Submitted to Bernoulli

Lévy processes as weak limits of rough Heston models

ALESSANDRO BONDI[1,a] and MARTIN FORDE[2,b,c]

[1]*Department of AI, Data and Decision Sciences, Luiss University Rome,* [a]*abondi@luiss.it*
[2]*Department of Mathematics, King's College London,* [b]*martin.forde@kcl.ac.uk,* [c]*https://martinforde.github.io*

We show weak convergence of the time-t marginals for the integrated variance in a re-scaled rough Heston model to those of an Inverse Gaussian Lévy process. This shows we can obtain such a limit without having to impose that the true Hurst exponent H for the model is $\frac{1}{2}$ as in [6], or that $H \searrow -\frac{1}{2}$ as in [3], so the result potentially has increased financial relevance. We later extend the analysis to the case where V has jumps, showing weak convergence of the finite-dimensional distributions of the integrated variance to a deterministic time-change of the first-passage time process to lower barriers for a more general class of spectrally positive Lévy processes. This convergence result is then strengthened to a functional setting, namely on the space of càdlàg functions on the non-negative half-line endowed with the M_1 topology.

Keywords: Affine Volterra processes with jumps; Fast Mean reversion; Lévy processes hitting times; Rough Heston model; Volterra Integral equations

1. Introduction

Stochastic Volterra equations (SVEs) arise in a variety of areas of applied mathematics, as they provide a natural framework for modeling systems with memory and irregular behavior through the presence of kernels that drive the dynamics. For instance, they appear as scaling limits of branching processes in chemical and biological interaction models [36], or of Hawkes processes in mathematical finance [25]. Volterra-type Lévy processes are also used in the stochastic modeling of energy markets; see [12]. In general, SVEs generate processes that are neither Markov processes nor semimartingales; nevertheless, they provide a theoretically convenient framework that is currently being explored in several directions, some of which are not primarily motivated by applications. For instance, general results on the theory of SVEs, in both the continuous and jump settings, can be found in [1,4,8,10,15]. The invariance theory for SVEs has also attracted considerable attention in recent years; see [5,9,16]. We also refer to [14,22,31,40] for studies of SVEs in infinite dimension.

In this work, our analysis starts from a Volterra square-root diffusion $V=(V_t)_{t\geq 0}$, which is a non-negative process satisfying the dynamics

$$V_t = V_0 + \int_0^t K(t-s)\Big(\lambda(\theta - V_s)ds + \sigma\sqrt{V_s}dW_s\Big), \qquad \text{where} \quad K(t) = \frac{1}{\Gamma(\alpha)}t^{\alpha-1}, \text{ for } \alpha \in \Big(\frac{1}{2}, 1\Big). \tag{1}$$

Here $V_0 \geq 0$, Γ is the Gamma function, W is a standard one-dimensional Brownian motion and λ, θ and σ are given positive parameters. Notice that, in the limiting case $\alpha = 1$, (1) reduces to the classical CIR process, see [21]. Weak existence results for (1) can be found in, e.g., [8].

The process V in (1) has trajectories that are rougher than those of Brownian motion; more precisely, they are $(H-\varepsilon)$−Hölder continuous, as is the case for fractional Brownian motion (see, e.g., Theorem

3.2 in [33]). Here $H = \alpha - \frac{1}{2}$ is known as the Hurst index. In the last decade, V has been extensively investigated, especially in the mathematical finance literature, as it describes the variance in the so-called rough Heston stochastic volatility model; see [26,27]. This model was introduced in [33], where the authors show, using C-tightness arguments, that it arises naturally as weak limit of a small jump high-frequency market microstructure model driven by two nearly unstable Hawkes processes. One of the main reasons for the popularity of this model is its affine nature; see [8,27,29]. This property allows to express the characteristic function of the log stock price and the integrated variance in rough Heston-type models in terms of solutions to deterministic non-linear Volterra integral equations (VIEs); see also [1,15,17,22] for extensions allowing jumps in V. This formulation enables accurate option pricing even for very small values of H close to 0 (and even for $H = 0$) by solving VIEs numerically via an Adams scheme. In particular, it avoids Monte Carlo techniques, which are known to perform poorly when $H \ll 1$, both in terms of bias and sample variance.

In this paper, we study the integrated variance process $A = (A_t)_{t\geq 0}$, defined by $A_t = \int_0^t V_s\, ds$. Our goal is to determine weak scaling limits for A and to characterize these limits as the laws of certain time-changed Lévy processes. The results we obtain extend the existing literature on weak limits for the integrated variance in rough Heston-type models. They also apply to extensions of the dynamics in (1) that exhibit jumps, in the spirit of [1,15,17], for which – to the best of our knowledge – no similar results have been established so far. Specifically, we consider dynamics that include

$$V_t = V_0 + \int_0^t K(t-s)\Big(\lambda(\theta - V_s)ds + \sigma\sqrt{V_s}dW_s + d\tilde{J}_s\Big), \tag{2}$$

where K is the same fractional kernel as in (1) and $\tilde{J}_\cdot = \int_0^{(\cdot)}\int_{\mathbb{R}_+} x(N(dx,dt) - V_t\nu(dx)dt)$. Here, $N(dx,dt)$ is an integer-valued random measure with compensator $V_t\nu(dx)dt$, for a given non-negative measure ν with positive support such that $\nu(\{0\}) = 0$ and $\int_{\mathbb{R}_+} x^2\nu(dx) < \infty$. Weak existence results for (2) are established in [1].

In the continuous setting (see (1)), significant contributions in this research direction can be found in [3,6]. More precisely, [6] show that a re-scaled standard Markov Heston model with fast mean-reversion and large vol-of-vol (via an H parameter which is not the Hurst exponent) tends weakly on path space to one of three different models (either Black-Scholes, a Normal Inverse Gaussian or a Normal Lévy model), depending on whether their H parameter is $>$, $=$, or $< -\frac{1}{2}$. [3] obtain a similar result without Markovian approximations but instead letting $H \searrow -\frac{1}{2}$ for the so-called *hyper-rough Heston model* (see also Section 5 in [28] and Section 7 in [1] for more on this model), and exploiting Dirac-type behaviour in their Lemma 2.4 (see Appendix D here for a short summary/formal derivation of their result).

One of the contributions of this work is to fill the gap between [6] and [3] by showing that a similar result holds for any $H \in (0, \frac{1}{2})$ (in particular, our limit regime with $H = \frac{1}{2}$ corresponds to the regime in Eq. (0.3) of [6], where their $H = -\frac{1}{2}$). This will follow as a consequence of a more general result presented in Section 3 (see Theorem 3.7), where we establish weak convergence on path space for an extended model that allows positive jumps in the dynamics of V, see (2). In this case, relying on the affine structure of the model, using the Laplace transform of the hitting time to a lower barrier for a spectrally positive Lévy process, we find that the limiting process for the integrated variance is a deterministic time-change of the first passage time process to lower barriers for a more general class of Lévy processes. The convergence is proved using a compactness argument with the Kolmogorov-Riesz-Fréchet theorem that is detailed in Appendix B.

Although we establish and focus here only on theoretical results, they may prove useful for applications that will be investigated in future research. For instance, in the continuous case (1), by the stochastic Fubini and Dubins–Schwarz theorems, it is well known that the integrated variance process satisfies an equation of the form

$$A_t = G_0(t) + \int_0^t \kappa(t-s) B_{A_s}\, ds,$$

for some Brownian motion B, where the deterministic curve G_0 and the kernel κ are related to the coefficients in (1) (see, e.g., [1,3] and Theorem 3.1 in [34]). This is a non-linear pathwise Volterra integral equation (VIE) for A in terms of the a.s. $(\frac{1}{2}-\varepsilon)-$Hölder continuous function $B_{(\cdot)}$, and it remains well defined even if $\kappa \in L^1(0,T)$; this also allows us to consider the so-called hyper-rough regime $H \in (-\frac{1}{2}, 0]$. If we discretize this equation for A and rewrite it in terms of the final (discrete-time) increments of A and B (see Appendix D for details), then an independent sequence of Inverse Gaussian random variables can be used to perform an approximate Monte Carlo simulation of A (see Algorithm 1 in [2]), which is naturally suited for the regimes $H \ll 1$ and $H \in (-\frac{1}{2}, 0]$.

The paper is organized as follows. In Section 2, we introduce a rescaling for the continuous dynamics in (1) to prove the convergence of the time-t marginals of the integrated variance to an Inverse Gaussian process. This is established in Theorem 2.1, which is derived independently of the other results in the paper, as it addresses a relevant example in the literature, namely weak scaling limits for the integrated variance in rough Heston-type models with continuous trajectories, using an original strategy. Moreover, its proof is simpler than those of the other results, as it relies only on asymptotics for deterministic VIEs from [28]. The scaling technique in Section 2 is then extended in Section 3 to affine SVEs with jumps that include (2) (see also Eqs. (11)-(13)). Theorem 3.5 establishes the convergence of the finite-dimensional distributions of the corresponding integral process to a Lévy subordinator with a deterministic time change. In Subsection 3.1, this convergence result is strengthened to a functional setting. More specifically, by combining tightness criteria from [39] with Theorem 3.5, which identifies the limiting process, in Theorem 3.7 we prove weak convergence of the laws of the re-scaled integral processes to the law of the aforementioned time-changed Lévy subordinator on the space of càdlàg functions on the non-negative half-line endowed with the M_1 topology (see [39] for further details on this and related topological spaces, also used in [3,6]). In the appendices, we prove some results needed for our analysis of the jump case in Section 3. More precisely, Appendix A studies the well-posedness of the VIE needed to describe the finite-dimensional moment generating function of the integrated variance. Appendix B contains the proof of the main technical lemma of the paper, Lemma 3.4, which, together with the results in Appendix C, enables us to identify the limiting law of the re-scaled integrated variance. Finally, Appendix D contains a formal derivation of the main idea in [3].
Due to its interpretation as the variance process in rough Heston-type models, we will often refer to V in (1) (or, in the jump case, in (2); see Section 3) simply as the “variance process”.

2. Marginal Integrated Variance Asymptotics in a Re-Scaled Rough Heston Model

We consider processes defined on (possibly different) probability spaces $(\Omega, \mathcal{F}, \mathbb{Q})$, each equipped with a filtration $(\mathcal{F}_t)_{t\geq 0}$ satisfying the usual conditions. For the main convergence result of this section, see Theorem 2.1 below, we recall the following definition.

Definition 2.1. *A random variable X is said to have an* Inverse Gaussian (IG) *distribution with parameters $\gamma > 0$ and $\delta > 0$, denoted by $X \sim \mathrm{IG}(\gamma, \delta)$, if its probability density function is*

$$f_X(x) = \left(\frac{\delta}{2\pi x^3}\right)^{1/2} \exp\left\{-\frac{\delta(x-\gamma)^2}{2\gamma^2 x}\right\}, \quad x > 0.$$

The moment generating function of $X \sim \mathrm{IG}(\gamma, \delta)$ is given by

$$\mathbb{E}[e^{pX}] = \exp\left\{\frac{\delta}{\gamma}\left(1 - \sqrt{1 - \frac{2\gamma^2 p}{\delta}}\right)\right\}, \quad p < \frac{\delta}{2\gamma^2}, \tag{3}$$

which shows that the $\mathrm{IG}(\gamma, \delta)$ distribution is infinitely divisible. An IG process with parameters (γ, δ) is a (càdlàg) Lévy process $L = (L_t)_{t\geq 0}$ such that $L_1 \sim \mathrm{IG}(\gamma, \delta)$. This process exists and is unique in law; see Theorem 7.10 in [37]. As is well known (see, for instance, Example 1.3.21 in [11]), an $\mathrm{IG}(\gamma, \delta)$ process is given by

$$L_t = \inf\left\{s \geq 0 : B_s + \frac{\sqrt{\delta}}{\gamma} s = \sqrt{\delta} t\right\}, \tag{4}$$

where B is a standard one-dimensional Brownian motion.

In the next result, Theorem 2.1, we show that the time-t marginals of the integrated solution to a re-scaled version of (1) converge to those of an IG Lévy process. This result is of interest as it demonstrates that such an IG process can be obtained without requiring $H = \alpha - \frac{1}{2}$, where α denotes the exponent in the kernel K driving (1), to be equal to $\frac{1}{2}$ as in [6], or to tend to $-\frac{1}{2}$ as in [3]. Therefore, it extends the existing literature on scaling limits of rough Heston models.

The proof relies on the affine structure of (1), which enables us to express the Laplace transform of the integrated variance via deterministic Riccati–Volterra equations. This is a well-known property of affine Volterra processes; see, e.g., [8], which generalizes the corresponding classical results for standard affine processes, for instance those in [24]. Thanks to a suitable rescaling, in the proof of Theorem 2.1 we show that the moment generating function of the integrated variance converges to that of an IG process, which implies convergence of the time-t marginals.

The rescaling introduced in Theorem 2.1 will then be generalized in Section 3 to the case with jumps. In particular, as a corollary of Theorems 3.5 and 3.7, the convergence in Theorem 2.1 will be strengthened to a functional setting; see Remark 3.5.

Theorem 2.1. *Consider a re-scaled rough Heston model for the variance process $V^\varepsilon = (V_t^\varepsilon)_{t\geq 0}$:*

$$V_t^\varepsilon = V_0 + \frac{1}{\Gamma(\alpha)} \int_0^t (t-s)^{\alpha-1} \left(\frac{1}{\varepsilon}\lambda(\theta - V_s^\varepsilon) ds + \frac{1}{\varepsilon}\sigma\sqrt{V_s^\varepsilon} dW_s\right), \tag{5}$$

where $V_0 \geq 0$, W is a standard one-dimensional Brownian motion, $\alpha \in (\frac{1}{2}, 1)$ and $\lambda, \theta, \sigma > 0$. Then, for every $t > 0$ fixed, $A_t^\varepsilon = \int_0^t V_s^\varepsilon ds$ tends weakly, as $\varepsilon \to 0$, to the time-t marginal of an Inverse Gaussian Lévy process with parameters $(\theta, \sigma^{-2}\lambda^2\theta^2)$ which does not depend on $H = \alpha - \frac{1}{2}$.

Proof. First, we notice that weak existence of a non-negative solution to (5) is established in, for instance, Theorem 7.1 of [8]. Let $I^\alpha(f)(t) = \frac{1}{\Gamma(\alpha)} \int_0^t (t-s)^{\alpha-1} f(s) ds$ denote the αth-order fractional integral of a function f for $\alpha \in (\frac{1}{2}, 1)$. Then for $p > 0$ (which will be sufficient for our purposes when we invoke a classical weak convergence result from [13] below), if we define $A_t^\varepsilon = \int_0^t V_s^\varepsilon ds$, by the arguments in, e.g., Section 7 in [8]

$$\mathbb{E}[e^{-pA_t^\varepsilon}] = e^{V_0 I^{1-\alpha}\phi_\varepsilon(t) + \frac{1}{\varepsilon}\lambda\theta I^1\phi_\varepsilon(t)}, \tag{6}$$

where ϕ_ε is the unique continuous solution of the non-linear Volterra integral equation (VIE):

$$\phi_\varepsilon(t) = \frac{1}{\Gamma(\alpha)}\int_0^t (t-s)^{\alpha-1}\Big(-p - \frac{1}{\varepsilon}\lambda\phi_\varepsilon(s) + \frac{1}{\varepsilon^2}\frac{1}{2}\sigma^2\phi_\varepsilon(s)^2\Big)ds.$$

In particular, note that ϕ_ε also depends on p.

Now let $\phi_\varepsilon(t) = \varepsilon\psi(\varepsilon^q t)$. Then, also using the change of variables $\varepsilon^q s = u$,

$$\begin{aligned}\varepsilon\psi(\varepsilon^q t) &= \frac{1}{\Gamma(\alpha)}\int_0^t (t-s)^{\alpha-1}\Big(-p - \lambda\psi(\varepsilon^q s) + \frac{1}{2}\sigma^2\psi(\varepsilon^q s)^2\Big)ds \\ &= \frac{\varepsilon^{-q(\alpha-1)}}{\Gamma(\alpha)}\int_0^{\varepsilon^q t} (\varepsilon^q t - u)^{\alpha-1}\Big(-p - \lambda\psi(u) + \frac{1}{2}\sigma^2\psi(u)^2\Big)du\,\varepsilon^{-q}.\end{aligned}$$

Setting $\varepsilon^q t \mapsto t$, we see that

$$\varepsilon\psi(t) = \frac{\varepsilon^{-q\alpha}}{\Gamma(\alpha)}\int_0^t (t-u)^{\alpha-1}F(\psi(u))du, \tag{7}$$

where $F(w) = -p - \lambda w + \frac{1}{2}\sigma^2 w^2$. If now we let $q = -\frac{1}{\alpha}$, the VIE (7) is independent of ε, so

$$\phi_\varepsilon(t) = \varepsilon\psi\left(\frac{t}{\varepsilon^{\frac{1}{\alpha}}}\right). \tag{8}$$

Hence for every $t > 0$, from Lemma 4.5 in [28]

$$\lim_{\varepsilon\to 0}\frac{1}{\varepsilon}\phi_\varepsilon(t) = \lim_{t\to\infty}\psi(t) \quad = \quad \psi(\infty) \quad = \quad \frac{1}{\sigma^2}\Big[\lambda - \sqrt{\lambda^2 + 2p\sigma^2}\Big] \quad =: \quad U_1(p) \tag{9}$$

for $p > 0$. More precisely, Lemma 4.5 in [28] can be applied to $-\psi$, which solves the following VIE on $\mathbb{R}_+$:

$$-\psi(t) = \frac{1}{\Gamma(\alpha)}\int_0^t (t-u)^{\alpha-1}F_1(-\psi(u))du, \quad \text{with } F_1(w) = p - \lambda w - \frac{1}{2}\sigma^2 w^2.$$

Indeed, since $F_1(0) > 0$ and F_1 is analytic and decreasing on $\mathbb{R}_+$, the aforementioned result in [28] implies that $-\psi(t)$ is monotonically increasing and converges as $t \longrightarrow \infty$ to the positive root of F_1, which coincides with $-U_1(p)$. Then for the exponent in (6), knowing from (8), (9) and the monotonicity of ψ that $|\frac{1}{\varepsilon}\phi_\varepsilon(t)| \leq -U_1(p)$ for every $t \geq 0$ and $\varepsilon > 0$, by the dominated convergence theorem and (9) we see that

$$V_0 I^{1-\alpha}\phi_\varepsilon(t) + \frac{\lambda\theta}{\varepsilon}I^1\phi_\varepsilon(t) = \frac{V_0}{\Gamma(1-\alpha)}\int_0^t (t-s)^{-\alpha}\phi_\varepsilon(s)ds + \lambda\theta\int_0^t \frac{1}{\varepsilon}\phi_\varepsilon(s)ds \longrightarrow 0 + \lambda\theta U_1(p)t$$

as $\varepsilon \longrightarrow 0$. By (3), $\lambda\theta U_1(p)t$ is the log moment generating function of an Inverse Gaussian Lévy process with parameters $(\theta, \sigma^{-2}\lambda^2\theta^2)$ at time t. Then from e.g. the solution to Problem 30.4 on Page 573 in [13], we conclude that A_t^ε tends weakly to the time-t marginal law of an $\mathrm{IG}(\theta, \sigma^{-2}\lambda^2\theta^2)$ process. ☐

3. Adding jumps into V^{ε}

We now assume that the rescaled variance process $V^{\varepsilon} = (V^{\varepsilon}_t)_{t\geq 0}$ satisfies the following SVE of convolution-type with jumps:

$$V^{\varepsilon}_t = V_0 + \frac{1}{\Gamma(\alpha)} \int_0^t \frac{1}{\varepsilon}(t-s)^{\alpha-1}\Big(\lambda(\theta - V^{\varepsilon}_s)ds + \sigma\sqrt{V^{\varepsilon}_s}\,dW_s + d\tilde{J}^{\varepsilon}_s\Big), \quad \mathbb{P}\otimes dt - \text{a.e.}, \tag{10}$$

where $\alpha \in (\frac{1}{2}, 1)$, $d\tilde{J}^{\varepsilon}_t = \int_{\mathbb{R}_+} x(N^{\varepsilon}(dx,dt) - V^{\varepsilon}_t \nu(dx)dt)$ and $N^{\varepsilon}(dx,dt)$ is an integer-valued random measure with compensator $V^{\varepsilon}_t \nu(dx)dt$; ν only has positive support with $\nu(\{0\}) = 0$ and $\int_{\mathbb{R}_+} x^2 \nu(dx) < \infty$, so $\tilde{J}^{\varepsilon}$ has positive-only jumps. The weak existence of a non-negative solution to (10) is established in Theorem 2.13 of [1], while uniqueness in law in the space $L^2_{\text{loc}}(\mathbb{R}_+)$ follows from Corollary 15 of [15]. Notice that V^{ε} is a rescaled version of the variance process in the rough Hawkes–Heston model by [17]. In the continuous case $\nu(dx) = 0$ (hence $\tilde{J}^{\varepsilon} = 0$), (10) reduces to (5) in Theorem 2.1 of Section 2.

For our analysis, it is convenient to rewrite the dynamics in (10) in an equivalent form that does not explicitly involve the fractional kernel $t \mapsto \frac{1}{\Gamma(\alpha)}t^{\alpha-1}$. More precisely, arguing as in Appendix A of [15] (see, in particular, the first equation following (A.3) therein), we find that V^{ε} solves the following SVE:

$$V^{\varepsilon}_t = V_0 - \lambda(V_0 - \theta)\int_0^t \kappa_{\varepsilon}(s)ds + \int_0^t \kappa_{\varepsilon}(t-s)(\sigma\sqrt{V^{\varepsilon}_s}\,dW_s + d\tilde{J}^{\varepsilon}_s), \quad \mathbb{P}\otimes dt - \text{a.e.} \tag{11}$$

Here, the kernel κ_{ε} is defined by

$$\kappa_{\varepsilon}(t) = \frac{1}{\varepsilon}t^{\alpha-1}E_{\alpha,\alpha}\Big(-\frac{\lambda}{\varepsilon}t^{\alpha}\Big), \quad \text{with } \alpha \in \Big(\frac{1}{2}, 1\Big) \text{ and } \lambda > 0,$$

where $E_{\alpha,\beta}(z)$ denotes the Mittag-Leffler function. The critical observation for the arguments that follow is that

$$f^{\alpha,\lambda}(t) = \lambda t^{\alpha-1}E_{\alpha,\alpha}(-\lambda t^{\alpha})$$

is a probability density, so $\lambda\kappa_{\varepsilon}(\cdot)$ has Dirac-type behaviour as $\varepsilon \to 0$. We also mention that $\int_t^{\infty} f^{\alpha,\lambda}(s)\,ds \underset{t\to\infty}{\sim} \frac{1}{\lambda\Gamma(1-\alpha)}t^{-\alpha}$, which implies that

$$\int_t^{\infty} \lambda\kappa_{\varepsilon}(s)ds \underset{\epsilon\to 0}{\sim} \frac{1}{\lambda\Gamma(1-\alpha)}\varepsilon t^{-\alpha} \tag{12}$$

for $t > 0$ (see Appendix A.1 of [27] for details on these points).

Motivated by (11), in this section we consider a more general predictable, non-negative variance process $V^{\varepsilon} = (V^{\varepsilon}_t)_{t\geq 0}$ with trajectories in $L^1_{\text{loc}}(\mathbb{R}_+)$ which satisfies

$$V^{\varepsilon}_t = \xi^{\varepsilon}_0(t) + \int_0^t \kappa_{\varepsilon}(t-s)(\sigma\sqrt{V^{\varepsilon}_s}\,dW_s + d\tilde{J}^{\varepsilon}_s), \quad \mathbb{P}\otimes dt - \text{a.e.}, \tag{13}$$

where $\xi^{\varepsilon}_0 \in L^1_{\text{loc}}(\mathbb{R}_+)$ is a given deterministic initial variance curve. Note that we require V^{ε} to be non-negative in order to consider the square root in (13), while the predictability of V^{ε} with locally integrable paths ensures that the stochastic integrals in (13) are properly defined. We refer to [1] for weak existence results for (13), see also [4].

The process V^ε here is an affine Volterra process with jumps and falls under the framework of [15] (see also [17]). In particular, Lemma 1 in [15] establishes the following integrability property of V^ε that we will use for our analysis:

$$\mathbb{E}\left[\int_0^T V_t^\varepsilon\, dt\right] < \infty, \quad T > 0. \tag{14}$$

We also refer to Lemma 12 in [15] for a stronger L^2−type integrability result which applies to our dynamics in (13) when $\xi_0^\varepsilon \in L^2_{\text{loc}}(\mathbb{R}_+)$. From (14), we see that

$$\mathbb{E}\left[\int_0^T \left(\int_{\mathbb{R}_+} |x|^2 \nu(dx)\right) V_t^\varepsilon\, dt\right] < \infty, \quad T > 0,$$

and hence

$$\tilde{J}^\varepsilon = \int_0^\cdot \int_{\mathbb{R}_+} x(N^\varepsilon(dx, dt) - V_t^\varepsilon \nu(dx)dt) \text{ is a square-integrable martingale on } [0, T], \\ \text{for every } T > 0. \tag{15}$$

For our asymptotic analysis, we require the following assumption.

Assumption 3.1. *We assume that $\xi_0^\varepsilon(\cdot)$ is non-negative, uniformly bounded and continuous and $\xi_0^\varepsilon(\cdot)$ tends pointwise to a bounded continuous function $\xi_0^0(\cdot)$ as $\varepsilon \to 0$.*

We observe that Assumption 3.1 is satisfied when we specify ξ_0^ε as in (11). Indeed, in this case, $\xi_0^\varepsilon(t) = V_0 - \lambda(V_0 - \theta)\int_0^t \kappa_\varepsilon(s)ds$ and, from (12) and the fact that $\lambda\kappa_\varepsilon(\cdot)$ is a probability density,

$$\xi_0^\varepsilon(t) \underset{\varepsilon\to 0}{\longrightarrow} \theta, \quad t > 0.$$

In the following remark, we present another standard setting in which Assumption 3.1 is satisfied. In Remark 3.2, we provide an alternative formulation of (13), known as the forward variance formulation, which is commonly used in the mathematical finance literature.

Remark 3.1. *If ξ_0 in (13) is independent of ε and given exogenously, then by Appendix A in [15] we can construct an SVE driven by a fractional kernel as in (10) equivalent to (13) by solving the linear VIE: $g_0 - \lambda\int_0^\cdot \kappa_\varepsilon(\cdot - s)g_0(s)ds + \lambda\theta\int_0^\cdot \kappa_\varepsilon(s)ds = \xi_0(\cdot)$, and then using the resulting function g_0 in place of V_0 in (10). Specifically, letting $f(\cdot) = \xi_0(\cdot) - \lambda\theta\int_0^\cdot \kappa_\varepsilon(s)ds$, we can re-write the VIE as: $g_0 - \lambda\int_0^\cdot \kappa_\varepsilon(\cdot - s)g_0(s)ds = f$, which has solution $g_0(\cdot) = f(\cdot) - \int_0^\cdot f(\cdot - s)r(s)ds$. Here r is the resolvent of the 2nd kind of $-\lambda\kappa_\varepsilon$ (which will depend on ε in general); see Section 2.3 in [30].*

Remark 3.2. *The forward variance $\xi_t^\varepsilon(u) := \mathbb{E}[V_u^\varepsilon|\mathcal{F}_t]$ corresponding to the variance process V^ε in (13) satisfies*

$$d\xi_t^\varepsilon(u) = \kappa_\varepsilon(u - t)(\sigma\sqrt{V_t^\varepsilon}dW_t + d\tilde{J}_t^\varepsilon), \quad u > t, \tag{16}$$

see also the rough Hawkes–Heston model in [17]. Note that, in the absence of jumps, (16) is the usual equation for the forward variance under the standard rough Heston model, see e.g. [26] or Proposition 2.2 in [28]. The model in (16) can be viewed as a generalized rough Heston model in the spirit of [17], where the mean-reversion speed, vol-of-vol, and jump-intensity all scale as $\frac{1}{\varepsilon}$.

To prove the main result of this section, Theorem 3.5 below, we rely on the affine structure of V^ε, which enables us to express the Laplace transform of its convolution with locally bounded functions via deterministic non-linear Riccati–Volterra equations. The following lemma establishes the well-posedness of the equations required in our analysis.

Lemma 3.2. *Consider a non-positive locally bounded function* $f \in L^\infty_{\text{loc}}(\mathbb{R}_+;\mathbb{R}_-)$. *Define the map* $G\colon \mathbb{R}_+ \times \mathbb{R}_- \to \mathbb{R}$ *by*

$$G(s,w) = f(s) + \frac{1}{2}\sigma^2 w^2 + V_1(w), \quad \textit{where } V_1(w) = \int_{\mathbb{R}_+} (e^{wx} - 1 - wx)\nu(dx), \tag{17}$$

for $(s,w) \in \mathbb{R}_+ \times \mathbb{R}_-$. *Then, for every* $\varepsilon > 0$, *there exists a unique continuous non-positive solution* $\psi_\varepsilon\colon \mathbb{R}_+ \to \mathbb{R}_-$ *to the Riccati–Volterra equation*

$$\psi_\varepsilon(t) = \int_0^t \kappa_\varepsilon(t-s) G(s, \psi_\varepsilon(s)) ds, \quad t \geq 0. \tag{18}$$

Proof. See Appendix A. We notice that V_1 in (17) is well-defined, i.e., $V_1(w) < \infty$ for every $w \leq 0$, since $|e^{wx} - 1 - wx| \leq w^2 |x|^2$ for any $x \in \mathbb{R}_+$. □

Formally, the asymptotic solution to (18) comes from considering its Dirac limit as $\varepsilon \to 0$:

$$\psi_0(t) = \frac{1}{\lambda} G(t, \psi_0(t))$$

(recall from above that $\lambda\kappa_\varepsilon(\cdot)$ has Dirac-type behaviour as $\varepsilon \to 0$, so $\lambda\kappa_\varepsilon(t-s)$ will be concentrated at $s = t$). Re-arranging terms here, we obtain our conjecture limit equation:

$$f(t) - \lambda\psi_0(t) + \bar{G}(\psi_0(t)) = -\lambda\psi_0(t) + G(t, \psi_0(t)) = 0, \quad t \geq 0, \tag{19}$$

where G is defined as in (17) and $\bar{G}(w) = \frac{1}{2}\sigma^2 w^2 + \int_{\mathbb{R}_+} (e^{xw} - 1 - xw)\nu(dx)$ for $w \leq 0$. In the next lemma, we prove that (19) is well-posed.

Lemma 3.3. *For any* $f \in L^\infty_{\text{loc}}(\mathbb{R}_+;\mathbb{R}_-)$, *there exists a unique solution* ψ_0 *to* (19). *Furthermore,* $\psi_0 \in L^\infty_{\text{loc}}(\mathbb{R}_+;\mathbb{R}_-)$.

Proof. Since, by the definition in (17), $G(t,0) = f(t) \leq 0$ and the function $w \mapsto G(t,w) - \lambda w$ is continuous and decreasing on $\mathbb{R}_-$, with $G(t,w) - \lambda w \to \infty$ as $w \to -\infty$, we see that there exists a unique non-positive solution $\psi_0\colon \mathbb{R}_+ \to \mathbb{R}_-$ to (19). In particular, since f is locally bounded on $\mathbb{R}_+$, ψ_0 is locally bounded on $\mathbb{R}_+$, as well (i.e., $\psi_0 \in L^\infty_{\text{loc}}(\mathbb{R}_+;\mathbb{R}_-)$). □

We now provide the main technical lemma needed in our analysis. It clarifies in which sense the solution ψ_0 to (19) can be interpreted as the asymptotic solution to (18).

Lemma 3.4. *Consider the solutions* $(\psi_\varepsilon)_\varepsilon$ *and* ψ_0 *to* (18) *and* (19), *respectively. Then, for every* $T > 0$,

$$\lim_{\varepsilon \to 0} \psi_\varepsilon = \psi_0 \quad \text{in } L^1(0,T).$$

Proof. See Appendix B. □

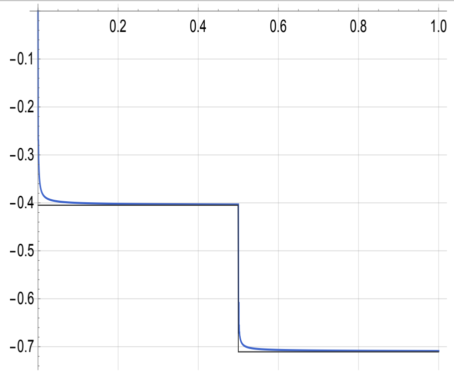

Figure 1. Here we have plotted ψ_ε in Eq. (20) (in blue) and ψ_0 in Eq. (19) (grey dashed), using an Adams scheme with 2000 time steps with $\varepsilon = .01$, $H = 0.2$, $\sigma = .4$, $\lambda = 1$, $T = 1$, $f(s) = -\frac{1}{2}(1_{\{s<1/2\}} + 1_{\{s\leq 1\}})$ and $\nu(x) = \frac{Ce^{-Mx}}{x^{1+Y}} 1_{\{x>0\}}$ for $C = 1$, $M = 3$ and $Y = 1.5$. Notice that such a $\nu(x)$ is the Lévy measure of a one-sided tempered stable (CGMY) process. We see convergence to ψ_0 (see e.g. [18] for details on refinements to Adams schemes). Numerically solving the VIE in Eq. (18) for $\varepsilon \ll 1$ appears to be much harder due to the Dirac nature of the kernel.

Remark 3.3. *A variation of constants argument, see Remark 5 in [15] and also Lemma 4.4 in [8], shows the equivalence between* (18) *and*

$$\varepsilon\psi_\varepsilon(\tau) = \int_0^\tau K(\tau-s)\Big(f(s) - \lambda\psi_\varepsilon(s) + \frac{1}{2}\sigma^2\psi_\varepsilon(s)^2 + V_1(\psi_\varepsilon(s))\Big)ds, \quad \tau \geq 0, \tag{20}$$

where $K(t) = \frac{1}{\Gamma(\alpha)}t^{\alpha-1}$ *is the fractional kernel. According to Lemma 3.4, the limiting solution as* $\varepsilon \to 0$ *is* ψ_0 *(we test this numerically in Figure 1).*

We now state the main theorem of this section. We note that this result, which concerns the convergence of the finite-dimensional distributions of the integrated variance process $A^\varepsilon_\cdot = \int_0^{(\cdot)} V^\varepsilon_s ds$, will be strengthened in Theorem 3.7 of Subsection 3.1, where weak convergence in the path space of càdlàg functions endowed with the M_1 topology is established. As in the proof of Theorem 2.1 in Section 2, the idea behind Theorem 3.5 consists in exploiting Lemma 3.4 to study the limit of the moment generating function of the finite-dimensional distributions of the re-scaled integrated variance process.

Theorem 3.5. *The finite-dimensional distributions of* $A^\varepsilon_\cdot = \int_0^{(\cdot)} V^\varepsilon_s ds$ *tend weakly to those of a time-changed Lévy process* $X_{g(\cdot)}$*, where* $X_t = \inf\{s : Z_s < -t\}$*,* $Z = (Z_t)_{t\geq 0}$ *is a Lévy process with Lévy triple* $(-\lambda, \sigma^2, \nu)$*, and* $g(t) = \lambda\int_0^t \xi_0^0(u)du$*.*

Remark 3.4. *X is a Lévy subordinator, see e.g. Theorem 46.2 in [37]; in particular, if we let* $(\tilde\lambda, 0, \tilde\nu)$ *denote the Lévy triple of X with respect to the truncation function* $h(x) \equiv 0$*, then X has no Gaussian component,* $\tilde\nu(\mathbb{R}_-) = 0$ *and* $\int_{[0,1]} x\tilde\nu(dx) < \infty$*, see e.g. Theorem 21.5 in [37]. Moreover, if* $\tilde\nu(dx)$ *has a density with respect to Lebesgue measure, i.e.* $\tilde\nu(dx) = \tilde\nu(x)dx$*, then the Lévy Khintchine-type formula in Theorem 25.17 in [37] for* X_T *is:*

$$\mathbb{E}[e^{-pX_T}] = e^{T(-\tilde\lambda p + \int_0^\infty (e^{-px}-1)\tilde\nu(x)dx)}$$

for $p \geq 0$. Thus $\tilde{V}(-p) := \log \mathbb{E}[e^{-pX_1}] = -\tilde{\lambda} p + \int_0^\infty (e^{-px} - 1)\tilde{\nu}(x)dx$, and differentiating both sides with respect to p, we obtain:

$$\tilde{V}'(-p) = \tilde{\lambda} + \int_0^\infty x e^{-px} \tilde{\nu}(x) dx$$

for $p \geq 0$, so in principle we can recover $\tilde{\nu}$ from $\tilde{V}$ by Laplace inversion.

Proof. Recall that the dynamics of the variance process V^ε are given in (13). Let $f : [0,\infty) \to (-\infty, 0]$ be a locally bounded function on $\mathbb{R}_+$ (i.e., $f \in L^\infty_{\text{loc}}(\mathbb{R}_+; \mathbb{R}_-)$). Then from Theorem 5 in [15] and Lemma 6.1 in [1] (see also Lemma 3.2 in [17]), we know that for every $T > 0$

$$M_t := e^{\int_0^t f(T-s) V_s^\varepsilon ds + G_t}, \quad t \in [0,T] \tag{21}$$

is a martingale if we set

$$G_t = \int_0^{T-t} G(u, \psi_\varepsilon(u))\, \xi_t^\varepsilon(T-u) du, \tag{22}$$

where G is defined in (17), $\xi_t^\varepsilon(u) = \mathbb{E}[V_u^\varepsilon | \mathcal{F}_t]$, $u > t$, is the forward variance (see (16) in Remark 3.2) and ψ_ε satisfies the corresponding non-linear Riccati–Volterra integral equation (18). Existence and uniqueness for ψ_ε are established in Lemma 3.2 above. Since $G_T = 0$ we see that

$$M_t = \mathbb{E}[M_T | \mathcal{F}_t] = \mathbb{E}\Big[e^{\int_0^T f(T-s) V_s^\varepsilon ds} | \mathcal{F}_t\Big].$$

In particular, at $t = 0$, taking the expected value we have

$$\mathbb{E}\Big[e^{\int_0^T f(T-s) V_s^\varepsilon ds}\Big] = e^{\int_0^T G(s, \psi_\varepsilon(s))\, \xi_0^\varepsilon(T-s) ds}. \tag{23}$$

As a corollary of Lemma 3.4, we see that

$$\begin{aligned} &\int_0^T |G(s, \psi_\varepsilon(s)) - G(s, \psi_0(s))| ds \\ &\quad \leq \int_0^T \Big(\frac{1}{2}\sigma^2 |\psi_\varepsilon(s) + \psi_0(s)| + |\tilde{h}(\psi_\varepsilon(s), \psi_0(s))|\Big) |\psi_\varepsilon(s) - \psi_0(s)| ds \\ &\quad \leq K_1 \int_0^T |\psi_\varepsilon(s) - \psi_0(s)| ds \xrightarrow[\varepsilon \to 0]{} 0, \end{aligned} \tag{24}$$

where $\tilde{h}$ is the function defined in (A-6), and

$$K_1 := \sup_{\varepsilon > 0} \Big\| \frac{1}{2}\sigma^2 |\psi_\varepsilon + \psi_0| + |\tilde{h}(\psi_\varepsilon, \psi_0)| \Big\|_{L^\infty(0,T)};$$

since ψ_0 is locally bounded on $\mathbb{R}_+$ by Lemma 3.3, K_1 is finite by (B-1) and $\tilde{h}$ is continuous, see Appendix A (in particular, the proof of Lemma A.3). Then

$$\Big| \int_0^T \Big(G(s, \psi_\varepsilon(s)) \xi_0^\varepsilon(T-s) - G(s, \psi_0(s)) \xi_0^0(T-s) \Big) ds \Big|$$

$$\leq \int_0^T |G(s,\psi_\varepsilon(s)) - G(s,\psi_0(s))|\, \xi_0^0(T-s)ds$$
$$+ \left(\sup_{\varepsilon>0} \|G(\cdot,\psi_\varepsilon)\|_{L^\infty(0,T)}\right) \int_0^T |\xi_0^\varepsilon(T-s) - \xi_0^0(T-s)|ds$$

where the last two terms tend to zero as $\varepsilon \to 0$ by (24) and Assumption 3.1. Hence from (19) and (23) we see that

$$\lim_{\varepsilon\to 0} \mathbb{E}\left[e^{\int_0^T f(T-s)V_s^\varepsilon ds}\right] = \lim_{\varepsilon\to 0} e^{\int_0^T G(s,\psi_\varepsilon(s))\xi_0^\varepsilon(T-s)ds}$$
$$= e^{\int_0^T G(s,\psi_0(s))\xi_0^0(T-s)ds} \;=\; e^{\lambda\int_0^T \psi_0(s)\xi_0^0(T-s)ds}. \tag{25}$$

We now characterize the process which has (25) as its Laplace transform.
For a Lévy process $Z = (Z_t)_{t\geq 0}$ with Lévy triple $(-\lambda, \sigma^2, \nu)$, $e^{uZ_t - \Lambda(u)t}$ is an $\mathcal{F}_t^Z$−martingale for $u \leq 0$, where

$$\Lambda(u) = -\lambda u + \frac{1}{2}\sigma^2 u^2 + V_1(u); \tag{26}$$

this is a consequence of the stationary and independent increments property, together with Theorem 25.17 in [37] (see (C-4) below for an analogous argument). From Proposition C.1 in Appendix C applied to the spectrally negative process $-Z$ (see also e.g. Theorem 46.3 in [37] or Eq. (2.5) in [35] for an alternative proof), we know that

$$\mathbb{E}[e^{-qH_a}] \;=\; e^{-\Lambda^{-1}(q)a} \;=\; e^{\Lambda^{-1}(q)|a|} \tag{27}$$

for $q \geq 0$ and $a \leq 0$, where $H_a := \inf\{t \geq 0 : Z_t < a\}$ and $\Lambda^{-1}(q) \leq 0$ denotes the inverse function of Λ.

Now let $X_t = H_{-t}$, which is a Lévy subordinator, see Remark 3.4 and the references therein. Then from the i.i.d. property for Lévy processes, (27) and the fundamental theorem of calculus, considering a right-continuous non-positive piecewise constant function f we have, for any continuously differentiable non-decreasing function g starting from 0,

$$\mathbb{E}\left[e^{\int_0^T f(T-s)d(X_{g(s)})}\right] \;=\; e^{\int_0^T \Lambda^{-1}(|f(T-s)|)g'(s)ds} \;=\; e^{\int_0^T \Lambda^{-1}(-f(s))g'(T-s)ds}$$
$$=\; e^{\int_0^T \psi_0(s)g'(T-s)ds}, \tag{28}$$

where in the last step we use that $\Lambda^{-1}(-f(\cdot))$ is $\psi_0(\cdot)$ by Eqs. (19) and (26) (see also Lemma 15.1 in [20], which is used in [3]).

At this point, for $0 \leq s_1 < \cdots < s_n < T$, choosing a right-continuous map f such that

$$f(T-s) = (u_1+u_2+\ldots+u_n)1_{\{0\leq s\leq s_1\}} + (u_2+\ldots+u_n)1_{\{s_1<s\leq s_2\}} + \ldots + u_n 1_{\{s_{n-1}<s\leq s_n\}}, \quad s \in [0,T]$$

with $u_1, u_2, \ldots, u_n \leq 0$, we see that

$$\int_0^T f(T-s)V_s^\varepsilon ds \;=\; u_1 A_{s_1}^\varepsilon + \ldots + u_n A_{s_n}^\varepsilon, \quad \int_0^T f(T-s)dX_{g(s)} \;=\; u_1 X_{g(s_1)} + \ldots + u_n X_{g(s_n)},$$

where $A_t^\varepsilon = \int_0^t V_s^\varepsilon ds$. Therefore by (25) and (28) (again by Problem 30.4 in [13]) we deduce that the finite-dimensional distributions of $A_\cdot^\varepsilon = \int_0^{(\cdot)} V_s^\varepsilon ds$ converge weakly to those of the time-changed Lévy process $X_{g(t)}$ with $g'(t) = \lambda\xi_0^0(t)$ and $g(0) = 0$. The proof is now complete. □

3.1. Weak functional convergence in the M_1 topology

The purpose of this subsection is to strengthen the convergence result of the finite-dimensional distributions established in Theorem 3.5 to a functional setting. More precisely, recalling the process $X_{g(\cdot)}$ and the integrated variances $A_\cdot^\varepsilon = \int_0^{(\cdot)} V_s^\varepsilon ds$ in the statement of Theorem 3.5, our aim is to prove the weak convergence in law of $A_\cdot^\varepsilon$ to $X_{g(\cdot)}$ as $\varepsilon \to 0$ in the path-space of càdlàg functions $D(\mathbb{R}_+;\mathbb{R})$, endowed with the M_1 topology. Notice that we work with trajectories defined on the whole non-negative half-line, instead of fixing a finite time horizon as in [3] or [6], where the M_1 topology is also used. We consider $D(\mathbb{R}_+;\mathbb{R})$ instead of the space of continuous functions $C(\mathbb{R}_+;\mathbb{R})$ – where $(A_\cdot^\varepsilon)_\varepsilon$ take values – because $X_{g(\cdot)}$ is a càdlàg process that, by Theorems 21.1 and 21.3 in [37], exhibits jumps with positive probability (unless in the trivial case $\xi_0^0(t) = 0$ for a.e. $t \geq 0$). The reason we focus on the M_1 topology rather than the arguably more common J_1 topology is that, as a consequence of, e.g., Proposition 2.1 a) in Chapter VI of [32], $C(\mathbb{R}_+;\mathbb{R})$ is a closed subset of $D(\mathbb{R}_+;\mathbb{R})$ in the J_1 topology. By the Portmanteau theorem, any weak limit probability of $(A_\cdot^\varepsilon)_\varepsilon$ is then supported on $C(\mathbb{R}_+;\mathbb{R})$, which again excludes the time-changed subordinator $X_{g(\cdot)}$.

Before stating our main result in Theorem 3.7, we recall some important definitions and properties of M_1. A comprehensive analysis of this and other topologies on spaces of càdlàg functions can be found in [39], see in particular Chapter 12. We start by presenting the definition of the topology $M_1^{(T)}$ on the space $D([0,T];\mathbb{R})$ of real-valued càdlàg functions defined on $[0,T]$, for every $T > 0$.

Definition 3.1. *Fix $T > 0$. Given $x \in D([0,T];\mathbb{R})$, the* thin graph *of x is defined by*

$$\Gamma_x = \{(z,t) \in \mathbb{R} \times [0,T] : z \in [x(t-), x(t)]\}, \quad \text{where } x(0-) = x(0).$$

The couple $(\Gamma_x, \leq)$ is a totally ordered set, where $\leq$ is the relation given by:

$$(z_1,t_1) \leq (z_2,t_2) \qquad \Longleftrightarrow \qquad \Big[t_1 < t_2\Big] \quad \text{or} \quad \Big[t_1 = t_2 \ \text{ and } \ |x(t_1-) - z_1| \leq |x(t_1-) - z_2|\Big].$$

A parametric representation *of x is a continuous nondecreasing function (u,r) mapping $[0,1]$ onto $(\Gamma_x, \leq)$; the set of such parametric representations is denoted by $\Pi(x)$.*
For every $x, y \in D([0,T];\mathbb{R})$, define the metric d_T as follows:

$$d_T(x,y) = \inf_{(u_1,r_1)\in\Pi(x),(u_2,r_2)\in\Pi(y)} \max\{\|u_1 - u_2\|_1, \|r_1 - r_2\|_1\}, \tag{29}$$

where $\|\cdot\|_1$ denotes the uniform norm in $[0,1]$. The $M_1^{(T)}$ topology is the topology generated by d_T.

We refer the reader to Theorem 12.3.1 in [39] for the proof that d_T is indeed a metric on $D([0,T];\mathbb{R})$.

In the sequel, for every $T > 0$, we denote by $r_T \colon D(\mathbb{R}_+;\mathbb{R}) \to D([0,T];\mathbb{R})$ the restriction operator given by $[r_T(x)](t) = x(t)$ for $t \in [0,T]$. This operator enables us to introduce the M_1 topology on $D(\mathbb{R}_+;\mathbb{R})$, hence for functions defined on the non-negative half-line; see also Section 12.9 in [39].

Definition 3.2. *The M_1 topology on $D(\mathbb{R}_+;\mathbb{R})$ is the topology generated by the metric*

$$d_{M_1}(x,y) = \int_0^\infty e^{-T} \min\left\{d_T(r_T(x), r_T(y)), 1\right\} dT, \tag{30}$$

where, for every $T > 0$, d_T is defined in (29).

Despite the metric d_T in (29) being incomplete, the space $(D([0,T];\mathbb{R}), M_1^{(T)})$ is Polish, see Section 12.8 in [39]. It follows that also $(D(\mathbb{R}_+;\mathbb{R}), M_1)$ is Polish, and we refer to Theorem 2.6 in [38] for an analogous argument for the J_1 topology. As a consequence of the definition of d_{M_1} in (30) and Theorem 12.3.2 in [39], the M_1 topology is coarser than the J_1 topology. Since they both provide Polish (hence Lusin) spaces, by Theorem 11.5.3 in [39] they generate the same $\sigma-$algebra, which coincides with the classical $\sigma-$algebra generated by the evaluation maps on $D(\mathbb{R}_+;\mathbb{R})$, see also Theorem 1.14 in Chapter VI of [32].

Denote by $\mathbb{P}^\varepsilon$ and $\mathbb{P}$ the probability measures induced by $A^\varepsilon_\cdot$ and $X_{g(\cdot)}$ on $\sigma(M_1)$, respectively. The next lemma enables us to reduce the study of the weak convergence of $\mathbb{P}^\varepsilon \implies \mathbb{P}$ on $(D(\mathbb{R}_+;\mathbb{R}), M_1)$ from an infinite to an arbitrary finite time horizon. More precisely, we have the following.

Lemma 3.6. *For any $T > 0$, let $\mathbb{P}r_T^{-1}$ be the pushforward probability measure on $\sigma(M_1^{(T)})$ of $\mathbb{P}$ by the restriction operator r_T, with an analogous notation for $\mathbb{P}^\varepsilon$. Then $\mathbb{P}^\varepsilon \implies \mathbb{P}$ on $(D(\mathbb{R}_+;\mathbb{R}), M_1)$ if and only if, for every $T > 0$, $\mathbb{P}^\varepsilon r_T^{-1} \implies \mathbb{P}r_T^{-1}$ on $(D([0,T];\mathbb{R}), M_1^{(T)})$.*

Proof. Since $X_{g(\cdot)}$ is a time-changed Lévy process with continuous time change g (we recall from the statement of Theorem 3.5 that $g(\cdot) = \lambda \int_0^{(\cdot)} \xi_0^0(u)du$), for any fixed $T > 0$,

$$\mathbb{P}\big(\{x \in D(\mathbb{R}_+;\mathbb{R}) : x(T) - x(T-) \neq 0\}\big) = 0.$$

Moreover, by Theorem 12.9.3 in [39], the restriction map r_T is continuous at all $x \in D(\mathbb{R}_+;\mathbb{R})$ that do not exhibit a jump at T. Combining these two facts, the proof can be completed using the same arguments (based on the continuous mapping and Portmanteau theorems) as in Theorem 2.8 of [38], which treats the case of the J_1 topology. □

At this point, we are ready to state the main result of this subsection. Its proof relies on Theorem 3.5, which ensures the convergence of the finite-dimensional distributions of $A^\varepsilon_\cdot$ to those of $X_{g(\cdot)}$, and tightness characterizations of sets of probability measures on $(D([0,T];\mathbb{R}), M_1^{(T)})$ for $T > 0$, established in Theorem 12.12.3 in [39]. These conditions are particularly simple to apply since the processes $(A^\varepsilon_\cdot)_{\varepsilon>0}$ are non-decreasing.

Theorem 3.7. *The integrated variance processes $A^\varepsilon_\cdot$ converge weakly to the time-changed Lévy process $X_{g(\cdot)}$ defined in Theorem 3.5 on $(D(\mathbb{R}_+;\mathbb{R}), M_1)$, that is, $\mathbb{P}^\varepsilon \implies \mathbb{P}$ on $(D(\mathbb{R}_+;\mathbb{R}), M_1)$ as $\varepsilon \to 0$.*

Proof. According to Lemma 3.6, it is sufficient to prove that

$$\mathbb{P}^\varepsilon r_T^{-1} \implies \mathbb{P}r_T^{-1} \text{ on } (D([0,T];\mathbb{R}), M_1^{(T)}), \text{ for every } T > 0. \tag{31}$$

In turn, by Theorem 3.5 and Theorem 11.6.6 in [39], (31) is satisfied if the family $(\mathbb{P}^\varepsilon r_T^{-1})_\varepsilon$ is tight on $(D([0,T];\mathbb{R}), M_1^{(T)})$. According to Theorem 12.12.3 in [39], sufficient (and necessary) conditions for tightness in the space $(D([0,T];\mathbb{R}), M_1^{(T)})$ are given by

$$\lim_{R\to\infty} \sup_{\varepsilon>0} \mathbb{P}\Big(\sup_{t\in[0,T]} |A_t^\varepsilon| > R\Big) = 0$$

and

$$\lim_{\delta\to 0} \sup_{\varepsilon>0} \mathbb{P}(w_T(A^\varepsilon_\cdot, \delta) > \eta) = 0, \quad \text{for any fixed } \eta > 0.$$

Here, denoting by $\mathrm{dist}(t_2,[t_1,t_3]) = \inf_{t\in[t_1,t_3]}|t_2-t|$ the distance between t_2 and the segment joining t_1 and t_3 (with $t_i \in \mathbb{R}$, $i=1,2,3$), we set, for any $\delta>0$,

$$\begin{aligned} w_T(A^\varepsilon_\cdot,\delta) &= \max\Big\{\sup_{0\le t\le T} w'_T(A^\varepsilon_\cdot,t,\delta), v_T(A^\varepsilon_\cdot,0,\delta), v_T(A^\varepsilon_\cdot,T,\delta)\Big\};\\ w'_T(A^\varepsilon_\cdot,t,\delta) &= \sup_{\max\{0,t-\delta\}\le t_1<t_2<t_3\le\min\{t+\delta,T\}} \mathrm{dist}\big(A^\varepsilon_{t_2},[A^\varepsilon_{t_1},A^\varepsilon_{t_3}]\big), \quad t\in[0,T];\\ v_T(A^\varepsilon_\cdot,t,\delta) &= \sup_{\max\{0,t-\delta\}\le t_1\le t_2\le\min\{t+\delta,T\}} |A^\varepsilon_{t_1}-A^\varepsilon_{t_2}|, \quad t\in[0,T]. \end{aligned}$$

Since the processes $A^\varepsilon_\cdot$ start from 0 and are non-decreasing, being obtained by integrating the non-negative variance processes $V^\varepsilon_\cdot$, as already observed in Section 4.3 of [3], these conditions simplify. In particular, the former reduces to

$$\lim_{R\to\infty}\sup_{\varepsilon>0}\mathbb{P}\big(A^\varepsilon_T>R\big)=0. \tag{32}$$

As for the latter, considering that, for every $t\in[0,T]$, $v_T(A^\varepsilon_\cdot,t,\delta) = A^\varepsilon_{\min\{t+\delta,T\}} - A^\varepsilon_{\max\{t-\delta,0\}}$ and

$$A^\varepsilon_{t_2}\in[A^\varepsilon_{t_1},A^\varepsilon_{t_3}] \text{ for all } t_1\le t_2\le t_3, \quad \text{hence } w'_T(A^\varepsilon_\cdot,t,\delta)=0,$$

it reads

$$\lim_{\delta\to0}\sup_{\varepsilon>0}\mathbb{P}\Big(\max\big\{A^\varepsilon_\delta, A^\varepsilon_T-A^\varepsilon_{T-\delta}\big\}>\eta\Big)=0, \quad \text{for all fixed } \eta>0. \tag{33}$$

To complete the proof, it will then be sufficient to verify (32) and (33). For every $\varepsilon>0$ and $T\ge0$, integrating (13) on $[0,T]$, an application of the stochastic Fubini theorem yields (see Lemma 3.2 in [4], and also Lemma 2.1 in [1])

$$A^\varepsilon_T = \int_0^T \xi^\varepsilon_0(s)ds + \int_0^T \kappa_\varepsilon(T-s)\Big(\sigma\int_0^s\sqrt{V^\varepsilon_r}dW_r+\tilde{J}^\varepsilon_s\Big)ds,$$

where we recall that $\tilde{J}^\varepsilon=(\tilde{J}^\varepsilon_t)_{t\ge0}$ is the process defined in (15). Taking expectations,

$$\mathbb{E}[A^\varepsilon_T] = \int_0^T \xi^\varepsilon_0(s)ds + \int_0^T \kappa_\varepsilon(T-s)\mathbb{E}\Big[\sigma\int_0^s\sqrt{V^\varepsilon_r}dW_r+\tilde{J}^\varepsilon_s\Big]ds = \int_0^T\xi^\varepsilon_0(s)ds, \tag{34}$$

where the last step follows from the martingale property of the processes $\int_0^\cdot\sqrt{V^\varepsilon_r}dW_r$ and $\tilde{J}^\varepsilon$, ensured by (14) and (15), respectively. Since by Assumption 3.1 the maps $\xi^\varepsilon_0(\cdot)$ are uniformly bounded, by Markov's inequality, (34) implies (32). As for (33), it follows by the same arguments, once we notice that, by (34),

$$\sup_{\varepsilon>0}\max\Big\{\mathbb{E}[A^\varepsilon_\delta],\mathbb{E}[A^\varepsilon_T-A^\varepsilon_{T-\delta}]\Big\}\le\delta\sup_{\varepsilon>0,t\in[0,T]}|\xi^\varepsilon_0(t)|\longrightarrow0 \quad \text{as } \delta\to0.$$

Therefore, the proof is complete. □

Remark 3.5. *If we consider the continuous solution $V^\varepsilon_\cdot$ to* (10) *without jumps, i.e., $\nu(dx)=0$ (hence $\tilde{J}^\varepsilon=0$), then $V^\varepsilon_\cdot$ coincides with* (5) *in Theorem 2.1 of Section 2. Since* (10) *is equivalent to* (11)*, which in turn is a particular case of* (13)*, an application of Theorem 3.5 yields that the finite-dimensional* *distributions of $A^{\varepsilon}_{\cdot} = \int_0^{(\cdot)} V^{\varepsilon}_s \, ds$ converge to those of $X_{\lambda\theta(\cdot)}$, where $X_{\cdot}$ is the Lévy subordinator given by*

$$X_t = \inf\left\{s \geq 0 : B_s + \frac{\lambda}{\sigma} s = \frac{1}{\sigma} t\right\},$$

for a standard one-dimensional Brownian motion B. By (4), we infer that $X_{\lambda\theta(\cdot)}$ is an IG *process with parameters $(\theta, \sigma^{-2}\lambda^2\theta^2)$. Therefore, in the case of continuous dynamics, Theorem 3.5 extends Theorem 2.1 by proving the convergence of the finite-dimensional distributions, rather than only that of the marginals at a fixed time t. Furthermore, Theorem 3.7 implies that the laws of $A^{\varepsilon}_{\cdot}$ converge, as $\varepsilon \to 0$, to the law of $X_{\lambda\theta(\cdot)}$ weakly on the path space $(D(\mathbb{R}_+;\mathbb{R}), M_1)$.*

Appendix A: Proof of Lemma 3.2

We first recall the Riccati–Volterra integral equation in (17)-(18), which we re-write as

$$\begin{aligned}\psi_{\varepsilon}(t) &= \int_0^t \kappa_{\varepsilon}(t-s) f(s) ds + \int_0^t \kappa_{\varepsilon}(t-s)\left(\frac{1}{2}\sigma^2\psi_{\varepsilon}^2(s) + \int_{\mathbb{R}_+} (e^{x\psi_{\varepsilon}(s)} - 1 - x\psi_{\varepsilon}(s))\nu(dx)\right) ds \\ &= (\kappa_{\varepsilon} * f)(t) + \left(\kappa_{\varepsilon} * \left(\frac{1}{2}\sigma^2\psi_{\varepsilon}^2 + V_1(\psi_{\varepsilon})\right)\right)(t), \quad t \geq 0. \end{aligned} \tag{A-1}$$

By Theorem 3.1, Chapter 5 in [30], κ_{ε} is completely monotone, and by Theorem 2.2, Chapter 2 in [30], $t \mapsto (\kappa_{\varepsilon} * f)(t)$ is continuous on $\mathbb{R}_+$ (as $f \in L^{\infty}_{\text{loc}}(\mathbb{R}_+;\mathbb{R}_-)$). Moreover, the function $\widetilde{G}$ defined by the relation

$$\widetilde{G}(w) - \frac{1}{2}\sigma^2 w^2 = \begin{cases} 0, & w > 0 \\ \displaystyle\int_{\mathbb{R}_+} (e^{xw} - 1 - xw)\nu(dx), & w \leq 0 \end{cases}$$

is continuous and non-negative on $\mathbb{R}$. Then by Theorem 1.1, Chapter 12 in [30], there exists a continuous noncontinuable local solution $\widetilde{\psi}_{\varepsilon}$ of the equation

$$\widetilde{\psi}_{\varepsilon}(t) = (\kappa_{\varepsilon} * f)(t) + (\kappa_{\varepsilon} * \widetilde{G}(\widetilde{\psi}_{\varepsilon}))(t), \quad t \in [0, T_{\max}), \tag{A-2}$$

for some $T_{\max} > 0$. Then, noting that $\widetilde{G} = \frac{1}{2}\sigma^2(\cdot)^2 + V_1(\cdot)$ on $\mathbb{R}_-$, from the following lemma we know that $\widetilde{\psi}_{\varepsilon}$ also solves (A-1) on $[0, T_{\max})$:

Lemma A.1. *$\widetilde{\psi}_{\varepsilon}$ is non-positive.*

Proof. For convenience we define $j : \mathbb{R} \to \mathbb{R}_-$ by

$$j(w) = \begin{cases} \displaystyle\frac{1}{w}\int_{\mathbb{R}_+} (e^{wx} - 1 - wx)\nu(dx), & w < 0, \\ 0, & w \geq 0; \end{cases} \tag{A-3}$$

(which is continuous and non-positive), so $\frac{1}{2}\sigma^2 w^2 + w \cdot j(w) = \widetilde{G}(w)$ for $w \in \mathbb{R}$.

Then, for every $T \in (0, T_{\max})$, from (A-2)

$$\widetilde{\psi}_{\epsilon}(t) = (\kappa_{\varepsilon} * f)(t) + \int_0^t \kappa_{\varepsilon}(t-s)\left(\frac{1}{2}\sigma^2\widetilde{\psi}_{\varepsilon}(s) + j(\widetilde{\psi}_{\varepsilon}(s))\right)\widetilde{\psi}_{\varepsilon}(s) ds, \quad t \in [0, T].$$

By Remark B.6 in [7] (which allows to consider a possibly discontinuous function f) and recalling that $f \leq 0$, this reformulation enables us to use Theorem C.1 in [7] and conclude that $\widetilde{\psi}_\varepsilon \leq 0$ in $[0, T_{\max})$, as T is arbitrary. □

Thanks to the discussion above and Lemma A.1, we can consider a continuous noncontinuable non-positive local solution ψ_ε to (A-1). In the next lemma, we show that ψ_ε is in fact a global solution, i.e., it is defined on the entire $\mathbb{R}_+$.

Lemma A.2. *Any continuous noncontinuable $\mathbb{R}_-$−valued solution ψ_ε to (A-1) is globally defined on $\mathbb{R}_+$.*

Proof. Recalling the definition of V_1 in (17), we notice that $\frac{1}{2}\sigma^2\psi_\varepsilon^2 + V_1(\psi_\varepsilon) \geq 0$ on $[0, T_{\max})$, so (A-1) yields

$$(\kappa_\varepsilon * f)(t) \leq \psi_\varepsilon(t) \leq 0, \quad t \in [0, T_{\max}). \tag{A-4}$$

Since $\kappa_\varepsilon * f$ is continuous and dominates ψ_ε on $[0, T_{\max})$, ψ_ε cannot explode to $-\infty$ at $T_{\max}$. Then, given that by Theorem 1.1, Chapter 12 in [30],

$$\limsup_{t \to T_{\max}} |\psi_\varepsilon(t)| = \infty \quad \text{if } T_{\max} < \infty,$$

we conclude that $T_{\max} = \infty$. □

To complete the proof of Lemma 3.2, it only remains to establish uniqueness for (A-1). We do this in the following lemma.

Lemma A.3. *If ψ_1 and ψ_2 are two continuous $\mathbb{R}_-$−valued global solutions of (A-1), then $\psi_1 = \psi_2$ on $\mathbb{R}_+$.*

Proof. Consider ψ_1 and ψ_2 as in the statement of the lemma. Then setting $\delta = \psi_1 - \psi_2$,

$$\begin{aligned}\delta(t) = \int_0^t \kappa_\varepsilon(t-s)\Bigg(&\frac{1}{2}\sigma^2(\psi_1(s) + \psi_2(s))\delta(s) \\ &+ \int_{\mathbb{R}_+}\Big(e^{x\psi_1(s)} - e^{x\psi_2(s)} - x(\psi_1(s) - \psi_2(s))\Big)\nu(dx)\Bigg)ds.\end{aligned} \tag{A-5}$$

Now let $\mathbb{R}_-^2 = \{(w_1, w_2) \in \mathbb{R}^2, w_1 \leq 0 \text{ and } w_2 \leq 0\}$ denote the negative quadrant of the plane, and define the auxiliary function $\tilde{h} \colon \mathbb{R}_-^2 \to \mathbb{R}$ by

$$\tilde{h}(w_1, w_2) = \begin{cases} \dfrac{1}{w_1 - w_2}\displaystyle\int_{\mathbb{R}_+}(e^{xw_1} - e^{xw_2} - x(w_1 - w_2))\nu(dx), & w_1 \neq w_2, \\ \displaystyle\int_{\mathbb{R}_+} x(e^{xw_1} - 1)\nu(dx), & \text{otherwise}. \end{cases} \tag{A-6}$$

The map $\tilde{h}$ is non-positive on its domain $\mathbb{R}_-^2$, because $w \mapsto e^{xw} - xw$ is nonincreasing on $\mathbb{R}_-$ for every $x \in \mathbb{R}_+$. By the dominated convergence theorem, $\tilde{h}$ is continuous on $\mathbb{R}_-^2 \setminus \{(w, w) : w \leq 0\}$. To prove the continuity at the points $(w_1, w_2) \in \mathbb{R}_-^2$ with $w_1 = w_2$, consider two sequences $(w_{1,n})_n, (w_{2,n})_n \subset \mathbb{R}_-$ such that

$$w_{1,n} \neq w_{2,n} \quad \text{and} \quad \lim_{n\to\infty} w_{1,n} = w_1 = w_2 = \lim_{n\to\infty} w_{2,n}.$$

Without loss of generality, suppose that $w_{1,n} > w_{2,n}$ for every $n \in \mathbb{N}$. We then compute, using the inequality $|e^u - 1 - u| \le |u|^2$ for $u \in \mathbb{R}_-$,

$$\begin{aligned}
&\left|\tilde{h}(w_{1,n}, w_{2,n}) - \tilde{h}(w_1, w_2)\right| \\
&\quad\le \frac{1}{|w_{1,n} - w_{2,n}|}\int_{\mathbb{R}_+} e^{xw_{1,n}}\left|e^{x(w_{2,n}-w_{1,n})} - 1 - x(w_{2,n} - w_{1,n})\right|\nu(dx) + \int_{\mathbb{R}_+} x\left|e^{xw_{1,n}} - e^{xw_1}\right|\nu(dx) \\
&\quad\le \left(\int_{\mathbb{R}_+} |x|^2\nu(dx)\right)|w_{2,n} - w_{1,n}| + \mathrm{o}(1) \underset{n\to\infty}{\longrightarrow} 0.
\end{aligned}$$

Considering that $\tilde{h}(w_{1,n}, w_{1,n}) \to \tilde{h}(w_1, w_2)$ by dominated convergence, the previous computations prove the continuity of $\tilde{h}$ at (w_1, w_2) with $w_1 = w_2$.

From the definition of $\tilde{h}$, we have

$$\int_{\mathbb{R}_+}\left(e^{x\psi_1(s)} - e^{x\psi_2(s)} - x(\psi_1(s) - \psi_2(s))\right)\nu(dx) = \tilde{h}(\psi_1(s), \psi_2(s))\delta(s), \quad s \ge 0,$$

hence (from (A-5)) δ solves the linear VIE:

$$\delta(t) = \int_0^t \kappa_\varepsilon(t-s)\left(\frac{1}{2}\sigma^2(\psi_1(s) + \psi_2(s)) + \tilde{h}(\psi_1(s), \psi_2(s))\right)\delta(s)ds.$$

This equation admits $\delta \equiv 0$ as its unique solution by the first part of Theorem C.1 in [7], whence we conclude that $\psi_1 = \psi_2$. The proof is now complete. ☐

Appendix B: Proof of Lemma 3.4

We recall that f is a locally bounded non-positive function defined on $\mathbb{R}_+$, which we also write as $f \in L^\infty_{\text{loc}}(\mathbb{R}_+; \mathbb{R}_-)$. Moreover, for every $\varepsilon > 0$, ψ_ε is the unique solution of the deterministic VIE (17)-(18), whose well-posedness is established in Lemma 3.2. The arguments that follow also rely on the following property of the kernel κ_ε, stated at the beginning of Section 3:

$$\|\kappa_\varepsilon\|_{L^1(0,\infty)} = \int_0^\infty \kappa_\varepsilon(t)dt = \frac{1}{\lambda}.$$

B.1. Relative compactness in L^1

Fix $T > 0$ and a sequence $(\varepsilon_n)_n \subset (0, \infty)$ which converges to 0.

Lemma B.1. *$(\psi_{\varepsilon_n})_n$ admits a convergent subsequence in $L^1(0, T)$.*

Proof. $(\psi_{\varepsilon_n})_n$ is bounded in $L^1(0,T)$ because (from (A-4)) we know that

$$\|\psi_\varepsilon\|_{L^\infty(0,T)} \le \|\kappa_\varepsilon * f\|_{L^\infty(0,T)} \le \|\kappa_\varepsilon\|_{L^1(0,T)}\|f\|_{L^\infty(0,T)} \le \frac{1}{\lambda}\|f\|_{L^\infty(0,T)}. \tag{B-1}$$

Now define the extended maps $\bar{\psi}_n : \mathbb{R} \to \mathbb{R}_-$ by

$$\bar{\psi}_n(t) = \begin{cases} \psi_{\varepsilon_n}(t), & t \in [0,T], \\ 0, & \text{otherwise.} \end{cases}$$

To prove the lemma, by the Kolmogorov-Riesz-Fréchet theorem (see e.g. Theorem 4.26 in [19]), it suffices to show that

$$\lim_{h\to 0} \|\tau_h\bar{\psi}_n - \bar{\psi}_n\|_{L^1(\mathbb{R})} = 0 \quad \text{uniformly in } n, \tag{B-2}$$

where τ_h is the translation operator defined by $\tau_h g(x) = g(x+h)$ for an arbitrary function $g : \mathbb{R} \to \mathbb{R}$.

Consider the case $h > 0$; when $h < T$, by (A-1), recalling (A-3) and (A-6),

$$\begin{aligned} \tau_h\bar{\psi}_n(t) - \bar{\psi}_n(t) = (\kappa_{\varepsilon_n} * f)(t+h) - (\kappa_{\varepsilon_n} * f)(t) \qquad\qquad & \\ + \int_t^{t+h} \kappa_{\varepsilon_n}(s)\Big(\frac{1}{2}\sigma^2\bar{\psi}_n(t+h-s) + j(\bar{\psi}_n(t+h-s))\Big)\bar{\psi}_n(t+h-s)ds & \\ + \int_0^t \kappa_{\varepsilon_n}(s)\phi_n(t-s;h)(\tau_h\bar{\psi}_n(t-s) - \bar{\psi}_n(t-s))ds & \\ := \mathbf{I}_{n,h}(t) + \mathbf{II}_{n,h}(t) + (\kappa_{\varepsilon_n} * (\phi_n(\cdot;h)(\tau_h\bar{\psi}_n - \bar{\psi}_n)))(t), \quad t \in [0, T-h], & \end{aligned}$$

where $\mathbf{I}_{n,h}(t)$ and $\mathbf{II}_{n,h}(t)$ refer to the first and second lines respectively on the right hand side here, and

$$\phi_n(t;h) = \frac{1}{2}\sigma^2(\tau_h\bar{\psi}_n(t) + \bar{\psi}_n(t)) + \tilde{h}(\tau_h\bar{\psi}_n(t), \bar{\psi}_n(t)), \quad t \in \mathbb{R}. \tag{B-3}$$

Hence $\chi = \tau_h\bar{\psi}_n - \bar{\psi}_n - \mathbf{I}_{n,h} - \mathbf{II}_{n,h}$ solves the linear VIE

$$\chi = \kappa_{\varepsilon_n} * (\phi_n(\cdot;h)(\tau_h\bar{\psi}_n - \bar{\psi}_n)) = \kappa_{\varepsilon_n} * (\phi_n(\cdot;h)\chi + \phi_n(\cdot;h)(\mathbf{I}_{n,h} + \mathbf{II}_{n,h}))$$

on the interval $[0, T-h]$. $\mathbf{I}_{n,h}$, $\mathbf{II}_{n,h}$ and $\phi_n(\cdot;h)$ are continuous on $[0, T-h]$, so (given that $\phi_n(\cdot;h)$ is non-positive), Theorem C.3 in [7] implies that

$$|(\tau_h\bar{\psi}_n - \bar{\psi}_n)(t)| \le |\mathbf{I}_{n,h}(t)| + |\mathbf{II}_{n,h}(t)| + (\kappa_{\varepsilon_n} * |\phi_n(\cdot;h)(\mathbf{I}_{n,h} + \mathbf{II}_{n,h})|)(t), \quad t \in [0, T-h]. \tag{B-4}$$

We compute

$$\begin{aligned} \int_0^{T-h}\left(\int_t^{t+h} \kappa_{\varepsilon_n}(s)|f(t+h-s)|ds\right)dt &= \int_0^{T-h}\left(\int_0^T 1_{\{s<t+h\}}1_{\{s>t\}}\kappa_{\varepsilon_n}(s)|f(t+h-s)|ds\right)dt \\ &= \int_0^T \kappa_{\varepsilon_n}(s)\left(\int_0^{T-h} 1_{\{t<s\}}1_{\{t>s-h\}}|f(t+h-s)|dt\right)ds \quad \text{(by Tonelli)} \\ &\le \int_0^h \kappa_{\varepsilon_n}(s)\left(\int_0^s |f(t+h-s)|dt\right)ds + \int_h^T \kappa_{\varepsilon_n}(s)\left(\int_{s-h}^s |f(t+h-s)|dt\right)ds \end{aligned}$$

$$\le \frac{2}{\lambda} h \|f\|_{L^\infty(0,T)}, \tag{B-5}$$

where h appears in the final term since both inner integrals have range $\le h$. Thus, denoting by $\bar{f} = f 1_{[0,T]} \in L^\infty(\mathbb{R})$ and using Theorem 2.2, Chapter 2 in [30],

$$\begin{aligned}
&\int_0^{T-h} |\mathbf{I}_{n,h}(t)|dt \\
&\quad\le \int_0^{T-h}\int_0^t \Big(\kappa_{\varepsilon_n}(s)|(\tau_h f - f)(t-s)|ds\Big)dt + \int_0^{T-h}\Big(\int_t^{t+h} \kappa_{\varepsilon_n}(s)|f(t+h-s)|ds\Big)dt \\
&\quad\le \Big(\int_0^{T-h} \kappa_{\varepsilon_n}(s)ds\Big)\|\tau_h \bar{f} - \bar{f}\|_{L^1(\mathbb{R})} + \frac{2}{\lambda} h \|f\|_{L^\infty(0,T)} \le \frac{1}{\lambda}\|\tau_h \bar{f} - \bar{f}\|_{L^1(\mathbb{R})} + \frac{2}{\lambda} h \|\bar{f}\|_{L^\infty(\mathbb{R})}.
\end{aligned}$$

Since these estimates do not depend on $n \in \mathbb{N}$, the continuity of the translation in $L^1(\mathbb{R})$ (see, for instance, Lemma 4.3 in [19]) yields that

$$\lim_{h\to 0+} \int_0^{T-h} |\mathbf{I}_{n,h}(t)|dt = 0 \quad \text{uniformly in } n. \tag{B-6}$$

Given that $(\bar{\psi}_n)_n$ is bounded in $L^\infty(\mathbb{R})$ by (B-1) and $j(\cdot)$ (defined in (A-3)) is continuous on $\mathbb{R}$, the same computations as in (B-5) (but with 1 in place of f) show that

$$\lim_{h\to 0+} \int_0^{T-h} |\mathbf{II}_{n,h}(t)|dt = 0 \quad \text{uniformly in } n. \tag{B-7}$$

Then (again by (B-1) and the continuity of $\tilde{h}$), there exists a constant $C > 0$ such that

$$|\phi_n(t;h)| \le C, \qquad t,\ h \in \mathbb{R},\ n \in \mathbb{N},$$

where $\phi_n(\cdot;h)$ is defined in (B-3). Therefore

$$\int_0^{T-h} |(\kappa_{\varepsilon_n} * |\phi_n(\cdot;h)(\mathbf{I}_{n,h} + \mathbf{II}_{n,h})|)(t)|dt \le C\frac{1}{\lambda}\int_0^{T-h} (|\mathbf{I}_{n,h}(t)| + |\mathbf{II}_{n,h}(t)|)dt,$$

whence (by (B-6) and (B-7)),

$$\lim_{h\to 0+} \int_0^{T-h} |(\kappa_{\varepsilon_n} * |\phi_n(\cdot;h)(\mathbf{I}_{n,h} + \mathbf{II}_{n,h})|)(t)|dt = 0 \quad \text{uniformly in } n. \tag{B-8}$$

Combining (B-6), (B-7) and (B-8) in (B-4) we deduce that

$$\lim_{h\to 0+} \|\tau_h \bar{\psi}_n - \bar{\psi}_n\|_{L^1(0,T-h)} = 0 \quad \text{uniformly in } n. \tag{B-9}$$

On the interval $[-h, 0]$ the maps $\bar{\psi}_n$ equal 0, hence, by (B-1),

$$\|\tau_h \bar{\psi}_n - \bar{\psi}_n\|_{L^1(-h,0)} = \int_{-h}^0 |\tau_h \bar{\psi}_n(t)|dt \le \Big(\sup_n \|\bar{\psi}_n\|_{L^\infty(\mathbb{R})}\Big) h \xrightarrow[h\to 0+]{} 0 \quad \text{uniformly in } n.$$

In a similar way, on the interval $[T-h,T]$

$$\|\tau_h\bar{\psi}_n - \bar{\psi}_n\|_{L^1(T-h,T)} = \int_{T-h}^{T} |\bar{\psi}_n(t)|dt \le \Big(\sup_n \|\bar{\psi}_n\|_{L^\infty(\mathbb{R})}\Big)h \xrightarrow[h\to 0+]{} 0 \quad \text{uniformly in } n.$$

The three previous equations yield

$$\lim_{h\to 0+} \|\tau_h\bar{\psi}_n - \bar{\psi}_n\|_{L^1(\mathbb{R})} = 0 \quad \text{uniformly in } n.$$

When $h<0$, assuming without loss of generality that $|h|<T$ we can simply write

$$\begin{aligned}\|\tau_h\bar{\psi}_n - \bar{\psi}_n\|_{L^1(\mathbb{R})} &= \|\bar{\psi}_n\|_{L^1(0,|h|)} + \|\tau_h\bar{\psi}_n\|_{L^1(T,T+|h|)} + \|\bar{\psi}_n - \tau_h\bar{\psi}_n\|_{L^1(|h|,T)} \\ &\le 2\Big(\sup_n \|\bar{\psi}_n\|_{L^\infty(\mathbb{R})}\Big)|h| + \|\tau_{|h|}\bar{\psi}_n - \bar{\psi}_n\|_{L^1(0,T-|h|)} \xrightarrow[h\to 0-]{} 0 \quad \text{uniformly in } n,\end{aligned}$$

where we use (B-9) for the last limit. Therefore (B-2) is verified and the proof is complete. □

B.2. Characterization of the limit points of ψ_ε

Lemma B.2. *For every $T>0$ and $g\in L^1(0,T)$,*

$$\lim_{\varepsilon\to 0}\int_0^T \Big|(\kappa_\varepsilon * g)(t) - \frac{1}{\lambda}g(t)\Big|dt = 0,$$

i.e. $\kappa_\varepsilon * g$ converges to $\frac{1}{\lambda}g$ in $L^1(0,T)$ as $\varepsilon\to 0$.

Proof. Let $c>0$. By the continuity of the translation in $L^1(\mathbb{R})$ (see e.g. Lemma 4.3 in [19]), there exists an $\eta=\eta(c)\in(0,T)$ such that, defining $\bar{g}=g1_{[0,T]}\in L^1(\mathbb{R})$, $\int_0^T|\bar{g}(t-s)-\bar{g}(t)|dt<c$, and hence

$$\int_s^T |g(t-s)-g(t)|dt < c,$$

for $s\in(0,\eta)$. Then from Tonelli's theorem and some straightforward manipulations,

$$\begin{aligned}\int_0^T \Big|(\kappa_\varepsilon * g)(t) - \frac{1}{\lambda}g(t)\Big|dt &\le \int_0^T \bigg(\int_0^t \kappa_\varepsilon(s)|g(t-s)-g(t)|ds + |g(t)|\Big(\frac{1}{\lambda} - \int_0^t \kappa_\varepsilon(s)ds\Big)\bigg)dt \\ &= \bigg\{\int_0^\eta + \int_\eta^T\bigg\}\kappa_\varepsilon(s)\Big(\int_s^T |g(t-s)-g(t)|dt\Big)ds + \int_0^T |g(t)|\Big(\frac{1}{\lambda} - \int_0^t \kappa_\varepsilon(s)ds\Big)dt \\ &\le \frac{1}{\lambda}c + 2\|g\|_{L^1(0,T)}\int_\eta^T \kappa_\varepsilon(s)ds + \int_0^T |g(t)|\Big(\frac{1}{\lambda} - \int_0^t \kappa_\varepsilon(s)ds\Big)dt,\end{aligned}$$

and hence (by the dominated convergence theorem and (12)),

$$\limsup_{\varepsilon\to 0}\int_0^T \Big|(\kappa_\varepsilon * g)(t) - \frac{1}{\lambda}g(t)\Big|dt \le \frac{1}{\lambda}c.$$

Since c can be chosen arbitrarily small the proof is complete. □

Now recall the $\varepsilon = 0$ solution ψ_0 defined in (19), see Lemma 3.3. Then we have the following:

Lemma B.3. *Consider* $T > 0$ *and a sequence* $(\varepsilon_n)_n \subset (0,\infty)$ *which converges to* 0. *Suppose that there exists a non-positive function* $\bar{\psi} \in L^1(0,T) \cap L^\infty(0,T)$ *such that* $\psi_{\varepsilon_n} \to \bar{\psi}$ *in* $L^1(0,T)$. *Then* $\bar{\psi} = \psi_0$ *a.e. in* $(0,T)$.

Proof. Recall our original VIE in (A-1): $\psi_\varepsilon = \kappa_\varepsilon * f + \kappa_\varepsilon * \bar{G}(\psi_\varepsilon)$. Multiplying by λ, taking the difference with (19), and adding and subtracting $\lambda(\kappa_{\varepsilon_n} * \bar{G}(\bar{\psi}))(t)$, we see that

$$\begin{aligned}&\lambda(\psi_{\varepsilon_n}(t) - \psi_0(t))\\ &\quad = \lambda(\kappa_{\varepsilon_n} * f)(t) - f(t) + \lambda(\kappa_{\varepsilon_n} * (\bar{G}(\psi_{\varepsilon_n}) - \bar{G}(\bar{\psi})))(t) + \lambda(\kappa_{\varepsilon_n} * \bar{G}(\bar{\psi}))(t) - \bar{G}(\psi_0(t))\end{aligned} \tag{B-10}$$

for every $t \in [0,T]$. By Lemma B.2, considering that $\bar{G}(\bar{\psi}(\cdot))$ belongs to $L^1(0,T)$ because $\bar{\psi} \in L^1(0,T) \cap L^\infty(0,T) = L^\infty(0,T)$,

$$\lim_{n\to\infty} \|\lambda(\kappa_{\varepsilon_n} * f) - f\|_{L^1(0,T)} = 0 \qquad \text{and} \qquad \lim_{n\to\infty} \|\lambda(\kappa_{\varepsilon_n} * \bar{G}(\bar{\psi})) - \bar{G}(\bar{\psi}(\cdot))\|_{L^1(0,T)} = 0,$$

and, using the map $\tilde{h}$ defined in (A-6) in the proof of Lemma A.3,

$$\begin{aligned}\lambda\|\kappa_{\varepsilon_n} * (\bar{G}(\psi_{\varepsilon_n}) - \bar{G}(\bar{\psi}))\|_{L^1(0,T)} \quad &\le \quad \lambda\|\kappa_{\varepsilon_n}\|_{L^1(0,T)}\|\bar{G}(\psi_{\varepsilon_n}) - \bar{G}(\bar{\psi})\|_{L^1(0,T)}\\ \le \sup_{n\in\mathbb{N}} &\left(\frac{1}{2}\sigma^2\|\psi_{\varepsilon_n} + \bar{\psi}\|_{L^\infty(0,T)} + \|\tilde{h}(\psi_{\varepsilon_n}, \bar{\psi})\|_{L^\infty(0,T)}\right)\|\psi_{\varepsilon_n} - \bar{\psi}\|_{L^1(0,T)}\\ \le &\left(\frac{1}{2}\sigma^2\left(\frac{1}{\lambda}\|f\|_{L^\infty(0,T)} + \|\bar{\psi}\|_{L^\infty(0,T)}\right) + \sup_{n\in\mathbb{N}}\|\tilde{h}(\psi_{\varepsilon_n}, \bar{\psi})\|_{L^\infty(0,T)}\right)\|\psi_{\varepsilon_n} - \bar{\psi}\|_{L^1(0,T)} \underset{n\to\infty}{\longrightarrow} 0.\end{aligned}$$

Note that $\tilde{h}$ is continuous, and hence bounded in compact sets, and since ψ_{ε_n} and $\bar{\psi}$ are (uniformly) bounded, they take value in a compact set (ball), a.e., so the supremum in the final line is finite. Thus, from (B-10) we deduce that

$$\lambda(\bar{\psi}(t) - \psi_0(t)) = \bar{G}(\bar{\psi}(t)) - \bar{G}(\psi_0(t)) = \left(\frac{1}{2}\sigma^2(\bar{\psi}(t) + \psi_0(t)) + \tilde{h}(\bar{\psi}(t), \psi_0(t))\right)(\bar{\psi}(t) - \psi_0(t)),$$

for a.e. $t \in (0,T)$.

This implies that $\bar{\psi} = \psi_0$ a.e. in $(0,T)$. Indeed, if there exists a subset $N \subset (0,T)$ with positive Lebesgue measure where $\bar{\psi} \neq \psi_0$, then dividing the previous equation by $\bar{\psi} - \psi_0$ gives

$$\lambda = \frac{1}{2}\sigma^2(\bar{\psi}(t) + \psi_0(t)) + \tilde{h}(\bar{\psi}(t), \psi_0(t)) < 0 \qquad \text{a.e. in } N,$$

which is a contradiction since $\lambda > 0$. The proof is now complete. ☐

B.3. Conclusion

From Lemma B.1, we know that every sequence $(\psi_{\varepsilon_n})_n$ of solutions to (A-1) (where $(\varepsilon_n)_n \subset (0,\infty)$ converges to 0 as $n \to \infty$) admits a convergent subsequence $(\psi_{\varepsilon_{n_k}})_k$ in $L^1(0,T)$. Since $(\psi_\varepsilon)_{\varepsilon>0}$ is a bounded family of (continuous) non-positive functions in $L^\infty(0,T)$, see (B-1), the limit point of this subsequence belongs to $L^1(0,T) \cap L^\infty(0,T)$ and is non-positive, as well.

By Lemma B.3 in Subsection B.2, there exists a unique possible non-positive $L^1(0,T)-$limit point for $(\psi_{\varepsilon_{n_k}})_k$ in $L^1(0,T) \cap L^\infty(0,T)$: ψ_0, the unique non-positive solution of (19). Therefore, by the subsequence convergence principle we conclude that

$$\lim_{\varepsilon \to 0} \psi_\varepsilon = \psi_0 \quad \text{in } L^1(0,T).$$

Appendix C: Laplace transform of hitting time to an upper barrier for a spectrally negative Lévy process

Let X be a spectrally negative one-dimensional Lévy process, i.e. $\nu_X(0,\infty) = 0$, where ν_X is the Lévy measure associated with X, and assume $X_0 = 0$. Suppose that ν_X satisfies

$$\int_{(-\infty,-1)} |x| \nu_X(dx) < \infty; \tag{C-1}$$

by general properties of Lévy processes (see e.g. Theorem 25.3 in [37]), (C-1) ensures that $\mathbb{E}[|X_t|] < \infty$ for every $t > 0$.

In the next proposition, we establish a formula for the Laplace transform of the first hitting time of X to upper (non-negative) barriers.

Proposition C.1. *Consider a spectrally negative one-dimensional Lévy process X with Lévy measure ν_X satisfying* (C-1)*, with X not identically zero. Suppose that*

$$\gamma := \mathbb{E}[X_1] \geq 0.$$

For every $b \geq 0$, denote by τ_b the first hitting time of X to b, i.e. $\tau_b = \inf\{t \geq 0 : X_t > b\}$, and define the function $V : \mathbb{R}_+ \to \mathbb{R}_+$ by

$$V(p) := \frac{1}{2}(\sigma_X)^2 p^2 + \gamma p + \int_{\mathbb{R}_-} (e^{px} - 1 - px) \nu_X(dx), \quad p \geq 0,$$

where $\sigma_X^2 \geq 0$ denotes the Gaussian component of X. Then for all $q \geq 0$

$$\mathbb{E}[e^{-q\tau_b}] = e^{-bV^{-1}(q)}, \tag{C-2}$$

where V^{-1} is the inverse of V.

Proof. By Theorem 25.17 in [37], for all $p \geq 0$ we have

$$\log \mathbb{E}[e^{pX_t}] = tV(p), \quad t \geq 0. \tag{C-3}$$

Thus, V is the logarithmic moment generating function (or cgf) of X_1 on $\mathbb{R}_+$. It then follows from Lemma 2.2.5 in [23] that V is convex. Moreover, V is continuous and differentiable, with

$$V'(p) = \gamma + (\sigma_X)^2 p + \int_{\mathbb{R}_-} x(e^{px} - 1) \nu_X(dx), \quad p \geq 0.$$

Since $V' > 0$ on $(0,\infty)$, V is increasing on $\mathbb{R}_+$ and $\lim_{p\to\infty} V(p) = \infty$.

From the stationary and independent increments property one can verify that $M_t := e^{pX_t - V(p)t}$ is an $\mathcal{F}_t^X$−martingale. Indeed, for $0 \le s \le t$,

$$\mathbb{E}[M_t|\mathcal{F}_s^X] = \mathbb{E}\Big[e^{p(X_t - X_s)}|\mathcal{F}_s^X\Big]e^{pX_s - V(p)t} = \mathbb{E}\Big[e^{pX_{t-s}}\Big]e^{pX_s - V(p)t} = M_s, \tag{C-4}$$

where we use (C-3) for the third equality.
Now choose $p > 0$. Then applying the Optional Stopping Theorem to the bounded stopping time $t \wedge \tau_b$ we have

$$1 \;=\; \mathbb{E}[M_{t\wedge\tau_b}(1_{\{\tau_b \le t\}} + 1_{\{\tau_b > t\}})] = \mathbb{E}[e^{pb - V(p)\tau_b}\, 1_{\{\tau_b \le t\}}] + \mathbb{E}[e^{pX_t - V(p)t} 1_{\{\tau_b > t\}}]\,.$$

Here for the second equality we use that $X_{\tau_b} = b$ when $\tau_b < \infty$ ($\mathbb{P}$−a.s.), because X can only have negative jumps. Using the monotone convergence theorem and that $\lim_{t\to\infty} 1_{\{\tau_b \le t\}} = 1_{\{\tau_b < \infty\}}$ for the left term, and the bounded convergence theorem for the right term (with the bound e^{pb}, since $V > 0$ on $(0, \infty)$), we can take the limit as $t \longrightarrow \infty$ and take the limit inside the expectation to obtain

$$1 = \mathbb{E}[e^{pb - V(p)\tau_b}\, 1_{\{\tau_b < \infty\}}].$$

V is a bijection from $\mathbb{R}_+$ onto itself (since $\gamma \ge 0$), so we can re-write this as

$$\mathbb{E}[e^{-q\tau_b} 1_{\{\tau_b < \infty\}}] = e^{-bV^{-1}(q)}, \quad q > 0. \tag{C-5}$$

Letting $q \searrow 0$ and using the bounded convergence theorem again, we see that

$$\mathbb{P}(\tau_b < \infty) = \mathbb{E}[1_{\{\tau_b < \infty\}}] = e^{-bV^{-1}(0+)},$$

where $V^{-1}(0+) = \lim_{q\searrow 0} V^{-1}(q)$. Since V^{-1} is continuous on $\mathbb{R}_+$, we deduce that $V^{-1}(0+) = V^{-1}(0) = 0$. Consequently, $\tau_b < \infty$ $\mathbb{P}$−a.s. and (C-5) becomes (C-2), completing the proof. □

Remark C.1. *When $\gamma > 0$, for every $b \ge 0$ the finiteness of the stopping time τ_b can be directly inferred from the LLN in Theorem 36.5 of [37].*

Appendix D: Brief formal derivation of the main idea in [3]

Consider a family of hyper-rough Heston models (with zero mean-reversion for simplicity) for which the quadratic variation of the log stock price satisfies

$$\langle \log S^n \rangle_t \;=\; X_t^n = V_0 t + \Big(H_n + \frac{1}{2}\Big)\sigma \int_0^t (t-s)^{H_n - \frac{1}{2}} W_{X_s^n}\, ds$$

for $H_n \in (-\frac{1}{2}, 1)$. We refer the reader to [34], Section 7 in [1] and Section 5 in [28] for more on this model. From Lemma 2.4 in [3], we formally expect that

$$\lim_{H_n \searrow -\frac{1}{2}} \Big(H_n + \frac{1}{2}\Big)\sigma \int_0^t (t-s)^{H_n - \frac{1}{2}} W_{X_s^n}\, ds \;=\; \sigma W_{X_t}\,,$$

where X is the weak limit of X^n, so we expect X to satisfy

$$X_t = V_0 t \,+\, \sigma W_{X_t}\,. \tag{D-1}$$

Now let

$$Y_t = -t + \sigma W_t \tag{D-2}$$

and set $\tilde{X}_t = H_{-V_0 t}$, where $H_b = \inf\{t : Y_t = b\}$. Then setting $t \mapsto \tilde{X}_t$ in (D-2), we see that

$$-V_0 t = -\tilde{X}_t + \sigma W_{\tilde{X}_t} \tag{D-3}$$

i.e. $\tilde{X}$ satisfies the same equation as X_t in (D-1). Hence (using the notation/setup in Lemma 2.3 in [3], i.e. $c = -V_0$, $b = \sigma$ and $a = -1$), we deduce that X is an Inverse Gaussian Lévy process with parameters $(V_0, \frac{V_0^2}{\sigma^2})$.

To analyze this process with VIEs, using that $\frac{1}{\Gamma(\alpha)} = \alpha + O(\alpha^2)$ as $\alpha \to 0$ (i.e. as $H \to -\frac{1}{2}$), we see that the usual rough Heston VIE (with $\rho = 0$) takes the form

$$\begin{aligned}\phi(t) &= \frac{1}{\Gamma(\alpha)} \int_0^t (t-s)^{\alpha-1}\Big(-\frac{1}{2}(u^2+iu) + \frac{1}{2}\sigma^2\phi(s)^2\Big)ds \\ &= (1+O(\alpha))\alpha \int_0^t (t-s)^{\alpha-1}\Big(-\frac{1}{2}(u^2+iu) + \frac{1}{2}\sigma^2\phi(s)^2\Big)ds \;\;\to\;\; -\frac{1}{2}(u^2+iu) + \frac{1}{2}\sigma^2\phi(t)^2\end{aligned}$$

as $\alpha \to 0$ (again using Lemma 2.4 in [3]), which is just an algebraic equation for ϕ. If we ignore the linear term in u for simplicity (i.e. ignore the drift of the log stock price), then the (relevant) solution to this equation is $\phi(t) = \frac{1}{\sigma^2}(1 - \sqrt{1+\sigma^2 u^2})$, i.e. the smaller root as in the proof of Theorem 2.1.

References

[1] Abi Jaber, E., “Weak existence and uniqueness for affine stochastic Volterra equations with L^1-kernels”, *Bernoulli, 27*(3), 1583-1615, 2021.

[2] Abi Jaber, E. and E. Attal, “Simulating integrated Volterra square-root processes and Volterra Heston models via Inverse Gaussian”, Preprint arXiv:2504.19885, 2025.

[3] Abi Jaber, E., E. Attal and M. Rosenbaum, “From Hyper Roughness to Jumps as $H \to -\frac{1}{2}$”, to appear in *The Annals of Applied Probability*.

[4] Abi Jaber, E., C. Cuchiero, M. Larsson and S. Pulido, “A weak solution theory for stochastic Volterra equations of convolution type”, *The Annals of Applied Probability, 31*(6), 2924-2952, 2021.

[5] Abi Jaber, E., C. Cuchiero, L. Pelizzari, S. Pulido and S. Svaluto-Ferro, “Polynomial Volterra processes”, *Electronic Journal of Probability, 29*, 1-37, 2024.

[6] Abi Jaber, E. and N. De Carvalho, “Reconciling rough volatility with jumps”, *SIAM Journal on Financial Mathematics, 15*(3), 785-823, 2024.

[7] Abi Jaber, E. and O. El Euch, “Multifactor approximation of rough volatility models”, *SIAM Journal on Financial Mathematics, 10*(2), 309-349, 2019.

[8] Abi Jaber, E., M. Larsson and S. Pulido, “Affine Volterra processes”, *The Annals of Applied Probability, 29*(5), 3155-3200, 2019.

[9] Alfonsi, A., “Nonnegativity preserving convolution kernels. Application to Stochastic Volterra Equations in closed convex domains and their approximation”, *Stochastic Processes and their Applications, 181*, 104535, 2025.

[10] Alfonsi, A. and G. Szulda, “On non-negative solutions of stochastic Volterra equations with jumps and non-Lipschitz coefficients”, *Bernoulli, 31*(4), 2890-2915, 2025.

[11] Applebaum, D., “*Lévy processes and stochastic calculus*”, Cambridge University Press, 2009.

[12] Barndorff-Nielsen, O. E., F. E. Benth and A. E. Veraart, "Modelling energy spot prices by volatility modulated Lévy-driven Volterra processes", *Bernoulli, 19*(3), 803-845, 2013.

[13] Billingsley, P., "*Probability and measure*", John Wiley&Sons, 2nd edition, 1986.

[14] Bondi, A. and F. Flandoli, "On the Kolmogorov equation associated with Volterra equations and Fractional Brownian Motion", *Journal of Functional Analysis*, 111234, 2025.

[15] Bondi, A., G. Livieri and S. Pulido, "Affine Volterra processes with jumps", *Stochastic Processes and their Applications, 168*, 1042-64, 2024.

[16] Bondi, A. and S. Pulido, "Feller's test for explosions of stochastic Volterra equations", to appear in *Electronic Journal of Probability*.

[17] Bondi, A., S. Pulido and S. Scotti, "The rough Hawkes Heston stochastic volatility model", *Mathematical Finance, 34*(4), 1197-1241, 2024.

[18] Boyarchenko, S. and S. Levendorskii, "Correct implied volatility shapes and reliable pricing in the rough Heston model", Preprint arXiv:2412.16067, 2024.

[19] Brézis, H., "*Functional analysis, Sobolev spaces and partial differential equations*" (Vol. 2, No. 3), New York: Springer, 2011.

[20] Cont, R. and P. Tankov, "*Financial modelling with Jump Processes*", Chapman&Hall, 2004.

[21] Cox, J.C., J.E. Ingersoll and S.A. Ross, "A theory of term structure of interest rates", *Econometrica, 53*, 385-408, 1985.

[22] Cuchiero, C. and J. Teichmann. "Generalized Feller processes and Markovian lifts of stochastic Volterra processes: the affine case", *Journal of Evolution Equations, 20*(4), 1301-1348, 2020.

[23] Dembo, A. and O. Zeitouni, "*Large deviations techniques and applications*", Springer-Verlag, New York, 1998.

[24] Duffie, D., D. Filipovic and W. Schachermayer, "Affine processes and applications in finance", *The Annals of Applied Probability, 13*(3), 984-1053, 2003.

[25] El Euch, O., M. Fukasawa and M. Rosenbaum, "The microstructural foundations of leverage effect and rough volatility", *Finance and Stochastics, 12*(6), 241-280, 2018.

[26] El Euch, O. and M. Rosenbaum, "Perfect hedging in rough Heston models", *The Annals of Applied Probability, 28*(6), 3813-3856, 2018.

[27] El Euch, O. and M. Rosenbaum, "The characteristic function of Rough Heston models", *Mathematical Finance, 29*(1), 3-38, 2019.

[28] Forde, M., S. Gerhold and B. Smith, "Small-time, large-time and $H \to 0$ asymptotics for the rough Heston model", *Mathematical Finance, 31*(1), 203-241, 2021.

[29] Gatheral, J. and M. Keller-Ressel, "Affine forward variance models", *Finance and Stochastics, 23*, 501-533, 2019.

[30] Gripenberg, G., S.O. Londen and O. Staffans, "*Volterra integral and functional equations*" (No. 34), Cambridge University Press, 1990.

[31] Hamaguchi, Y., "Markovian lifting and asymptotic log-Harnack inequality for stochastic Volterra integral equations", *Stochastic Processes and their Applications, 178*, 2024, 104482.

[32] Jacod, J. and A. Shiryaev, "*Limit theorems for stochastic processes*", Vol. 288, Springer Science & Business Media, 2003.

[33] Jaisson, T. and M. Rosenbaum, "Rough fractional diffusions as scaling limits of nearly unstable heavy tailed Hawkes processes", *The Annals of Applied Probability, 26*(5), 2860-2882, 2016.

[34] Jusselin, P. and M. Rosenbaum, "No-arbitrage implies power-law market impact and rough volatility", *Mathematical Finance, 30*(4), 1309-1336, 2020.

[35] Kuznetsov, A., A. Kyprianou and V. Rivero, "The theory of scale functions for spectrally negative Lévy processes", *Lévy Matters II: Recent Progress in Theory and Applications: Fractional Lévy Fields, and Scale Functions,* 2013.

[36] Mytnik, L. and T. S. Salisbury, "Uniqueness for Volterra-type stochastic integral equations", Preprint arXiv:1502.05513, 2015.

[37] Sato, K., "*Lévy Processes and Infinitely Divisible Distributions*", Cambridge University Press, 1999.

[38] Whitt, W., "Some useful functions for functional limit theorems", *Mathematics of operations research, 5*(1), 67-85, 1980.

[39] Whitt, W., "*Stochastic-process limits: an introduction to stochastic-process limits and their application to queues*", New York, NY: Springer New York, 2002.

[40] Zhang, X., "Stochastic Volterra equations in Banach spaces and stochastic partial differential equation", *Journal of Functional Analysis, 258*(4), 1361-1425, 2010.