Reconciling rough volatility with jumps

Eduardo Abi Jaber, Nathan de Carvalho

Reconciling rough volatility with jumps

Eduardo Abi Jaber*[1] and Nathan De Carvalho†[2]

[1]Ecole Polytechnique, CMAP
[2]Engie Global Markets & Université Paris Cité and Sorbonne Université, LPSM

March 13, 2023

Abstract

We reconcile rough volatility models and jump models using a class of *reversionary Heston models* with fast mean reversions and large vol-of-vols. Starting from hyper-rough Heston models with a Hurst index $H \in (-1/2, 1/2)$, we derive a Markovian approximating class of one dimensional *reversionary Heston-type models*. Such proxies encode a trade-off between an exploding vol-of-vol and a fast mean-reversion speed controlled by a reversionary time-scale $\epsilon > 0$ and an unconstrained parameter $H \in \mathbb{R}$. Sending ϵ to 0 yields convergence of the *reversionary Heston model* towards different explicit asymptotic regimes based on the value of the parameter H. In particular, for $H \leq -1/2$, *the reversionary Heston model* converges to a class of Lévy jump processes of Normal Inverse Gaussian type. Numerical illustrations show that the *reversionary Heston model* is capable of generating at-the-money skews similar to the ones generated by rough, hyper-rough and jump models.

Mathematics Subject Classification (2010): 91G20, 60G22, 60G51

JEL Classification: G13, C63, G10.

Keywords: Stochastic volatility, Heston model, Normal Inverse Gaussian, rough Heston model, Ricatti equations

*eduardo.abi-jaber@polytechnique.edu. The first author is grateful for the financial support from the Chaires FiME-FDD, Financial Risks, Deep Finance & Statistics and Machine Learning and systematic methods in finance at Ecole Polytechnique.

†nathan.decarvalho@engie.com. The second author is grateful for the financial support provided by Engie Global Markets. We would like to thank Ryan McCrickerd for fruitful discussions.

Contents

Introduction

Since the 1987 financial crash, financial option markets have exhibited a notable implied volatility skew, especially for short-term maturities. This skew reflects the market's expectation of significant price movements on very short time scales in the underlying asset, which poses a challenge to traditional continuous models based on standard Brownian motion. To address this issue, the literature has developed several classes of models that capture the skewness in implied volatilities. Three prominent approaches are:

- conventional one-factor stochastic volatility models boosted with large mean-reversion and vol-of-vol. This class of models have been justified by several empirical studies that have identified the presence of very fast mean-reversion in the S&P volatility time series [10, 16, 26, 27] and by the fact that they are able to correct conventional models to reproduce the behavior of the at-the-money (ATM) skew for short maturities [38];

- jump diffusion models, especially the class of affine jump-diffusions for which valuation problems become (semi-)explicit using Fourier inversion techniques, see [19]. Such class of models incorporates occasional and large jumps to explain the skew observed implicitly on option markets, see [18], and [12] for an empirical analysis of the impact of adding jumps to stochastic volatility diffusion on the implied volatility surface;

- rough volatility models, where the volatility process is driven by variants of the Riemann-Liouville fractional Brownian motion

$$W_t^H = \frac{1}{\Gamma\left(H+1/2\right)} \int_0^t (t-s)^{H-1/2}\, dW_s, \quad t \geq 0, \tag{0.1}$$

  with W a standard Brownian motion and $H \in (0, 1/2)$ the Hurst index. Such models are able to reproduce the roughness of the spot variance's trajectories measured empirically [31, 15] together with the explosive behavior of the ATM-skew [11, 14, 20, 28, 3].

So far, in the mathematical finance community, jump diffusion models and rough volatility models have often been treated as distinct approaches, and, in some cases, they have even been opposed to each other, see for instance [14, Section 5.3.1]. However, on the one side, connections between rough volatility models and fast mean-reverting factors have been established in [1, 4, 5]. On the other side, jump models have been related

to fast regimes stochastic volatility models in [38, 37]. In parallel, from the empirical point of view, it can be very challenging for the human eye and for statistical estimators to distinguish between roughness, fast mean-reversions and jump-like behavior, as shown in [5, 17, 29].

The above suggests that rough volatility and jump models may not be that different after all. Our main motivation is to establish for the fist time in the literature a connection between rough volatility and jump models through conventional volatility models with fast mean-reverting regimes.

We aim to reconcile these two classes of models through the use of the celebrated conventional Heston model [33] but with a parametric specification which encodes a trade-off between a fast mean-reversion and a large vol-of-vol. We define the *reversionary Heston model* as follows:

$$dS_t^\epsilon = S_t^\epsilon \sqrt{V_t^\epsilon} \left(\rho dW_t + \sqrt{1-\rho^2} dW_t^\perp \right), \quad S_0^\epsilon = S_0, \tag{0.2}$$

$$dV_t^\epsilon = \left(\epsilon^{H-\frac{1}{2}} \theta - \epsilon^{-1} \left(V_t^\epsilon - V_0 \right) \right) dt + \epsilon^{H-\frac{1}{2}} \xi \sqrt{V_t^\epsilon} dW_t, \quad V_0^\epsilon = V_0, \tag{0.3}$$

where $\left(W, W^\perp\right)$ is a two-dimensional Brownian motion, $\theta \geq 0$, $S_0, \xi, V_0 > 0$, $\rho \in [-1, 1]$. The two crucial parameters here are the *reversionary time-scale* $\epsilon > 0$ and $H \in \mathbb{R}$. Such parametrizations nest as special cases the fast regimes extensively studied by Fouque et al. [26], Feng et al. [22], see also [25, Section 3.6], which correspond to the case $H = 0$; and also the regimes studied in [38, 37] for the case $H = -1/2$. Letting the parameter $H \in (\infty, 1/2]$ free in (0.3) introduces more flexibility in practice and leads to better fits with stable calibrated parameters across time as recently shown in [7]. In theory, it allows for a better understanding of the impact of the scaling in H on the limiting behavior of the model as $\epsilon \to 0$ as highlighted in this paper.

In a nutshell, we show that:

(i) for $H > -1/2$, the *reversionary Heston model* can be constructed as a proxy of rough and hyper-rough Heston models where $H \in (-1/2, 1/2]$ plays the role of the Hurst index,

(ii) for $H \leq -1/2$, as $\epsilon \to 0$, the *reversionary Heston model* converges towards Lévy jump processes of Normal Inverse Gaussian type with distinct regimes for $H = -1/2$ and $H < -1/2$ respectively,

(iii) the *reversionary Heston model* is capable of generating implied volatility surfaces and at-the-money (ATM) skews similar to the ones generated by rough, hyper-rough and jump models, and comes arbitrarily close to the ATM skew scaling as $\tau^{-0.5}$ for small τ that characterizes the market, contrary to widespread understanding.

Our results allow for a reconciliation between rough and jump models as they suggest that jump models and (hyper-)rough volatility models are complementary, and do not overlap. For $H > -1/2$, the *reversionary Heston model* can be interpreted as a proxy of rough and hyper-rough volatility models, while for $H \leq -1/2$, it can be interpreted as a proxy of jump models. Jump models actually start at $H = -1/2$ (and below), the first value for which hyper-rough volatility models can no-longer be defined.

More precisely, our argument is structured as follows. First, in Section 1, we show how the reversionary Heston model (0.2)-(0.3) can be obtained as a Markovian and semimartingale proxy of rough and hyper-rough Heston models [20, 35] with Hurst index $H \in (-1/2, 1/2)$. This is achieved using the resolvent of the first kind of the shifted fractional kernel.
Second, in Section 2, we derive the joint conditional characteristic functional of the log-price $\log S^\epsilon$ and the integrated variance $\bar{V}^\epsilon := \int_0^\cdot V_s^\epsilon ds$ in the model (0.2)–(0.3) in terms of a solution to a system of time-dependent Riccati ordinary differential equations; see Theorem 2.1. Compared to the literature, we provide a novel and concise proof for the existence and uniqueness of a global solution to such Riccati equations using the variation of constant formulas.
Finally, in Section 3, we establish the convergence of the log-price and the integrated variance $(\log S^\epsilon, \bar{V}^\epsilon)$ in the reversionary Heston model (0.2)-(0.3) towards a Lévy jump process (X, Y), as ϵ goes to 0. More precisely,

we show that the limit (X, Y) belongs to the class of Normal Inverse Gaussian - Inverse Gaussian (NIG-IG) processes which we construct from its Lévy exponent and we connect such class to first hitting-time representations in the same spirit of Barndorff-Nielsen [13]. Our main Theorem 3.4 provides the convergence of the finite-dimensional distributions of the joint process $(\log S^\epsilon, \bar{V}^\epsilon)$ through the study of the limiting behavior of the Riccati equations and hence the characteristic functional given in Theorem 2.1. Interestingly, the limiting behavior disentangles three different asymptotic regimes based on the values of H. The convergence of the integrated variance process is even strengthened to a functional weak convergence on the Skorokhod space of càdlàg paths on $[0, T]$ endowed with the M_1 topology. We stress that the usual J_1 topology is not useful here, since jump processes cannot be obtained as limits of continuous processes in the J_1 topology.

Related Literature. Convergence of the reversionary Heston models towards jump processes: our results clarify and extend the results of [38, 37], derived for the case $H = -1/2$, that establish and make clear the precise limiting connection between the Heston log-price process and the normal inverse-Gaussian (NIG) process of [13]. Connections between the long time behavior of the Heston log-price process and NIG distribution were first exposed in [24, 36] and were the main motivations behind the work of Mechkov [38].
Relevance of fast regimes in practice: the pricing of options near maturity is challenging because of the very steep slope of smiles observed on the market and Fouque et al. [26] showed that stochastic volatility should embed both a fast regime Ornstein-Uhlenbeck factor (see Remark 1.3 below) from which approximations of option prices can be derived using a singular perturbation expansion, and a slowly varying factor to be able to match options with long maturities. On the other hand, Feng et al. [22] considers a Heston model with a fast mean-reverting volatility and uses large deviation theory techniques to derive an approximation price for out-of-the-money vanilla options when the maturity is small, but large compared to the characteristic time-scale of the stochastic volatility factor. More recently an Ornstein-Uhlenbeck process with the same parametrization as in (0.3) has been used to construct the Quintic stochastic volatility model [8] to achieve remarkable joint fits of SPX and VIX implied volatilities, outperforming its rough and path-dependent counterparts as shown empirically in [7].

Notations. For $p \geq 1$, we denote by L^p_{loc} the space of measurable functions $f : \mathbb{R}_+ \to \mathbb{R}$ such that $\int_0^T |f(s)|^p ds < \infty$, for all $T > 0$. We will denote by $\sqrt{x}$ the principal square root of $x \in \mathbb{C}$, i.e. its argument lies within $(-\pi/2, \pi/2]$.

1 From rough Heston to reversionary Heston

In this section, we show how reversionary Heston models (0.2)-(0.3) can be seen as proxies of rough and hyper-rough Heston models whenever $H > -1/2$.

1.1 Rough and hyper-rough Heston

Let $\left(W, W^\perp\right)$ be a two-dimensional Brownian motion and set $B := \rho W + \sqrt{1-\rho^2} W^\perp$ with $\rho \in [-1, 1]$. We take as starting point a stochastic volatility model for an underlying asset P in terms of a time-changed Brownian motion:

$$dP_t = P_t dB_{\bar{U}_t}, \quad P_0 > 0, \tag{1.1}$$

for some non-decreasing continuous process $\bar{U}$. If $\bar{U}_t = \int_0^t U_s ds$, then U would correspond to the spot variance and $\bar{U}$ plays the role of the integrated variance. The hyper-rough Volterra Heston model introduced in [35] and studied further in [2, Section 7] assumes that the dynamics of the integrated variance is of the form

$$\bar{U}_t = \bar{G}_0(t) + \xi \int_0^t K_H(t-s) W_{\bar{U}_s} ds, \tag{1.2}$$

for a suitable continuous function $\bar{G}_0$, and $\xi > 0$, and K_H is the fractional kernel

$$K_H(t) = t^{H-1/2}, \quad t > 0, \tag{1.3}$$

for $H \in (-1/2, 1/2]$. The lower bound $H > -1/2$ ensures the L^1_{loc} integrability of the kernel K_H so that the stochastic convolution appearing in (1.2) is well-defined.

Any kernel K only in L^1_{loc} can be considered for the specification of the integrated variance in (1.2), and if furthermore the kernel happens to be in L^2_{loc}, the following lemma ensures the existence of a spot variance process.

Lemma 1.1 (Existence of spot variance)**.** *Let $K \in L^2_{loc}$ and $g_0 \in L^1_{loc}$. Assume there exists a non-decreasing adapted process $\bar{U}$ and a Brownian motion W such that*

$$\bar{U}_t = \int_0^t g_0(s)ds + \int_0^t K(t-s)W_{\bar{U}_s}ds, \tag{1.4}$$

with $\sup_{t\leq T} \mathbb{E}\left[\left|\bar{U}_t\right|\right] < \infty$, for all $T > 0$. Then, $\bar{U}_t = \int_0^t U_s ds$, where U is a non-negative weak solution to the following stochastic Volterra equation

$$U_t = g_0(t) + \int_0^t K(t-s)\sqrt{U_s}dW_s, \quad \mathbb{Q}\otimes dt - a.e. \tag{1.5}$$

Conversely, assume there exists a non-negative weak solution U to the stochastic Volterra equation (1.5) *such that $\sup_{t\leq T}\mathbb{E}\left[U_t^2\right] < \infty$, for all $T > 0$, then $\bar{U}$ solves* (1.4)*.*

Proof. This is obtained by an application of stochastic Fubini's theorem, see [2, Lemma 2.1]. □

Going back to the fractional case, if we restrict H in $(0, 1/2]$, then we have $K_H \in L^2_{loc}$. For $\bar{G}_0(t) := \int_0^t g_0(s)ds$, a direct application of Lemma 1.1 yields that the model (1.1)-(1.2) is equivalent to the rough Heston model of El Euch and Rosenbaum [20] written in spot-variance form

$$\begin{aligned} dP_t &= P_t\sqrt{U_t}dB_t, \\ U_t &= g_0(t) + \int_0^t K_H(t-s)\xi\sqrt{U_s}dW_s, \end{aligned}$$

for some initial input curve $g_0 : \mathbb{R}_+ \to \mathbb{R}$ ensuring the non-negativity of V. Two notable specifications of such admissible input curves are given by [4, Example 2.2] and read

$$\begin{cases} g_0 \text{ continuous and non-decreasing with } g_0 \geq 0, \\ \text{or} \\ g_0(t) = U_0 + \theta\int_0^t K_H(s)ds, \text{ for some } U_0, \theta \geq 0. \end{cases}$$

Moreover, for $H \in (0, 1/2]$ the sample paths of the spot variance U are locally Hölder continuous of any order strictly less than H, and consequently rougher than those of the standard Brownian motion, which corresponds to the case $H = 1/2$, justifying the denomination 'rough model'. The hyper-rough appellation corresponds to the case $H \in (-1/2, 0]$ for which the process $\bar{U}$ is continuous but no longer absolutely continuous. Indeed, in this case, one can show that the trajectories of U are nowhere differentiable, see [35, Proposition 4.6].

A key advantage of rough and hyper-rough Heston models is the semi-explicit knowledge of the characteristic function of the log-price modulo a deterministic Riccati Volterra convolution equation, as they belong the class of Affine Volterra processes [6, 2]. More precisely, for any $u = (u_1, u_2) \in \mathbb{C}^2$ satisfying

$$\Re(u_1) = 0, \quad \Re(u_2) \leq 0,$$

the joint Fourier–Laplace transform of $(\log P, \bar{U})$ is given by

$$\mathbb{E}\left[\exp\left(u_1 \log P_T + u_2\bar{U}_T\right)\right] = \exp\left(u_1 \log P_0 + \int_0^T R(\psi_H(T-s))d\bar{G}_0(s)\right),$$

for all $T \geq 0$, where ψ_H is the continuous solution to the following fractional Riccati–Volterra equation

$$\psi_H(t) = \int_0^t K_H(t-s)R(\psi_H(s))ds, \quad t \geq 0,$$
$$R(x) = \frac{1}{2}(u_1^2 - u_1) + u_2 + \rho\nu u_1 x + \frac{\nu^2}{2}x^2, \tag{1.6}$$

see [2, Section 7]. This allows fast pricing and calibration via Fourier inversion techniques. Compared to the conventional Heston model where the characteristic function is known explicitly, the solution to the Riccati Volterra equation is not explicitly known.

1.2 Deriving reversionary Heston as a proxy: ϵ-shifting the singularity

In both regimes, rough and hyper-rough, with the exception of $H = \frac{1}{2}$, the model is non-Markovian, non-semimartingale with singular kernels. From a practitioner standpoint, it is therefore natural to look for Markovian approximations by suitable smoothing of the singularity of the fractional kernel (1.3) sitting at the origin. In this section, we show how we can build a Markovian semi-martingale proxy of hyper-rough models. This is achieved using a two-step procedure.

First step: recover semimartingality by smoothing out the singularity of the fractional kernel K_H. We fix $\epsilon > 0$, and we consider the shifted fractional kernel

$$K_{H,\epsilon}(t) := (t+\epsilon)^{H-\frac{1}{2}}, \quad t > 0,$$

and the corresponding 'integrated variance' $\bar{U}^\epsilon$ given by

$$\bar{U}_t^\epsilon = \int_0^t g_0^\epsilon(s)ds + \xi \int_0^t K_{H,\epsilon}(t-s)W_{\bar{U}_s^\epsilon}ds,$$

with

$$g_0^\epsilon(t) = U_0 + \theta \int_0^t K_{H,\epsilon}(s)ds.$$

Note that now $K_{H,\epsilon}$ is in L^2_{loc} for any value of H, so that an application of Lemma 1.1 yields that $\bar{U}^\epsilon = \int_0^\cdot U_s^\epsilon ds$ where the spot variance U^ϵ solves the equation

$$dP_t^\epsilon = P_t^\epsilon \sqrt{U_t^\epsilon} dB_t$$
$$U_t^\epsilon = U_0 + \int_0^t K_{H,\epsilon}(t-s)\theta ds + \int_0^t K_{H,\epsilon}(t-s)\xi\sqrt{U_s^\epsilon}dW_s.$$

Moreover, since $K_{H,\epsilon}$ is continuously differentiable on $[0,T]$, denoting by $K'_{H,\epsilon}$ its derivative, we get that U^ϵ is a semimartingale with the following dynamics

$$dU_t^\epsilon = \left(K_{H,\epsilon}(0)\theta + \int_0^t K'_{H,\epsilon}(t-s)dZ_s^\epsilon\right)dt + K_{H,\epsilon}(0)\xi\sqrt{U_t^\epsilon}dW_t, \tag{1.7}$$

with

$$dZ_t^\epsilon = \theta dt + \xi\sqrt{U_t^\epsilon}dW_t.$$

Second step: recover a Markovian proxy thanks to the resolvent of the first kind. The only non-Markovian term in (1.7) is the term $\int_0^t K'_{H,\epsilon}(t-s)dZ_s^\epsilon$ appearing in the drift. Using the resolvent of the first kind of $K_{H,\epsilon}$ we will re-express this term in terms of a functional of the past of the process U^ϵ. For a kernel K, a resolvent of the first kind is a measure L on $\mathbb{R}_+$ of locally bounded variation such that

$$\int_{[0,t]} K(t-s)L(ds) = 1, \quad t \geq 0,$$

see [32, Definition 5.5.1]. A resolvent of the first kind does not always exist. We will make use of the notations $(f*g)(t) = \int_0^t f(t-s)g(s)ds$ and $(f*L)(t) = \int_{[0,t]} f(t-s)L(ds)$.

Lemma 1.2. *Fix $\epsilon > 0$ and $H \in (-1/2, 1/2)$. The kernel $K_{H,\epsilon}$ admits a resolvent of the first kind L_ϵ of the form*

$$L_\epsilon(dt) = \frac{\delta_0(dt)}{K_{H,\epsilon}(0)} + \ell_\epsilon(t)dt, \tag{1.8}$$

with ℓ_ϵ a locally integrable function. Moreover, the function $(K'_{H,\epsilon} * L_\epsilon)$ is continuously differentiable and it holds, for all $t \geq 0$, that

$$\int_0^t K'_{H,\epsilon}(t-s)dZ^\epsilon_s = -\left(\frac{1}{2} - H\right)\epsilon^{-1}\left(U^\epsilon_t - U_0\right) + \int_0^t \left(K'_{H,\epsilon} * L_\epsilon\right)'(t-s)\left(U^\epsilon_s - U_0\right)ds. \tag{1.9}$$

Proof. First, the existence of the resolvent is justified as follows. Given $H \in (-\frac{1}{2}, \frac{1}{2})$, $K_{H,\epsilon}$ is a positive completely monotone kernel[1] on $[0,T]$ so that an application of [32, Theorem 5.5.4] yields the existence of a resolvent of the first kind in the form (1.8) with ℓ_ϵ a completely monotone function. Convolving (1.8) with $K'_{H,\epsilon}$ one obtains that

$$(K'_{H,\epsilon} * L_\epsilon)(t) = \frac{K'_{H,\epsilon}(0)}{K_{H,\epsilon}(0)} + (K'_{H,\epsilon} * \ell_\epsilon)(t).$$

Since $K_{H,\epsilon}$ is twice continuously differentiable on $[0,T]$ and ℓ_ϵ is integrable, it follows that $(K'_{H,\epsilon} * \ell_\epsilon)$ is continuously differentiable and so is $(K'_{H,\epsilon} * L_\epsilon)$. Writing

$$\left(K'_{H,\epsilon} * L_\epsilon\right) = \left(K'_{H,\epsilon} * L_\epsilon\right)(0) + 1 * \left(K'_{H,\epsilon} * L_\epsilon\right)',$$

convolving on the left hand side by by $K_{H,\epsilon}$ combined with the associativity of the convolution operation and the fact that $(L_\epsilon * K_{H,\epsilon}) = 1$ yields:

$$K'_{H,\epsilon} * 1 = K'_{H,\epsilon} * (L_\epsilon * K_{H,\epsilon}) = (K'_{H,\epsilon} * L_\epsilon) * K_{H,\epsilon} = \left(K'_{H,\epsilon} * L_\epsilon\right)(0)(1 * K_{H,\epsilon}) + 1 * \left(K'_{H,\epsilon} * L_\epsilon\right)' * K_{H,\epsilon}.$$

And thus, we obtain almost everywhere with regards to the Lebesgue measure that:

$$K'_{H,\epsilon} = \left(K'_{H,\epsilon} * L_\epsilon\right)(0)K_{H,\epsilon} + (K'_{H,\epsilon} * L_\epsilon)' * K_{H,\epsilon}.$$

In addition, using (1.8), we notice that

$$\left(K'_{H,\epsilon} * L_\epsilon\right)(0) = \frac{K'_{H,\epsilon}(0)}{K_{H,\epsilon}(0)} = \left(H - \frac{1}{2}\right)\epsilon^{-1}.$$

Combining the above, we obtain that

$$\int_0^t K'_{H,\epsilon}(t-s)dZ^\epsilon_s = \left(H - \frac{1}{2}\right)\epsilon^{-1}\int_0^t K_{H,\epsilon}(t-s)dZ^\epsilon_s + \left((K'_{H,\epsilon} * L_\epsilon)' * (K_{H,\epsilon} * dZ^\epsilon_t)\right)_t,$$

which yields (1.9), after recalling that $U^\epsilon - U_0 = \int_0^\cdot K_{H,\epsilon}(\cdot - s)dZ^\epsilon_s$. □

With the help of the resolvent of the first kind, we were able to recover in the first term of (1.9), the first order mean-reversion scale of the fractional kernel, the second term depends on the whole past trajectory of U^ϵ.

We can now derive our Markovian proxy of the hyper-rough Heston model as follows: plugging the expression (1.9) in the drift of (1.7), recalling that $K_{H,\epsilon}(0) = \epsilon^{H-1/2}$ and dropping the non-Markovian term $\left(\left(K'_{H,\epsilon} * L_\epsilon\right)'\right)_\cdot * \left(U^\epsilon_\cdot - U^\epsilon_0\right)\Big)_t$, we arrive to the Markovian process:

$$d\tilde{V}^\epsilon_t = \left(\epsilon^{H-\frac{1}{2}}\theta - \left(\frac{1}{2} - H\right)\epsilon^{-1}\left(\tilde{V}^\epsilon_t - U_0)\right)\right)dt + \epsilon^{H-\frac{1}{2}}\xi\sqrt{\tilde{V}^\epsilon_t}dW_t, \quad \tilde{V}_0 = U_0.$$

Finally, re-scaling the mean-reversion speed from $\left(\frac{1}{2} - H\right)\epsilon^{-1}$ to ϵ^{-1} leads to our reversionary Heston model (0.2)–(0.3) where the parameter H becomes unconstrained. In the following section, we illustrate numerically the fact that such reversionary Heston model can be seen as a proxy of rough and hyper-rough Heston models.

[1]Recall that a function f is completely monotone if it is infinitely differentiable on $(0,\infty)$ such that $(-1)^n f^{(n)} \geq 0$, for all $n \geq 0$.

Remark 1.3. *Such proxy approximation can directly be applied to the Riemann-Liouville fractional Brownian motion defined in* (0.1) *to get the proxy:*

$$W_t^{H,\epsilon} = \epsilon^{H-\frac{1}{2}} \int_0^t e^{-\left(\frac{1}{2}-H\right)\epsilon^{-1}(t-s)} dW_s,$$

which is a mean-reverting Ornstein-Uhlenbeck process as long as $H < \frac{1}{2}$*, while the value* $H = \frac{1}{2}$ *yields back the standard Brownian motion.*
First, such Ornstein-Uhlenbeck process has been recently used to construct the Quintic stochastic volatility model [8] to achieve remarkable joint fits of SPX and VIX implied volatilies, outperforming even its rough and path-dependent counterparts as shown empirically in [7].
Furthermore, notice that the case $H = 0$ *degenerates into the fast scale volatility factor from Fouque et al. [26], with* $m = 0$*,* $\nu = 1$ *and their time-scale is twice the reversionary time-scale* ϵ*, and whose auto-correlation under the invariant distribution is given by*

$$\mathbb{E}\left[W_t^{0,\epsilon} W_s^{0,\epsilon}\right] = e^{-\frac{|t-s|}{2\epsilon}}.$$

Consequently, the reversionary time-scale ϵ *sets the speed of decay of the auto-correlation function of* $W^{H,\epsilon}$*.*

1.3 Numerical illustration

We now illustrate numerically that the reversionary Heston model (0.2)–(0.3) is able to reproduce shapes of implied volatility surfaces and at-the-money skew that are similar to the ones generated by rough and hyper-rough Heston models.

For this, we first generate implied volatility surfaces of the hyper-rough and rough Heston model via the Fourier-Cosine series expansion technique from [21], where we used the fractional Adams scheme described in [20] on the fractional Riccati equation (1.6) to compute the characteristic function of the (hyper-)rough Heston models. Three target smiles are generated with a (hyper-)rough Heston having parameters

$$\rho = -0.7, \quad \theta = 0.02, \quad \xi = 0.3, \quad U_0 = 0.02, \tag{1.10}$$

for $H \in \{0.1, 0, -0.05\}$.
For each of these smiles, we calibrate the parameters $\left(\hat{\epsilon}, \hat{H}\right)$ of the reversionary Heston model (0.2)-(0.3), while fixing the other parameters equal to those of the hyper-rough Heston's, by minimizing a weighted loss

$$\sum_{i,j} w_{i,j} \left(C_{rough\ Heston}(T_i, K_j) - C^{\hat{\epsilon},\hat{H}}_{reversionary\ Heston}(T_i, K_j)\right)^2.$$

The reversionary Heston prices are also obtained by Fourier-Cosine expansion of the characteristic function. In contrast to the rough Heston models, the characteristic function is known explicitly, see Corollary 2.2 below. After calibration, we obtain the following parameters

Target (hyper-)rough Heston	**Calibrated reversionary Heston**	
H	$\hat{\epsilon}$	$\hat{H}$
0.1	0.10183756	-0.29183935
0	0.06258637	-0.33057822
-0.05	0.05932449	-0.38692275

Table 1: Calibrated values of $\left(\hat{\epsilon}, \hat{H}\right)$ of the reversionary Heston model to the hyper-rough Heston volatility surfaces. The other parameters are fixed as in (1.10).

The resulting At-The-Money (ATM) skews between 1 week and 1 year are shown on Figure 2. The implied volatility surfaces for the case $H = 0.1$ is illustrated on Figure 1. The fit of the smiles for $H = 0$ and

$H = -0.05$ are deferred to Appendix B, see Figures 6 and 7. The graphs show that the reversionary Heston model seems to be able to generate similar shapes of the implied volatility surfaces of rough and hyper-rough models and very steep skews even in the hyper rough regimes $H \leq 0$.

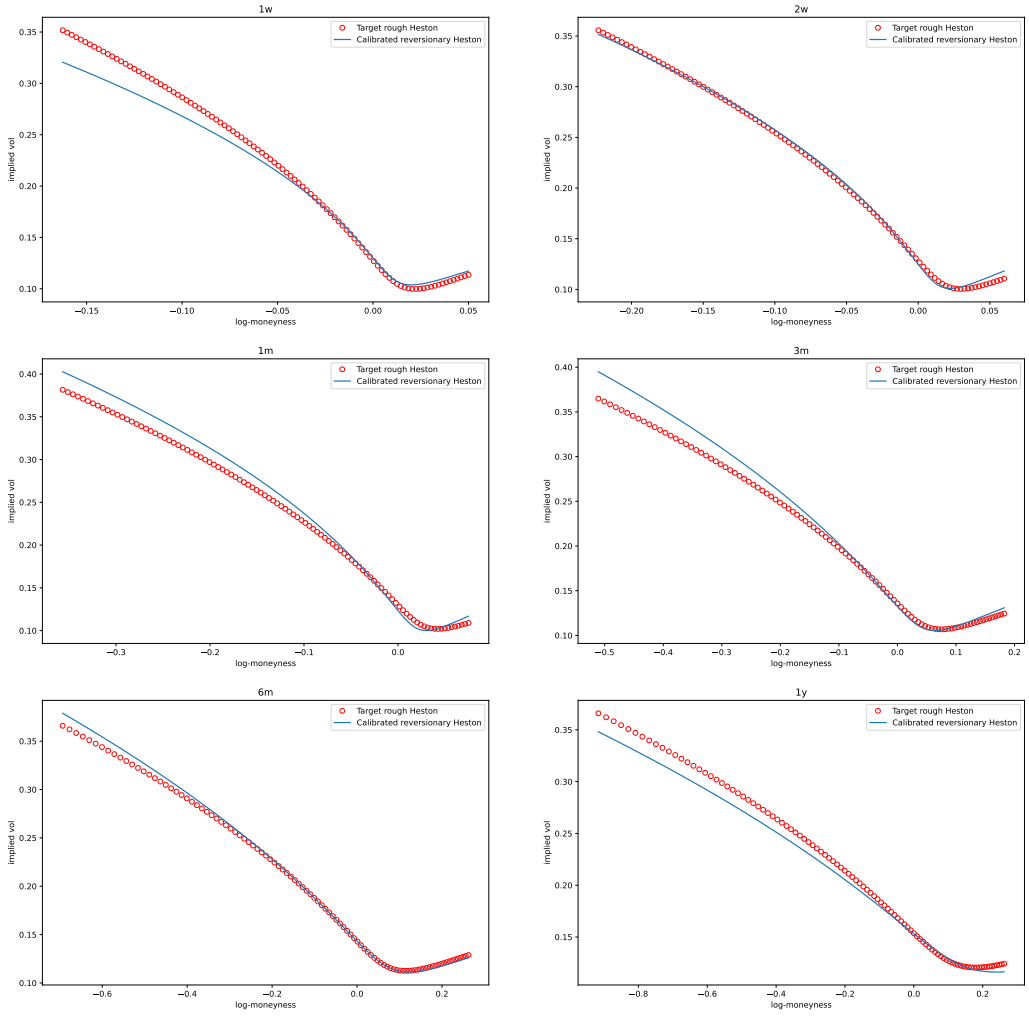

Figure 1: Smiles comparison between target rough Heston with parameters (1.10), with $H = 0.1$, and reversionary Heston with calibrated parameters from the first row of Table (1.3) for different maturities from one week to one year.

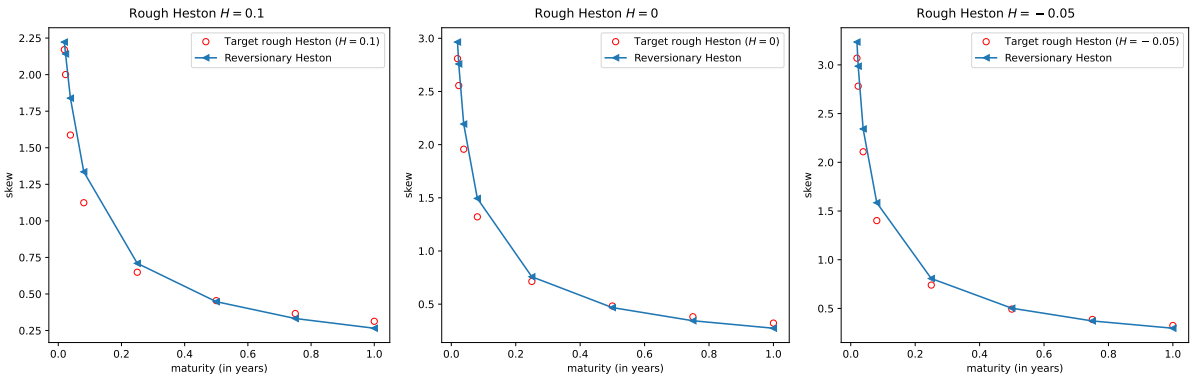

Figure 2: Resulting ATM skew $\left\{|\partial_k \sigma_{implicit}(k,T)|_{k=0}\right\}_T$ comparison between target rough Heston with parameters (1.10) and reversionary Heston with calibrated parameters given in Table (1.3) for different maturities from one week to one year.

2 The joint characteristic functional of reversionary Heston

The following theorem provides the joint conditional characteristic functional of the log-price $\log S^\epsilon$ and the integrated variance $\bar{V}^\epsilon := \int_0^\cdot V_s^\epsilon ds$ in the model (0.2)–(0.3) in terms of a solution to a system of time-dependent Riccati ordinary differential equations.

Theorem 2.1 (Joint characteristic functional)**.** *Let $f, g : [0,T] \to \mathbb{C}$ be measurable and bounded functions such that*

$$\Re g + \frac{1}{2}\left((\Re f)^2 - \Re f\right) \leq 0. \tag{2.1}$$

Then, the joint conditional characteristic functional of $(\log S^\epsilon, \bar{V}^\epsilon)$ is given by

$$\mathbb{E}\left[\exp\left(\int_t^T f(T-s)d\log S_s^\epsilon + \int_t^T g(T-s)d\bar{V}_s^\epsilon\right)\middle|\mathscr{F}_t\right] = \exp\left(\phi_\epsilon(T-t) + \epsilon^{\frac{1}{2}-H}\psi_\epsilon(T-t)V_t^\epsilon\right), \quad t \leq T, \tag{2.2}$$

where $(\phi_\epsilon, \psi_\epsilon)$ is the solution to the following system of time-dependent Riccati equations

$$\phi_\epsilon'(t) = \left(\theta + \epsilon^{-H-\frac{1}{2}}V_0\right)\psi_\epsilon(t), \quad \phi_\epsilon(0) = 0, \tag{2.3}$$

$$\begin{aligned}\psi_\epsilon'(t) = {} & \epsilon^{H-\frac{1}{2}}\frac{\xi^2}{2}\psi_\epsilon^2(t) + \left(\rho\xi\epsilon^{H-\frac{1}{2}}f(t) - \epsilon^{-1}\right)\psi_\epsilon(t) \\ & + \epsilon^{H-\frac{1}{2}}\left(g(t) + \frac{f^2(t) - f(t)}{2}\right), \qquad \psi_\epsilon(0) = 0.\end{aligned} \tag{2.4}$$

Proof. The proof is given in Section 2.2 below. □

Before proving the result, we note that in the case f and g are constant, one recovers the usual formula for the characteristic function of the Heston [33] model, where the solution $(\phi_\epsilon, \psi_\varepsilon)$ of (2.3)-(2.4) is explicit as stated in the following corollary.

Corollary 2.2 (Explicit marginals)**.** *Let $u, v \in \mathbb{R}$ and set $f(t) = iu$ and $g(t) = iv$, for all $t \in [0,T]$. Then, the solution $(\phi_\epsilon, \psi_\epsilon)$ to the Riccati equations (2.3)-(2.4) is explicitly given by*

$$\phi_\epsilon(t) = \left(\epsilon^{-\frac{1}{2}-H}\theta + \epsilon^{-1-2H}V_0\right)\xi^{-2}\left(\left(1 - i\rho\epsilon^{H+\frac{1}{2}}\xi u - d\right)t - 2\epsilon\ln\left(\frac{1 - ge^{-\epsilon^{-1}td}}{1-g}\right)\right),$$

$$\psi_\epsilon(t) = \epsilon^{-H-\frac{1}{2}}\xi^{-2}\left(1 - i\rho\epsilon^{H+\frac{1}{2}}\xi u - d\right)\frac{1 - e^{-\epsilon^{-1}td}}{1 - ge^{-\epsilon^{-1}td}},$$

with

$$g := \frac{1 - i\rho\epsilon^{H+\frac{1}{2}}\xi u - d}{1 - i\rho\epsilon^{H+\frac{1}{2}}\xi u + d}, \quad d := \sqrt{\left(1 - i\rho\epsilon^{H+\frac{1}{2}}\xi u\right)^2 - 2\left(\epsilon^{H+\frac{1}{2}}\xi\right)^2\left(iv - \frac{u^2 + iu}{2}\right)}, \quad \Re d > 0.$$

Consequently, the conditional joint characteristic function of $(\log S_T^\epsilon, \int_t^T V_s^\epsilon ds)$ *is given by*

$$\mathbb{E}\left[\exp\left(iu\log S_T^\epsilon + iv\int_t^T V_s^\epsilon ds\right)\middle|\mathscr{F}_t\right] = \exp\left(iu\log S_t^\epsilon + \phi_\epsilon(T-t) + \epsilon^{\frac{1}{2}-H}\psi_\epsilon(T-t)V_t^\epsilon\right), \quad t \le T. \tag{2.5}$$

Proof. For the explicit derivation of the formulas, see for example [30, Chapter 2]. An application of Theorem 2.1 yields the result. □

Remark 2.3. *Such formulas for* ϕ_ϵ *avoid branching issues as described in [9].*

The rest of the section if dedicated to the proof of Theorem 2.1. We first study the existence of a solution to time-dependent Riccati ODEs for which equation (2.4) is a particular case, and provide some of their properties in Section 2.1. We complete the proof of Theorem 2.1 in Section 2.2.

2.1 Time-dependent Riccati ODEs: existence and uniqueness

In this section, we consider a generic class of time-dependent Ricatti equations that encompass equation (2.4), in the form

$$\psi'(t) = a(t)\psi^2(t) + b(t)\psi(t) + c(t), \quad \psi(0) = u_0, \quad t \le T, \tag{2.6}$$

with $u_0 \in \mathbb{C}$ and $a, b, c : [0,T] \to \mathbb{C}$ three measurable and bounded functions. We say that $\psi : [0, t^*] \to \mathbb{C}$ for some $t^* \in (0,T]$ is a local extended solution to (2.6) with some initial condition $\psi(0) = u_0 \in \mathbb{C}$ if, almost everywhere on $[0,t^*]$, it is continuously differentiable and satisfies the relations in (2.6). The extended solution is global if $t^* = T$.

The presence of the squared non-linearity in (2.6) precludes the application of the celebrated Cauchy-Lipschitz theorem and can lead to explosive solutions in finite time. Compared to the related literature on similar Riccati equations [23, Lemma 2.3 and Section B], we provide a concise and simplified proof for the existence and uniqueness of a global extended solution to the Riccati equation (2.6) using a variation of constant formula under the following assumption on the coefficients (a,b,c) and the initial condition u_0:

$$\Im(a(t)) = 0, \quad a(t) > 0, \quad \Re(c(t)) + \frac{\Im(b(t))^2}{4a(t)} \le 0, \quad \Re(u_0) \le 0, \quad t \le T. \tag{2.7}$$

The following theorem gives the existence and uniqueness of a solution to the Riccati equation (2.6).

Theorem 2.4 (Existence and uniqueness for the Riccati)**.** *Let* $u_0 \in \mathbb{C}$ *with* $\Re(u_0) \le 0$ *and* $a, b, c : [0,T] \to \mathbb{C}$ *be measurable and bounded functions satisfying* (2.7)*. Then, there exists a unique extended solution* $\psi : [0,T] \to \mathbb{C}$ *to the Ricatti equation* (2.6) *such that*

$$\Re(\psi(t)) \le 0, \quad t \le T, \tag{2.8}$$

and

$$\sup_{t \le T}|\psi(t)| < \infty. \tag{2.9}$$

Proof. For the existence part, we proceed in two steps. First, we start by arguing the existence of a local solution using Carathéodory's theorem. For this we rely on [32, Chapter 12, Section 2], using the notations therein (see equation (1.7) for example), we consider the integral equation

$$\psi(t) = \psi(0) + \int_0^t g(t, s, \psi(s))\, ds, \quad t \ge 0, \tag{2.10}$$

where the operator g is defined by

$$g(t,s,\psi(s)) := a(s)x(s)^2 + b(s)x(s) + c(s).$$

Let D be an open, connected subset of $\mathbb{R}^+ \times \mathbb{C}$ that contains $(0,\psi(0))$. Define

$$T_\infty := \sup\left\{t \in \mathbb{R}^+ | C_{\psi(0),D}\left([0,T),\mathbb{C}\right) \neq \emptyset\right\},$$

where

$$C_{\psi(0),D}\left([0,T),\mathbb{C}\right) := \left\{\phi \in C\left(D,\mathbb{C}\right) | \phi(0) = \psi(0) \text{ and } (t,\phi(t)) \in D \text{ for } t \in [0,T)\right\}.$$

An application of [32, Theorem 2.6] yields the existence of a unique non-continuable solution to (2.10) that satisfies $(t,\psi(t)) \in D$ on the interval $[0,T_\infty)$ such that either $T_\infty = T$ or $\lim\limits_{t\to T_\infty} |\psi(t)| = \infty$. Indeed, the assumptions (i) to (v) of [32, Theorem 2.6] are readily satisfied by boundedness and integrability of a, b and c and the fact that g does not depend on t and satisfies the Carathéodory conditions.
Second, we argue that

$$\sup_{t\leq T_\infty} |\psi(t)| < \infty, \tag{2.11}$$

which would then yield $T_\infty = T$ and the existence of a global solution ψ. Let $t < T_\infty$. We start by showing that $\Re(\psi(t)) \leq 0$. Indeed, taking real parts in (2.6), $\psi_{\mathrm{r}} := \Re\left(\psi\right)$ satisfies the following equation on $[0,T_\infty)$:

$$\psi_{\mathrm{r}}'(s) = \left\{a(s)\Re\left(\psi(s)\right) + \Re\left(b(s)\right)\right\}\psi_{\mathrm{r}}(s) + d(s),$$

where $d(s) = -a(s)\left(\Im\left(\psi(s)\right) + \frac{\Im(b(s))}{2a(s)}\right)^2 + \Re\left(c(s)\right) + \frac{\Im^2(b(s))}{4a(s)} \leq 0$ thanks to condition (2.7), after a completion of squares. The variation of constant for ψ_{r} then yields

$$\psi_{\mathrm{r}}(t) = e^{\int_0^t (a(u)\Re(\psi(u)) + \Re(b(u)))du}\Re(u_0) + \int_0^t d(s)e^{\int_s^t (a(u)\Re(\psi(u)) + \Re(b(u)))du}ds \leq 0,$$

since the exponential is positive and $d(s) \leq 0$, and $\Re(u_0) \leq 0$ by assumption. This shows that $\Re(\psi) \leq 0$, on $[0,T_\infty)$. Finally, an application of a similar variation of constants formula on equation (2.6) leads to

$$\psi(t) = e^{\int_0^t (a(u)(\psi(u)) + \Re(b(u)))du}u_0 + \int_0^t e^{\int_s^t (a(u)(\psi(u)) + (b(u)))du}c(s)ds,$$

so that taking the module together with the triangle inequality and the fact that $\Re(\psi) \leq 0$ on $[0,T_\infty)$, yields

$$\begin{aligned}
|\psi(t)| &\leq \left|e^{\int_0^t (a(u)\psi(u) + (b(u)))du}\right| |u_0| + \int_0^t \left|e^{\int_s^t (a(u)\psi(u) + (b(u)))du}\right| |c(s)|ds \\
&= e^{\int_0^t (a(u)\Re(\psi(u)) + \Re(b(u)))du}|u_0| + \int_0^t e^{\int_s^t (a(u)\Re(\psi(u)) + \Re(b(u)))du}|c(s)|ds \\
&\leq e^{\int_0^t \Re(b(u))du}|u_0| + \int_0^T e^{\int_s^t \Re(b(u))du}|c(s)|ds \\
&\leq C\left(|u_0| + \int_0^T |c(s)|ds\right),
\end{aligned}$$

where $C = \sup_{s,s'\in[0,T]^2} e^{\int_s^{s'} \Re(b(u))du}$ does not depend on t and is finite by boundedness of b. This shows (2.11) as needed. Combining the above we obtain the existence of a solution ψ on $[0,T]$ satisfying (2.8) and (2.9).
To argue uniqueness, assume there are two such extended solutions ψ_1 and ψ_2 that satisfy (2.6). Then,

$$\left(\psi_2 - \psi_1\right)'(t) = \left(a(t)\left(\psi_2 + \psi_1\right)(t) + b(t)\right)\left(\psi_2 - \psi_1\right)(t), \quad \left(\psi_2 - \psi_1\right)(0) = 0, \quad t \leq T,$$

which yields

$$\left|\psi_2 - \psi_1\right|(t) \leq \int_0^t \left|a(s)\left(\psi_2 + \psi_1\right)(s) + b(s)\right| \left|\psi_2 - \psi_1\right|(s)ds \leq c\int_0^t \left|\psi_2 - \psi_1\right|(s)ds, \quad t \leq T,$$

for some $c > 0$ by boundedeness of (ψ_1, ψ_2, a, b) using (2.11), so that the uniqueness is obtained from Gronwall's lemma. □

2.2 Proof of Theorem 2.1

We first argue the existence of a solution to the system of Riccati equations (2.3)-(2.4). Let us rewrite the Ricatti ODE from (2.4) as

$$\psi_\epsilon'(t) = a_\epsilon \psi_\epsilon^2(t) + b_\epsilon(t)\psi_\epsilon(t) + c_\epsilon(t), \quad \psi_\epsilon(0) = 0, \quad t \leq T, \tag{2.12}$$

where we defined

$$\begin{cases} a_\epsilon & := \epsilon^{H-\frac{1}{2}} \frac{\xi^2}{2} \\ b_\epsilon(t) & := \epsilon^{H-\frac{1}{2}} \rho \xi f(t) - \epsilon^{-1} \\ c_\epsilon(t) & := \epsilon^{H-\frac{1}{2}} \left[g(t) + \frac{f^2(t) - f(t)}{2} \right]. \end{cases} \tag{2.13}$$

Since condition (2.1) ensures

$$\Re\left(c_\epsilon\right) + \frac{\Im\left(b_\epsilon\right)^2}{4a_\epsilon} = \epsilon^{H-1/2} \left(\Re g + \frac{1}{2} \left(\left(\Re f\right)^2 - \Re f \right) + \left(\rho^2 - 1\right) \left(\Im f\right)^2 \right) \leq 0,$$

then conditions (2.7) are readily satisfied and consequently Theorem 2.4 yields the existence and uniqueness of a solution $\psi_\epsilon : [0,T] \to \mathbb{C}$ to the Ricatti ODE (2.4) such that

$$\Re(\psi_\epsilon(t)) \leq 0, \quad t \leq T.$$

The function ϕ_ϵ defined in integral form as

$$\phi_\epsilon(t) = \left(\theta + \epsilon^{-H-\frac{1}{2}} V_0\right) \int_0^t \psi_\epsilon(s) ds, \quad t \leq T,$$

solves (2.3).
We now prove the expression for the charateristic functional (2.2). Define the following process M:

$$\begin{aligned} M_t &= \exp(U_t), \\ U_t &= \phi_\epsilon\left(T-t\right) + \epsilon^{\frac{1}{2}-H} \psi_\epsilon\left(T-t\right) V_t^\epsilon + \int_0^t f(T-s) d\log S_s^\epsilon + \int_0^t g(T-s) d\bar{V}_s^\epsilon. \end{aligned}$$

In order to obtain (2.2), it suffices to show that M is a martingale. Indeed, if this is the case, and after observing that the terminal value of M is given by

$$M_T = \exp\left(\int_0^T f(T-s) d\log S_s^\epsilon + \int_0^T g(T-s) d\bar{V}_s^\epsilon \right),$$

recall that $\phi_\epsilon(0) = \psi_\epsilon(0) = 0$, we obtain

$$\mathbb{E}\left[\exp\left(\int_0^T f(T-s) d\log S_s^\epsilon + \int_0^T g(T-s) d\bar{V}_s^\epsilon \right) \middle| \mathscr{F}_t \right] = \mathbb{E}\left[M_T | \mathscr{F}_t \right] = M_t = \exp\left(U_t\right),$$

which yields (2.2). We now argue that M is a martingale. We first show that M is a local martingale using Itô formula. The dynamics of M read

$$dM_t = M_t \left(dU_t + \frac{1}{2} d\langle U \rangle_t \right),$$

with

$$\begin{aligned} dU_t = &\left\{ \phi_\epsilon'(T-t) - \left(\theta + \epsilon^{-\frac{1}{2}-H} V_0\right) \psi_\epsilon(T-t) + \left(-\epsilon^{\frac{1}{2}-H} \psi_\epsilon'(T-t) - \epsilon^{-\frac{1}{2}-H} \psi_\epsilon(T-t) + g(T-t) - \frac{f(T-t)}{2} \right) V_t^\epsilon \right\} dt \\ &+ \left(\xi \psi_\epsilon(T-t) + \rho f(T-t)\right) \sqrt{V_t^\epsilon} dW_t + \sqrt{1-\rho^2} f(T-t) \sqrt{V_t^\epsilon} dW_t^\perp. \end{aligned}$$

This yields that the drift in dM_t/M_t is given by

$$\phi'_\epsilon(T-t) - \left(\theta + \epsilon^{-\frac{1}{2}-H}V_0\right)\psi_\epsilon(T-t)$$
$$+ \left(-\epsilon^{\frac{1}{2}-H}\psi'_\epsilon(T-t) + \frac{\xi^2}{2}\left(\psi_\epsilon(T-t)\right)^2 + \left(\rho\xi f(T-t) - \epsilon^{-\frac{1}{2}-H}\right)\psi_\epsilon(T-t) + g(T-t) + \frac{f^2(T-t) - f(T-t)}{2}\right)V_t^\epsilon$$

which is equal to 0 from the Riccati equations (2.3) and (2.4). This shows that M is a local martingale. To argue that M is a true martingale, we note that $\Re(\psi_\epsilon) \leq 0$ which implies $\Re(\phi_\epsilon) \leq 0$, so that

$$\begin{aligned}
\Re(U_t) &\leq \int_0^t \Re(f(T-s))d\log S_s^\epsilon + \int_0^t \Re(g(T-s))d\bar{V}_s^\epsilon \\
&= \int_0^t \left(\Re(g(T-s)) - \frac{1}{2}\Re(f(T-s))\right)V_s^\epsilon ds + \int_0^t \Re(f(T-s))\sqrt{V_s^\epsilon}dB_s \\
&\leq -\frac{1}{2}\int_0^t \Re(f(T-s))^2 V_s^\epsilon ds + \int_0^t \Re(f(T-s))\sqrt{V_s^\epsilon}dB_s =: \tilde{U}_t,
\end{aligned}$$

where the last inequality follows from (2.1). It follows that

$$|M_t| = \exp(\Re(U_t)) \leq \exp(\tilde{U}_t),$$

where the process $\exp(\tilde{U})$ is a true martingale, see [6, Lemma 7.3]. This shows that M is a true martingale, being a local martingale bounded by a true martingale, see [34, Lemma 1.4], which concludes the proof.

3 From reversionary Heston to jump processes

In this section, we establish the convergence of the log-price and the integrated variance $(\log S^\epsilon, \bar{V}^\epsilon)$ in the reversionary Heston model (0.2)-(0.3) towards a Lévy jump process (X, Y), as ϵ goes to 0. More precisely, the limit (X, Y) belongs to the class of Normal Inverse Gaussian - Inverse Gaussian (NIG-IG) processes defined as follows.

Definition 3.1 (NIG-IG process)**.** *Fix $\alpha \geq |\beta| \geq 0$, $\delta, \lambda > 0$ and $\mu \in \mathbb{R}$. We say that $(X_t, Y_t)_{t\geq 0}$ is a Normal Inverse Gaussian - Inverse Gaussian (NIG-IG) process with parameters $(\alpha, \beta, \delta, \mu, \lambda)$ if it is a two-dimensional homogeneous Lévy process with càdlàg sample paths, starting from $(X_0, Y_0) = (0, 0)$ almost surely, with Lévy exponent η defined by*

$$\eta(u, v) := \left[i\mu u + \delta\left(\sqrt{\alpha^2 - \beta^2} - \sqrt{\alpha^2 - 2i\lambda v - (\beta + iu)^2}\right)\right], \quad u, v \in \mathbb{R}, \tag{3.1}$$

i.e. the joint characteristic function is given by

$$\mathbb{E}\left[\exp\left(iuX_t + ivY_t\right)\right] = \exp\left(\eta(u, v)t\right), \quad u, v \in \mathbb{R}, \quad t \leq T. \tag{3.2}$$

In order to justify the existence of such a class of Lévy processes, one needs to justify that η given in (3.1) is indeed the logarithm of a characteristic function associated to an infinitely divisible distribution, see [41, Corollary 11.6]. This is the object of the following lemma, which also provides the link with first-hitting times and subordinated processes.

Lemma 3.2 (Representation using subordination)**.** *Let $\alpha \geq |\beta| \geq 0$, $\delta, \lambda > 0$, $\mu \in \mathbb{R}$ and $(\widetilde{W}, \widetilde{W}^\perp)$ be a two dimensional Brownian motion. Let $(\Lambda_t)_{t\in[0,T]}$ be the first hitting-time process defined as*

$$\Lambda_t := \inf\left\{s \geq 0 : \sqrt{\alpha^2 - \beta^2}s + \widetilde{W}_s \geq \delta t\right\}, \quad t \in [0, T], \tag{3.3}$$

and define Z as the following shifted subordinated process

$$Z_t = \mu t + \beta\Lambda_t + \widetilde{W}^\perp_{\Lambda_t}, \quad t \in [0, T]. \tag{3.4}$$

Then,

$$\mathbb{E}\left[\exp\left(iuZ_t + iv\lambda\Lambda_t\right)\right] = \exp(\eta(u,v)t), \quad u,v \in \mathbb{R}, \quad t \in [0,T].$$

In particular, η given by (3.1) is the logarithm of the characteristic function of the joint random variable $(Z_1, \lambda\Lambda_1)$ which is infinitely divisible.

Proof. Fix $t \in [0,T]$. By construction, it is well-known that Λ_t has an Inverse Gaussian distribution if $\alpha > |\beta|$ with parameters $IG\left(\frac{\delta t}{\sqrt{\alpha^2-\beta^2}}, \delta^2 t\right)$, and in the drift-free case $\alpha = |\beta|$, Λ_t follows a Lévy distribution with parameters $Lévy\left(0, \delta^2 t\right)$ (see [13] and Definition A.3 in the Appendix). Now conditional on Λ_t, Z_t is Gaussian with parameters $\mathcal{N}\left(\mu t + \beta\Lambda_t, \Lambda_t\right)$ and using the tower property of conditional expectation, we get for the first case that

$$\begin{aligned}
\mathbb{E}\left[\exp\left(iuZ_t + iv\lambda\Lambda_t\right)\right] &= \mathbb{E}\left[\mathbb{E}\left[\exp\left(iuZ_t\right)\middle|\,\Lambda_t\right]\exp\left(iv\lambda\Lambda_t\right)\right] \\
&= \mathbb{E}\left[\exp\left(iu\left(\mu t + \beta\Lambda_t\right) - \frac{\Lambda_t u^2}{2} + iv\lambda\Lambda_t\right)\right] \\
&= \exp\left(iu\mu t\right)\mathbb{E}\left[\exp\left(\left(iu\beta - \frac{u^2}{2} + iv\lambda\right)\Lambda_t\right)\right] \\
&= \exp\left(iu\mu t + \delta t\sqrt{\alpha^2-\beta^2}\left(1 - \sqrt{1 - \frac{2}{\alpha^2-\beta^2}\left(iu\beta - \frac{u^2}{2} + iv\lambda\right)}\right)\right) \\
\text{i.e. } \mathbb{E}\left[\exp\left(iuZ_t + iv\lambda\Lambda_t\right)\right] &= \exp\left(iu\mu t + \delta t\left(\sqrt{\alpha^2-\beta^2} - \sqrt{\alpha^2 - 2i\lambda v - (\beta+iu)^2}\right)\right) = \exp(\eta(u,v)t), \quad u,v \in \mathbb{R},
\end{aligned}$$

where we used Definition A.3 to get the fourth equality, noting that $\Re\left(i\left(u\beta + v\lambda\right) - \frac{u^2}{2}\right) \leq 0$. Similar computations yield the result for the case $\alpha = |\beta|$. Furthermore, we will say that that the random variable $(Z_1, \lambda\Lambda_1)$ follows a NIG-IG distribution with parameters $(\alpha, \beta, \mu, \delta, \lambda)$ (see Definition A.3 in the Appendix). Such distribution is infinitely divisible because if $(X_1, Y_1), \cdots, (X_m, Y_m)$ are independent NIG-IG random variables with common parameters (α, β, λ) and individual (μ_i, δ_i), for $i = 1, \cdots, m$, then $(X, Y) := (\sum_{i=1}^m X_i, \sum_{i=1}^m Y_i)$ is again NIG-IG-distributed with parameters $(\alpha, \beta, \sum_{i=1}^m \mu_i, \sum_{i=1}^m \delta_i, \lambda)$. □

The appellation NIG-IG for the couple (X, Y) in Definition 3.1 is justified as follows:

- Y is an Inverse Gaussian process first derived by Schrödinger [42] which can be checked either by recovering the Inverse Gaussian distribution with parameters $IG\left(\frac{\lambda\delta}{\sqrt{\alpha^2-\beta^2}}, \lambda\delta^2\right)$ after setting $u = 0$ in (3.2); or by using the representation as a first passage-time in (3.3). It is worth pointing that, for $\alpha = |\beta|$, one recovers the well-known Lévy distribution for the first-passage of a Brownian motion with parameters $Lévy\left(0, \lambda\delta^2\right)$. The *Lévy* distribution can be seen as a special case of the *Inverse Gaussian* distribution.
- X is the celebrated Normal Inverse Gaussian process of Barndorff-Nielsen [13], with parameters $NIG\left(\alpha, \beta, \mu, \delta\right)$, which can be checked by setting $v = 0$ in (3.2) or by using the representation as subordinated Brownian motion with an Inverse Gaussian subordinator as in (3.4).

In addition, we allow in Definition 3.1 the parameter α to be equal to ∞ in the following sense:

Remark 3.3 (Normal process)**.** *Considering the set of parameters*

$$(\alpha, \beta, \delta, \mu, \lambda) = \left(\alpha, 0, \sigma^2\alpha, \mu, 1\right),$$

a second order Taylor expansion, as $\alpha \to \infty$, of the square root yields

$$\mathbb{E}\left[\exp\left(iuX_t + ivY_t\right)\right] = \exp\left(\left[i\mu u - \sigma^2\left(\frac{u^2}{2} + iv\right)\right]t\right), \quad u,v \in \mathbb{R}, \quad t \leq T,$$

which is equivalent to the normal-deterministic process defined by

$$(X_t, Y_t)_{t\in[0,T]} = \left(\mu t + \sigma\widetilde{W}_t, \sigma^2 t\right)_{t\in[0,T]} \overset{d}{=} \left(\mu t + \widetilde{W}_{\sigma^2 t}, \sigma^2 t\right)_{t\in[0,T]}.$$

We will consider that such (degenerate) process is a particular case of Definition 3.1 with parameters denoted by

$$\left(\alpha, 0, \sigma^2\alpha, \mu, 1\right)\big|_{\alpha\to\infty}.$$

We are now in place to state our main convergence theorem. Theorem 3.4 provides the convergence of the finite-dimensional distributions of the joint process $(\log S^\epsilon, \bar{V}^\epsilon)$ through the study of the limiting behavior of the characteristic functional given in Theorem 2.1. Interestingly, the limiting behavior disentangles three different asymptotic regimes based on the values of H that can be seen intuitively on the level of the Riccati equation (2.4) as follows. Applying the variation of constants on ψ_ϵ, we get:

$$\psi_\epsilon(t) = \epsilon^{H+1/2}\int_0^t K_\epsilon(t-s)F(s,\psi_\epsilon(s))ds, \quad t \le T, \tag{3.5}$$

$$F(s,u) := \frac{\xi^2}{2}u^2 + \rho\xi f(s)u + g(s) + \frac{f^2(s) - f(s)}{2}, \tag{3.6}$$

with K_ϵ the kernel defined by

$$K_\epsilon(t) = \epsilon^{-1}e^{-\epsilon^{-1}t}, \quad t \ge 0.$$

Assuming that ψ_ϵ converges to some ψ_0 and observing that K_ϵ plays the role of the Dirac delta as $\epsilon \to 0$, one expects $\int_0^t K_\epsilon(t-s)F(s,\psi_\epsilon(s))ds \to F(t,\psi_0(t))$ in (3.5), the pre-factor $\epsilon^{H+1/2}$ suggests then three different limiting regimes with respect to H that can be characterized through the functions F and ψ_0:

$$\begin{cases} \psi_0(t) = 0, & \text{if } H > -1/2, \\ \psi_0(t) = F(t,\psi_0(t)), & \Re(\psi_0(t)) \le 0, \quad \text{if } H = -1/2, \\ 0 = F(t,\psi_0(t)), & \Re(\psi_0(t)) \le 0, \quad \text{if } H < -1/2. \end{cases} \tag{3.7}$$

The function ψ_0 in (3.7) is even explicitly given by

$$\psi_0(t) = \begin{cases} 0, & \text{if } H > -1/2, \\ \xi^{-2}\left(1 - \rho\xi f(t) - \sqrt{\left(1-\rho\xi f(t)\right)^2 - 2\xi^2\left(g(t) + \frac{f^2(t)-f(t)}{2}\right)}\right), & \text{if } H = -1/2, \\ -\xi^{-1}\left(\rho f(t) + \sqrt{f(t)\left(1-(1-\rho^2)f(t)\right) - 2g(t)}\right), & \text{if } H < -1/2, \end{cases} \tag{3.8}$$

see Lemma A.1 below. Furthermore, the convergence of the integrated variance process is strengthened to a functional weak convergence on the Skorokhod space $(D(0,T), M_1)$ of càdlàg paths on $[0,T]$ endowed with the strong M_1 topology, see Section 3.3.1 below. Such topology is weaker and less restrictive than the commonly used uniform or J_1 topologies which share the property that a jump in a limiting process can only be approximated by jumps of comparable size at the same time or, respectively, at nearby times. On the contrary, the M_1 topology of Skorokhod [43] captures approximations of unmatched jumps, which in our case, will allow us to prove the convergence of the stochastic process $\bar{V}^\epsilon$ with continuous sample trajectories towards a Lévy process with càdlàg sample trajectories. The statement is now made rigorous in the following theorem.

Theorem 3.4 (Convergence towards NIG-IG processes)**.** *Let $f, g : [0,T] \to \mathbb{C}$ be bounded and measurable such that $\Re f = \Re g = 0$ and such that ψ_0 defined in* (3.8) *has bounded variations. Then, based on the value of H, we obtain different explicit asymptotic formulas for the characteristic functional given in Theorem 2.1:*

$$\lim_{\epsilon\to 0}\mathbb{E}\left[\exp\left(\int_0^T f(T-s)d\log S_s^\epsilon + \int_0^T g(T-s)d\bar{V}_s^\epsilon\right)\right] = \exp\left(\phi_0(T)\right), \tag{3.9}$$

with

$$\phi_0(T) := \begin{cases} V_0 \int_0^T h(s)ds, & \text{if } H > -1/2, \\ (\theta + V_0)\,\xi^{-2}\left(T - \int_0^T \left(\rho\xi f(s) + \sqrt{(1-\rho\xi f(s))^2 - 2\xi^2 h(s)}\right) ds\right), & \text{if } H = -1/2, \\ -\theta\xi^{-1}\int_0^T \left(\rho f(s) + \sqrt{\rho^2 f^2(s) - 2h(s)}\right) ds, & \text{if } H < -1/2, \end{cases} \tag{3.10}$$

where $h(s) := g(s) + \frac{f^2(s)-f(s)}{2}$. *In particular for* $\rho \in (-1,1)$, *as* $\epsilon \to 0$, *the finite-dimensional distributions of the joint process* $(\log \frac{S^\epsilon}{S_0}, \bar{V}^\epsilon)$ *converge to the finite-dimensional distributions of a NIG-IG process* (X,Y) *in the sense of Definition 3.1 with the following parameters depending on the value of* H*:*

$$(\alpha,\beta,\delta,\mu,\lambda) := \begin{cases} \left(\alpha, 0, V_0\alpha, -\frac{V_0}{2}, 1\right)\big|_{\alpha\to\infty}, & \text{if } H > -1/2, \\ \left(\frac{1}{2}\frac{\sqrt{(\xi-2\rho)^2+4(1-\rho^2)}}{\xi(1-\rho^2)}, -\frac{1}{2}\frac{\xi-2\rho}{\xi(1-\rho^2)}, \sqrt{1-\rho^2}(\theta+V_0)\xi^{-1}, -\rho\,(\theta+V_0)\,\xi^{-1}, \frac{1}{1-\rho^2}\right), & \text{if } H = -1/2, \\ \left(\frac{1}{2(1-\rho^2)}, -\frac{1}{2(1-\rho^2)}, \sqrt{1-\rho^2}\theta\xi^{-1}, -\rho\theta\xi^{-1}, \frac{1}{1-\rho^2}\right), & \text{if } H < -1/2, \end{cases} \tag{3.11}$$

where θ, S_0, ξ *and* V_0 *are the same from* (0.2)-(0.3). *Furthermore, the process* $\bar{V}^\epsilon$ *converges weakly towards* Y *on the space* $(D(0,T), M_1)$, *as* $\epsilon \to 0$.

Proof. The convergence of the characteristic functional in (3.9) is established in Section 3.1 (Lemmas 3.7 and 3.8). This implies the convergence of the finite-dimensional distributions of $(\log S^\epsilon, \bar{V}^\epsilon)$ as detailed in Section 3.2. Finally, the weak convergence of $\bar{V}^\epsilon$ on $(D(0,T), M_1)$ is proved in Section 3.3. □

Remark 3.5 (An interesting interpretation)**.** *The interpretation of the convergence results becomes even more interesting when combined with Section 1. In Section 1, for* $H > -1/2$, *the reversionary Heston model* $(\log S^\epsilon, \bar{V}^\epsilon)$ *is constructed as a proxy of rough and hyper-rough Heston models. Theorem 3.4 shows that the limiting regime for* $H > -1/2$ *is a (degenerate) Black-Scholes regime, cf.Remark 3.3, whereas, for* $H \le -1/2$ *one obtains the convergence of the reversionary regimes towards (non-degenerate) jump processes with distinct regimes between* $H = -1/2$ *and* $H < -1/2$, *see Corollary 3.6 below. This suggests that jump models and (hyper-)rough volatility models are complementary, and do not overlap. For* $H > -1/2$ *the reversionary model can be interpreted as a proxy of rough and hyper-rough volatility models, while for* $H \le -1/2$ *it can be interpreted as a proxy of jump models. Jump models actually start at* $H = -1/2$ *(and below), the first value for which hyper-rough volatility models can no-longer be defined.*

In Figures 3 and 4, we plot respectively the convergence of the smiles and the skew of the reversionary Heston model $(\log S^\epsilon, \bar{V}^\epsilon)$ for the case $H = -1/2$ towards the Normal Inverse Gaussian model. The volatility surface is obtained by applying Fourier inversion formulas on the corresponding characteristic functions. Similar to Figures 1 and 2, the graphs show that the fast parametrizations introduced in the Heston model are able to reproduce very steep skews for the implied volatility surface.

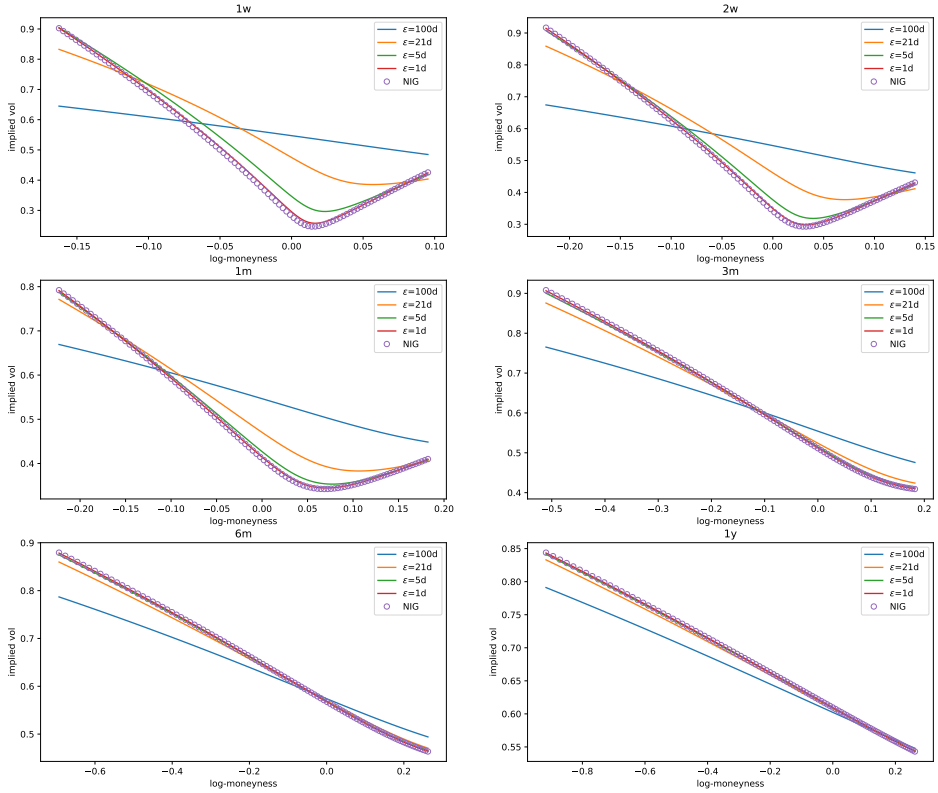

Figure 3: Reversionary Heston smiles for different maturities from one week to one year. Parameters are: $S_0 = 100, \rho = -0.7, \theta = 0.3$, $\xi = 0.8$, $V_0 = 0.3$ and $H = -1/2$, and the reversionary time-scale is varied from one hundred days to one day.

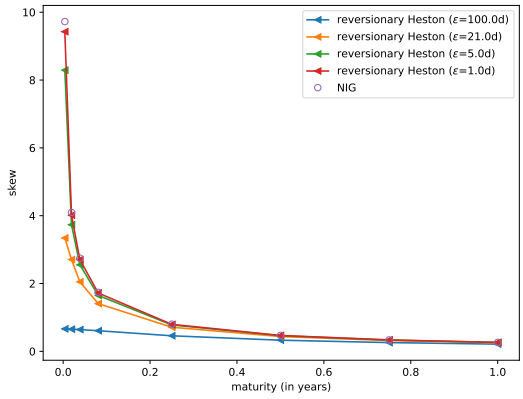

Figure 4: ATM skew of reversionary Heston and its asymptotic NIG law in the regime $H = -0.5$. Parameters are: $S_0 = 100, \rho = -0.7, \theta = 0.3$, $\xi = 0.8$, $V_0 = 0.3$ and the reversionary time-scale is varied from one hundred days to one day.

In the case $(f,g) = (iu, iv)$, with $u, v \in \mathbb{R}$, the asymptotic marginals of reversionary Heston expressed in Corollary 3.6 below are obtained as a direct consequence of the convergence Theorem 3.4.

Corollary 3.6 (Explicit asymptotic marginals)**.** *Based on the value of $H \in \mathbb{R}$, the pair of normalized log price and integrated variance $\left(\log \frac{S_T^\epsilon}{S_0}, \bar{V}_T^\epsilon\right)$ has distinct asymptotic marginals as the reversionary time-scale ϵ goes to zero given by:*

1. *$H > -1/2$, i.e. Black Scholes-type asymptotic regime (BS regime)*

$$\mathbb{E}\left[iu \log \frac{S_T^\epsilon}{S_0} + iv\bar{V}_T^\epsilon\right] \underset{\epsilon\to 0}{\longrightarrow} \exp\left\{-\frac{V_0}{2}\left(u^2 - 2i\left(v - \frac{u}{2}\right)\right)T\right\}. \tag{3.12}$$

2. *$H = -1/2$, i.e. Normal Inverse Gaussian-type asymptotic regime (NIG regime)*

$$\mathbb{E}\left[iu \log \frac{S_T^\epsilon}{S_0} + iv\bar{V}_T^\epsilon\right] \underset{\epsilon\to 0}{\longrightarrow} \exp\left\{(\theta + V_0)\,\xi^{-2}\left(1 - i\rho\xi u - \sqrt{(1 - i\rho\xi u)^2 - 2\xi^2\left(iv - \frac{u^2 + iu}{2}\right)}\right)T\right\}. \tag{3.13}$$

3. *$H < -1/2$, i.e. Normal Lévy-type asymptotic regime (NL regime)*

$$\mathbb{E}\left[iu \log \frac{S_T^\epsilon}{S_0} + iv\bar{V}_T^\epsilon\right] \underset{\epsilon\to 0}{\longrightarrow} \exp\left\{-\theta\xi^{-1}\left(i\rho u + \sqrt{(1-\rho^2)\,u^2 - 2i\left(v - \frac{u}{2}\right)}\right)T\right\}. \tag{3.14}$$

In Figure 5, we illustrate numerically the convergence of the characteristic function in all three regimes.

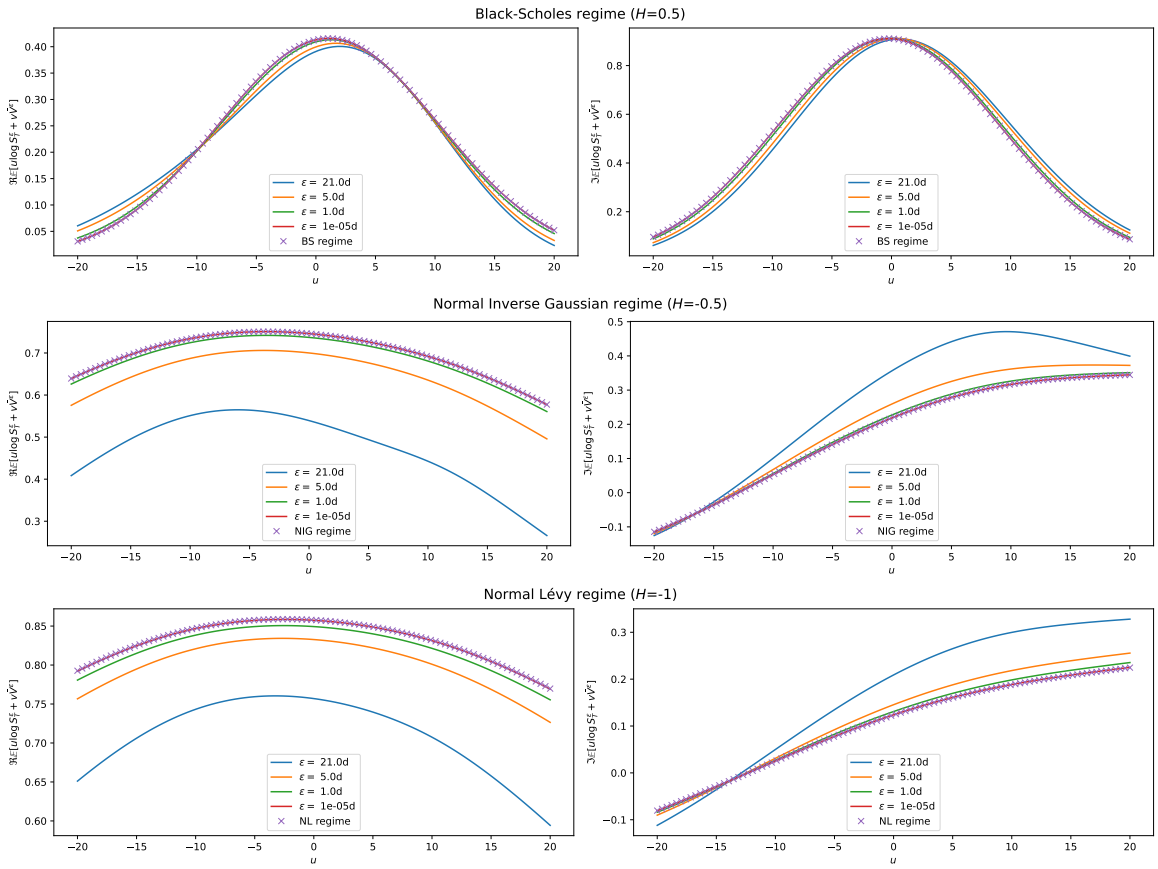

Figure 5: Convergence of reversionary Heston's joint characteristic function (2.5) in all three regimes whose asymptotic joint characteristic functions are given respectively in (3.12) for the Black-Scholes regime, in (3.13) for the Normal Inverse Gaussian regime and in (3.14) for the Normal-Lévy regime. v is fixed at 100 and model parameters are $\rho = -0.7, \theta = 0.3$, $\xi = 0.8$, $V_0 = 0.3$ and the reversionary time-scale is varied from one 21 days to $1e-5$ day.

The rest of the section is dedicated to the proof of Theorem 3.4.

3.1 Convergence of the joint characteristic functional

In this section we prove the convergence of the characteristic functional of $(\log \frac{S^\epsilon}{S_0}, \bar{V}^\epsilon)$ as ϵ goes to 0 stated in Theorem 3.4. For this, we fix $f, g : [0,T] \to \mathbb{C}$ bounded and measurable such that $\Re f = \Re g = 0$. We note that (2.1) is trivially satisfied so that an application of Theorem 2.1, with $t = 0$, yields that

$$\mathbb{E}\left[\exp\left(\int_0^T f(T-s)d\log S_s^\epsilon + \int_0^T g(T-s)d\bar{V}_s^\epsilon\right)\right] = \exp\left(\phi_\epsilon(T) + \epsilon^{\frac{1}{2}-H}\psi_\epsilon(T)V_0\right),$$

where $(\phi_\epsilon, \psi_\epsilon)$ solve (2.3)-(2.4). We start by showing that the second term in the exponential $\epsilon^{1/2-H}\psi_\epsilon$ goes to 0 for any value of $H \in \mathbb{R}$ in the following lemma.

Lemma 3.7. *For $\epsilon > 0$, let ψ_ϵ be a solution the time-dependent Ricatti ODE (2.4) such that $\Re(\psi^\epsilon) \leq 0$ with $f, g : [0,T] \to \mathbb{C}$ bounded and measurable functions such that $\Re f = \Re g = 0$. Then,*

$$|\psi_\epsilon(t)| \leq C\epsilon^{H+\frac{1}{2}}\left(1 - e^{-\epsilon^{-1}t}\right), \quad t \leq T,$$

for some constant C independent of ϵ. In particular, we have the uniform convergence

$$\lim_{\epsilon\to 0}\sup_{t\leq T}\epsilon^{1/2-H}|\psi_\epsilon(t)| = 0, \quad H \in \mathbb{R}.$$

Proof. The variation of constants applied to the differential equation (2.12) yields:

$$\psi_\epsilon(t) = \int_0^t c_\epsilon(s) e^{\int_s^t (a_\epsilon \psi_\epsilon(u) + b_\epsilon(u))du} ds,$$

with $a_\epsilon, b_\epsilon, c_\epsilon$ defined as in (2.13). Given that f and g are both bounded on $[0,T]$, we fix $C \geq 0$ such that $|c_\epsilon(u)| \leq C\epsilon^{H-\frac{1}{2}}, u \leq T$. Note that $\Re b_\epsilon = -\epsilon^{-1}$, having also $\Re(\psi_\epsilon) \leq 0$ and $a_\epsilon > 0$, we get the desired upper bound on the solution. The convergence result follows immediately. □

The first term ϕ_ϵ, however, yields different limits based on the value of H. Consequently, we will study in Lemma 3.8 the convergence of the following quantity

$$\phi_\epsilon(t) = \int_0^t \left(\theta + \epsilon^{-H-\frac{1}{2}} V_0\right) \psi_\epsilon(s) ds,$$

for different regimes of H.

Lemma 3.8. *We have the convergence*

$$\lim_{\epsilon \to 0} \phi_\epsilon(T) = \phi_0(T),$$

where $\phi_0(T)$ is given by (3.10).

Proof. **Case $H > -1/2$.** In this case, the solution ψ_ϵ converges uniformly to zero from the upper bound given in Lemma 3.7 which, combined with the expression of F in (3.6), yields

$$\lim_{\epsilon \to 0} F(s, \psi_\epsilon(s)) = F(s,0) = g(s) + \frac{f^2(s) - f(s)}{2}, \quad s \leq T.$$

Furthermore, integrating the variation of constants expression (3.5) leads to

$$\begin{aligned}
\int_0^t \epsilon^{-H-1/2} \psi_\epsilon(u) du &= \int_0^t \left\{ \int_0^u K_\epsilon(u-s) F(s, \psi_\epsilon(s)) ds \right\} du \\
&= \int_0^t \left\{ \int_s^t K_\epsilon(u-s) du \right\} F(s, \psi_\epsilon(s)) ds \\
&= \int_0^t \left(1 - e^{-\epsilon^{-1}(t-s)}\right) F(s, \psi_\epsilon(s)) ds,
\end{aligned}$$

where we used Fubini for the second equality as the integrated quantity is bounded and measurable. Now, given the function $s \mapsto \left(1 - e^{-\epsilon^{-1}(t-s)}\right) F(s, \psi_\epsilon(s))$ is uniformly bounded in ϵ by a constant on $[0,T]$, and that it converges pointwise to $s \mapsto F(0,s)$, then an application of Lebesgue's Dominated Convergence Theorem yields

$$\int_0^t \epsilon^{-H-1/2} \psi_\epsilon(u) du \underset{\epsilon \to 0}{\longrightarrow} \int_0^t \left(g(s) + \frac{f^2(s) - f(s)}{2} \right) ds,$$

hence the resulting convergence

$$\phi_\epsilon(t) \underset{\epsilon \to 0}{\longrightarrow} V_0 \int_0^t \left(g(s) + \frac{f^2(s) - f(s)}{2} \right) ds.$$

Case $H = -1/2$. Now fix $\epsilon > 0$ and using the second equation in (3.7), observe that

$$\begin{aligned}
\psi'_\epsilon(t) &= -\epsilon^{-1} \psi_\epsilon(t) + \epsilon^{-1} \left(F(t, \psi_\epsilon(t)) - F(t, \psi_0(t)) + \psi_0(t) \right) \\
&= -\epsilon^{-1} \left(\psi_\epsilon(t) - \psi_0(t) \right) + \epsilon^{-1} \beta_\epsilon(t) \left(\psi_\epsilon(t) - \psi_0(t) \right), \\
\text{with } \beta_\epsilon(t) &:= \frac{\xi^2}{2} \left(\psi_\epsilon(t) + \psi_0(t) \right) + \rho \xi f(T-t). \qquad (3.15)
\end{aligned}$$

Since ψ_0 has bounded variations by assumption, the complex-valued Riemann-Stieltjes integral on continuous functions f against ψ_0 is well-defined, and satisfies an integration by part formula such that

$$\int_0^t f(s)d\psi_0(s) = f(t)\psi_0(t) - f(0)\psi_0(0) - \int_0^t \psi_0(s)df(s),$$

see Theorems A.1 and A.2 from [39]. Define $\Delta_\epsilon(t) := (\psi_\epsilon(t) - \psi_0(t))e^{t/\epsilon}$. Then, it follows that

$$\begin{aligned} d\Delta_\epsilon(t) &= (\psi_\epsilon'(t)dt - d\psi_0(t))e^{t/\epsilon} + \epsilon^{-1}(\psi_\epsilon(t) - \psi_0(t))e^{t/\epsilon}dt \\ &= \epsilon^{-1}\beta_\epsilon(t)\Delta_\epsilon(t) - e^{t/\epsilon}d\psi_0(t), \end{aligned}$$

and applying the variation of constants formula leads to

$$\Delta_\epsilon(t) = e^{\epsilon^{-1}\int_0^t \beta_\epsilon(r)dr}\Delta_\epsilon(0) - \int_0^t e^{\epsilon^{-1}\int_s^t \beta_\epsilon(r)dr}e^{\epsilon^{-1}s}d\psi_0(s),$$

so that, recalling that $\psi_\epsilon(0) = 0$,

$$\begin{aligned} \psi_\epsilon(t) - \psi_0(t) &= -e^{\epsilon^{-1}\int_0^t \beta_\epsilon(r)dr}e^{-\epsilon^{-1}t}\psi_0(0) - \int_0^t e^{\epsilon^{-1}\int_s^t \beta_\epsilon(r)dr}e^{-\epsilon^{-1}(t-s)}d\psi_0(s) \\ &=: \mathbf{I}_\epsilon(t) + \mathbf{II}_\epsilon(t). \end{aligned}$$

We now prove successively that $|\int_0^T \mathbf{I}_\epsilon(t)dt| \to 0$ and $|\int_0^T \mathbf{II}_\epsilon(t)dt| \to 0$.

- Given that $\Re\beta_\epsilon \leq 0$, then

$$\left|e^{\epsilon^{H-1/2}\int_0^t \beta_\epsilon(r)dr}e^{-\epsilon^{-1}t}\psi_0(0)\right| \leq e^{-\epsilon^{-1}t}\left|\psi_0(0)\right| \to 0, \quad t \in (0,T),$$

  so that $\mathbf{I}_\epsilon(t)$ converges pointwise to 0 on $(0,T)$ and is dominated by an integrable function, hence $|\int_0^T \mathbf{I}_\epsilon(t)dt| \to 0$ by Lebesgue's dominated convergence theorem.

- Regarding the second term, we have

$$\begin{aligned} \left|\int_0^T \int_0^t e^{\epsilon^{H-1/2}\int_s^t \beta_\epsilon(r)dr}e^{-\epsilon^{-1}(t-s)}d\psi_0(s)dt\right| &\leq \int_0^T \int_0^t e^{-\epsilon^{-1}(t-s)}\left|d\psi_0(s)\right|dt \\ &= \int_0^T \int_s^T e^{-\epsilon^{-1}(t-s)}dt\left|d\psi_0(s)\right| \\ &= \int_0^T \epsilon\left(1 - e^{-\epsilon^{-1}(T-s)}\right)\left|d\psi_0(s)\right| \end{aligned}$$

  where the inequality comes from Theorem A.4 from [39] and using again that $\Re\beta_\epsilon \leq 0$, with the positive measure on the right-hand side being the total variation measure defined as in Theorem 6.2 from [40], and we used Fubini-Lebesgue to get the first equality. Noting the point-wise convergence of the function $f_\epsilon : s \mapsto \epsilon\left(1 - e^{-\epsilon^{-1}(T-s)}\right)$ to zero and its uniform boundedness in ϵ by a constant on $[0,T]$, the dominated convergence theorem applied to the total variation measure proves the result.

Thus, we obtained

$$\int_0^T \psi_\epsilon(t)dt \underset{\epsilon\to 0}{\longrightarrow} \int_0^T \psi_0(t)dt,$$

where ψ_0 satisfies the second equation in (3.7), hence we get

$$\phi_\epsilon(T) \underset{\epsilon\to 0}{\longrightarrow} (\theta + V_0)\int_0^T \psi_0(s)ds,$$

which is the desired convergence.

Case $H < -1/2$. Define ψ_0 as the root with non-positive real part from the third equation in (3.7), recall Lemma A.1. Fixing again $\epsilon > 0$, we have

$$\begin{aligned}\psi_\epsilon'(t) &= -\epsilon^{-1}\psi_\epsilon(t) + \epsilon^{H-1/2}\left(F(t,\psi_\epsilon(t)) - F(t,\psi_0(t))\right)\\ &= -\epsilon^{-1}\psi_\epsilon(t) + \epsilon^{H-1/2}\beta_\epsilon(t)\left(\psi_\epsilon(t) - \psi_0(t)\right),\end{aligned}$$

with β_ϵ given by (3.15). Similarly the case $H < -1/2$, computing the differential of $\Delta_\epsilon(t) := (\psi_\epsilon(t) - \psi_0(t))e^{t/\epsilon}$ and applying the variation of constants formula leads to

$$\Delta_\epsilon(t) = e^{\epsilon^{H-1/2}\int_0^t \beta_\epsilon(r)dr}\Delta_\epsilon(0) - \int_0^t e^{\epsilon^{H-1/2}\int_s^t \beta_\epsilon(r)dr} e^{s/\epsilon}\left(d\psi_0(s) + \epsilon^{-1}\psi_0(s)ds\right),$$

so that

$$\begin{aligned}\psi_\epsilon(t) - \psi_0(t) &= -e^{\epsilon^{H-1/2}\int_0^t \beta_\epsilon(r)dr}e^{-\epsilon^{-1}t}\psi_0(0) - \int_0^t e^{\epsilon^{H-1/2}\int_s^t \beta_\epsilon(r)dr}e^{-\epsilon^{-1}(t-s)}\left(d\psi_0(s) + \epsilon^{-1}\psi_0(s)ds\right)\\ &= -e^{\epsilon^{H-1/2}\int_0^t \beta_\epsilon(r)dr}e^{-\epsilon^{-1}t}\psi_0(0) - \int_0^t e^{\epsilon^{H-1/2}\int_s^t \beta_\epsilon(r)dr}e^{-\epsilon^{-1}(t-s)}d\psi_0(s)\\ &\quad - \int_0^t e^{\epsilon^{H-1/2}\int_s^t \beta_\epsilon(r)dr}e^{-\epsilon^{-1}(t-s)}\epsilon^{-1}\psi_0(s)ds\\ &=: \mathbf{I}_\epsilon(t) + \mathbf{II}_\epsilon(t) + \mathbf{III}_\epsilon(t).\end{aligned}$$

We already have from the previous case $H = -1/2$ that, as $\epsilon \to 0$, both integrals $\int_0^T \mathbf{I}_\epsilon(t)dt$, $\int_0^T \mathbf{II}_\epsilon(t)dt$ converge to 0, all that remains to show consequently is that $\int_0^T \mathbf{III}_\epsilon(t)dt$ converges to 0 too.

- A finer upper bound on $\Re\beta_\epsilon$ is required to deal with this third term. We already know from Theorem 2.4 that $\Re\psi_\epsilon \le 0$, and by definition of ψ_0, we get the following bound

$$\Re\beta_\epsilon = \frac{\xi^2}{2}\left(\Re\psi_\epsilon + \Re\psi_0\right) \le \frac{\xi^2}{2}\Re\psi_0 \le 0.$$

  Set

$$E := \{s \in [0,T], (f(s),g(s)) \neq (0,0)\},$$

  and from (3.8), we know that $\psi_0 = 0$ on E^C while lemma A.2 yields $\Re\psi_0 \neq 0$ on E so that we can bound $\mathbf{III}_\epsilon$ as follows

$$\begin{aligned}|\mathbf{III}_\epsilon(t)| &\le \int_{[0,t]\cap E} \epsilon^{-1}|\psi_0(s)|e^{\epsilon^{H-1/2}\int_s^t \Re\beta_\epsilon(r)dr}e^{-\epsilon^{-1}(t-s)}ds\\ &\le \int_{[0,t]\cap E} \epsilon^{-1}\frac{|\psi_0(s)|}{-\epsilon^{H-1/2}\frac{\xi^2}{2}\Re\psi_0(s)}\left(-\epsilon^{H-1/2}\frac{\xi^2}{2}\Re\psi_0(s)e^{\int_s^t \epsilon^{H-1/2}\frac{\xi^2}{2}\Re\psi_0(r)dr}\right)ds\\ &\le \int_{[0,t]\cap E} \epsilon^{-H-1/2}2\xi^{-2}\frac{|\psi_0|}{-\Re\psi_0}\left(-\epsilon^{H-1/2}\frac{\xi^2}{2}\Re\psi_0(s)e^{\int_s^t \epsilon^{H-1/2}\frac{\xi^2}{2}\Re\psi_0(r)dr}\right)ds,\end{aligned}$$

  and an application of lemma A.2 yields the existence of a finite positive constant $C > 0$ such that

$$\frac{|\psi_0|}{-\Re\psi_0} \le C, \quad \forall s \in E,$$

  so that

$$\begin{aligned}|\mathbf{III}_\epsilon(t)| &\le C\epsilon^{-H-1/2}\int_{[0,t]\cap E}\left(-\epsilon^{H-1/2}\frac{\xi^2}{2}\Re\psi_0(s)e^{\int_s^t \epsilon^{H-1/2}\frac{\xi^2}{2}\Re\psi_0(r)dr}\right)ds\\ &\le C\epsilon^{-H-1/2}\left(1 - e^{\int_0^t \epsilon^{H-1/2}\frac{\xi^2}{2}\Re\psi_0(r)dr}\right) \underset{\epsilon\to 0}{\longrightarrow} 0,\end{aligned}$$

since $-H-1/2>0$ so that $\epsilon^{-H-1/2}\to 0$ as $\epsilon\to 0$. Thus $\mathbf{III}_\epsilon$ is dominated by a finite constant C independent of ϵ (which is integrable on $[0,T]$) and Lebesgue's dominated convergence theorem yields that $\int_0^T \mathbf{III}_\epsilon(t)dt$ converges to 0.

Consequently, we obtained

$$\int_0^T \psi_\epsilon(t)dt \underset{\epsilon\to 0}{\longrightarrow} \int_0^T \psi_0(t)dt,$$

which then yields

$$\epsilon^{-H-1/2}\int_0^T \psi_\epsilon(t)dt \underset{\epsilon\to 0}{\longrightarrow} 0,$$

and finally we get

$$\phi_\epsilon(T) \underset{\epsilon\to 0}{\longrightarrow} \theta\int_0^T \psi_0(s)ds.$$

□

3.2 Convergence of the finite-dimensional distributions towards NIG-IG

In this section, we prove the second part of Theorem 3.4, that is the convergence of the finite-dimensional distributions of $\left(\log\frac{S^\epsilon}{S_0},\bar{V}^\epsilon\right)$ towards those of a NIG-IG process (X,Y) in the sense of Definition 3.1 with parameters $(\alpha,\beta,\mu,\delta,\lambda)$ as in (3.11) depending on the regime of H. Let $d\in\mathbb{N}^*$ and take $0=:t_0<t_1<\cdots<t_d\le T$ to be d distinct times of the time interval $[0,T]$ and $(u_k,v_k)_{k\in\{1,\dots,d\}}\in\left(\mathbb{R}^2\right)^d$. We will prove that

$$\mathbb{E}\left[\exp\left(i\sum_{k=1}^d u_k\log\frac{S^\epsilon_{t_k}}{S_0}+i\sum_{k=1}^d v_k\bar{V}^\epsilon_{t_k}\right)\right]\underset{\epsilon\to 0}{\longrightarrow}\mathbb{E}\left[\exp\left(i\sum_{k=1}^d u_kX_{t_k}+i\sum_{k=1}^d v_kY_{t_k}\right)\right]. \tag{3.16}$$

First, we recover the finite-dimensional distributions of $\left(\log\frac{S^\epsilon_T}{S_0},\bar{V}^\epsilon_T\right)$ from (3.9) by setting the bounded and measurable functions f and g to be equal to

$$f(s):=i\sum_{k=1}^d \mathbb{1}_{[t_{k-1},t_k)}(T-s)\sum_{k=j}^d u_j \quad\text{and}\quad g(s):=i\sum_{k=1}^d \mathbb{1}_{[t_{k-1},t_k)}(T-s)\sum_{k=j}^d v_j.$$

Notice indeed that

$$\begin{aligned}i\sum_{k=1}^d u_k\log\frac{S^\epsilon_{t_k}}{S_0}+i\sum_{k=1}^d v_k\bar{V}^\epsilon_{t_k} &= i\sum_{k=1}^d\left(\log\frac{S^\epsilon_{t_k}}{S_0}-\log\frac{S^\epsilon_{t_{k-1}}}{S_0}\right)\sum_{j=k}^d u_j+i\sum_{k=1}^d\left(\bar{V}^\epsilon_{t_k}-\bar{V}^\epsilon_{t_{k-1}}\right)\sum_{j=k}^d v_j\\ &= \int_0^T f(T-s)d\log S^\epsilon_s+\int_0^T g(T-s)d\bar{V}^\epsilon_s,\end{aligned}$$

and that the corresponding ψ_0 defined in (3.7) has bounded variations (being piece-wise constant for this choice of f and g), so that an application of the convergence of the characteristic functional in (3.9) yields

$$\mathbb{E}\left[\exp\left(i\sum_{k=1}^d u_k\log\frac{S^\epsilon_{t_k}}{S_0}+i\sum_{k=1}^d v_k\bar{V}^\epsilon_{t_k}\right)\right]\longrightarrow\exp\left(\phi_0(T)\right),$$

with

$$\phi_0(T)=\begin{cases}V_0\sum_{k=1}^d(t_k-t_{k-1})(i\bar{v}_k-\frac{\bar{u}_k^2+i\bar{u}_k}{2}), & \text{if } H>-1/2,\\ (\theta+V_0)\,\xi^{-2}\sum_{k=1}^d(t_k-t_{k-1})\left(1-\rho\xi i\bar{u}_k+\sqrt{(1-\rho\xi i\bar{u}_k)^2-2\xi^2(i\bar{v}_k-\frac{\bar{u}_k^2+i\bar{u}_k}{2})}\right), & \text{if } H=-1/2,\\ -\theta\xi^{-1}\sum_{k=1}^d(t_k-t_{k-1})\left(\rho i\bar{u}_k+\sqrt{-\rho^2\bar{u}_k^2-2(i\bar{v}_k-\frac{\bar{u}_k^2+i\bar{u}_k}{2})}\right), & \text{if } H<-1/2,\end{cases}$$

where we defined $\bar{u}_k := \sum_{j=k}^d u_j$, $\bar{v}_k := \sum_{j=k}^d v_j$.

Second, we identify such $\phi_0(T)$ with the corresponding finite-dimensional distributions of the NIG-IG process (X, Y) with parameters $(\alpha, \beta, \mu, \delta, \lambda)$ as in (3.11) depending on the regime of H. We denote by η its Lévy exponent, recall (3.1), and we write

$$\begin{aligned}
\mathbb{E}\left[\exp\left(i\sum_{k=1}^d u_k X_{t_k} + i\sum_{k=1}^d v_k Y_{t_k}\right)\right] &= \mathbb{E}\left[\exp\left(i\sum_{k=1}^d \left(X_{t_k} - X_{t_{k-1}}\right)\sum_{j=k}^d u_j + i\sum_{k=1}^d \left(Y_{t_k} - Y_{t_{k-1}}\right)\sum_{j=k}^d v_j\right)\right] \\
&= \prod_{k=1}^d \mathbb{E}\left[\exp\left(\left(X_{t_k} - X_{t_{k-1}}\right) i\sum_{j=k}^d u_j + \left(Y_{t_k} - Y_{t_{k-1}}\right) i\sum_{j=k}^d v_j\right)\right] \\
&= \prod_{k=1}^d \mathbb{E}\left[\exp\left(X_{t_k - t_{k-1}} i\sum_{j=k}^d u_j + Y_{t_k - t_{k-1}} i\sum_{j=k}^d v_j\right)\right] \\
&= \exp\left(\sum_{k=1}^d \left(t_k - t_{k-1}\right)\eta\left(\bar{u}_k, \bar{v}_k\right)\right)
\end{aligned}$$

using respectively telescopic summation, the independence of increments, the fact that $(X_{t_2}, Y_{t_2}) - (X_{t_1}, Y_{t_1}) \overset{law}{=} (X_{t_2 - t_1}, Y_{t_2 - t_1})$ and the definition of the characteristic function of $\left(X_{t_k - t_{k-1}}, Y_{t_k - t_{k-1}}\right)$ for all $k \in 1, \ldots, d$ to get the successive equalities. Using the parameters $(\alpha, \beta, \mu, \delta, \lambda)$ as in (3.11) it is immediate to see that

$$\sum_{k=1}^d \left(t_k - t_{k-1}\right)\eta\left(\bar{u}_k, \bar{v}_k\right) = \phi_0(T),$$

hence the desired convergence (3.16).

3.3 Weak-convergence of the integrated variance process for the M_1 topology

In this section, we prove the weak convergence stated in Theorem 3.4 of the integrated variance process $\bar{V}^\epsilon$ with sample paths in $C\left([0,T], \mathbb{R}^+\right)$ to the Lévy process $\bar{V}^0$ whose Lévy exponent is given by $v \mapsto \eta(0, v)$ with η defined in (3.1) and with sample paths valued in the càdlàg functional space $D := \mathscr{D}\left([0,T], \mathbb{R}^+\right)$ endowed with Skorokhod's Strong M1 (SM_1) topology. There will be two subsections: first, in Section 3.3.1 we recall briefly the definition of the Strong M1 (SM_1) topology as well as some associated convergence results, and then we prove the tightness of the integrated variance in Section 3.3.2.

3.3.1 Reminder on the SM_1 topology and conditions for convergence

We recall succinctly key definitions and convergence theorems for the M_1 topology. We refer the reader to the key reference book Whitt [44, Chapter 12] for more details. For $x \in D$, we define the thin graph of x as

$$\Gamma_x := \left\{(z,t) \in \mathbb{R} \times [0,T] : z \in \left[x(t^-), x(t)\right]\right\},$$

where, for $t \in [0,T]$, $[x(t^-), x(t)]$ denotes the standard segment $\{\alpha x(t^-) + (1-\alpha)\, x(t), 0 \leq \alpha \leq 1\}$, which is different from a singleton at discontinuity points of the càdlàg sample trajectory x. We denote $Disc(x)$ the set of such instants. Define on Γ_x the strong order relation as follows: $(z_1, t_1) \leq (z_2, t_2)$ if either $t_1 < t_2$ or $t_1 = t_2$ and $\left|x(t_1^-) - z_1\right| \leq \left|x(t_1^-) - z_2\right|$. Furthermore define a strong parametric representations of x as a continuous non-decreasing (with respect to the previous order relation) function $(\hat{x}, \hat{r})$ mapping $[0,1]$ into Γ_x such that $(\hat{x}(0), \hat{r}(0)) = (x(0), 0)$ and $(\hat{x}(T), \hat{r}(T)) = (x(T), T)$. We say the component $\hat{r}$ scales the time interval $[0,T]$ to $[0,1]$ while $\hat{x}$ time-scales x, and we denote by $\Pi_s(x)$ the set of all strong parametric representations of x. Finally, the strong M1 topology is the topology induced by the metric d_s defined as

$$d_s\left(x_1, x_2\right) := \underset{(\hat{x}^i, \hat{r}^i) \in \Pi_s(x^j), j=1,2}{inf} \left\{\left|\hat{x}^1 - \hat{x}^2\right| \vee \left|\hat{r}^1 - \hat{r}^2\right|\right\}.$$

Below we mention briefly five theorems in a row that eventually yields criteria to prove the desired convergence.

Theorem 3.9. *$D := D([0,T], \mathbb{R}^+)$ endowed with the SM_1 topology is Polish, i.e. metrizable as a complete separable metric space.*

Theorem 3.10 (Prohorov's theorem)**.** *Let (S,m) be a metric space. If a subset A in $\mathscr{P}(S)$ is tight, then it's relatively compact. On the other hand, if the subset A is relatively compact and the topological space is Polish, then A is tight.*

The first two theorems ensure that, since (D, SM_1) is Polish, proving relative compactness of any family of probability measures on such space is sufficient to ensure the existence of a convergent sub-sequence. Finally the convergence of finite-dimensional laws allows to uniquely determine its limit as formulated in the following theorem.

Theorem 3.11 (Criteria for convergence in distribution in (D, SM_1))**.** *Let $((X_n)_{n\in\mathbb{N}}, X)$ be random functions defined on a common filtered probability space $(\Omega, (\mathscr{F}_t)_t, \mathbb{P})$. Then, $X_n \Rightarrow X$ in D for the SM_1 topology if the following conditions hold:*

- *$(X_n)_{n\in\mathbb{N}}$ is tight with regards to the SM_1 topology.*
- *The finite dimensional distributions of X_n converge to those of X on T_X, where:*

$$T_X := \{t > 0, \mathbb{P}(t \in Disc(X)) = 0\} \cup \{T\}.$$

To conclude this section, we recall a characterization of tightness for a sequence of probability measures.

Theorem 3.12 (Characterization of tightness)**.** *The sequence of probability measures $\{\mathbb{P}_n\}_{n\geq 1}$ on (D, SM_1) is tight if and only if:*

(i) $\forall \bar{\varepsilon} > 0, \exists c < \infty, \forall n \geq 1, \mathbb{P}_n(\{x \in D : ||x|| > c\}) < \bar{\varepsilon}$

(ii) $\forall \bar{\varepsilon} > 0, \forall \eta > 0, \exists \delta > 0, \forall n \geq 1, \mathbb{P}_n(\{x \in D : w'(x,\delta) \geq \eta\}) < \bar{\varepsilon}$

Where we defined for $x \in D$, $t \in [0,T]$ and $\delta > 0$:

$$\begin{aligned}
||x|| &:= \sup_{t\in[0,T]} |x_t|, \\
w'(x,\delta) &:= w(x,\delta) \vee \bar{v}(x,0,\delta) \vee \bar{v}(x,T,\delta), \\
w(x,\delta) &:= \sup_{t\in[0,T]} w_S(x,t,\delta), \\
w_S(x,t,\delta) &:= \sup_{0\vee t-\delta \leq t_1 < t_2 < t_3 \leq (t+\delta)\wedge T} |x(t_2) - [x(t_1), x(t_3)]|, \\
\bar{v}(x,t,\delta) &:= \sup_{0\vee t-\delta \leq t_1 \leq t_2 \leq (t+\delta)\wedge T} |x(t_1) - x(t_2)|.
\end{aligned}$$

3.3.2 Convergence of the integrated variance process

We already proved in Section 3.2 the convergence of finite dimensional distributions as ϵ goes to zero of $\bar{V}^\epsilon$ toward those of either the deterministic linear, or the Inverse Gaussian, or the Lévy process denoted Y depending respectively on the value of the Hurst index with Lévy exponent $\eta(0,\cdot)$ from (3.1) with the respective parameters given in Theorem 3.4. Consequently, all that remains to prove is the tightness of the family of processes $\left(\bar{V}^\epsilon\right)_{\epsilon>0}$ for the SM_1 topology to get the desired convergence result as a direct consequence of Theorem 3.11. We will apply the characterization Theorem 3.12 of tightness in SM_1 to conclude, and more precisely, we will see that the criteria of tightness within the SM_1 topology simplifies greatly for almost surely non-decreasing and continuous stochastic processes in general.

Fix $\bar{\varepsilon} > 0$. Since, for all $\epsilon > 0$, $\bar{V}^\epsilon$ is almost surely non-decreasing and non-negative, we have that, for all $\omega \in \Omega$

$$\begin{aligned} ||\bar{V}^\epsilon(\omega)|| &= \bar{V}^\epsilon_T(\omega), \\ w'(\bar{V}^\epsilon(\omega)), \delta) &= 0 \vee \bar{V}^\epsilon_\delta(\omega) \vee \left(\bar{V}^\epsilon_T(\omega) - \bar{V}^\epsilon_{T-\delta}(\omega)\right), \quad \delta > 0. \end{aligned}$$

This yields that, for a threshold $c > 0$ big enough, the probability in condition (i) on the measures $\mathbb{P}_\epsilon := \mathbb{P}_{\left(\bar{V}^\epsilon_t\right)^{-1}}$ reduces to

$$\mathbb{P}_\epsilon\left(\{x \in D, ||x|| > c\}\right) = \mathbb{P}\left(\left\{\omega : ||\bar{V}^\epsilon_\cdot(\omega)|| > c\right\}\right) = \mathbb{P}\left(\bar{V}^\epsilon_T > c\right) \leq \sup_{\epsilon > 0} \mathbb{P}\left(\bar{V}^\epsilon_T > c\right) < \bar{\varepsilon},$$

where the last inequality is satisfied by tightness of the family $\left(\bar{V}^\epsilon_T\right)_{\epsilon>0}$ of random variables in $\mathbb{R}$ which is obtained as a direct consequence of Lévy's continuity theorem, recall that $\left(\bar{V}^\epsilon_T\right)_{\epsilon>0}$ has been shown to converge in Section 3.1. This yields (i). In addition, regarding the second condition (ii), set an arbitrary $\eta > 0$, then the oscillation function w' simplifies to

$$\mathbb{P}_\epsilon\left(\{x \in D : w'(x, \delta) \geq \eta\}\right) = \mathbb{P}\left(0 \vee \left|\bar{V}^\epsilon_\delta\right| \vee \left|\bar{V}^\epsilon_T - \bar{V}^\epsilon_{T-\delta}\right| \geq \eta\right) < \bar{\varepsilon},$$

where we take δ small enough to ensure the last inequality by stochastic continuity of $\bar{V}^\epsilon$.

Remark 3.13. *Since* $\log S^\epsilon = -\frac{1}{2}\bar{V}^\epsilon + \rho W_{\bar{V}^\epsilon} + \sqrt{1-\rho^2} W^\perp_{\bar{V}^\epsilon}$*, and composition is not continuous in* (D, M_1) *(see [37, Theorem 4.1]), we cannot expect the tightness of the log price within* SM_1*.*

A Some lemmas

Lemma A.1. *(Uniqueness of the complex root with a non-positive real part) Take* $\xi > 0$*. For all* f*,* g *bounded and measurable with* $\Re f = \Re g = 0$*,* $t \in [0, T]$ *and* $\rho \in [-1, 1]$*, both polynomials*

$$\begin{aligned} P(X) &:= \frac{\xi^2}{2} X^2 - \left(1 - \rho\xi f(t)\right) X + g(t) + \frac{f^2(t) - f(t)}{2} \\ Q(X) &:= \frac{\xi^2}{2} X^2 + \rho\xi f(t) X + g(t) + \frac{f^2(t) - f(t)}{2} \end{aligned}$$

admit exactly two roots with respective real parts of strict opposite signs if $(f(t), g(t)) \neq (0, 0)$*, and if* $(f(t), g(t)) = (0, 0)$*, then the polynomial* P *has roots* 0 *and* $\frac{2}{\xi^2}$*, while* Q *has* 0 *as a double root.*

Proof. Let us detail the proof for P, similar arguments will apply to Q. By d'Alembert-Gauss theorem, the polynomial P admits exactly two roots expressed as:

$$\left\{\xi^{-2}\left(1 - \rho\xi f(t) \pm \sqrt{\left(1 - \rho\xi f(t)\right)^2 - 2\xi^2\left(g(t) + \frac{f^2(t) - f(t)}{2}\right)}\right)\right\},$$

where we take the principal square root in the expression above, i.e.with non-negative real-part. Consequently, the roots have real parts

$$\left\{\xi^{-2}\left(1 \pm \Re\left(\sqrt{\left(1 - \rho\xi f(t)\right)^2 - 2\xi^2\left(g(t) + \frac{f^2(t) - f(t)}{2}\right)}\right)\right)\right\},$$

so that it remains to show $\left|\Re\left(\sqrt{\left(1 - \rho\xi f(t)\right)^2 - 2\xi^2\left(g(t) + \frac{f^2(t)-f(t)}{2}\right)}\right)\right| > 1$.

Denote $\delta = a + ib$, $a, b \in \mathbb{R}$ such that $\delta^2 = \left(1 - \rho\xi f(t)\right)^2 - 2\xi^2\left(g(t) + \frac{f^2(t)-f(t)}{2}\right)$, then it follows that a and b satisfy

$$\begin{cases} a^2 - b^2 &= 1 + \left(1-\rho^2\right)\left(\xi \Im f(t)\right)^2, \\ ab &= -\left(\rho\xi\Im f(t) + \xi^2\left(\Im g(t) - \frac{\Im f(t)}{2}\right)\right). \end{cases} \tag{A.1}$$

If $\rho \neq \pm 1$, then the result is immediate from the first inequality in (A.1), while if $\rho = \pm 1$, then $|a| = \sqrt{1+b^2}$ and b cannot be zero, otherwise $\delta^2 = a^2 = 1 - 2i\left(\rho\xi\Im f(t) + \xi^2\left(\Im g(t) - \frac{\Im f(t)}{2}\right)\right)$ which cannot be the case, since $a \in \mathbb{R}$ and $(\Im f(t), \Im g(t)) \neq (0,0)$. □

Lemma A.2. *Let f and g be bounded measurable functions such that $\Re f = \Re g = 0$. Then there exists a finite positive constant C such that the ratio*

$$\gamma(s) := \frac{|\psi_0(s)|}{-\Re\psi_0(s)} \leq C, \forall s \in E,$$

where the set E is given by

$$E := \{s \in [0,T], (f(s), g(s)) \neq (0,0)\},$$

and ψ_0 is given in (3.8) *in the case $H < -1/2$.*

Proof. We start by explicitly computing the real and imaginary parts of ψ_0 in the case $H < -1/2$, whose expression is given anew by

$$\psi_0(s) = -\xi^{-1}\left(\rho f(s) + \sqrt{f(s)\left(1 - \left(1-\rho^2\right)f(s)\right) - 2g(s)}\right), \quad s \in [0,T].$$

Set the real functions a and b such that, for any $s \in [0,T]$

$$a(s) + ib(s) = \sqrt{f(s)\left(1 - \left(1-\rho^2\right)f(s)\right) - 2g(s)},$$

square the above equality, identify the real and imaginary parts, find a unambiguously on E (which imposes $a \neq 0$) as a root to a fourth-degree polynomial knowing the square roots in (3.7) are principal (i.e. with positive real parts), then deduce the associated b, such that

$$\begin{aligned} a &= \sqrt{\frac{1}{2}\left(\left(1-\rho^2\right)\left(\Im f\right)^2 + \sqrt{\left(1-\rho^2\right)^2\left(\Im f\right)^4 + \left(\Im f - 2\Im g\right)^2}\right)}, \\ b &= \frac{\frac{1}{2}\Im f - \Im g}{a}. \end{aligned}$$

Consequently, we get explicitly on E

$$\psi_0 = \Re\psi_0 + i\Im\psi_0,$$

with

$$\begin{aligned} \Re\psi_0 &= -\xi^{-1}a, \\ \Im\psi_0 &= -\xi^{-1}\left(\rho\Im f + \frac{\frac{1}{2}\Im f - \Im g}{a}\right). \end{aligned}$$

Rewrite γ as

$$\gamma = \sqrt{1+\tilde{\gamma}^2}, \quad \tilde{\gamma} := \frac{\Im\psi_0}{\Re\psi_0},$$

and we can readily discard the case $\rho = 0$ and $\Im f - 2\Im g = 0$, since $\Im\psi_0 = 0$ in that case and the ratio simplifies into 1 which yields the result. Assume from now on that $\rho \neq 0$ or $\Im f - 2\Im g \neq 0$. We can write

$$\tilde{\gamma} = \mathbf{A}\left(\Im f, \Im g\right) + \mathbf{B}\left(\Im f, \Im g\right),$$

with

$$\mathbf{A}(\Im f, \Im g) := \frac{\rho \Im f}{\sqrt{\frac{1}{2}\left((1-\rho^2)(\Im f)^2 + \sqrt{(1-\rho^2)^2(\Im f)^4 + (\Im f - 2\Im g)^2}\right)}},$$

$$\mathbf{B}(\Im f, \Im g) := \frac{\frac{1}{2}\Im f - \Im g}{\frac{1}{2}\left((1-\rho^2)(\Im f)^2 + \sqrt{(1-\rho^2)^2(\Im f)^4 + (\Im f - 2\Im g)^2}\right)},$$

so that there are three remaining cases for which it is sufficient to show that both $\mathbf{A}(\Im f, \Im g)$ and $\mathbf{B}(\Im f, \Im g)$ are bounded to conclude the proof.

- Case $\rho \neq 0$ and $\Im f - 2\Im g = 0$, then

$$\mathbf{A}(\Im f, \Im g) = \frac{\rho}{\sqrt{1-\rho^2}},$$
$$\mathbf{B}(\Im f, \Im g) = 0,$$

  are both bounded, recall $\rho \in (-1, 1)$.

- Case $\rho = 0$ and $\Im f - 2\Im g \neq 0$, then

$$\mathbf{A}(\Im f, \Im g) = 0,$$
$$\mathbf{B}(\Im f, \Im g) = \frac{\frac{1}{2}\Im f - \Im g}{\frac{1}{2}\left((\Im f)^2 + \sqrt{(\Im f)^4 + (\Im f - 2\Im g)^2}\right)},$$

  and $\mathbf{B}(\Im f, \Im g)$ is bounded since $(x, y) \mapsto \mathbf{B}(x, y)$ is continuous on any compact set of $\mathbb{R}^2 \backslash \{(0,0)\}$ (recall both $\Im f$ and $\Im g$ are bounded) and has a finite limit at $(0, 0)$, valued 1, indeed

$$\mathbf{B}(x, y) = \frac{1}{(1-\rho^2)\frac{x^2}{x-2y} + \sqrt{(1-\rho^2)^2 \frac{x^4}{(x-2y)^2} + 1}} \underset{\substack{(x,y)\to(0,0) \\ x-2y\neq 0}}{\longrightarrow} 1.$$

- Case $\rho \neq 0$ and $\Im f - 2\Im g \neq 0$, then

$$\mathbf{A}(\Im f, \Im g) = \frac{\rho \Im f}{\sqrt{\frac{1}{2}\left((1-\rho^2)(\Im f)^2 + \sqrt{(1-\rho^2)^2(\Im f)^4 + (\Im f - 2\Im g)^2}\right)}},$$

$$\mathbf{B}(\Im f, \Im g) = \frac{\frac{1}{2}\Im f - \Im g}{\frac{1}{2}\left((1-\rho^2)(\Im f)^2 + \sqrt{(1-\rho^2)^2(\Im f)^4 + (\Im f - 2\Im g)^2}\right)},$$

  are both bounded by continuity of $(x, y) \mapsto \mathbf{A}(x, y)$ and $(x, y) \mapsto \mathbf{B}(x, y)$ on any compact set of $\mathbb{R}^2 \backslash \{(0,0)\}$ (recall both $\Im f$ and $\Im g$ are bounded) and both functions have a finite limit at $(0, 0)$, valued 0 and 1 respectively, obtained with similar arguments as in the previous case.

□

Definition A.3. *(Inverse Gaussian, Lévy and NIG-IG distributions)*

- *We say X follows a Normal Inverse distribution, denoted $X \hookrightarrow IG(\mu, \lambda)$ if its probability density writes*

$$f(x) = \sqrt{\frac{\lambda}{2\pi x^3}} \exp\left(-\frac{\lambda(x-\mu)^2}{2\mu^2 x}\right),$$

  where $\mu \in \mathbb{R}$, $\lambda > 0$, or equivalently if the following equality holds true

$$\mathbb{E}[\exp(wX)] = \exp\left(\frac{\lambda}{\mu}\left(1 - \sqrt{1 - \frac{2\mu^2 w}{\lambda}}\right)\right), \quad w \in \mathbb{C}, \quad \Re w \leq 0.$$

- *We say τ follows a Lévy distribution, denoted $\tau \hookrightarrow$ Lévy(μ, c), if its probability density writes*

$$\sqrt{\frac{c}{2\pi}} \frac{e^{-\frac{c}{2(x-\mu)}}}{(x-\mu)^{3/2}},$$

where $\mu \in \mathbb{R}$, $c > 0$, or equivalently if the following equality holds true

$$\mathbb{E}\left[\exp\left(w\tau\right)\right] = \exp\left(\mu w - \sqrt{-2cw}\right), \quad w \in \mathbb{C}, \quad \Re w \leq 0.$$

- *We say (X, Y) follows a Normal Inverse Gaussian - Inverse Gaussian distribution, denoted $(X, Y) \hookrightarrow$ NIG-IG$(\alpha, \beta, \mu, \delta, \lambda)$, if its characteristic function writes*

$$\mathbb{E}\left[\exp\left(iuX + ivY\right)\right] = \exp\left[i\mu u + \delta\left(\sqrt{\alpha^2 - \beta^2} - \sqrt{\alpha^2 - 2i\lambda v - (\beta + iu)^2}\right)\right], \quad u, v \in \mathbb{R},$$

B Additional plots

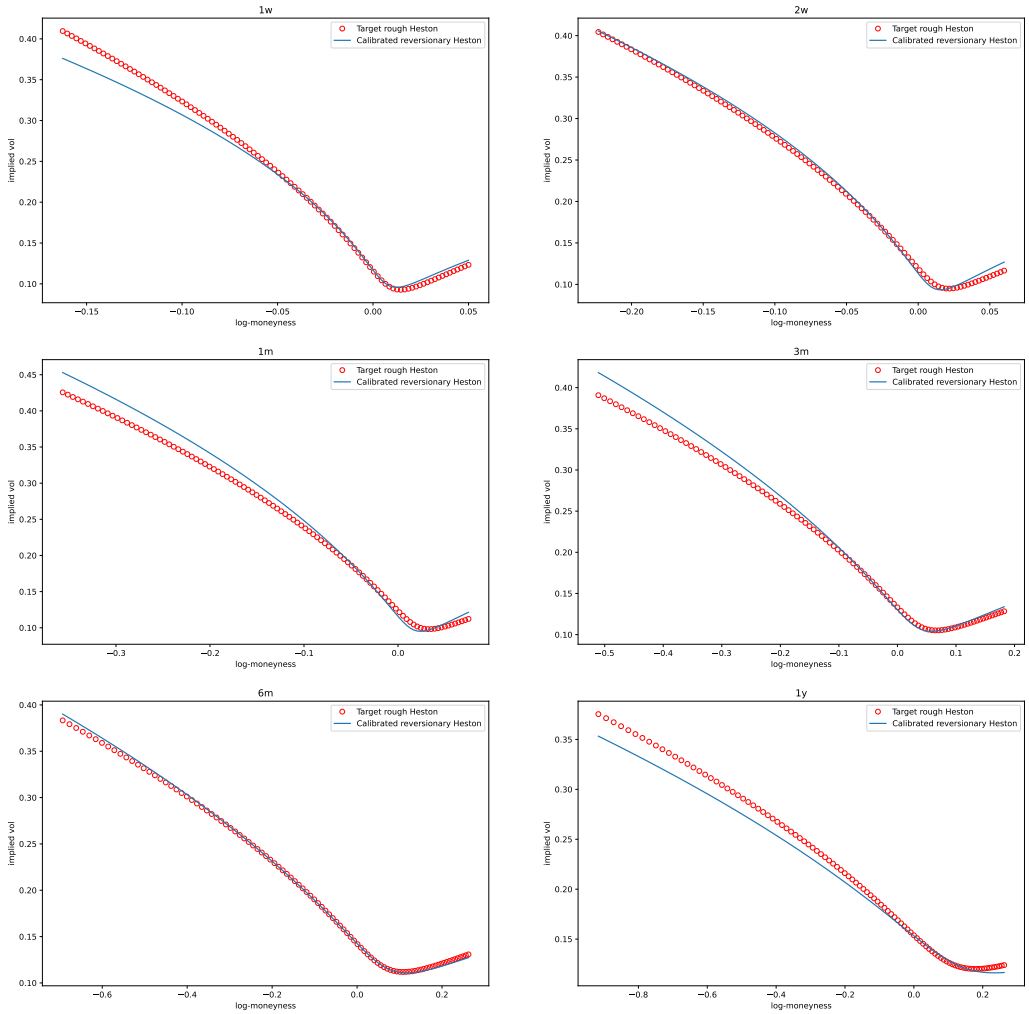

Figure 6: Smiles comparison between target rough Heston with parameters (1.10), with $H = 0$, and reversionary Heston with calibrated parameters from the second row of Table (1.3) for different maturities from one week to one year.

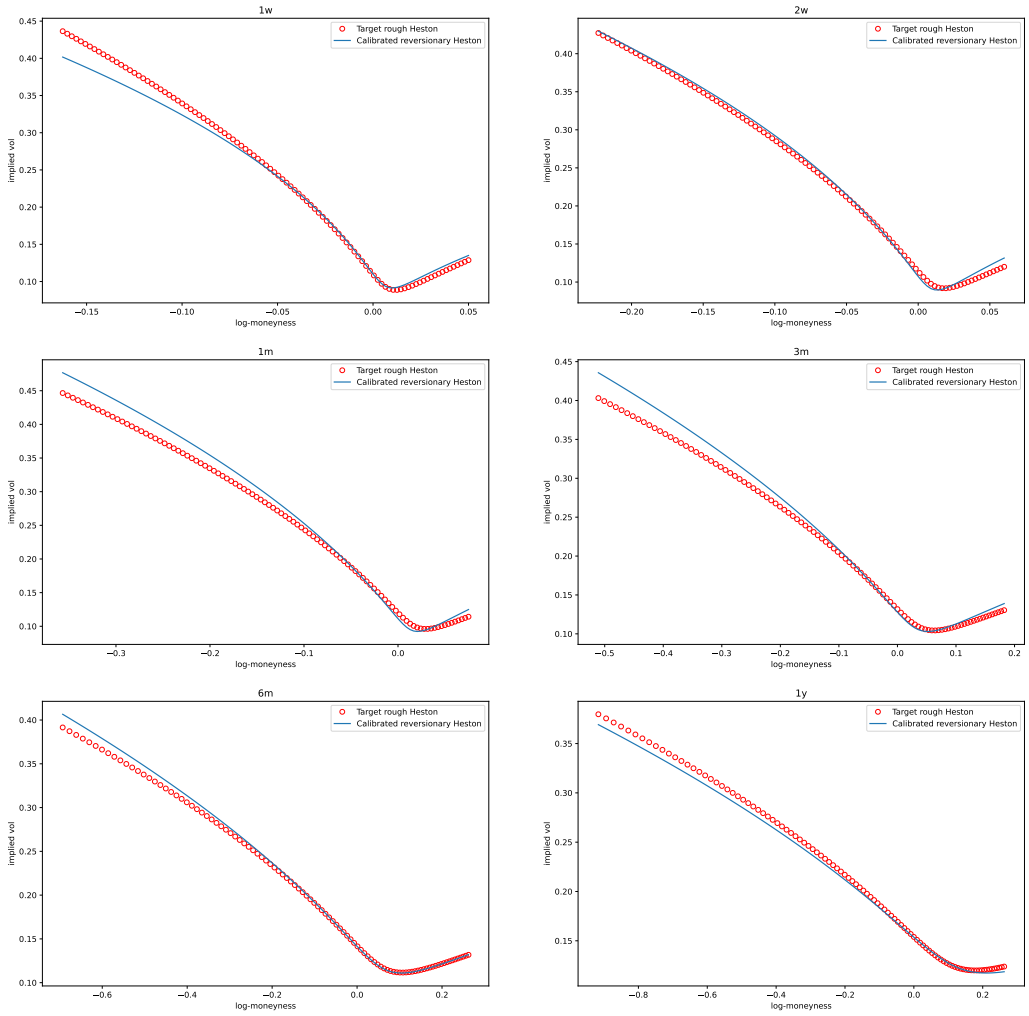

Figure 7: Smiles comparison between target rough Heston with parameters (1.10), with $H = -0.05$, and reversionary Heston with calibrated parameters from the third row of Table (1.3) for different maturities from one week to one year.

References

[1] Eduardo Abi Jaber. Lifting the Heston model. *Quantitative Finance*, 19(12):1995–2013, 2019.

[2] Eduardo Abi Jaber. Weak existence and uniqueness for affine stochastic Volterra equations with L^1-kernels. 2021.

[3] Eduardo Abi Jaber. The characteristic function of gaussian stochastic volatility models: an analytic expression. *Finance and Stochastics*, 26(4):733–769, 2022.

[4] Eduardo Abi Jaber and Omar El Euch. Markovian structure of the Volterra Heston model. *Statistics & Probability Letters*, 149:63–72, 2019.

[5] Eduardo Abi Jaber and Omar El Euch. Multifactor approximation of rough volatility models. *SIAM Journal on Financial Mathematics*, 10(2):309–349, 2019.

[6] Eduardo Abi Jaber, Martin Larsson, and Sergio Pulido. Affine Volterra processes. *The Annals of Applied Probability*, 29(5):3155–3200, 2019.

[7] Eduardo Abi Jaber, Camille Illand, and Shaun Li. Joint SPX-VIX calibration with Gaussian polynomial volatility models: deep pricing with quantization hints. *arXiv preprint arXiv:2212.08297*, 2022.

[8] Eduardo Abi Jaber, Camille Illand, and Shaun Li. The quintic Ornstein-Uhlenbeck volatility model that jointly calibrates SPX & VIX smiles. *arXiv preprint arXiv:2212.10917*, 2022.

[9] Schoutens W. Albrecher H., Mayer P. and Tistaert J. The little Heston trap. *Wilmott magazine*, (1): 83–92, 2007.

[10] Sassan Alizadeh, Michael W Brandt, and Francis X Diebold. Range-based estimation of stochastic volatility models. *The Journal of Finance*, 57(3):1047–1091, 2002.

[11] Elisa Alos, Jorge A León, and Josep Vives. On the short-time behavior of the implied volatility for jump-diffusion models with stochastic volatility. *Available at SSRN 1002308*, 2006.

[12] Gurdip Bakshi, Charles Cao, and Zhiwu Chen. Empirical performance of alternative option pricing models. *The Journal of Finance*, 52(5):2003–2049, 1997.

[13] Ole Barndorff-Nielsen. Normal inverse gaussian distributions and stochastic volatility modelling. *Scandinavian Journal of Statistics*, 24(1):1–13, 1997.

[14] Christian Bayer, Peter Friz, and Jim Gatheral. Pricing under rough volatility. *Quantitative Finance*, 16(6):887–904, 2016.

[15] Mikkel Bennedsen, Asger Lunde, and Mikko S Pakkanen. Decoupling the Short- and Long-Term Behavior of Stochastic Volatility. *Journal of Financial Econometrics*, 01 2021.

[16] Mikhail Chernov, A Ronald Gallant, Eric Ghysels, and George Tauchen. Alternative models for stock price dynamics. *Journal of Econometrics*, 116(1-2):225–257, 2003.

[17] Rama Cont and Purba Das. Rough volatility: fact or artefact? *arXiv preprint arXiv:2203.13820*, 2022.

[18] Rama Cont and Peter Tankov. *Financial modelling with jump processes*. Chapman and Hall/CRC, 2003.

[19] James Darrel Duffie, Jun Pan, and Kenneth Singleton. Transform analysis and asset pricing for affine jump-diffusions. *Econometrica*, 68(6):1343–1376, 2002.

[20] Omar El Euch and Mathieu Rosenbaum. The characteristic function of rough Heston models. *Mathematical Finance*, 29(1):3–38, 2019.

[21] Oosterlee C.W. Fang F. A novel pricing method for European options based on Fourier-cosine series expansions. *Siam Journal on Scientific Computing*, 31, 2009.

[22] Jin Feng, Marting Forde, and Jean-Pierre Fouque. Short maturity asymptotics for a fast mean-reverting Heston stochastic volatility model. *SIAM journal on Financial Mathematics*, 1, 2010.

[23] Damir Filipovic and Eberhard Mayerhofer. Affine diffusion processes: Theory and applications. *arXiv preprint arXiv:0901.4003*, 2009.

[24] Martin Forde and Antoine Jacquier. The large-maturity smile for the Heston model. *Finance and Stochastics*, 15(4):755–780, 2011.

[25] Jean-Pierre Fouque, George Papanicolaou, and K Ronnie Sircar. *Derivatives in financial markets with stochastic volatility*. Cambridge University Press, 2000.

[26] Jean-Pierre Fouque, George Papanicolaou, Ronnie Sircar, and Knut Solna. Multiscale stochastic volatility asymptotics. *Multiscale Modeling & Simulation*, 2(1):22–42, 2003.

[27] Jean-Pierre Fouque, George Papanicolaou, Ronnie Sircar, and Knut Solna. Short time-scale in s&p500 volatility. *Journal of Computational Finance*, 6(4):1–24, 2003.

[28] Masaaki Fukasawa. Asymptotic analysis for stochastic volatility: martingale expansion. *Finance and Stochastics*, 15:635–654, 2011.

[29] Matthieu Garcin and Martino Grasselli. Long versus short time scales: the rough dilemma and beyond. *Decisions in economics and finance*, 45(1):257–278, 2022.

[30] Jim Gatheral. *The Volatility Surface - A Practicioner's Guide*, volume 159. Wiley, 2006.

[31] Jim Gatheral, Thibault Jaisson, and Mathieu Rosenbaum. Volatility is rough. *Quantitative finance*, 18 (6):933–949, 2018.

[32] Londen S-O. Gripenberg, G. and Staffans O. *Volterra Integral and Functional Equations*. Cambridge University Press, 1990.

[33] 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.

[34] Robert A Jarrow. *Continuous-time asset pricing theory*. Springer, 2018.

[35] Paul Jusselin and Mathieu Rosenbaum. No-arbitrage implies power-law market impact and rough volatility. *Mathematical Finance*, 30(4):1309–1336, 2020.

[36] Martin Keller-Ressel. Moment explosions and long-term behavior of affine stochastic volatility models. *Mathematical Finance: An International Journal of Mathematics, Statistics and Financial Economics*, 21(1):73–98, 2011.

[37] Ryan McCrickerd. Pathwise volatility: Cox-Ingersoll-Ross initial-value problems and their fast reversion exit-time-limits. 2019.

[38] Serguei Mechkov. Fast-reversion limit of the Heston model. *Available at SSRN 2418631*, 2015.

[39] Hugh L. Montgomery and Robert C. Vaughan. *Multiplicative nmber theory I: classical theory*. Cambridge University Press, 2007.

[40] Walter Rudin. *Real and complex analysis*. Cambridge University Press, 1966.

[41] Ken-iti Sato. *Lévy process and Infinitely Divisible Distributions*. Cambridge University Press, 1999.

[42] Erwin Schrödinger. Zur theorie der fall- und steigversuche an teilchen mit brownscher bewegung. *Physikalische Zeitschrift*, 16:289–295, 1915.

[43] Anatoli V. Skorokhod. Limit theorems for stochastic processes. *Theory of Probability and Its Applications*, 1(3):261–290, 1956.

[44] Ward Whitt. *Stochastic-process limits*. Springer Series in Operations Research, 2002.