Lifting the Heston model

Eduardo Abi Jaber*

AXA Investment Managers, Multi Asset Client Solutions, Quantitative Research,
6 place de la Pyramide, 92908 Paris - La Défense, France.

Université Paris-Dauphine, PSL University, CNRS, CEREMADE, 75016 Paris, France.

October 12, 2018

Abstract

How to reconcile the classical Heston model with its rough counterpart? We introduce a lifted version of the Heston model with n multi-factors, sharing the same Brownian motion but mean reverting at different speeds. Our model nests as extreme cases the classical Heston model (when $n = 1$), and the rough Heston model (when n goes to infinity). We show that the lifted model enjoys the best of both worlds: Markovianity and satisfactory fits of implied volatility smiles for short maturities with very few parameters. Further, our approach speeds up the calibration time and opens the door to time-efficient simulation schemes.

Keywords: Stochastic volatility, implied volatility, affine Volterra processes, Riccati equations, rough volatility.

1   Introduction

Conventional one-dimensional continuous stochastic volatility models, including the renowned Heston model [25]:

$$dS_t = S_t\sqrt{V_t}dB_t, \quad S_0 > 0, \tag{1.1}$$

$$dV_t = \lambda(\theta - V_t)dt + \nu\sqrt{V_t}dW_t, \quad V_0 \geq 0, \tag{1.2}$$

have struggled in capturing the risk of large price movements on a short timescale. In the pricing world, this translates into failure to reproduce the at-the-money skew observed in the market as illustrated on the following figure.
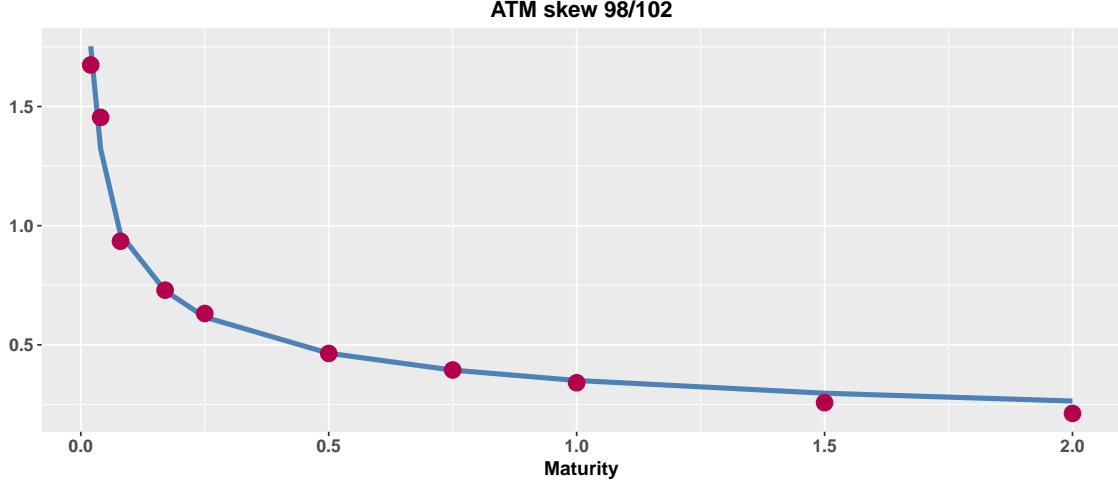

Figure 1: Term structure of the at-the-money skew for the S&P index on June 20, 2018 (red dots) and a power-law fit $t \to 0.35 \times t^{-0.41}$.

In view of improving the overall fit, several directions have been considered over the past decades. Two of the most common extensions are adding jumps [11, 22] and stacking additional random factors [7, 20], in order to jointly account for short and long timescales. While the two approaches have structural differences, they both suffer from the curse of dimensionality, as more parameters are introduced, slowing down the calibration process. Recently, rough volatility models have been introduced as a fresh substitute with remarkable fits of the implied volatility surface, see [23, 5, 17]. The rough variance process involves a one-dimensional Brownian motion, keeps the number of parameters small and enjoys continuous paths. However, the price to pay is that rough volatility models leave the realm of semimartingale and Markovian models, which makes pricing and hedging a challenging task, while degrading the calibration time. Here, the curse of dimensionality hits us straight in the face in the non-Markovianity of the process. Indeed, the rough model can be seen as an infinite dimensional Markovian model, as shown in [2, 13].

Going back to the standard Heston model (1.1)-(1.2), despite its lack of fit for short maturities, it remains increasingly popular among practitioners. This is due to its high tractability, by virtue of the closed form solution of the characteristic function, allowing for fast pricing and calibration by Fourier inversion techniques [9, 18]. Recently, El Euch and Rosenbaum [16] combined the tractability of the Heston model with the flexibility of rough volatility models, to elegantly concoct a rough counterpart of (1.1)-(1.2), dubbed the rough Heston model. More precisely, the rough model is constructed by replacing the variance process (1.2) by a fractional square-root process as follows

$$dS_t = S_t\sqrt{V_t}dB_t, \quad S_0 > 0, \tag{1.3}$$

$$V_t = V_0 + \frac{1}{\Gamma(H+1/2)} \int_0^t (t-s)^{H-1/2} \left(\lambda(\theta - V_s)ds + \nu\sqrt{V_s}dW_s \right), \tag{1.4}$$

where $H \in (0, 1/2]$ has a physical interpretation, as it measures the regularity of the sample paths of V, see [23, 6], the case $H = 1/2$ corresponding to the standard Heston model. More precisely, the sample paths of V are locally Hölder continuous of any order strictly less than H. As for the standard Heston model, the characteristic function of the log-price is known, but only up to the solution of a certain fractional Riccati Volterra equation. Indeed, both models belong to the tractable and unifying class of affine Volterra processes introduced in [3]. The following table summarizes the characteristics of the two models.

Characteristics	Heston	Rough Heston
Markovian	✓	✗
Semimartingale	✓	✗
Simulation	Fast	Slow
Affine Volterra process	✓	✓
Characteristic function	Closed	Fractional Riccati
Calibration	Fast	Slower
Fit short maturities	✗	✓
Regularity of sample paths	$H = 0.5$	$0 < H \leq 0.5$

Table 1: Summary of the characteristics of the models.

In the present paper, we introduce a conventional multi-factor continuous stochastic volatility model: *the lifted Heston model*. The variance process is constructed as a weighted sum of n factors, driven by the same one-dimensional Brownian motion, but mean reverting at different speeds, in order to accommodate a full spectrum of timescales. At first glance, the model seems over-parametrized, with already $2n$ parameters for the mean reversions and the weights. Inspired by the approximation results of [1], we provide a good parametrization of these $2n$ parameters in terms of one single parameter H, which is nothing else but the Hurst index of a limiting rough Heston model (1.3)-(1.4), obtained after sending the numbers of factors to infinity.

The lifted model not only nests as extreme cases the classical Heston model (when $n = 1$) and the rough Heston model (when n goes to infinity), but also enjoys the best of both worlds: the flexibility of rough volatility models, and the Markovianity of their conventional counterparts. Further, the model remains tractable, as it also belongs to the class of affine Volterra processes. Here, the characteristic function of the log-price is known up to a solution of a finite system of Riccati ordinary differential equations. From a practical viewpoint, we demonstrate that the *lifted Heston model*:

- reproduces the same volatility surface as the rough Heston model for maturities ranging from one week to two years,

- mimics the explosion of the at-the-money skew for short maturities,

- calibrates twenty times faster than its rough counterpart,

- is easier to simulate than the rough model.

All in all, the *lifted Heston model* can be more easily implemented than its rough counterpart, while still retaining the precision of implied volatility fits of the rough Heston model. Further, the *lifted Heston model* is able to generate a volatility surface, which cannot be generated by the classical Heston model, with only one additional parameter. Finally, the stock price and the variance process enjoy continuous paths and only depend on a two-dimensional Brownian motion, leading to simple and feasible hedging strategies.

The paper is outlined as follows. In Section 2 we introduce our *lifted Heston model* and provide its existence, uniqueness and its affine Fourier-Laplace transform. Exploiting the limiting rough model, we proceed in Section 3 to a reduction of the number of parameters to calibrate. Numerical experiments for the model, with $n = 20$ factors, are illustrated in Section 4, both for calibration and simulation. Finally, some technical material is postponed to Appendices A-C.

2 The lifted Heston model

We fix $n \in \mathbb{N}$ and we define the *lifted Heston model* as a conventional stochastic volatility model, with n factors for the variance process:

$$dS_t^n = S_t^n \sqrt{V_t^n} dB_t, \quad S_0^n > 0, \tag{2.1}$$

$$V_t^n = g_0^n(t) + \sum_{i=1}^n c_i^n U_t^{n,i}, \tag{2.2}$$

$$dU_t^{n,i} = \left(-x_i^n U_t^{n,i} - \lambda V_t^n\right) dt + \nu \sqrt{V_t^n} dW_t, \quad U_0^{n,i} = 0, \quad i = 1, \ldots, n, \tag{2.3}$$

with parameters the function g_0^n, $\lambda, \nu \in \mathbb{R}_+$, $c_i^n, x_i^n \geq 0$, for $i = 1, \ldots, n$, and $B = \rho W + \sqrt{1 - \rho^2} W^\perp$, with $(W, W^\perp)$ a two dimensional Brownian motion on a fixed filtered probability space $(\Omega, \mathcal{F}, \mathbb{F} := (\mathcal{F}_t)_{t \geq 0}, \mathbb{Q})$, with $\rho \in [-1, 1]$.

We stress that all the factors $(U^{n,i})_{1 \leq i \leq n}$ start from zero[1] and share the same dynamics, with the same one-dimensional Brownian motion W, except that they mean revert at different speeds $(x_i^n)_{1 \leq i \leq n}$. Further, the deterministic input curve g_0^n allows one to plug-in initial term-structure curves. More precisely, taking the expectation in (2.2) leads to the following relation

$$\mathbb{E}[V_t^n] + \lambda \sum_{i=1}^n c_i^n \int_0^t e^{-x_i^n(t-s)} \mathbb{E}[V_s^n] ds = g_0^n(t), \quad t \geq 0.$$

In practice, the forward variance curve, up to a horizon $T > 0$, can be extracted from variance swaps observed in the market and then plugged-in in place of $(\mathbb{E}[V_t^n])_{t \leq T}$ in the previous expression. For a suitable choice of continuous curves g_0^n, for instance if

$$g_0^n \text{ is non-decreasing such that } g_0^n(0) \geq 0, \tag{2.4}$$

or

$$g_0^n : t \to V_0 + \sum_{i=1}^n c_i^n \int_0^t e^{-x_i^n(t-s)} \theta(s) ds, \text{ with } V_0, \theta \geq 0, \tag{2.5}$$

there exists a unique continuous $\mathbb{F}$-adapted strong solution $(S^n, V^n, (U^{n,i})_{1 \leq i \leq n})$ to (2.1)-(2.3), such that $V_t^n \geq 0$, for all $t \geq 0$, and S^n is a $\mathbb{F}$-martingale. We refer to Appendix A for more details and the exact definition of the set of admissible input curves g_0^n.

Since our main objective is to compare the lifted model to other existent models, we will restrict to the case of input curves of the form

$$g_0^n : t \to V_0 + \theta \sum_{i=1}^n c_i^n \int_0^t e^{-x_i^n(t-s)} ds, \quad \text{with } V_0, \theta \geq 0. \tag{2.6}$$

Setting $n = 1$, $c_1^1 = 1$ and $x_1^1 = 0$, the *lifted Heston model* degenerates into the standard Heston model (1.1)-(1.2). So far, the multi-factor extensions of the standard Heston model have been considered by stacking additional square-root processes as in the double Heston model[2] of [10]

---

[1] Notice that the initial value of the variance process V^n is $g_0^n(0)$.

[2] The double Heston model is defined in [10] as follows

$$dS_t^n = S_t^n \left(\sqrt{U_t^1} dB_t^1 + \sqrt{U_t^2} dB_t^2\right), \tag{2.7}$$

$$dU_t^i = \lambda_i(\theta_i - U_t^i) dt + \nu_i \sqrt{U_t^i} dW_t^i, \quad U_0^i \geq 0, \quad i \in \{1, 2\}, \tag{2.8}$$

where $B^i = \rho_i W^i + \sqrt{1 - \rho_i^2} W^{i,\perp}$ with $\rho_i \in [-1, 1]$ and $(W^1, W^2, W^{1,\perp}, W^{2,\perp})$ a four-dimensional Brownian motion.

and the multi-scale Heston model of [19], or by considering a Wishart matrix-valued process as in [14]. In both cases, the dimension of the driving Brownian motion for the variance process, along with the number of parameters, grows with the number of factors. Clearly, the *lifted Heston model* differs from these extensions, one can compare (2.1)-(2.3) for $n = 2$ with (2.7)-(2.8).

Just like the classical Heston model, the *lifted Heston model* remains tractable. Specifically, fix $u \in \mathbb{C}$ such that $\mathrm{Re}(u) \in [0, 1]$. By virtue of Appendix B, the Fourier-Laplace transform of the log-price is exponentially affine with respect to the factors $(U^{n,i})_{1 \le i \le n}$:

$$\mathbb{E}\left[\exp\left(u \log S_t^n\right) \mid \mathcal{F}_t\right] = \exp\left(\phi^n(t, T) + u \log S_t^n + \sum_{i=1}^{n} c_i^n \psi^{n,i}(T - t) U_t^{n,i}\right), \qquad (2.9)$$

for all $t \le T$, where $(\psi^{n,i})_{1 \le i \le n}$ solves the following n-dimensional system of Riccati ordinary differential equations

$$(\psi^{n,i})' = -x_i^n \psi^{n,i} + F\left(u, \sum_{j=1}^{n} c_j^n \psi^{n,j}\right), \quad \psi^{n,i}(0) = 0, \quad i = 1, \ldots, n, \qquad (2.10)$$

with

$$F(u, v) = \frac{1}{2}(u^2 - u) + (\rho\nu u - \lambda)v + \frac{\nu^2}{2}v^2,$$

and

$$\phi^n(t, T) = \int_0^{T-t} F\left(u, \sum_{i=1}^{n} c_i^n \psi^{n,i}(s)\right) g_0^n(T - s)ds, \quad t \le T.$$

In particular, for $t = 0$, since $U_0^{n,i} = 0$ for $i = 1, \ldots, n$, the unconditional Fourier-Laplace transform reads

$$\mathbb{E}\left[\exp\left(u \log S_t^n\right)\right] = \exp\left(u \log S_0^n + \int_0^T F\left(u, \sum_{i=1}^{n} c_i^n \psi^{n,i}(s)\right) g_0^n(T - s)ds\right). \qquad (2.11)$$

A similar formula holds for the Fourier-Laplace transform of the joint process $(\log S^n, V^n)$ with integrated log-price and variance, we refer to the Appendix B for the precise expression.

Consequently, the Fourier-Laplace transform of the *lifted Heston model* is known in closed-form, up to the solution of a deterministic n-dimensional system of ordinary differential equations (2.10), which can be solved numerically. Once there, standard Fourier inversion techniques can be applied on (2.11) to deduce option prices. This is illustrated in the following sections.

3   Parameter reduction and the choice of the number of factors

In this section, we proceed to a reduction of the number of parameters to calibrate. Our inspiration stems from rough volatility. In a first step, for every n, we provide a parametrization of the weights and the mean reversions $(c_i^n, x_i^n)_{1 \le i \le n}$ in terms of the Hurst index H of a limiting rough volatility model and one additional parameter r_n. Then, we specify the number of factors n and the value of the additional parameter r_n so that the lifted model reproduces the same volatility surface as the rough Heston model for maturities ranging from one week up to two years, while calibrating twenty times faster than its rough counterpart. Benchmarking against rough volatility models is justified by the fact that one of the main strengths of these models is their ability to achieve better fits of the implied volatility surface than conventional stochastic volatility models. This has been illustrated on real market data in [5, 17]. Finally, for the sake of completeness, we provide a comparison with the standard Heston model.

3.1 Parametrization in terms of the Hurst index

For an initial input curve of the form (2.6), the *lifted Heston model* (2.1)-(2.3) has the same five parameters $(V_0, \theta, \lambda, \nu, \rho)$ of the Heston model, plus $2n$ additional parameters for the weights and the mean reversions $(c_i^n, x_i^n)_{1 \le i \le n}$.[3] At first sight, the model seems to suffer from the curse of dimensionality, as it requires the calibration of $(2n + 5)$ parameters. This is where the exciting theory of rough volatility finally comes into play. Inspired by the approximation result [1, Theorem 3.5], we suggest to use a parametrization of $(c_i^n, x_i^n)_{1 \le i \le n}$ in terms of two well-chosen parameter. By doing so, we reduce the $2n$ additional parameters to calibrate to only two effective parameters.

Qualitatively, we choose the weights and mean reversions $(c_i^n, x_i^n)_{1 \le i \le n}$ in such a way that sending the number of factors $n \to \infty$ would yield the convergence of the *lifted Heston model* towards a rough Heston model (1.3)-(1.4), with parameters $(V_0, \theta, \lambda, \nu, \rho, H)$. The additional parameter $H \in (0, 1/2)$ is the so-called Hurst index of the limiting fractional variance process (1.4), and it measures the regularity of its sample paths.

More precisely, for a fixed even number of factors n, we fix $r_n > 1$ and we consider the following parametrization for the weights and the mean reversions

$$c_i^n = \frac{(r_n^{1-\alpha} - 1)r_n^{(\alpha-1)(1+n/2)}}{\Gamma(\alpha)\Gamma(2-\alpha)} r_n^{(1-\alpha)i} \quad \text{and} \quad x_i^n = \frac{1-\alpha}{2-\alpha} \frac{r_n^{2-\alpha} - 1}{r_n^{1-\alpha} - 1} r_n^{i-1-n/2}, \quad i = 1, \ldots, n, \quad (3.1)$$

where $\alpha := H + 1/2$ for some $H \in (0, 1/2)$.[4]

If in addition, the sequence $(r_n)_{n \ge 1}$ satisfies

$$r_n \downarrow 1 \quad \text{and} \quad n \ln r_n \to \infty, \quad \text{as } n \to \infty, \quad (3.2)$$

then, Theorem A.2 in the Appendix ensures the convergence of the lifted model towards the rough Heston model, as n goes to infinity. We refer to Appendix A for more details.

In order to visualize this convergence, we first define the following sequence

$$r_n = 1 + 10\, n^{-0.9}, \quad n \ge 1, \quad (3.3)$$

which clearly satisfies (3.2). Then, we generate our benchmark implied volatility surface,

$$\text{for 9 maturities } T \in \{\text{1w, 1m, 2m, 3m, 6m, 9m, 1y, 1.5y, 2y}\}, \quad (3.4)$$

$$\text{with up to 80 strikes } K \text{ per maturity}, \quad (3.5)$$

with a rough Heston model[5] with parameters $\Theta_0 := (V_0, \theta, \lambda, \nu, \rho, H)$ given by

$$V_0 = 0.02, \quad \theta = 0.02, \quad \lambda = 0.3, \quad \nu = 0.3, \quad \rho = -0.7 \quad \text{and} \quad H = 0.1. \quad (3.6)$$

---

[3]If one chooses g_0^n to match the forward variance curve, then, the parameters (V_0, θ) can be eliminated from both models.

[4]This corresponds to equation (3.6) in [1] with the geometric partition $\eta_i^n = r_n^{i-n/2}$ for $i = 0, \ldots, n$, which is in the spirit of [8] for the approximation of the factional Brownian motion.

[5]The implied volatility surface is generated by first solving numerically the corresponding fractional Riccati equations with the Adams Predictor-Corrector scheme [15] with 200 time steps, see [17, Appendix A] for more details. Then, call prices are computed via the cosine method [18] for the inversion of the characteristic function. We note that other Fourier inversion techniques can be used for the second step, for instance, the Carr-Madan method [9], as done in [17]. As illustrated in [18], for the same level of accuracy, the cosine method is approximately 20 times faster than the Carr-Madan method, and needs drastically less evaluation points of the characteristic function $(E\left[\exp\left(u_i \log S_t^n\right)\right])_{i \in \mathcal{I}}$ ($|\mathcal{I}| = 160$ for the cosine methods and $|\mathcal{I}| = 4096$ for the Carr-Madan method). This latter point is crucial in our case since, for every $i \in \mathcal{I}$, evaluation of $E\left[\exp\left(u_i \log S_t^n\right)\right]$ requires a numerical discretization of the corresponding Riccati equation.

The generated implied volatility is kept fixed and is denoted by $\sigma_\infty(K,T;\Theta_0)$, for every pair (K,T) in (3.4)-(3.5).

Then, for each $n \in \{10, 20, 50, 100, 500\}$, we generate the implied volatility surface of the *lifted Heston model*[6] with n-factors, with the same set of parameters Θ_0 as in (3.6), and (3.3) plugged in (3.1). For each n, the generated surface is denoted by $\sigma_n(K,T;r_n,\Theta_0)$, for every pair (K,T) in (3.4)-(3.5).

Because the sequence $(r_n)_{n\geq 1}$ defined in (3.3) satisfies condition (3.2), as n grows,

$$\sigma_n(K,T;r_n,\Theta_0) \to \sigma_\infty(K,T;\Theta_0),$$

by virtue of Theorem A.2 in the Appendix. This convergence phenomenon is illustrated in Figure 2 below for two maturity slices, one week and one year.
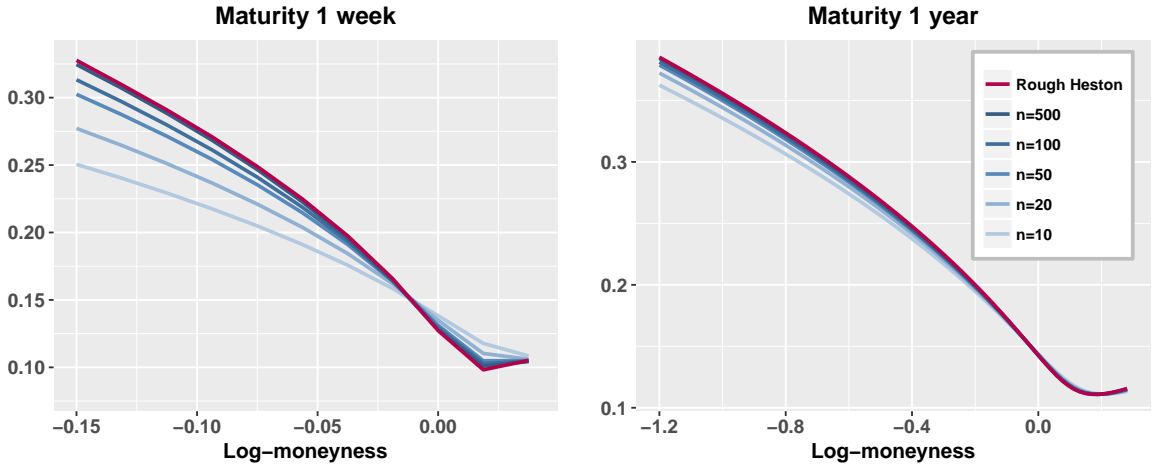

Figure 2: Convergence of the implied volatility surface of the lifted model $\sigma_n(k,T;r_n,\Theta_0)$, with $r_n = 1 + 10\, n^{-0.9}$, towards its rough counterpart $\sigma_\infty(k,T;\Theta_0)$, illustrated on two maturities slices $T \in \{1\text{ week}, 1\text{ year}\}$. Here $k := \ln(K/S_0)$ stands for the log-moneyness.

In view of assessing the proximity between the implied volatility surface $\sigma_n(K,T;r_n,\Theta_0)$ of the *lifted Heston model* and that of the rough Heston model $\sigma_\infty(K,T;\Theta_0)$, we compute the mean squared error (MSE) between the two volatility surfaces defined as follows

$$\frac{1}{\sum_{(K',T')} w(K',T')} \sum_{(K,T)} w(K,T)(\sigma_n(K,T;r_n,\Theta_0) - \sigma_\infty(K,T;\Theta_0))^2,$$

where we sum over all pairs (K,T) as in (3.4)-(3.5). Here, w stands for a matrix of weights, where we put more weight on options near the money and with short time to maturity (one could also set $w(K,T) = 1$ for all (K,T)).

The corresponding mean squared errors of Figure 2 are reported in Table 2 below, along with the computational time[7] for generating the whole volatility surface, for all pairs (K,T) as in

---

[6] The implied volatility surface is generated by first solving numerically the n-dimensional Riccati equations with the explicit-implicit scheme (C.2) detailed in the Appendix with a number of time steps $N = 300$. As before, the call prices are then computed via the cosine method [18] for the inversion of the characteristic function.

[7] All cpu times are computed on a laptop with Intel core i7 processor at 2.2GHz and 16GB of memory. The code, written in R, is far from being optimized.

(3.4)-(3.5), that is, for 9 maturities slices with up to 80 strikes per maturity.[8]

	n	$r_n = 1 + 10\,n^{-0.9}$	Time (seconds)	MSE
Lifted Heston	10	2.26	3.9	1.20e-03
	20	1.67	4.4	1.85e-04
	50	1.3	5.2	6.81e-05
	100	1.16	6.6	2.54e-05
	500	1.04	17.4	3.66e-06
Rough Heston	$n \to \infty$	$r_n \downarrow 1$	106.8	

Table 2: Convergence of the lifted model towards its rough counterpart for $r_n = 1 + 10\,n^{-0.9}$, with the corresponding computational time in seconds for generating the implied volatility surface (3.4)-(3.5).

All in all, we notice that the number of effective parameters remains constant and does not depend on the number of factors n. This has to be contrasted with the usual multi-factor extensions: the double Heston model (2.7)-(2.8) already has 10 parameters $(U_0^i, \theta_i, \lambda_i, \nu_i, \rho_i)_{i \in \{1,2\}}$, the multi-scale model of [19] also suffers from over-parametrization.

In the subsequent subsection, we will explain how to fix n and r_n, so that the parameters to calibrate are reduced to only six effective parameters $(V_0, \theta, \lambda, \nu, \rho, H)$, one additional parameter than the standard Heston model!

3.2 Practical choice of n and r_n

We suggest to fix the following values

$$n = 20 \quad \text{and} \quad r_{20} = 2.5. \tag{3.7}$$

Our choice will be based on the numerical comparison with the rough Heston model of the previous section.

We start by explaining our choice for the number of factors n in (3.7). Based on Table 2, we choose n with a good trade-off between time-efficiency and proximity to the rough volatility surface. Fixing $n = 20$ seems to be a good choice. Visually, as already shown on Figure 2, the two implied volatility slices have almost identical shapes. Whence, one would expect that by letting the parameters r_{20} free, one could achieve a perfect fit of the rough surface with only $n = 20$ factors. This can be formulated as follows: keeping the six parameters of the lifted model fixed as in (3.6), can one find $r_{20}^*(\Theta_0) > 1$ such that

$$\sigma_{20}(K, T; r_{20}^*(\Theta_0), \Theta_0) \approx \sigma_\infty(K, T; \Theta_0), \quad \text{for all } K, T?$$

The next subsection provides a positive answer.

3.2.1 Mimicking roughness by increasing r_{20}

First, one needs to understand the influence of the parameter r_n on the *lifted Heston model*. Increasing r_n has the effect of boosting the parameters $(c_i^n, x_i^n)_{1 \le i \le n}$ in (3.1), leading to an

---

[8]One cannot draw definite quantitative conclusions regarding the comparison between the computational times of the lifted surface and the one of the rough surface. Indeed, one needs a more careful study of the discretization errors of the corresponding Riccati equations before comparing the computational times needed to reach the same level of accuracy. We omit to do so here. However, even if one reduces the number of time steps from 200 to 150 in the Adams scheme, it still takes 67.2 seconds to compute the rough surface. Recall that we used $N = 300$ time steps for the n-dimensional Riccati equation of the lifted model. In any case, it should be clear that solving the 20-dimensional Riccati equations is considerably faster then solving the fractional Riccati equation.

increase of the vol-of-vol parameter of the lifted model given by $\nu \sum_{i=1}^{n} c_i^n$, together with faster mean-reversions $(x_i^n)_{1 \le i \le n}$ for the factors. In analogy with conventional stochastic volatility models, such as the standard Heston model (1.1)-(1.2), increasing the vol-of-vol parameter together with the speed of mean reversion yields a steeper skew at the short-maturity end of the volatility surface. Consequently, increasing the parameter r_n in the lifted model should steepen the implied volatility slice for short-maturities. Figure 3 below confirms that this is indeed the case when one increases the value of r_{20} from 1.67 to 2.8, for the 20-dimensional lifted model, as the two slices now almost perfectly match:
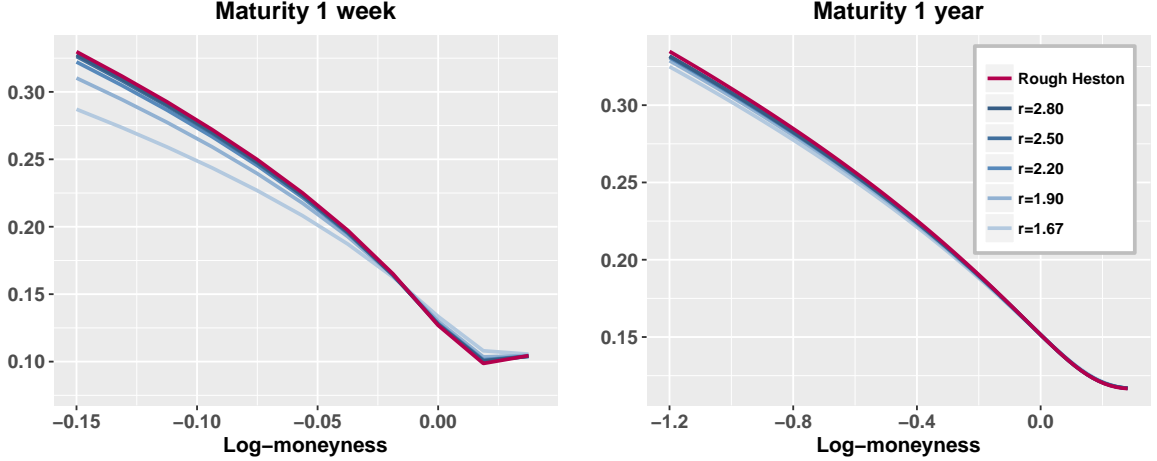

Figure 3: Implied volatility of the 20-dimensional lifted model $\sigma_{20}(k, T; r_{20}, \Theta_0)$, for different values of r_{20} ranging from 1.67 to 2.8, and the rough surface $\sigma_{\infty}(k, T; \Theta_0)$, for two maturities slices $T \in \{1 \text{ week}, 1 \text{ year}\}$.

The corresponding mean squared errors of Figure 3 are collected in Table 3 below.

Lifted Heston	$(n = 20)$
r_{20}	MSE
1.67	1.85e-04
1.90	4.16e-05
2.20	8.72e-06
2.50	3.64e-06
2.80	2.81e-06

Table 3: Mean squared errors between the 20-dimensional lifted model $\sigma_{20}(k, T; r_{20}, \Theta_0)$ and the rough model $\sigma_{\infty}(k, T; \Theta_0)$, for different values of r_{20}.

Because r_n has to converge to 1, when n goes to infinity, recall (3.2), we seek to keep r_n as small as possible. For $n = 20$, fixing $r_{20}^*(\Theta_0) = 2.5$ yields already satisfactory results, improving the mean squared error of 1.85e-04 in Table 2 to 3.64e-06. Further, this choice yields the same order of precision as with $n = 500$ factors given in Table 2.

Before moving to a physical justification of the choice of r_{20}, we proceed to the full calibration of the *lifted Heston model* with $n = 20$ and $r_{20} = 2.5$ to the rough volatility surface $\sigma_{\infty}(K, T; \Theta_0)$. That is, we let the six effective parameters $(V_0, \theta, \lambda, \nu, \rho, H)$ of the lifted model free. The calibrated values $\hat{\Theta}_0 := (\hat{V}_0, \hat{\theta}, \hat{\lambda}, \hat{\nu}, \hat{\rho}, \hat{H})$, provided in Table 4, agree with (3.6). At the visual

level, as shown on Figure 10 in the Appendix, the calibrated lifted surface is indistinguishable from the rough surface $\sigma_\infty(K, T; \Theta_0)$ for all maturities ranging from one week to two years, with a mean squared error of order 4.01e-07.

Parameters	Calibrated values
$\hat{V}_0$	0.02012504
$\hat{\theta}$	0.02007956
$\hat{\lambda}$	0.29300681
$\hat{\nu}$	0.30527694
$\hat{\rho}$	-0.70241116
$\hat{H}$	0.09973346

Table 4: Calibrated *lifted Heston model* parameters.

We now provide another physical justification for the choice of r_{20} based on an infinite-dimensional Markovian representation of the limiting rough variance process (1.4) due to [2], which we recall in the following remark.

Remark 3.1 (Representation of the limiting rough process). *The fractional kernel appearing in the limiting rough process* (1.4) *admits the following Laplace representation*

$$\frac{t^{H-1/2}}{\Gamma(H+1/2)} = \int_0^\infty e^{-xt}\mu(dx), \quad \text{with} \ \mu(dx) = \frac{x^{-H-1/2}}{\Gamma(1/2-H)\Gamma(H+1/2)},$$

so that the stochastic Fubini theorem, after setting $V_0 \equiv 0$ in (1.4), *leads to*

$$V_t = \int_0^\infty U_t(x)\mu(dx), \quad x > 0,$$

where, for all $x > 0$,

$$U_t(x) := \int_0^t e^{-x(t-s)}\left(\lambda(\theta - V_s)ds + \nu\sqrt{V_s}dW_s\right).$$

This can be seen as the mild formulation of the following stochastic partial differential equation

$$dU_t(x) = \left(-xU_t(x) + \lambda\left(\theta - \int_0^\infty U_t(y)\mu(dy)\right)\right)dt + \nu\sqrt{\int_0^\infty U_t(y)\mu(dy)}dW_t, \tag{3.8}$$

$$U_0(x) = 0, \quad x > 0. \tag{3.9}$$

Whence, the rough process can be reinterpreted as a superposition of infinitely many factors $(U_.(x))_{x>0}$ sharing the same dynamics but mean reverting at different speeds $x \in (0, \infty)$. We refer to [2] for the rigorous treatment of this representation. One makes the following observations:

- *multiple timescales are naturally encoded in rough volatility models, which can be a plausible explanation for their ability to achieve better fits than conventional one-dimensional models,*

- *the largest mean reversions going to infinity characterize the factors responsible of the roughness of the process.*

10

We notice that for the lifted model, the mean reversions in (3.1) satisfy

$$x_i^n \geq r_n^{i-1-n/2}, \quad i = 1, \ldots, n.$$

Therefore, based on Remark 3.1, for $n = 20$, one would like to force x_{20}^{20} to be large enough in order to mimic roughness and account for very short timescales, while having x_1^{20} small enough to accommodate a whole palette of timescales. Setting

$$r_{20} \approx 2.5,$$

would cover mean reversions between 10^{-4} and 10^4.

The previous justification suggests that once $n = 20$ is fixed, one can choose r_{20} independently of the parameters Θ. The next experiment shows that this is indeed the case.

3.2.2   Robustness of r_{20}: a numerical test

Throughout this section, we fix the three parameters $V_0, \theta = 0.02$ and $\lambda = 0$. In order to verify experimentally the robustness of $r_{20} = 2.5$, we proceed as follows.

1. Simulate $M = 500$ set of parameters $(\Theta_k := (0.02, 0.02, 0, \nu_k, \rho_k, H_k))_{k=1,\ldots,M}$ uniformly distributed with the following bounds

$$0.05 \leq \nu \leq 0.5, \quad -0.9 \leq \rho \leq -0.5, \quad 0.05 \leq H \leq 0.2.$$

2. For each $k = 1, \ldots, M$:

   (a) Generate the rough volatility surface $\sigma_\infty(K, T; \Theta_k)$, for all pairs (T, K) in (3.4)-(3.5),

   (b) Generate the lifted volatility surface $\sigma_{20}(K, T; r_{20} = 2.5, \Theta_k)$, for all pairs (T, K) in (3.4)-(3.5),

   (c) Compute the mean squared error between the two volatility surfaces:

   $$\text{MSE}_k := \frac{1}{\sum_{(K',T')} w(K', T')} \sum_{(K,T)} w(K, T)(\sigma_{20}(K, T; r_{20} = 2.5, \Theta_k) - \sigma_\infty(K, T; \Theta_k))^2.$$

The scatter plot and the empirical distribution of the mean squared error $(\text{MSE}_k)_{k=1,\ldots,M}$ are illustrated in Figure 4 below.
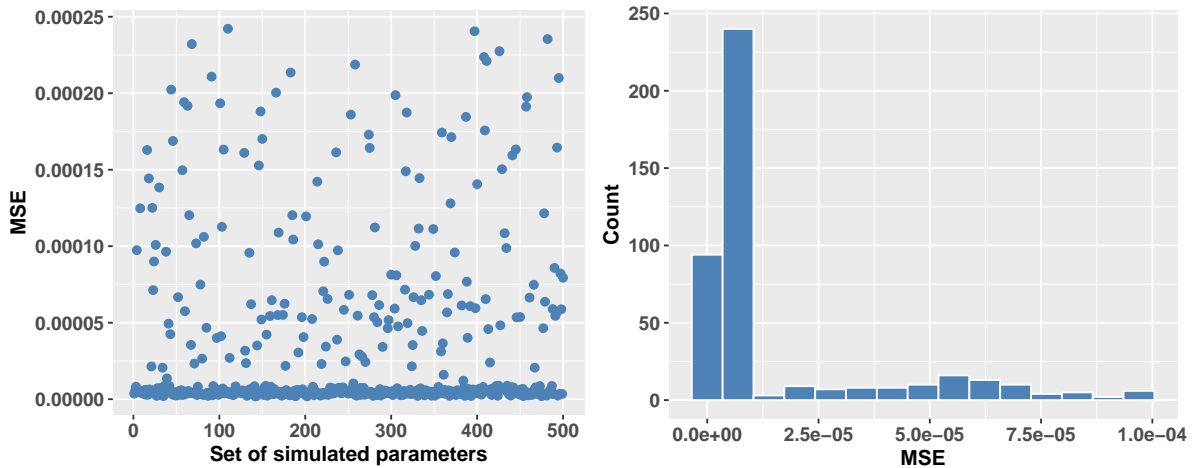

Figure 4:   Scatter plot (left) and empirical distribution (right) of the mean squared error $(\text{MSE}_k)_{k=1,\ldots,M}$ of the $M = 500$ simulated set of parameters $(\Theta_k)_{k=1,\ldots,M}$.

The first twenty values of the simulated set of parameters with the corresponding mean squared error are provided in Table 8 in the Appendix. We observe that the lifted surfaces are quite close to the rough surface, for any value of the simulated parameters. This is confirmed by Table 5 below, where we collect the descriptive statistics of the computed mean squared errors $(\mathrm{MSE}_k)_{k=1,\ldots,M}$.

	MSE
Minimum	1.81e-06
1st Quantile	3.83e-06
Median	5.48e-06
3rd Quantile	4.91e-05
Maximum	2.42e-04

Table 5: Descriptive statistics of the mean squared error $(\mathrm{MSE}_k)_{k=1,\ldots,M}$ of the $M = 500$ simulated set of parameters $(\Theta_k)_{k=1,\ldots,M}$.

We now show that the mean squared errors can be improved by letting the three parameters (ν, ρ, H) of the lifted model free. Specifically, consider the worst mean squared error of Table 5

$$\max_{\Theta_k} \mathrm{MSE}_k = 2.42\text{e-}04, \tag{3.10}$$

which is attained for the set of parameters Θ_{101} with

$$\nu_{101} = 0.1537099, \quad \rho_{101} = -0.8112745 \ \text{ and } \ H_{101} = 0.1892725.$$

Keeping the first three parameters fixed $V_0, \theta = 0.02$ and $\lambda = 0$, we proceed to the calibration of the lifted model to the rough surface $\sigma_\infty(K, T; \Theta_{101})$. The calibration yields

$$\hat{\nu} = 0.1647801, \quad \hat{\rho} = -0.7961080 \ \text{ and } \ \hat{H} = 0.1957235,$$

improving the previous mean squared error (3.10) to 1.62e-06. This shows that, by fine tuning the parameters of the lifted model, for any rough volatility surface $\sigma_\infty(K, T; \Theta)$ with a realistic set of parameters Θ, one can find a set of parameters $\hat{\Theta}$, not too far from Θ, such that

$$\sigma_{20}(K, T; r_{20} = 2.5, \hat{\Theta}) \approx \sigma_\infty(K, T; \Theta), \quad \text{for any pair } (K, T) \text{ in } (3.4)\text{-}(3.5).$$

To sum up, we showed so far that the *lifted Heston model*, with $n = 20$ and $r_{20} = 2.5$, is able to produce the same volatility surfaces of the rough Heston model, for any realistic set of parameters, for maturities ranging between one week and two years. Consequently, it can be used directly to fit real market data instead of the rough Heston model.

Why is it more convenient to use the lifted Heston model rather than its rough counterpart?

On the one hand, it speeds-up calibration time. Indeed, solving numerically the 20-dimensional system of Riccati ordinary differential equations (2.10) is up to twenty times faster than the Adams scheme for the fractional Riccati equation. On the other hand, the lifted model remains Markovian and semimartingale, which opens the door to time-efficient recursive simulation schemes for pricing and hedging more complex exotic options. Before testing the lifted model in practice, we compare it to the standard Heston model.

3.3 Comparison with the standard Heston model

For the sake of comparison, we calibrate a standard Heston model (1.1)-(1.2) to the full rough volatility surface $\sigma_\infty(K, T; \Theta_0)$, with Θ_0 as in (3.6). Recall that the standard Heston model corresponds to the case $n = 1$, $x_1^1 = 0$ and $c_1^1 = 1$. The calibrated parameters of the standard Heston are provided in Table 6 below. We observe that the calibrated values of $(\hat{V}_0, \hat{\theta}, \hat{\rho})$ have the same magnitude as the ones of (3.6). This is not surprising since these parameters have the same interpretation in the two models: the first two parameters $(\hat{V}_0, \hat{\theta})$ govern the level of the term structure of forward variance at time 0 while ρ dictates the leverage effect between the stock price and its variance.

Parameters	Calibrated values
$\hat{V}_0$	0.019841
$\hat{\theta}$	0.032471
$\hat{\lambda}$	3.480784
$\hat{\nu}$	0.908037
$\hat{\rho}$	-0.710067

Table 6: Calibrated Heston model parameters.

Despite the extreme values of the calibrated mean reversion and vol-of-vol parameters $(\hat{\lambda}, \hat{\nu})$, the Heston model is not able to reproduce the steepness of the skew for short maturities as shown on Figure 11 in the Appendix, with a mean squared error of order 2.06e-03. For long maturities, the fit is fairly good.

In order to compare our findings with the observed stylized fact of Figure 1, we plot on Figure 5 below the term structure of the at-the-money skew of the three models: the rough Heston with parameters as in (3.6), the calibrated *lifted Heston* model of Table 4 and the calibrated Heston model of Table 6. The Heston model fails in reproducing the explosive behavior of the term structure of the at-the-money skew observed in the market. On the contrary, this feature is captured by the lifted and rough counterparts. For long maturities, all three model have the same behavior.
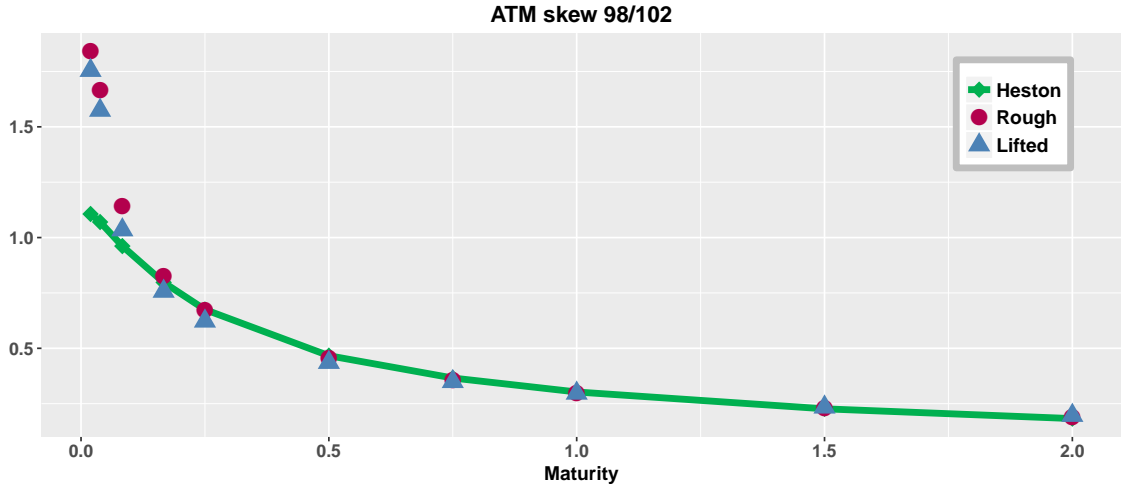

Figure 5: Term structure of the at-the-money skew of the rough Heston model $\sigma_\infty(K, T; \Theta_0)$ of (3.6) (red circles), the calibrated *lifted Heston model* $\sigma_{20}(K, T; r_{20} = 2.5, \hat{\Theta}_0)$ of Table 4 (blue

triangles) and the calibrated Heston model of Table 6 (green line).

In the sequel, we will show that, for $n = 20$ factors, the *lifted Heston model* provides an appealing trade-off between consistency with market data and tractability. We stress that $r_{20} = 2.5$ is kept fixed in the lifted model, which now has only six effective parameters to calibrate $(V_0, \theta, \lambda, \nu, \rho, H)$. Again, in practice, V_0 and Θ_0 can be eliminated by specifying the initial forward variance curve as input and λ can be set to 0, as mean reversions at different speeds are naturally encoded in the lifted model through the family $(x_i^n)_{1 \leq i \leq n}$. By doing so, one reduces the effective number of parameters to only three (ν, ρ, H), as already done in [17] for the rough Heston model.

4 Calibration on market data and simulation

In this section, we fix the number of factors to $n = 20$ and set $r_{20} = 2.5$ in (3.1). We demonstrate that the *lifted Heston model*:

- captures the explosion of the at-the-money skew observed in the market,
- is easier to simulate than the rough model,
- tricks the human eye as well as the statistical estimator of the Hurst index.

4.1 Calibration to the at-the-money skew

Going back to real market data, we calibrate the lifted model to the at-the-money skew of Figure 1. Keeping the parameters $V_0 = 0.02$, $\theta = 0.02$ and $\lambda = 0$ fixed, the calibrated parameters are given by

$$\hat{\nu} = 0.3161844, \quad \hat{\rho} = -0.6852625 \quad \text{and} \quad \hat{H} = 0.1104290. \tag{4.1}$$
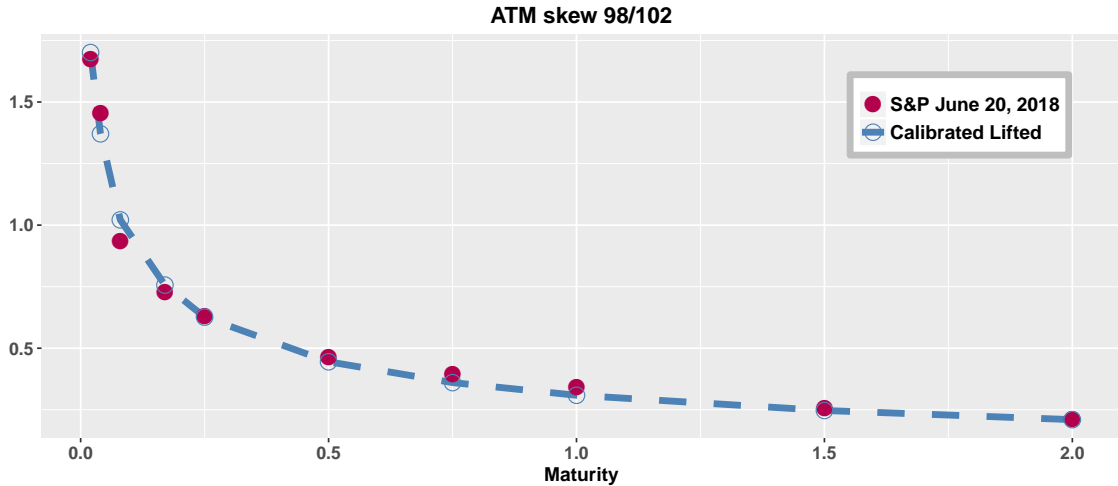
The fit is illustrated on Figure 6 below.

Figure 6: Term structure of the at-the-money skew for the S&P index on June 20, 2018 (red dots) and for the lifted model with calibrated parameters (4.1) (blue circles with dashed line).

We notice the calibrated value $\hat{H}$ in (4.1) is coherent with the value $(0.5 - 0.41) = 0.09$, which can be read off the power-law fit of Figure 1. Consequently, in the pricing world, the parameter

14

H quantifies the explosion of the at-the-money skew through a power-law $t \to Ct^{0.5-H}$, see also [21].

We discuss briefly the simulation procedure of our lifted model in the next subsection.

4.2 Simulation and estimated roughness

Until now, there is no existing scheme to simulate the variance process (1.4) of the rough Heston model, the crux resides in the non-Markovianity of the variance process, the singularity of the kernel and the square-root dynamics. In contrast, numerous approximation schemes have been developed for the simulation of the standard square-root process (1.2), see [4, Chapters 3 and 4] and the references therein. Because the *lifted Heston model* (2.1)-(2.3) is a Markovian and semimartingale model, one can adapt standard recursive Euler-Maruyama schemes to simulate the variance process V^n first, and then the stock price S^n. For $T > 0$, we consider the modified explicit-implicit scheme (C.3)-(C.4) detailed in the Appendix for the variance process V^n.

We observe on Figure 7 below that the factors $(U^{20,i})_{1 \le i \le 20}$ are highly correlated. We can distinguish between the short-term factors with fast mean reversions, responsible of the 'roughness', and the long-term factors, with slower mean reversions, determining the level of the variance process. The variance process is then obtained by aggregating these factors with respect to (2.2). We also notice that some of the factors $(U^{n,i})_{1 \le i \le n}$ become negative, but that the aggregated process V^n remains nonnegative at all time.

Remark 4.1 (Nonnegativity of the variance process)**.** *Looking at the stochastic differential equation* (2.2)-(2.3), *it is not straightforward at all why V^n should stay nonnegative at all time, even for the zero initial curve $g_0 \equiv 0$. Indeed, some of the factors $(U^{n,i})_{1 \le i \le n}$ may become negative, but surprisingly enough, their aggregated sum V^n remains nonnegative, at all time. This is due to a very special underlying structure: equations* (2.2)-(2.3) *can be recast as a stochastic Volterra equation of convolution type for a suitable kernel, we refer to Appendix A for more details.*
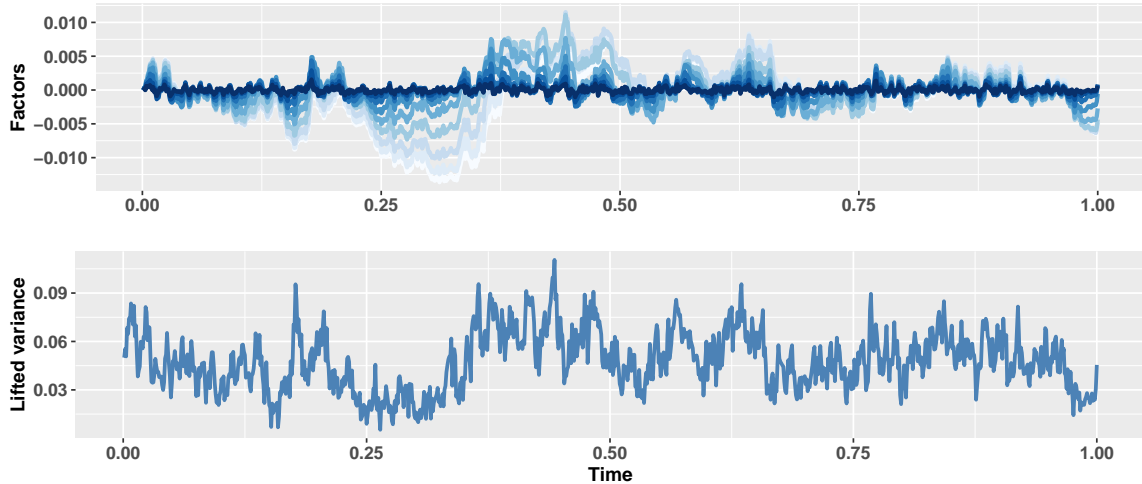

Figure 7: One sample path of the simulated factors $(U^{20,i})_{1 \le i \le 20}$ with blue intensity proportional to the speed of mean reversions $(x_i^n)_{1 \le i \le 20}$ (upper) and the corresponding aggregated variance process V^n (lower) with parameters $V_0 = 0.05$, $\theta = 0.05$, $\lambda = 0.3$, $\nu = 0.1$ and $H = 0.1$ for a time step of 0.001 and $T = 1$.

Visually, the sample path of the variance process seems rougher than the one of a standard Brownian motion. As shown on Figure 8 below, at the daily timescale, the simulated volatility process of the *lifted Heston model* not only tricks the human eye, but also misleads the statistical estimator of the Hurst index constructed in [6]. Specifically, the estimator recognizes a semimartingale model for the simulated volatility of the Heston model, with an estimated $\hat{H}$ close to 0.5. However, it fails to do so for the lifted model, the estimator displays $\hat{H} = 0.18$. The lifted model is therefore capable of mimicking, up to some extent, the 'roughness' of the volatility observed on the market, at least at the daily timescale. This should be paralleled with the explosive-like behavior of the at-the-money skew encountered earlier on Figures 5-6. Stated otherwise, if one is only provided the lower graph of Figure 8, one cannot say if the path has been generated by a rough volatility model with Hurst index $H = 0.18$ or by our lifted model with $H = 0.1$. As the step size of the discretization scheme goes to 0, the estimated H of the lifted model has to converge to 0.5, since V^n is a semimartingale, and therefore has the same regularity as a standard Brownian motion. The convergence is illustrated on Figure 12 below.

On another note, the upper graph of Figure 8 highlights the physical interpretation of the parameter H as it measures the roughness of the empirical realized volatility. Indeed, empirical studies on a very wide range of assets volatility time-series in [23, 6] revealed that the dynamics of the log-volatility are close to that of a fractional Brownian motion with a 'universal' Hurst parameter H of order 0.1, from intra-day up to daily timescales.
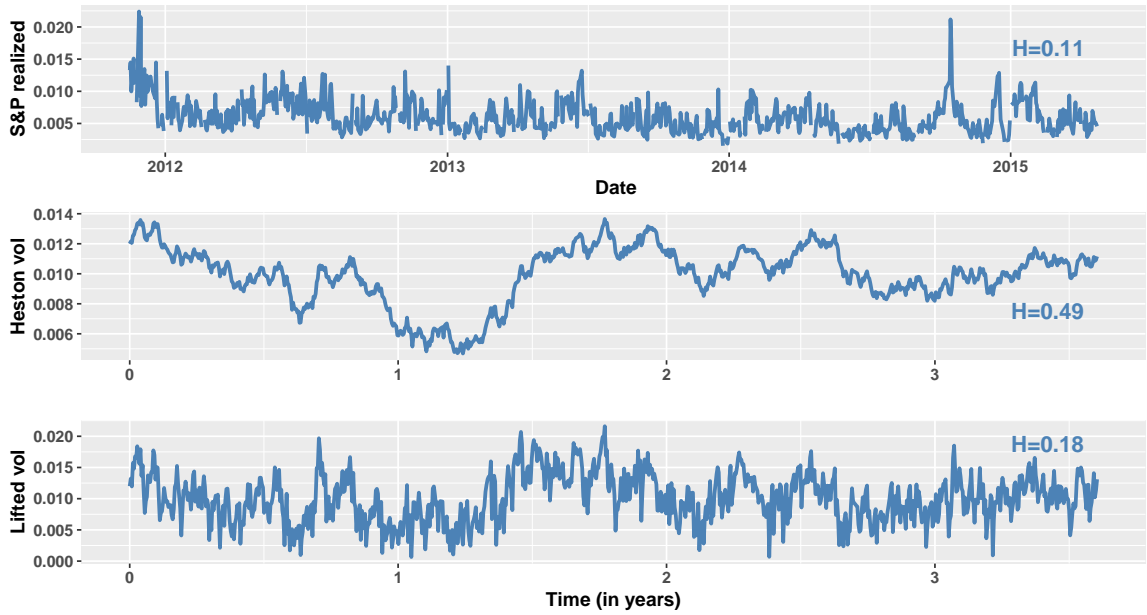

Figure 8: Estimated Hurst index of: the realized volatility of the S&P[(a)] (upper), a sample path of the volatility process in the Heston model (middle), and a sample path of the volatility process in the the lifted model with $H = 0.1$ (lower). The simulation is run with $N = 250$ time steps for each year.
[(a)]The realized volatility data series can be downloaded from https://realized.oxford-man.ox.ac.uk/.

5   Conclusion

We introduced *the lifted Heston model*, a conventional multi-factor stochastic volatility model, where the factors share the same one-dimensional Brownian motion but mean revert at different

speeds corresponding to different timescales. The model nests as extreme cases the standard Heston model (for $n = 1$ factor), and the rough Heston model (when n goes to infinity). Inspired by rough volatility models, we provided a good parametrization of the model reducing the number of parameters to calibrate: the model has only one additional effective parameter than the standard Heston model, independently of the number of factors. The first five parameters have the same interpretation as in the standard Heston model, whereas the additional one has a physical interpretation as it is linked to the regularity of the sample paths and the explosion of the at-the-money skew.

This sheds some new light on the reason behind the remarkable fits of rough volatility models. Indeed, a rough variance process can be seen as a superposition of infinitely many factors sharing the same one-dimensional Brownian motion but mean reverting at different speeds ranging from 0^+ to ∞. Each factor corresponds to a certain timescale. Therefore, time multi-scaling is naturally encoded in rough volatility models, which explains why these models are able to jointly handle different maturities in a satisfactory fashion.[9]

Finally, Table 7 below compares the characteristics of the three different models. As it can be seen, the *lifted Heston model* possesses an appealing trade-off between flexibility and tractability!

	Stochastic volatility models		
Characteristics	Heston	Rough Heston	Lifted Heston
Markovian	✓	✗	✓
Semimartingale	✓	✗	✓
Simulation	Fast	Slow	Fast
Affine Volterra process	✓	✓	✓
Characteristic function	Closed	Fractional Riccati	n-Riccati
Calibration	Fast	Slower	20x rough[(a)]
Fit short maturities	✗	✓	✓
Estimated daily regularity	$H \approx 0.5$	$H \approx 0.1$	$H \approx 0.2$

Table 7: Summary of the characteristics of the different models. [(a)]for $n = 20$.

A   Existence and uniqueness

In the sequel, the symbol $*$ stands for the convolution operation, that is $(f * \mu)(t) = \int_0^t f(t - s)\mu(ds)$ for any suitable function f and measure μ. For a right-continuous function f of locally bounded variation, we denote by df the measure induced by its distributional derivative, that is $f(t) = f(0) + \int_{(0,t]} df(s)$.

We provide in this appendix the strong existence and uniqueness of (2.1)-(2.3), for a fixed $n \in \mathbb{N}$. We start by noticing that (2.1) is equivalent to

$$S_t^n = \mathcal{E}\left(\int_0^t V_s^n dB_s\right), \quad t \geq 0,$$

---

[9]Multiple timescales in the volatility process have been identified in the literature, see for instance [20, Section 3.4].

where $\mathcal{E}$ is the Doléans-Dade exponential. Therefore, it suffices to prove the existence and uniqueness of (2.2)-(2.3). Formally, starting from a solution to (2.2)-(2.3), the variation of constants formula on (2.3) yields

$$U_t^{n,i} = \int_0^t e^{-x_i^n(t-s)} \left(-\lambda V_s^n ds + \nu \sqrt{V_s^n} dW_s \right), \quad i = 1, \ldots, n, \tag{A.1}$$

so that (2.2) reads

$$V_t^n = g_0^n(t) + \int_0^t K^n(t-s) \left(-\lambda V_s^n ds + \nu \sqrt{V_s^n} dW_s \right), \tag{A.2}$$

where K^n is the following completely monotone[10] kernel

$$K^n(t) = \sum_{i=1}^n c_i^n e^{-x_i^n t}, \quad t \geq 0. \tag{A.3}$$

Whence, if one proves the uniqueness of (A.2), then, uniqueness of (2.3) follows by virtue of (A.1). Conversely, if one proves the existence of a nonnegative solution V^n to (A.2), then, one can define $(U^{n,i})_{1 \leq i \leq n}$ as in (A.1), showing that $(V^n, (U^{n,i})_{1 \leq i \leq n})$ is a solution to (2.2)-(2.3). Therefore, the problem is reduced to proving the existence and uniqueness for the stochastic Volterra equation (A.2).

In [2], the existence of a nonnegative solution to (A.2) is proved, provided the initial input curve g_0^n satisfies a certain 'monotonicity' condition. This condition is related to the resolvent of the first kind L^n of the kernel (A.3), which is defined as the unique measure satisfying

$$\int_0^t K^n(t-s) L^n(ds) = 1, \quad t \geq 0.^{[11]}$$

More precisely, denoting by Δ_h the semigroup of right shifts acting on continuous functions, i.e. $\Delta_h f = f(h + \cdot)$ for $h \geq 0$, g_0^n should satisfy

$$\Delta_h g_0^n - (\Delta_h K^n * L^n)(0) g_0^n - d(\Delta_h K^n * L^n) * g_0^n \geq 0, \quad h \geq 0,^{[12]} \tag{A.4}$$

leading to the following definition of the set $\mathcal{G}_{K^n}$ of admissible input curves:

$$\mathcal{G}_{K^n} = \{ g_0^n \text{ Hölder continuous of any order less than } 1/2, \text{ satisfying (A.4) and } g_0^n(0) \geq 0 \}.$$

It is shown in [2, Example 2.2] that the two specifications of input curves (2.4)-(2.5) provided earlier satisfy (A.4).

We now provide the rigorous existence and uniqueness result.

Theorem A.1 (Existence and uniqueness). *Fix $n \in \mathbb{N}$, $S_0^n > 0$ and assume that $g_0^n \in \mathcal{G}_{K^n}$. Then, the stochastic differential equation (2.1)-(2.3) has a unique continuous strong solution $(S^n, V^n, (U^{n,i})_{1 \leq i \leq n})$ such that $V_t^n \geq 0$, for all $t \geq 0$, almost surely. Further, the process S^n is a martingale.*

---

[10] A function f is said to be completely monotone, if it is infinitely differentiable on $(0, \infty)$ such that $(-1)^p f^{(p)} \geq 0$, for all $p \in \mathbb{N}$.

[11] The existence of L^n is ensured by the complete monoticity of K^n, see [24, Theorem 5.5.4].

[12] One can show that $\Delta_h K^n * L^n$ is right-continuous and of locally bounded variation, thus the associated measure $d(\Delta_h K^n * L^n)$ is well defined.

Proof. By virtue of the variation of constants formula on the factors, the lifted Heston model is equivalent to a Volterra Heston model in the sense of [2] of the form

$$dS_t^n = S_t^n \sqrt{V_t^n} dB_t, \quad S_0^n > 0, \tag{A.5}$$

$$V_t^n = g_0^n(t) + \int_0^t K^n(t-s)\left(-\lambda V_s^n ds + \nu\sqrt{V_s^n}dW_s\right), \tag{A.6}$$

with K^n given by (A.3). Since K^n is locally Lipschitz and completely monotone, the assumptions of [2, Theorem 2.1] are met. Consequently, the stochastic Volterra equation (A.5)-(A.6) has a unique $\mathbb{R}_+^2$-valued weak continuous solution (S^n, V^n) on some filtered probability space $(\Omega^n, \mathcal{F}^n, (\mathcal{F}_t^n)_{t\geq 0}, \mathbb{Q}^n)$ for any initial condition $S_0^n > 0$ and admissible input curve $g_0^n \in \mathcal{G}_{K^n}$. Moreover, since K^n is differentiable, strong uniqueness is ensured by [1, Proposition B.3]. The claimed existence and uniqueness statement now follows from (A.1). Finally, the martingality of S^n follows along the lines of [3, Theorem 7.1(iii)]. $\qquad\square$

A.1 The rough limiting model

We now discuss the convergence of the *lifted Heston model* towards the rough Heston model (1.3)-(1.4), as the number of factors goes to infinity, we refer to [1] for more details. We fix $H \in (0, 1/2)$ and we denote by $K_H : t \to t^{H-\frac{1}{2}}/\Gamma(H+1/2)$ the fractional kernel of the rough Heston model appearing in (1.4). The kernel K_H can be re-expressed as a Laplace function

$$K_H(t) = \int_0^\infty e^{-xt}\mu(dx), \quad t \geq 0,$$

with $\mu(dx) = \frac{x^{-\alpha}}{\Gamma(\alpha)\Gamma(1-\alpha)}$ and $\alpha = H + 1/2$. On the one hand, for a fixed n, the parametrization (3.1) is linked to μ as follows:

$$c_i^n = \int_{\eta_{i-1}^n}^{\eta_i^n} \mu(dx), \quad x_i^n = \frac{1}{c_i^n}\int_{\eta_{i-1}^n}^{\eta_i^n} \mu(dx), \quad i = 1, \ldots, n, \tag{A.7}$$

where $\eta_i^n = r_n^{i-n/2}$, for $i = 0, \ldots, n$. We will show that, under (3.2),

$$K^n \to K_H, \quad \text{as } n \text{ goes to infinity}, \quad \text{in the } L^2 \text{ sense}. \tag{A.8}$$

On the other hand, for each $n \in \mathbb{N}$, we have proved the existence of a solution to (A.2). One would therefore expect from (A.8) the convergence of the sequence of solutions of (A.6) towards the solution of (1.4). This is indeed the case, as illustrated by the following theorem, which adapts [1, Theorem 3.5] to the geometric partition.

Theorem A.2 (Convergence towards the rough Heston model)**.** *Consider a sequence $(r_n)_{n\geq 1}$ satisfying (3.2), and set g_0^n as in (2.6) and $(c_i^n, x_i^n)_{1\leq i\leq n}$ as in (3.1), for every even $n = 2p$, with $p \geq 1$. Assume $S_0^n = S_0$, for all n, then, the sequence of solutions $(S^n, V^n)_{n=2p, p\geq 1}$ to (2.1)-(2.2) converges weakly, on the space of continuous functions on $[0, T]$ endowed with the uniform topology, towards the rough Heston model (1.3)-(1.4), for any $T > 0$.*

We will only sketch the proof for the L^2 convergence of the kernels (A.8), in order to highlight the small adjustments that one needs to make to the proof of [1, Theorem 3.5]. Indeed, since $\eta_0^n \neq 0$ in our case, [1, Theorem 3.5] cannot be directly applied, compare with [1, Assumption 3.1] where the left-end point of the partition is zero. The following lemma adapts [1, Proposition 3.3] to the geometric partition. The rest of the proof of Theorem A.2 follows along the lines of [1, Theorem 3.5] by making the same small adjustments highlighted below, mainly to treat the integral chunk between $[0, \eta_0^n]$.

Lemma A.3 (Convergence of K^n towards K_H). *Let $(r_n)_{n\geq 1}$ as in (3.2), and $(c_i^n, x_i^n)_{1\leq i\leq n}$ given by (3.1). Define K^n by (A.3), then,*

$$\|K^n - K_H\|_{L^2(0,T)} \to 0, \quad as \ n \to \infty, \tag{A.9}$$

for all $T > 0$.

Proof. Set $\eta_i^n = r_n^{i-n/2}$, for $i = 0, \ldots, n$. Using (A.7), we start by decomposing $(K_H - K^n)$ as follows

$$
\begin{aligned}
K_H - K^n &= \int_0^\infty e^{-x(\cdot)} \mu(dx) - \sum_{i=1}^n c_i^n e^{-x_i^n(\cdot)} \\
&= \int_0^{\eta_0^n} e^{-x(\cdot)} \mu(dx) + \left(\sum_{i=1}^n \int_{\eta_{i-1}^n}^{\eta_i^n} \left(e^{-x(\cdot)} - e^{-x_i^n(\cdot)} \right) \mu(dx) \right) + \int_{\eta_n^n}^\infty e^{-x(\cdot)} \mu(dx) \\
&:= J_1^n + J_2^n + J_3^n,
\end{aligned}
$$

so that

$$\|K_H - K^n\|_{L^2(0,T)} \leq I_1^n + I_2^n + I_3^n,$$

with $I_k^n = \|J_k^n\|_{L^2(0,T)}$, for $k = 1, 2, 3$. We now prove that each $I_k^n \to 0$, as n tends to ∞. Relying on a second order Taylor expansion, along the lines of the proof of [12, Proposition 7.1], we get the following bound

$$\left| \int_{\eta_{i-1}^n}^{\eta_i^n} \left(e^{-xt} - e^{-x_i^n t} \right) \mu(dx) \right| \leq C t^2 r_n^{1/2} (r_n - 1)^2 \int_{\eta_{i-1}^n}^{\eta_i^n} (1 \wedge x^{-1/2}) \mu(dx), \quad t \leq T,$$

for all $i = 1, \ldots, n$, where C is a constant independent of n, i and t. Summation over $i = 1, \ldots, n$ leads to

$$I_2^n \leq C \frac{T^{5/2}}{\sqrt{5}} r_n^{1/2} (r_n - 1)^2 \int_0^\infty (1 \wedge x^{-1/2}) \mu(dx),$$

so that $I_2^n \to 0$, as $n \to \infty$, by virtue of the first condition in (3.2). On another note,

$$I_1^n \leq \int_0^{\eta_0^n} \mu(dx) = \frac{(\eta_0^n)^{1-\alpha}}{\Gamma(\alpha)\Gamma(2-\alpha)} = \frac{r_n^{-(1-\alpha)n/2}}{\Gamma(\alpha)\Gamma(2-\alpha)} \to 0, \quad \text{when } n \to \infty,$$

thanks to the second condition in (3.2). Similarly,

$$I_3^n \leq \int_{\eta_n^n}^\infty \sqrt{\frac{1 - e^{-2xT}}{2x}} \mu(dx) \leq \frac{r_n^{(1/2-\alpha)n/2}}{\Gamma(\alpha)\Gamma(1-\alpha)(1/2-\alpha)} \to 0, \quad \text{when } n \to \infty.$$

Combining the above leads to (A.9). $\qquad\square$

B   The full Fourier-Laplace transform

We provide the full Fourier-Laplace transform for the joint process $X^n := (\log S^n, V^n)$ extending (2.9). The formula can be used to price path-dependent options on the stock price S^n and the variance process V^n.

Once again, this is a particular case of [2, Section 4], by observing that K^n defined in (A.3) is the Laplace transform of the following nonnegative measure

$$\mu^n(dx) = \sum_{i=1}^n c_i^n \delta_{x_i^n}(dx).$$

20

Fix row vectors $u = (u_1, u_2) \in \mathbb{C}^2$ and $f \in L^1_{\text{loc}}(\mathbb{R}_+, (\mathbb{C}^2))$ such that

$$\text{Re}\,(u_1 + 1 * f_1) \in [0, 1], \quad \text{Re}\,u_2 \leq 0 \quad \text{and} \quad \text{Re}\,f_2 \leq 0,$$

then, it follows from [2, Remark 4.3] with $\mu = \sum_{i=1}^n c_i^n \delta_{x_i^n}$ that the Fourier-Laplace transform of $X^n = (\log S^n, V^n)$ is exponentially affine with respect to the family $(U^{n,i})_{1 \leq i \leq n}$,

$$\mathbb{E}\left[\exp\left(uX_T^n + (f * X^n)_T\right) \;\Big|\; \mathcal{F}_t\right] = \exp\left(\phi^n(t, T) + \psi_1(T - t)\log S_t^n + \sum_{i=1}^n c_i^n \psi_2^{n,i}(T - t)U_t^{n,i}\right),$$

for all $t \leq T$, where $(\psi_1, (\psi_2^{n,i})_{1 \leq i \leq n})$ are the unique solutions of the following system of Riccati ordinary differential equations

$$\psi_1 = u_1 + 1 * f_1,$$

$$(\psi_2^{n,i})' = -x_i^n \psi_2^{n,i} + F\left(\psi_1, \sum_{j=1}^n c_j^n \psi_2^{n,j}\right), \quad \psi_2^{n,i}(0) = u_2, \quad i = 1, \ldots, n,$$

with

$$F(\psi_1, \psi_2) = f_2 + \frac{1}{2}\left(\psi_1^2 - \psi_1\right) + (\rho\nu\psi_1 - \lambda)\psi_2 + \frac{\nu^2}{2}\psi_2^2$$

and

$$\phi^n(t, T) = u_2 g_0^n(T) + \int_0^{T-t} F\left(\psi_1, \sum_{i=1}^n c_i^n \psi_2^{n,i}(s)\right) g_0^n(T - s)ds + \int_0^t f(T - s)X_s ds, \quad t \leq T.$$

C   Discretization schemes

C.1   Riccati equations

The aim of this section is to design an approximation scheme of the n-dimensional Riccati system of equations (2.10). In order to gain some insights, consider first the case where $F \equiv 0$ so that (2.10) reduces to

$$(\psi^{n,i})' = -x_i^n \psi^{n,i}, \quad i = 1, \ldots, n, \tag{C.1}$$

and the solution is given by

$$\psi^{n,i}(t) = \psi^{n,i}(0)e^{-x_i^n t}, \quad i = 1, \ldots, n.$$

One could start with an explicit Euler scheme for (C.1), that is

$$\hat{\psi}_{t_{k+1}}^{n,i} = \hat{\psi}_{t_k}^{n,i} - x_i^n \Delta t \hat{\psi}_{t_k}^{n,i} = (1 - x_i^n \Delta t)\,\hat{\psi}_{t_k}^{n,i}, \quad i = 1, \ldots, n,$$

for a regular time grid $t_k = (kT)/N$ for all $k = 1, \ldots, N$, where T is the terminal time, N the number of time steps and $\Delta t = T/N$. A sufficient condition for the stability of the scheme reads

$$\Delta t \leq \min_{1 \leq i \leq n} \frac{1}{x_i^n}.$$

Recall from (3.1) that x_n^n grows very large as n increases. For instance, for $n = 20$, $r_{20} = 2.5$ and $H = 0.1$, $x_n^n = 6417.74$. Consequently, if one needs to ensure the stability of the explicit scheme, one needs a very large number of time steps N. In contrast, the implicit Euler scheme

$$\hat{\psi}_{t_{k+1}}^{n,i} = \hat{\psi}_{t_k}^{n,i} - x_i^n \Delta t \, \hat{\psi}_{t_{k+1}}^{n,i}, \quad i = 1, \ldots, n,$$

is stable for any number of time steps N and reads

$$\hat{\psi}_{t_{k+1}}^{n,i} = \frac{1}{1 + x_i^n \Delta t} \hat{\psi}_{t_k}^{n,i}, \quad i = 1, \ldots, n.$$

For this reason, we consider the following explicit-implicit discretization scheme of the n-dimensional Riccati system of equations (2.10)

$$\hat{\psi}_0^{n,i} = 0, \quad \hat{\psi}_{t_{k+1}}^{n,i} = \frac{1}{1 + x_i^n \Delta t} \left(\hat{\psi}_{t_k}^{n,i} + \Delta t \, F\left(u, \sum_{j=1}^n c_j^n \hat{\psi}_{t_k}^{n,j}\right) \right), \quad i = 1, \ldots, n, \qquad \text{(C.2)}$$

for a regular time grid $t_k = k\Delta t$ for all $k = 1, \ldots, N$, with time step size $\Delta t = T/N$, terminal time T and number of time steps N. Alternatively, one could also consider the exponential scheme for the Riccati equations by replacing the term $1/(1 + x_i^n \Delta t)$ with $e^{-x_i^n \Delta t}$. One can also combine more involved discretization schemes for the explicit part involving the quadratic function F, for instance higher order Runge-Kutta methods can be used, see [26].

C.2 Stochastic process

Similarly, we suggest to consider the following modified explicit-implicit scheme for the variance process V^n:

$$\hat{V}_{t_k}^n = g_0^n(t_k) + \sum_{i=1}^n c_i^n \hat{U}_{t_k}^{n,i}, \quad \hat{U}_0^{n,i} = 0, \qquad \text{(C.3)}$$

$$\hat{U}_{t_{k+1}}^{n,i} = \frac{1}{1 + x_i^n \Delta t} \left(\hat{U}_{t_k}^{n,i} - \lambda \hat{V}_{t_k}^n \Delta t + \nu \sqrt{\left(\hat{V}_{t_k}^n\right)^+} \left(W_{t_{k+1}} - W_{t_k}\right) \right), \quad i = 1, \ldots, n, \qquad \text{(C.4)}$$

for a regular time grid $t_k = k\Delta t$, $k = 1 \ldots N$, $\Delta t = T/N$ and $(W_{t_{k+1}} - W_{t_k}) \sim \mathcal{N}(0, \Delta t)$. Notice that we take the positive part $(\cdot)^+$ since the simulated process can become negative. Once there, simulating the spot-price process S^n is straightforward. We leave the theoretical study of convergence and stability for future work. Numerically, the scheme seems stable. Alternatively, one could also consider the exponential scheme for the stochastic process by replacing the term $1/(1 + x_i^n \Delta t)$ with $e^{-x_i^n \Delta t}$. As a final remark, one notices that (C.3)-(C.4) corresponds to the space-time discretization of the integro-differential stochastic partial differential equation (3.8)-(3.9). This is illustrated on Figure 9 below.
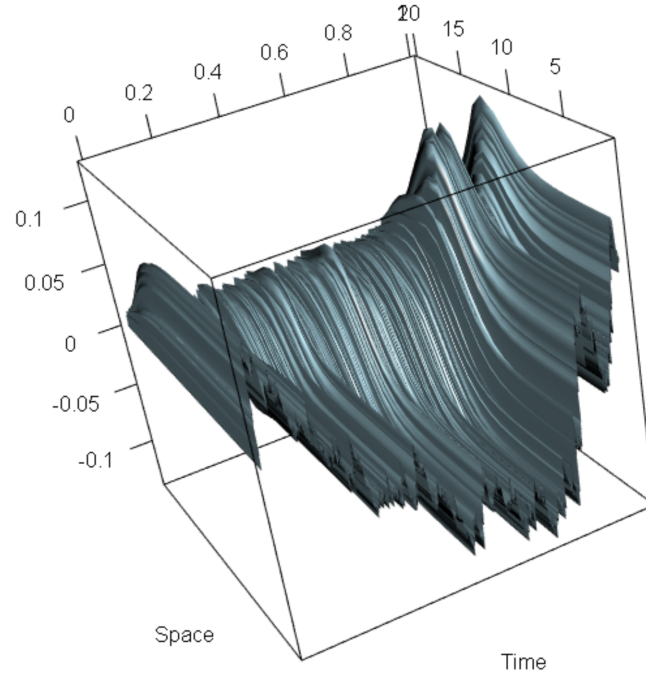
22

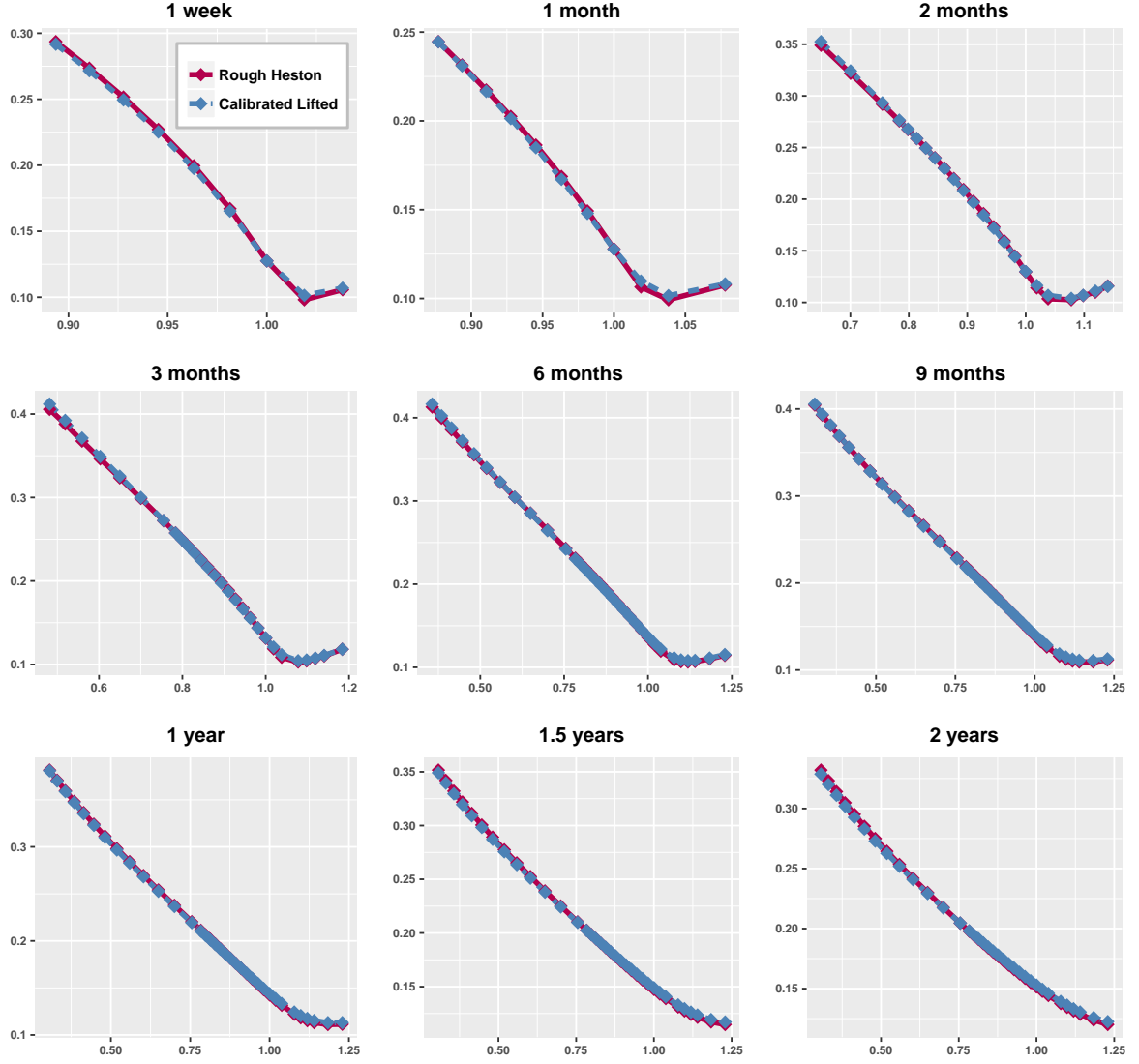Figure 9: Simulated path of the stochastic partial differential equation (3.8)-(3.9) by using the scheme (C.3)-(C.4).

Figure 10: Implied volatility surface of the rough Heston model $\sigma_\infty(K, T; \Theta_0)$ of (3.6) (red) and the calibrated lifted Heston model $\sigma_{20}(K, T; r_{20} = 2.5, \hat{\Theta}_0)$ of Table 4 (blue) for maturities ranging from 1 week to 2 years (MSE = 4.01e-07).
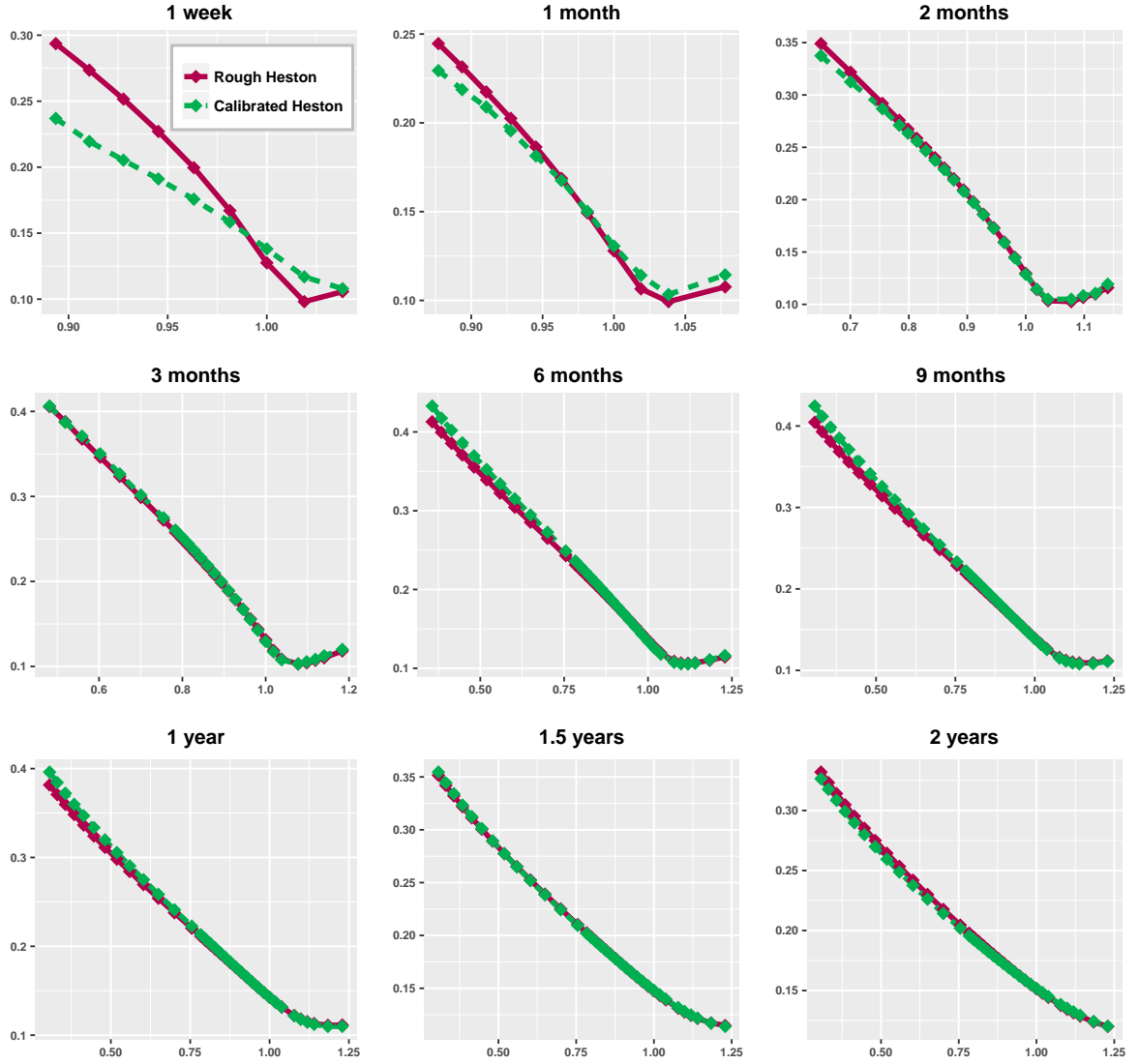
Figure 11: Implied volatility surface of the rough Heston model $\sigma_\infty(K, T; \Theta_0)$ (red) and the calibrated Heston model of Table 6 (green) for maturities ranging from 1 week to 2 years (MSE = 2.06e-03).
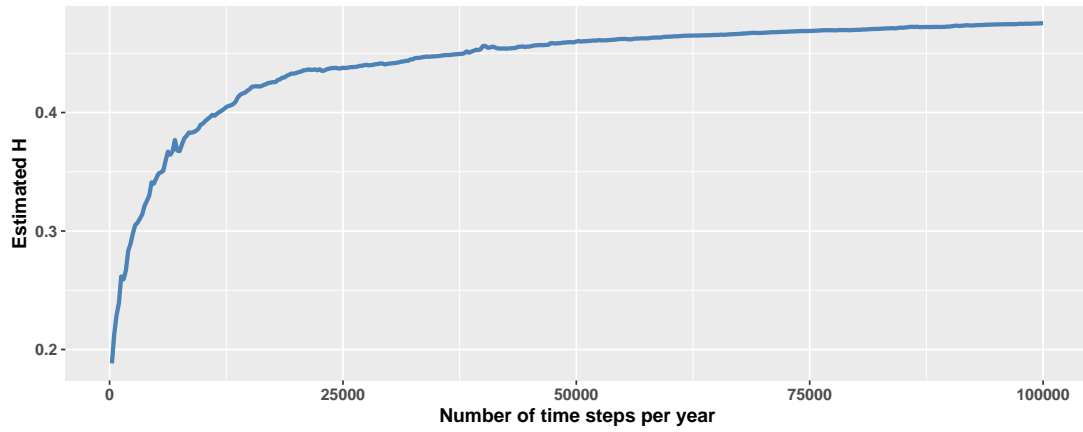
Figure 12: Estimated Hurst index of the simulated sample path of the volatility process in the lifted model with $H = 0.1$ as a function of the number of time steps per year.

ν	ρ	H	MSE
0.22	-0.67	0.09	3.63e-06
0.14	-0.54	0.19	5.34e-06
0.35	-0.65	0.19	8.17e-06
0.14	-0.83	0.06	9.74e-05
0.22	-0.59	0.15	4.60e-06
0.37	-0.50	0.12	4.55e-06
0.40	-0.53	0.11	4.56e-06
0.34	-0.85	0.08	3.45e-04
0.22	-0.89	0.09	1.25e-04
0.44	-0.76	0.11	2.79e-04
0.32	-0.70	0.12	4.56e-06
0.42	-0.63	0.08	5.22e-06
0.10	-0.61	0.17	3.69e-06
0.42	-0.64	0.11	4.81e-06
0.30	-0.69	0.17	5.96e-06
0.06	-0.71	0.17	2.98e-06
0.36	-0.71	0.16	6.14e-06
0.25	-0.80	0.18	1.63e-04
0.09	-0.77	0.06	2.87e-06
0.35	-0.74	0.13	1.44e-04

Table 8: **Robustness of** $r_{20} = 2.5$**:** First 20 values of the simulated parameters and the corresponding mean squared error between the implied volatility surface of the lifted model $\sigma_{20}(K, T; 2.5, \Theta_k)$ and the rough model $\sigma_\infty(K, T; \Theta_k)$, for $k = 1, \ldots, 20$.

References

[1] Eduardo Abi Jaber and Omar El Euch. Multi-factor approximation of rough volatility models. *arXiv preprint*, 2018.

[2] Eduardo Abi Jaber and Omar El Euch. Markovian structure of the Volterra Heston model. *arXiv preprint arXiv:1803.00477*, 2018.

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

[4] Aurélien Alfonsi. *Affine diffusions and related processes: simulation, theory and applications*, volume 6 of *Bocconi & Springer Series*. Springer, Cham; Bocconi University Press, Milan, 2015. ISBN 978-3-319-05220-5; 978-3-319-05221-2.

[5] Christian Bayer, Peter Friz, and Jim Gatheral. Pricing under rough volatility. *Quantitative Finance*, 16(6):887–904, 2016. doi: 10.1080/14697688.2015.1099717. URL http://dx.doi.org/10.1080/14697688.2015.1099717.

[6] Mikkel Bennedsen, Asger Lunde, and Mikko S. Pakkanen. Decoupling the short- and long-term behavior of stochastic volatility. *arXiv preprint arXiv:1610.00332*, 2016.

[7] Lorenzo Bergomi. Smile dynamics II. *Risk*, 18:67–73, 2005.

[8] Philippe Carmona, Laure Coutin, and Gérard Montseny. Approximation of some Gaussian processes. *Stat. Inference Stoch. Process.*, 3(1-2):161–171, 2000. ISSN 1387-0874. doi: 10.1023/A:1009999518898. URL http://dx.doi.org/10.1023/A:1009999518898. 19th "Rencontres Franco-Belges de Statisticiens" (Marseille, 1998).

[9] Peter Carr and Dilip Madan. Option valuation using the fast Fourier transform. *Journal of Computational Finance*, 2(4):61–73, 1999.

[10] Peter Christoffersen, Steven Heston, and Kris Jacobs. The shape and term structure of the index option smirk: Why multifactor stochastic volatility models work so well. *Management Science*, 55(12):1914–1932, 2009.

[11] Rama Cont and Peter Tankov. *Financial modelling with jump processes*, volume 2. CRC press, 2003.

[12] Laure Coutin and Monique Pontier. Approximation of the fractional brownian sheet via ornstein-uhlenbeck sheet. *ESAIM: Probability and Statistics*, 11:115–146, 2007.

[13] Christa Cuchiero and Josef Teichmann. Generalized Feller processes and Markovian lifts of stochastic Volterra processes: the affine case. *arXiv preprint arXiv:1804.10450*, 2018.

[14] José Da Fonseca, Martino Grasselli, and Claudio Tebaldi. A multifactor volatility heston model. *Quantitative Finance*, 8(6):591–604, 2008.

[15] Kai Diethelm, Neville J Ford, and Alan D Freed. A predictor-corrector approach for the numerical solution of fractional differential equations. *Nonlinear Dynamics*, 29(1-4):3–22, 2002.

[16] Omar El Euch and Mathieu Rosenbaum. The characteristic function of rough Heston models. *Mathematical Finance*, 2016.

[17] Omar El Euch, Jim Gatheral, and Mathieu Rosenbaum. Roughening heston. 2018.

[18] Fang Fang and Cornelis Oosterlee. A novel pricing method for european options based on fourier-cosine series expansions. *SIAM Journal on Scientific Computing*, 31(2):826–848, 2008.

[19] Jean-Pierre Fouque and Matthew Lorig. A fast mean-reverting correction to heston's stochastic volatility model. *SIAM Journal on Financial Mathematics*, 2(1):221–254, 2011.

[20] Jean-Pierre Fouque, George Papanicolaou, Ronnie Sircar, and Knut Sølna. *Multiscale stochastic volatility for equity, interest rate, and credit derivatives*. Cambridge University Press, 2011.

[21] Masaaki Fukasawa. Short-time at-the-money skew and rough fractional volatility. *Quantitative Finance*, 17(2):189–198, 2017.

[22] Jim Gatheral. *The volatility surface: a practitioner's guide*, volume 357. John Wiley & Sons, 2011.

[23] Jim Gatheral, Thibault Jaisson, and Mathieu Rosenbaum. Volatility is rough. *Quantitative Finance*, pages 1–17, 2018.

[24] Gustaf Gripenberg, Stig-Olof Londen, and Olof Staffans. *Volterra integral and functional equations*, volume 34 of *Encyclopedia of Mathematics and its Applications*. Cambridge University Press, Cambridge, 1990. ISBN 0-521-37289-5. doi: 10.1017/CBO9780511662805. URL http://dx.doi.org/10.1017/CBO9780511662805.

[25] Steven Heston. A closed-form solution for options with stochastic volatility with applications to bond and currency options. *Review of Financial Studies*, 6(2):327–343, 1993.

[26] John Denholm Lambert. *Numerical methods for ordinary differential systems: the initial value problem*. John Wiley & Sons, Inc., 1991.