X_{t_1}^\epsilon - X_{t_0}^\epsilon) + \cdots + u_{n-1}(X_T^\epsilon - X_{t_{n-1}}^\epsilon)\right)\right] = \\ &= \exp\left(\int_0^T g_0^\epsilon(T-s;u)\xi_{t_0}(s)ds\right), \end{aligned}$$

for any $u \in D$ and $\epsilon \leqslant 1/N$, where g_0^ϵ satisfies the iterated Riccati convolution equations (4.9) with $R(u,w) = R^\epsilon(u,w)$. By Corollary A.8 each of these equations can be transformed into a non-linear Volterra equation, whose solution depends continuously on (ϵ,t,u) by [GLS90, Thm. 13.1.1]. In addition, Lemma 4.1 yields the convergence $\frac{1}{\epsilon}R^\epsilon(u,w) \to R_V(u,cw)$. Hence we conclude – as in the proof of Proposition 4.2 – that $\frac{1}{\epsilon}g_0^\epsilon(t;u)$ converges, uniformly for (t,u) in compacts, to $g_0(t;u)$ as $\epsilon \to 0$, where $g_0(t;u)$ is the unique solution of the iterated Riccati convolution equations (4.9) with $R(u,w) = R_V(u,cw)$. Consider now the joint moment generating function $Z(u)$ of the increments $(X_{t_1} - X_{t_0}, X_{t_2} - X_{t_1}, \ldots, X_{t_n} - X_{t_{n-1}})$ of the AFV model with parameters a,ρ and kernel $\kappa_V = c\kappa$. The convergence $\frac{1}{\epsilon}g_0^\epsilon(t;u) \to g_0(t;u)$ together with Proposition 4.7 yields

$$Z^\epsilon(u) = \exp\left(\int_0^T g_0^\epsilon(T-s)\xi_0^\epsilon(s)ds\right) \to \exp\left(\int_0^T g_0(T-s;u)\xi_0(s)ds\right) = Z(u)$$

for all $u \in D$. By [Bil86, Thm. 29.4 and Prob. 30.4] convergence of the moment generating function on a set with non-empty interior implies convergence in distribution, and (4.12) follows. □

5 Summary and Conclusions

Starting from a generic formulation in forward variance form of a stochastic volatility model, we have given necessary and sufficient conditions for such a model to have an affine cumulant generating function (CGF). In addition, we have shown that this CGF can be expressed in terms of the unique global solution of a convolution Riccati equation. We have introduced the class of affine forward order flow intensity (AFI) models and have shown that these model also have an affine CGF, which can be expressed in terms of the unique global solution of a *generalized* convolution Riccati equation. We have further shown that affine forward variance models can be obtained as the high-frequency limit of appropriately rescaled AFI models. Finally, we have computed joint moment generating functions of the terminal log-stock price and forward variance curve in AFV and AFI models.

A Some results on Volterra equations with convex non-linearity

We show some results on Volterra equations with convex non-linearity, of the type appearing in Theorem 2.5 and 3.1. On the non-linearity we impose the following assumptions:

Assumption A.1. The function $H : (-\infty, w_{\max}] \to \mathbb{R}$ is continuously differentiable and convex with a unique root $H(w_*) = 0$ in $(-\infty, w_{\max}]$. Moreover, $H'(w_*) < 0$ and $H(w_{\max}) < 0$.

For a function H satisfying Assumption A.1, we set

$$w_0 = \operatorname{argmin}_{w \in (-\infty, w_{\max}]} H(w);$$

if the minimum is not unique (i.e., if H has a flat part), then w_0 shall denote the leftmost minimizer. Note that either

- $w_0 = w_{\max}$, in which case H is strictly decreasing on $(-\infty, w_{\max}]$; or
- $w_0 < w_{\max}$, in which case H is strictly decreasing on $(-\infty, w_0)$ and increasing on $[w_0, w_{\max}]$.

In any case, $w_* < w_0 \leqslant w_{\max}$ holds true. Also the following definition will be useful:

Definition A.2. Let H be a function satisfying Assumption A.1. The *decreasing envelope* of H is defined as

$$\overline{H} := \begin{cases} H(w), & w \leqslant w_0 \\ H(w_0), & w \in [w_0, w_{\max}] \end{cases}. \tag{A.1}$$

Clearly $\overline{H}$ also satisfies Assumption A.1, but is in addition decreasing and satisfies $\overline{H} \leqslant H$. Both Assumption A.1 and Definition A.2 are illustrated in Figure 1.

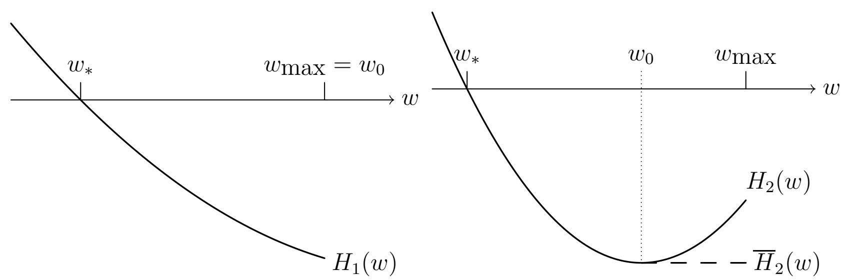

Figure 1: Illustration of two convex functions H_1, H_2 satisfying Assumption A.1. While H_1 is monotone decreasing, H_2 is not, and its decreasing envelope $\overline{H}_2$ is also shown.

Lemma A.3. *Let $H : (-\infty, w_{max}] \to \mathbb{R}$ be a convex function that satisfies Assumption A.1; in particular it has a root $H(w_*) = 0$. Then*

(a) For any $a \in (w_*, w_{max}]$ the function

$$w \mapsto Q_1(w, a) = -\int_w^a \frac{d\zeta}{H(\zeta)}, \tag{A.2}$$

maps $(w_*, a]$ onto $[0, \infty)$; is strictly decreasing, and has an inverse $Q_1^{-1}(r, a)$, which maps $[0, \infty)$ onto $(w_*, a]$.

(b) For any $a \in (-\infty, w_*)$ the function

$$w \mapsto Q_2(w, a) = \int_a^w \frac{d\zeta}{H(\zeta)}, \tag{A.3}$$

maps $[a, w_*)$ onto $[0, \infty)$; is strictly increasing, and has an inverse $Q_2^{-1}(r, a)$, which maps $[0, \infty)$ onto $[a, w_*)$.

Remark A.4. Analogous to (A.2), we denote by $\overline{Q}_1$ the function

$$w \mapsto \overline{Q}_1(w, a) = -\int_w^a \frac{d\zeta}{\overline{H}(\zeta)}, \tag{A.4}$$

where $\overline{H}$ is the decreasing envelope of H.

Proof. To show (a), note that the integrand $-1/H(\zeta)$ is strictly positive on (w_*, a). It follows that $Q_1(., a)$ is strictly increasing and maps $(w_*, a]$ *into* $[0, \infty)$. It remains to show that range of this map covers all of $[0, \infty)$. To this end, observe that by convexity we have

$$H(w) \geqslant H'(w_*)(w - w_*), \qquad \text{for all } w \in (-\infty, w_{\max}], \tag{A.5}$$

and $H'(w_*) < 0$. Thus, we obtain

$$\lim_{w \downarrow w_*} Q_1(w, a) = -\int_{w_*}^{a} \frac{d\zeta}{H(\zeta)} \geqslant -\frac{1}{H'(w_*)} \int_{w_*}^{a} \frac{d\zeta}{\zeta - w_*} = +\infty.$$

The proof of (b) is analogous; only the different sign of H on $(-\infty, w_*)$ has to be taken into account. ☐

Theorem A.5. *Let κ be a positive, continuous, and decreasing function in $L_1(\mathbb{R}_{\geqslant 0})$ and let H be a convex function that satisfies Assumption A.1; in particular $H(w_*) = 0$ is its unique root in $(-\infty, w_{max}]$. For any continuous function $a : \mathbb{R}_{\geqslant 0} \to (-\infty, w_{max}]$ consider the non-linear Volterra equation*

$$f(t) = a(t) + \int_0^t \kappa(t - s) H(f(s)) ds, \qquad t \in \mathbb{R}_{\geqslant 0}. \tag{A.6}$$

(a) If a is increasing with values in $(w_*, w_0]$ then (A.6) *has a unique global solution f which satisfies*

$$w_* < r_1(t) \leqslant f(t) < a(t), \qquad \forall\, t > 0, \tag{A.7}$$

where $r_1(t) = Q_1^{-1}\left(\int_0^t \kappa(s)ds, a(0)\right)$ and Q_1 is given by (A.2).

(b) If $a \equiv w_*$ then $f \equiv w_*$ is the unique global solution of (A.6)

(c) If a is decreasing with values in $(-\infty, w_*)$ then (A.6) *has a unique global solution f which satisfies*

$$a(t) < f(t) \leqslant r_2(t) < w_*, \qquad \forall\, t > 0, \tag{A.8}$$

where $r_2(t) = Q_2^{-1}\left(\int_0^t \kappa(s)ds, a(0)\right)$ and Q_2 is given by (A.3).

In addition, case (a) can be extended to the following more general statement:

(a') If a is increasing with values in $(w_*, w_{max}]$ then (A.6) *has a unique global solution f which satisfies*

$$w_* < \overline{r}_1(t) \leqslant f(t) < a(t), \qquad \forall\, t > 0, \tag{A.9}$$

where $\overline{r}_1(t) = \overline{Q}_1^{-1}\left(\int_0^t \kappa(s)ds, a(0)\right)$ and $\overline{Q}_1$ is given by (A.4).

Remark A.6. Clearly, if H is decreasing (and hence $w_0 = w_{\max}$), cases (a) and (a') coincide. In the general case (a) gives better bounds on f than (a'), but is more restrictive in its assumption on the function a.

Before proving the theorem, we add two Corollaries that are used in the proofs of Theorems 2.5, 3.1 and 4.9.

Corollary A.7. *Under the assumptions of Theorem A.5, consider the non-linear integral equation*

$$g(t) = H\left(a(t) + \int_0^t \kappa(t-s)g(s)ds\right), \qquad t \in \mathbb{R}_{\geqslant 0}. \tag{A.10}$$

(a) If a is increasing with values in $(w_*, w_0]$ then (A.10) *has a unique global solution g which satisfies*

$$H(a(t)) < g(t) \leqslant H(r_1(t)) < 0, \qquad \forall\, t > 0. \tag{A.11}$$

(b) If $a \equiv w_*$ then $g \equiv 0$ is the unique global solution of (A.10)*.*

(c) If a is decreasing with values in $(-\infty, w_*)$ then (A.10) *has a unique global solution g which satisfies*

$$0 < g(t) \leqslant H(r_2(t)) < H(a(t)), \qquad \forall\, t > 0. \tag{A.12}$$

In addition, case (a) can be extended to:

(a') If a is increasing with values in $(w_*, w_{max}]$ then (A.10) *has a unique global solution g which satisfies*

$$g(t) < 0, \qquad \forall\, t > 0. \tag{A.13}$$

In any of the above cases, $g(t) = H(f(t))$, where f is the solution of (A.6)*.*

Corollary A.8. *Let the assumptions of Theorem A.5 hold with $w_{max} = 0$. Let $\Delta > 0$ and let h be a piecewise continuous function from $[0, \Delta)$ to $\mathbb{R}_{\leqslant 0}$. Consider the non-linear integral equation*

$$g(t) = H\left(\int_0^t \kappa(t-s)g(s)ds\right), \qquad t \in [\Delta, \infty), \tag{A.14}$$

with initial condition

$$g(t) = h(t), \qquad t \in [0, \Delta).$$

If $w_* < \int_0^\Delta \kappa(\Delta - s)h(s)ds$, then (A.14) *has a unique global solution g taking values in $\mathbb{R}_{\leqslant 0}$, which satisfies*

$$w_* < \int_0^t \kappa(t-s)g(s)ds \quad \textit{for all } t \geqslant 0. \tag{A.15}$$

We start with the proof of Theorem A.5, which closely follow the account of Lakshmikantham's comparison method in [BS13, Sec. II.7].

Proof of Theorem A.5. Clearly, H can be extended to a continuous function on all of $\mathbb{R}$ and thus it follows from [GLS90, Thm. 12.1.1] that (A.6) has a local continuous solution f on an interval $[0, T_{\max})$ with $T_{\max} > 0$. In addition, $T_{\max}$ can be chosen maximal, in the sense that the solution cannot be continued beyond $[0, T_{\max})$.
Case (a): By assumption, a is increasing and takes values in $(w_*, w_0]$. Set

$$T_* := \inf \{t \in (0, T_{\max}) : f(t) = w_* \text{ or } f(t) = a(T_{\max})\} \tag{A.16}$$

and note that $T_* > 0$. From (A.6) it is clear that

$$f(t) = a(t) + \int_0^t \kappa(t-s) H(f(s)) ds < a(t) \leqslant a(T_{\max}), \quad \forall t \in [0, T_*), \tag{A.17}$$

i.e. the lower bound w_* in (A.16) is always hit before the upper bound $a(T_{\max})$. In addition, using that the kernel κ is decreasing, we obtain that

$$f(t) = a(t) + \int_0^t \kappa(t-s) H(f(s)) ds \leqslant a(T) + \int_0^t \kappa(T-s) H(f(s)) ds := v(t, T) \tag{A.18}$$

for all $0 \leqslant t \leqslant T \leqslant T_*$. The function $v(t, T)$, which we have just defined, satisfies

$$v(t, t) = f(t) \tag{A.19}$$
$$v(0, T) = a(T) \geqslant a(0) \tag{A.20}$$

and the differential inequality

$$\frac{\partial}{\partial t} v(t, T) = \kappa(T-t) H(f(t)) \geqslant \kappa(T-t) H(v(t, T)). \tag{A.21}$$

Here, where we have used (A.18) and the fact that H is decreasing on $(w_*, w_0]$. Together with the initial estimate (A.20), a standard comparison principle for differential inequalities (cf. [Wal96, II.§9]) yields

$$v(t, T) \geqslant r(t, T), \tag{A.22}$$

where

$$\frac{\partial}{\partial t} r(t, T) = \kappa(T-t) H(r(t, T)), \qquad r(0, T) = a(0). \tag{A.23}$$

We claim that the differential equation (A.23) is solved by

$$r(t, T) = Q_1^{-1} \left(\int_0^t \kappa(T-s) ds, a(0) \right). \tag{A.24}$$

Indeed, applying $Q_1(.,a(0))$ to both sides of (A.24) yields

$$\int_0^t \kappa(T-s)ds = Q_1(r(t,T),a(0)) = -\int_{r(t,T)}^{a(0)} \frac{d\zeta}{H(\zeta)}.$$

Taking $\frac{\partial}{\partial t}$-derivatives, we obtain

$$\kappa(T-t) = \frac{1}{H(r(t,T))}\frac{\partial}{\partial t}r(t,T)$$

which is equivalent to (A.23). From (A.18), (A.19) and (A.22) we obtain the bound

$$r_1(t) := \lim_{T \downarrow t} r(t,T) \leqslant \lim_{T \downarrow t} v(t,T) = f(t) \tag{A.25}$$

for all $t \in [0,T_*)$. This implies that

$$\lim_{t \to T_*} f(t) \geqslant r_1(T_*) > w_*, \tag{A.26}$$

which, in light of (A.16), means that $T_* = T_{\max}$, i.e. we have shown the bounds (A.7) to hold for all $t \in [0,T_{\max})$. However, by [GLS90, Thm. 12.1.1] $\lim_{t \to T_{\max}} |f(t)| = \infty$ whenever $T_{\max} < \infty$. We conclude that $T_{\max} = \infty$, and hence that f is a global solution of (A.6). Uniqueness follows from [GLS90, Thm. 13.1.2].

Case (b): By assumption, $a \equiv w_*$. Since $H(w_*) = 0$ it is clear that $f(t) \equiv w_*$ is a global solution of (A.6). Uniqueness follows from [GLS90, Thm. 13.1.2].

Case (c): By assumption, a is decreasing and takes values in $(-\infty, w_*]$. This case can be handled analogous to case (a) with the following adaptations: The inequality signs in equations (A.17) – (A.22) have to be reversed. In (A.24) Q_1 has to be substituted by Q_2 and also in (A.25) and (A.26) the inequalities have to be reversed.

Case (a'): The proof of Case (a) applies, except for the following modification: (A.21) holds only when $v(t,T) \leqslant w_0$, since H is decreasing only on $(-\infty, w_0]$. However, when $v(t,T) > w_0$, we can use the trivial estimate

$$\frac{\partial}{\partial t}v(t,T) = \kappa(T-t)H(f(t)) \geqslant \kappa(T-t)H(w_0),$$

which can be combined with (A.21) into

$$\frac{\partial}{\partial t}v(t,T) = \kappa(T-t)H(f(t)) \geqslant \kappa(T-t)\overline{H}(v(t,T)),$$

where $\overline{H}$ is the decreasing envelope of H from Definition A.2. The remaining proof of Case (a) applies after substituting H by $\overline{H}$ and Q_1 by $\overline{Q}_1$. □

Proof of Corollary A.7. Let f be the global solution of (A.6). Applying H to both sides of (A.6), we see that $g(t) := H(f(t))$ is a global solution of (A.10).

Next, we show uniqueness. To this end, assume that $\widetilde{g}$ is a local solution of (A.10) on $[0, T)$, different from g and define

$$\widetilde{f}(t) := a(t) + \int_0^t \kappa(t-s)\widetilde{g}(s)ds.$$

Clearly, $\widetilde{g}(t) = H(\widetilde{f}(t))$ on $[0, T)$, and hence $\widetilde{f}$ is a local solution of (A.6). By [GLS90, Thm. 13.1.2], this solution is unique, and we conclude that $\widetilde{f} = f$, and hence also $\widetilde{g} = g$. Finally, applying H – which is decreasing on $(-\infty, w_0]$ – to the inequalities (A.7) and (A.8) yields (A.11) and (A.12). In case (a') monotonicity of H is lost, but $H(w) < 0$ for all $w \in (w_*, w_{\max}]$ yields (A.13). □

Proof of Corollary A.8. Set

$$a(t) := \int_0^\Delta \kappa(t + \Delta - s)h(s)ds$$

and note that a is increasing with values in $(w_*, 0]$. Consider the non-linear Volterra equation

$$f(t) = a(t) + \int_0^t \kappa(t-s)H(f(s))ds, \tag{A.27}$$

which has a unique global solution f by Theorem A.5.(a) or (a'). For $t' \in \mathbb{R}_{\geqslant 0}$ set

$$g(t') = \begin{cases} H(f(t' - \Delta)), & t' \in [\Delta, \infty) \\ h(t') & t' \in [0, \Delta) \end{cases}.$$

For $t' \geqslant \Delta$ we have

$$\begin{aligned} g(t') = H(f(t' - \Delta)) = H\left(\int_0^\Delta \kappa(t' - s)h(s)ds + \int_\Delta^{t'} \kappa(t' - s)g(s)ds\right) = \\ = H\left(\int_0^{t'} \kappa(t' - s)g(s)ds\right), \end{aligned}$$

showing that g is a global solution of (A.14). From cases (a) or (a') of Theorem A.5, we obtain the bound

$$w_* < f(t' - \Delta) = \int_0^{t'} \kappa(t' - s)g(s)ds,$$

as claimed. To show uniqueness, assume that $\widetilde{g}$ is a solution of (A.14), different from g. Setting

$$\widetilde{f}(t) := a(t) + \int_0^t \kappa(t-s)\widetilde{g}(s + \Delta),$$

we see that $\widetilde{f}$ is a solution of (A.27) and conclude from Theorem A.5 that $\widetilde{f} = f$ and hence also $\widetilde{g} = g$. □

References

[BG12] Lorenzo Bergomi and Julien Guyon. Stochastic volatility's orderly smiles. *Risk*, 25(5):60–66, 2012.

[Bil86] Patrick Billingsley. *Probability and measure.* John Wiley & Sons, 2nd edition edition, 1986.

[BMM15] Emmanuel Bacry, Iacopo Mastromatteo, and Jean-François Muzy. Hawkes processes in finance. *Market Microstructure and Liquidity*, 1(01):1550005, 2015.

[BS13] Drumi D Bainov and Pavel S Simeonov. *Integral inequalities and applications*, volume 57. Springer Science & Business Media, 2013.

[DFS03] Darrell Duffie, Damir Filipović, and Walter Schachermayer. Affine processes and applications in finance. *Annals of applied probability*, pages 984–1053, 2003.

[EER16] Omar El Euch and Mathieu Rosenbaum. The characteristic function of rough Heston models. *arXiv preprint arXiv:1609.02108*, 2016.

[EMOT81] Arthur Erdélyi, Wilhelm Magnus, Fritz Oberhettinger, and Francesco G Tricomi. *Higher transcendental functions, Vol. III, based on notes left by Harry Bateman, reprint of the 1955 original.* Robert E. Krieger Publishing Co., Inc., Melbourne, Fla, 1981.

[ER17] Omar El Euch and Mathieu Rosenbaum. Perfect hedging in rough Heston models. *arXiv preprint arXiv:1703.05049*, 2017.

[Fil09] D. Filipovic. *Term-Structure Models.* Springer, 2009.

[GLS90] Gustaf Gripenberg, Stig-Olof Londen, and Olof Staffans. *Volterra integral and functional equations*, volume 34. Cambridge University Press, 1990.

[GRT92] A Ronald Gallant, Peter E Rossi, and George Tauchen. Stock prices and volume. *The Review of Financial Studies*, 5(2):199–242, 1992.

[Hes93] Steven L Heston. A closed-form solution for options with stochastic volatility with applications to bond and currency options. *The review of financial studies*, 6(2):327–343, 1993.

[HMS11] Hans J Haubold, Arak M Mathai, and Ram K Saxena. Mittag-leffler functions and their applications. *Journal of Applied Mathematics*, 2011, 2011.

[JLP17] Eduardo Abi Jaber, Martin Larsson, and Sergio Pulido. Affine volterra processes. *arXiv preprint arXiv:1708.08796*, 2017.

[JR15] Thibault Jaisson and Mathieu Rosenbaum. Limit theorems for nearly unstable hawkes processes. *The Annals of Applied Probability*, 25(2):600–631, 2015.

[JR16] Thibault Jaisson and Mathieu Rosenbaum. Rough fractional diffusions as scaling limits of nearly unstable heavy tailed hawkes processes. *The Annals of Applied Probability*, 26(5):2860–2882, 2016.

[KR11] Martin Keller-Ressel. Moment explosions and long-term behavior of affine stochastic volatility models. *Mathematical Finance*, 21(1):73–98, 2011.

[Pod98] Igor Podlubny. *Fractional differential equations: an introduction to fractional derivatives, fractional differential equations, to methods of their solution and some of their applications*, volume 198. Academic press, 1998.

[Roc70] R. Tyrrell Rockafellar. *Convex Analysis*. Princeton University Press, 1970.

[Ver12] Mark Veraar. The stochastic fubini theorem revisited. *Stochastics An International Journal of Probability and Stochastic Processes*, 84(4):543–551, 2012.

[Wal96] Wolfgang Walter. *Ordinary Differential Equations*. Springer, 6th edition, 1996.