Received: 9 April 2019 | Accepted: 14 September 2020

DOI: 10.1111/mafi.12290

ORIGINAL ARTICLE

MATHEMATICAL FINANCE An International Journal of Mathematics, Statistics and Financial Economics WILEY

# Small-time, large-time, and $H \to 0$ asymptotics for the Rough Heston model

**Martin Forde**[2] | **Stefan Gerhold**[1] | **Benjamin Smith**[2]

[1] TU Wien, Financial and Actuarial Mathematics, Vienna, Austria

[2] Department of Mathematics, King's College London, Strand, London, UK

**Correspondence**
Stefan Gerhold, TU Wien, Financial and Actuarial Mathematics, Wiedner Hauptstraße 8/105-1, A-1040 Vienna, Austria.
Email: sgerhold@fam.tuwien.ac.at

**Funding information**
Austrian Science Fund, Grant/Award Number: P 30750

**Abstract**

We characterize the behavior of the Rough Heston model introduced by Jaisson and Rosenbaum (2016, *Ann. Appl. Probab.*, *26*, 2860–2882) in the small-time, large-time, and $\alpha \to \frac{1}{2}$ (i.e., $H \to 0$) limits. We show that the short-maturity smile scales in qualitatively the same way as a general rough stochastic volatility model , and the rate function is equal to the Fenchel–Legendre transform of a simple transformation of the solution to the same Volterra integral equation (VIE) that appears in El Euch and Rosenbaum (2019, *Math. Financ.*, *29*, 3–38), but with the drift and mean reversion terms removed. The solution to this VIE satisfies a space–time scaling property which means we only need to solve this equation for the moment values of $p = 1$ and $p = -1$ so the rate function can be efficiently computed using an Adams scheme or a power series, and we compute a power series in the log-moneyness variable for the asymptotic implied volatility which yields tractable expressions for the implied vol skew and convexity which is useful for calibration purposes. We later derive a formal saddle point approximation for call options in the Forde and Zhang (2017) large deviations regime which goes to higher order than previous works for rough models. Our higher-order expansion captures the effect of both drift terms, and at leading order is of qualitatively the same form as the higher-order expansion for a general model which appears in Friz et al. (2018, *Math. Financ.*, *28*, 962–988). The limiting asymptotic smile in the large-maturity regime is obtained via a stability analysis of the fixed points of the VIE, and is the same as for the standard Heston model in Forde and Jacquier (2011, *Finance Stoch.*, *15*, 755–780). Finally, using Lévy's convergence theorem, we show that the log stock price $X_t$ tends weakly to a nonsymmetric random variable $X_t^{(\frac{1}{2})}$ as $\alpha \to \frac{1}{2}$ (i.e., $H \to 0$) whose moment generating function (MGF) is also the solution to the Rough Heston VIE with $\alpha = \frac{1}{2}$, and we show that $X_t^{(\frac{1}{2})}/\sqrt{t}$ tends weakly to a nonsymmetric random variable as $t \to 0$, which leads to a nonflat nonsymmetric asymptotic smile in the Edgeworth regime, where the log-moneyness $z = k\sqrt{t}$ as $t \to 0$, and we compute this asymptotic smile numerically. We also show that the third moment of the log stock price tends to a finite constant as $H \to 0$ (in contrast to the Rough Bergomi model discussed in Forde et al. (2020, Preprint) where the skew flattens or blows up) and the $V$ process converges on pathspace to a random tempered distribution which has the same law as the $H = 0$ hyper-rough Heston model, discussed in Jusselin and Rosenbaum (2020, *Math. Finance*, *30*, 1309–1336) and Abi Jaber (2019, *Ann. Appl. Probab.*, *29*, 3155–3200).

**KEYWORDS**
asymptotics, implied volatility, integral equations, random fields, Rough Heston model, Rough volatility

## 1 | INTRODUCTION

Jaisson and Rosenbaum (2016) introduced the Rough Heston stochastic volatility model and show that the model arises naturally as the large-time limit of a high-frequency market microstructure model driven by two nearly unstable self-exciting Poisson processes (otherwise known as Hawkes process) with a Mittag–Leffler kernel which drives buy and sell orders (a Hawkes process is a generalized Poisson process where the intensity is itself stochastic and depends on the jump history via the kernel). The microstructure model captures the effects of endogeneity of the market, no-arbitrage, buying/selling asymmetry, and the presence of metaorders. El Euch and Rosenbaum (2019) show that the characteristic function of the log stock price for the Rough

Heston model is the solution to a fractional Ricatti equation which is nonlinear (see also El Euch, Fukasawa, and Rosenbaum, 2018; El Euch and Rosenbaum, 2018), and the variance curve for the model evolves as $d\xi_u(t) = \kappa(u-t)\sqrt{V_t}dW_t$, where $\kappa(t)$ is the kernel for the $V_t$ process itself multiplied by a *Mittag–Leffler* function (see Proposition 2.2 for a proof of this). Theorem 2.1 in El Euch and Rosenbaum (2018) shows that a Rough Heston model conditioned on its history up to some time is still a Rough Heston model, but with a time-dependent mean reversion level $\theta(t)$ which depends on the history of the $V$ process. Using Fréchet derivatives, El Euch and Rosenbaum (2018) also show that one can replicate a call option under the Rough Heston model if we assume the existence a tradeable variance swap, and the same type of analysis can be done for the Rough Bergomi model using the Clark–Ocone formula from Malliavin calculus. See also Dandapani, Jusselin, and Rosenbaum (2019) who introduce the super Rough Heston model to incorporate the strong Zumbach effect as the limit of a market microstructure model driven by quadratic Hawkes process (this model is no longer affine and thus not amenable to the Volterra integral equation [VIE] techniques in this paper).

Gatheral and Keller-Ressel (2019) consider the more general class of affine forward variance (AFV) models of the form $d\xi_u(t) = \kappa(u-t)\sqrt{V_t}dW_t$ (for which the Rough Heston model is a special case). They show that AFV models arise naturally as the weak limit of a so-called affine forward intensity (AFI) model, where order flow is driven by two generalized Hawkes-type process with an arbitrary jump size distribution, and we exogenously specify the evolution of the conditional expectation of the intensity at different maturities in the future, akin to a variance curve model. The weak limit here involves letting the jump size tends to zero as the jump intensity tends to infinity in a certain way, and one can argue that an AFI model is more realistic than the bivariate Hawkes model in El Euch and Rosenbaum (2019), since the latter only allows for jumps of a single magnitude (which correspond to buy/sell orders). Using martingale arguments (which do not require considering a Hawkes process as in the aforementioned El Euch & Rosenbaum articles) they show that the mgf of the log stock price for the affine variance model satisfies a convolution Ricatti equation, or equivalently is a nonlinear function of the solution to a VIE.

Gerhold, Gerstenecker, and Pinter (2019) use comparison principle arguments for VIEs to show that the moment explosion time for the Rough Heston model is finite if and only if it is finite for the standard Heston model. Gerhold et al. (2019) also establish upper and lower bounds for the explosion time, and show that the critical moments are finite for all maturities, and formally derive refined tail asymptotics for the Rough Heston model using Laplace's method. A recent talk by Keller-Ressel (joint work with Majid) states an alternate upper bound for the moment explosion time for the Rough Heston model, based on a comparison with a (deterministic) time change of the standard Heston model, which they claim is usually sharper than the bound in Gerhold et al. (2019).

Jacquier & Pannier (2020) compute a small-time large deviation principle (LDP) on pathspace for a more general class of stochastic Volterra models in the same spirit as the classical Freidlin–Wentzell LDP for small-noise diffusion. More specifically, for a simple Volterra system of the form

$$Y_t = Y_0 + \int_0^t K_2(t-s)\zeta(Y_s)dW_s, \tag{1}$$

we have the corresponding deterministic system:

$$Y_t = Y_0 + \int_0^t K_2(t-s)\zeta(Y_s)v_s ds,$$

where $v \in L^2([0,T])$. When $K_2(t) = const.t^{H-\frac{1}{2}}$ the right term is proportional to the $\alpha$th fractional integral of $\zeta v$ (where $\alpha = H + \frac{1}{2}$), and in this case Jacquier & Pannier (2020) show that $Y_{\varepsilon(.)}$ satisfies an LDP as $\varepsilon \to 0$ with rate function

$$I_Y(\varphi) = \frac{1}{2}const. \times \int_0^T \left(\frac{D^\alpha(\varphi(.) - \varphi(0))(t)}{\zeta(\varphi(t))}\right)^2 dt$$

(see proposition 4.3 in Jacquier & Pannier, 2020) in terms of the rate function of the underlying Brownian motion which is well known from Schilder's theorem (one can also add drift terms into (1) which will not affect $I_Y$). The corresponding LDP for the log stock price is then obtained using the usual contraction principle method, so the rate function has a variational representation, and does not involve VIEs.

Corollary 7.1 in Friz, Gassiat, and Pigato (2018) provides a sharp small-time expansion in the Forde and Zhang (2017) large deviations regime (valid for $x$-values in some interval) for a general class of Rough Stochastic volatility models using regularity structures, which provides the next order correction to the leading order behavior obtained in Forde and Zhang (2017), and some earlier intermediate results in Bayer, Friz, Gulisashvili, Horvath, and Stemper (2018). Forde, Smith, and Viitasaari (2019) derive formal small-time Edgeworth expansions for the Rough Heston model by solving a nested sequence of linear VIEs. The Edgeworth-regime implied vol expansions in El Euch, Fukasawa, Gatheral, and Rosenbaum (2019) and Forde et al. (2019) both include an additional $O(T^{2H})$ term, which itself contains an at-the-money, convexity, and higher-order correction term, which are important effects to capture for these approximations to be useful in practice.

In this paper, we establish small- and large-time large deviation principles for the Rough Heston model, via the solution to a VIE, and we translate these results into asymptotic estimates for call options and implied volatility. The solution to the VIE satisfies a certain scaling property which means we only have to solve the VIE for the moment values of $p = +1$ and $-1$, rather than solving an entire family of VIEs. Using the Lagrange inversion theorem, we also compute the first three terms in the power series for the asymptotic implied volatility $\hat{\sigma}(x)$. We later derive formal asymptotics for the small-time moderate deviations regime and a formal saddle point approximation for European call options in the original Forde and Zhang (2017) large deviations regime which goes to higher order than previous works for rough models, and captures the effect of the mean reversion term and the drift of the log stock price, and we discuss practical issues and limitations of this result. Our higher-order expansion is of qualitatively the same form as the higher-order expansion for a general model in theorem 6 in Friz, Gassiat, et al. (2018) (their expansion is not known to hold for large $x$-values since in their more general setup there are additional complications with focal points, proving nondegeneracy, etc.). For the large time, large log-moneyness regime, we show that the asymptotic smile is the same as for the standard Heston model as in Forde and Jacquier (2011a), and we briefly outline how one could go about computing the next order term using a saddle point approximation, in the same spirit as Forde, Jacquier, and Mijatovic (2011).

In the final section, using Lévy's convergence theorem and result from Gripenberg, Londen, and Staffans (1990) on the continuous dependence of VIE solutions as a function of a parameter in the VIE, we show that the log stock price $X_t$ (for $t$ fixed) tends weakly as $\alpha \to \frac{1}{2}$ to a random variable $X_t^{(\frac{1}{2})}$ whose mgf is also the solution to the Rough Heston VIE with $\alpha = \frac{1}{2}$ and whose law is nonsymmetric when $\rho \neq 0$. From this we show that $X_t^{(\frac{1}{2})}/\sqrt{t}$ tends weakly to a nonsymmetric random variable as $t \to 0$, which leads to a nontrivial asymptotic smile in the Edgeworth (or central limit theorem) regime, where the log-moneyness scale as $z = k\sqrt{t}$ as $t \to 0$. We also show that the third moment of the log stock price for the driftless version of the model tends to a finite constant as $H \to 0$ (in contrast to the Rough Bergomi model discussed in Forde et al. (2020) where the skew flattens or blows up depending on the vol-of-vol parameter $\gamma$) and using the expression in Abi Jaber, Larsson, and Pulido (2019) for $\mathbb{E}(e^{\int_0^T f(T-t)V_t dt})$, we show that $V$ converges to a random tempered distribution whose characteristic functional also satisfies a nonlinear VIE and (from theorem 2.5 in Abi Jaber, 2019) this tempered distribution has the same law as the $H = 0$ hyper-rough Heston model.

## 2 | ROUGH HESTON AND OTHER VARIANCE CURVE MODELS—BASIC PROPERTIES

In this section, we recall the definition and basic properties and origins of the Rough Heston model, and more general affine and nonaffine forward variance models. Most of the results in this section are given in various locations in El Euch and Rosenbaum (2018, 2019) and Gatheral and Keller-Ressel (2019), but for pedagogical purposes we found it instructive to collate them together in one place.

Let $(\Omega, \mathcal{F}, \mathbb{P})$ denote a probability space with filtration $(\mathcal{F}_t)_{t\geq 0}$ which satisfies the usual conditions, and consider the Rough Heston model for a log stock price process $X_t$ introduced in Jaisson and Rosenbaum (2016):

$$dX_t = -\frac{1}{2}V_t dt + \sqrt{V_t}dB_t$$

$$V_t = V_0 + \frac{1}{\Gamma(\alpha)}\int_0^t (t-s)^{\alpha-1}\lambda(\theta - V_s)ds + \frac{1}{\Gamma(\alpha)}\int_0^t (t-s)^{\alpha-1}\nu\sqrt{V_s}dW_s \tag{2}$$

for $\alpha \in \left(\frac{1}{2}, 1\right)$, $\theta > 0$, $\lambda \geq 0$, and $\nu > 0$, where $W, B$ are two $\mathcal{F}_t$-Brownian motions with correlation $\rho \in (-1, 1)$. We assume $X_0 = 0$ and zero interest rate without loss of generality, since the law of $X_t - X_0$ is independent of $X_0$.

### 2.1 | Computing $\mathbb{E}(V_t)$

**Proposition 2.1.**

$$\mathbb{E}(V_t) = V_0 - (V_0 - \theta)\int_0^t f^{\alpha,\lambda}(s)ds, \tag{3}$$

*where* $f^{\alpha,\lambda}(t) := \lambda t^{\alpha-1}E_{\alpha,\alpha}(-\lambda t^{\alpha})$, *and* $E_{\alpha,\beta}(z) := \sum_{n=0}^{\infty}\frac{z^n}{\Gamma(\alpha n+\beta)}$ *denotes the* Mittag–Leffler function.

*Proof.* (See also p. 7 in Gatheral and Keller-Ressel (2019), and proposition 3.1 in El Euch and Rosenbaum (2018) for an alternate proof). Let $r(t) = f^{\alpha,\lambda}(t)$. Taking expectations of (2) and using that the expectation of the stochastic integral term is zero, we see that

$$\mathbb{E}(V_t) = V_0 + \frac{1}{\Gamma(\alpha)} \int_0^t (t-s)^{\alpha-1} \lambda(\theta - \mathbb{E}(V_s))dt\,. \tag{4}$$

Let $k(t) := \frac{\lambda t^{\alpha-1}}{\Gamma(\alpha)}$ and $f(t) := \mathbb{E}(V_t) - \theta$. Then we can rewrite (4) as

$$f(t) = (V_0 - \theta) - k * f(t)\,, \tag{5}$$

where $*$ denotes convolution. Now define the resolvent $r(t)$ as the unique function which satisfies $r = k - k * r$. Then, we claim that

$$f(t) = (V_0 - \theta) - r * (V_0 - \theta).$$

To verify the claim, we substitute this expression into (5) to get:

$$\begin{aligned}(V_0 - \theta) - k * [(V_0 - \theta) - r * (V_0 - \theta)] &= (V_0 - \theta) - (V_0 - \theta) * (k - k * r)(t) \\ &= (V_0 - \theta) - (V_0 - \theta) * r(t)\end{aligned}$$

so $(V_0 - \theta) - k * f(t) = (V_0 - \theta) - (V_0 - \theta) * r(t) = f(t)$, which is precisely the integral equation we are trying to solve. Taking Laplace transform of both sides of $k - k * r = r$ we obtain $\hat{r} = \hat{k} - \hat{k}\hat{r}$, which we can rearrange as

$$\hat{r} = \frac{\hat{k}}{1 + \hat{k}} = \frac{\lambda z^{-\alpha}}{1 + \lambda z^{-\alpha}} = \frac{\lambda}{z^\alpha + \lambda}$$

and the inverse Laplace transform of $\hat{r}$ is $r(t) = \lambda t^{\alpha-1} E_{\alpha,\alpha}(-\lambda t^\alpha)$. □

## 2.2 | Computing $\mathbb{E}(V_u|\mathcal{F}_t)$

Now let $\xi_t(u) := \mathbb{E}(V_u|\mathcal{F}_t)$. Then $\xi_t(u)$ is an $\mathcal{F}_t$-martingale, and

$$\xi_t(u) = V_0 + \frac{1}{\Gamma(\alpha)} \int_0^u (u-s)^{\alpha-1} \lambda(\theta - \mathbb{E}(V_s|\mathcal{F}_t)ds + \frac{1}{\Gamma(\alpha)} \int_0^t (u-s)^{\alpha-1} \nu \sqrt{V_s} dW_s.$$

If $\lambda = 0$, we can rewrite this expression as

$$d\xi_t(u) = \frac{1}{\Gamma(\alpha)} (u-t)^{\alpha-1} \sqrt{V_t} dW_t.$$

**Proposition 2.2.** *(see El Euch & Rosenbaum, 2019). For $\lambda > 0$,*

$$d\xi_t(u) = \kappa(u-t)\sqrt{V_t}dW_t = \kappa(u-t)\sqrt{\xi_t(t)}dW_t, \tag{6}$$

*where $\kappa$ is the inverse Laplace transform of $\hat{\kappa}(z) = \frac{\nu z^{-\alpha}}{1+z^{-\alpha}}$, which is given explicitly by*

$$\kappa(x) = \nu x^{\alpha-1} E_{\alpha,\alpha}(-\lambda x^{\alpha}) \sim \frac{1}{\Gamma(\alpha)} \nu x^{\alpha-1} \tag{7}$$

*as $x \to 0$ (see also p. 6 in Gatheral & Keller-Ressel (2019) and p. 29 in El Euch and Rosenbaum (2018)).*

*Proof.* See Appendix A. □

*Remark* 2.3. Integrating (6) and setting $u = t$ we see that

$$V_t = \xi_0(t) + \int_0^t \kappa(t-s)\sqrt{V_s}\,dW_s\,. \tag{8}$$

*Remark* 2.4. From (6), we see that $\xi_t(.)$ is Markov in $\xi_t(.)$. However, $V$ is not Markov in itself.

## 2.3 | Evolving the variance curve

We simulate the variance curve at time $t > 0$ using

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

and substituting the expression for $\xi_0(t) = \mathbb{E}(V_t)$ in (3) and the expression for $\kappa(t)$ in Proposition 2.2 (which are both expressed in terms of the Mittag–Leffler function).

## 2.4 | The characteristic function of the log stock price

From corollary 3.1 in El Euch and Rosenbaum (2019) (see also theorem 6 in Gerhold et al., 2019), we know that for all $t \geq 0$

$$\mathbb{E}(e^{pX_t}) = e^{V_0 I^{1-\alpha} f(p,t) + \lambda\theta I^1 f(p,t)} \tag{9}$$

for $p$ in some open interval $I \supset [0,1]$, where $f(p,t)$ satisfies

$$D^{\alpha} f(p,t) = \frac{1}{2}(p^2 - p) + (p\,\rho\nu - \lambda) f(p,t) + \frac{1}{2}\nu^2 f(p,t)^2 \tag{10}$$

with initial condition $f(p,0) = 0$, where $I^{\alpha} f$ denotes the fractional integral operator of order $\alpha$ (see, e.g., p. 16 in El Euch & Rosenbaum, 2019, for definition) and $D^{\alpha}$ denotes the fractional derivative operator of order $\alpha$ (see p. 17 in El Euch   Rosenbaum, 2019, for definition).

## 2.5 | The generalized time-dependent Rough Heston model and fitting the initial variance curve

If we now replace the constant $\theta$ with a time-dependent function $\theta(t)$, then

$$\mathbb{E}(V_t) = V_0 + \frac{1}{\Gamma(\alpha)} \int_0^t (t-s)^{\alpha-1} \lambda(\theta(s) - \mathbb{E}(V_s))dt,$$

which we can rearrange as

$$\mathbb{E}(V_t) - V_0 + \lambda I^{\alpha} \mathbb{E}(V_t) = \lambda I^{\alpha} \theta(t)$$

so to make this generalized model consistent with a given initial variance curve $\mathbb{E}(V_t)$, we set

$$\theta(t) = \frac{1}{\lambda} D^{\alpha}(\mathbb{E}(V_t) - V_0 + \lambda I^{\alpha} \mathbb{E}(V_t)) = \frac{1}{\lambda} D^{\alpha}(\mathbb{E}(V_t) - V_0) + \mathbb{E}(V_t)$$

(see also remark 3.2, theorem 3.2, and corollary 3.2 in El Euch & Rosenbaum, 2018).

## 2.6 | Other affine and nonaffine variance curve models

Another well-known (and nonaffine) variance curve model is the *Rough Bergomi* model, for which $d\xi_t(u) = \eta(u-t)^{H-\frac{1}{2}} \xi_t(u) dW_t$ or the standard Bergomi model (with mean reversion) for which $d\xi_t(u) = \eta e^{-\lambda(u-t)} \xi_t(u) dW_t$.

# 3 | SMALL-TIME ASYMPTOTICS

## 3.1 | Scaling relations

Let

$$d\tilde{X}_t^{\varepsilon} = \sqrt{\varepsilon}\sqrt{V_t^{\varepsilon}} dB_t, \tag{11}$$

which satisfies

$$\tilde{X}_t^{\varepsilon} \stackrel{(d)}{=} \tilde{X}_{\varepsilon t}.$$

Then, the characteristic function of $\tilde{X}_t$ for $\varepsilon = 1$ is

$$\mathbb{E}(e^{p\tilde{X}_t}) = e^{V_0 I^{1-\alpha} \psi(p,t)}, \tag{12}$$

where $\psi(p,t)$ satisfies

$$D^{\alpha}\psi(p,t) = \frac{1}{2}p^2 + p\rho\nu\psi(p,t) + \frac{1}{2}\nu^2\psi(p,t)^2 \tag{13}$$

with $\psi(p,0)=0$. We first recall that $D^{\alpha}\psi(p,t)=\frac{d}{dt}\frac{1}{\Gamma(1-\alpha)}\int_0^t \psi(p,s)(t-s)^{-\alpha}ds$. Then,

$$
\begin{aligned}
D^{\alpha}\psi(p,\varepsilon t) := (D^{\alpha}\psi)(p,\varepsilon t) &= \frac{1}{\varepsilon}\frac{d}{dt}\frac{1}{\Gamma(1-\alpha)}\int_0^{\varepsilon t}\psi(p,s)(\varepsilon t-s)^{-\alpha}ds \\
&= \frac{1}{\varepsilon}\frac{d}{dt}\frac{1}{\Gamma(1-\alpha)}\int_0^{t}\psi(p,\varepsilon u)(\varepsilon t-\varepsilon u)^{-\alpha}\varepsilon du \\
&= \varepsilon^{-\alpha}\frac{d}{dt}\frac{1}{\Gamma(1-\alpha)}\int_0^{t}\psi(p,\varepsilon u)(t-u)^{-\alpha}du \\
&= \varepsilon^{-\alpha}D^{\alpha}\psi(p,\varepsilon(.))(t).
\end{aligned}
$$

Combining this with (13) we see that

$$\varepsilon^{-\alpha}D^{\alpha}(\psi(p,\varepsilon.))(t)=\frac{1}{2}p^2+p\rho\nu\psi(p,\varepsilon t)+\frac{1}{2}\nu^2\psi(p,\varepsilon t)^2\,. \tag{14}$$

Setting $p\to\varepsilon^{\gamma}q$ and multiplying by $\varepsilon^{-2\gamma}$ we have

$$\varepsilon^{-\alpha-2\gamma}D^{\alpha}(\psi(\varepsilon^{\gamma}q,\varepsilon(.)))(t)=\frac{1}{2}q^2+q\rho\nu\varepsilon^{-\gamma}\psi(\varepsilon^{\gamma}q,\varepsilon t)+\frac{1}{2}\nu^2\varepsilon^{-2\lambda}\psi(\varepsilon^{\gamma}q,\varepsilon t)^2. \tag{15}$$

Now setting $\gamma=-\alpha$ we see that

$$D^{\alpha}(\varepsilon^{\alpha}\psi(\varepsilon^{-\alpha}q,\varepsilon(.)))(t)=\frac{1}{2}q^2+q\rho\nu\varepsilon^{\alpha}\psi(\varepsilon^{-\alpha}q,\varepsilon t)+\frac{1}{2}\nu^2\varepsilon^{2\alpha}\psi(\varepsilon^{-\alpha}q,\varepsilon t)^2 \tag{16}$$

with $\psi(\varepsilon^{-\alpha}q,0)=0$. Thus, we see that $\varepsilon^{\alpha}\psi(\varepsilon^{-\alpha}p,\varepsilon t)$ and $\psi(p,t)$ satisfy the same VIE with the same boundary condition, so

$$\psi(p,t)=\varepsilon^{\alpha}\psi(\varepsilon^{-\alpha}p,\varepsilon t). \tag{17}$$

From the form of the characteristic function in (12), the function $\Lambda(p,t):=I^{1-\alpha}\psi(p,t)$ is clearly of interest too. Using the scaling relation on $\psi(p,t)$:

$$I^{1-\alpha}\psi(p,\varepsilon t)=\frac{1}{\Gamma(1-\alpha)}\int_0^{\varepsilon t}(\varepsilon t-s)^{-\alpha}\psi(p,s)ds \tag{18}$$

$$=\frac{\varepsilon}{\Gamma(1-\alpha)}\int_0^{t}(\varepsilon t-\varepsilon u)^{-\alpha}\psi(p,\varepsilon u)du \tag{19}$$

$$=\frac{\varepsilon^{1-\alpha}}{\Gamma(1-\alpha)}\int_0^{t}(t-u)^{-\alpha}\varepsilon^{-\alpha}\psi(\varepsilon^{\alpha}p,u)du=\varepsilon^{-2H}I^{1-\alpha}\psi(\varepsilon^{\alpha}p,t). \tag{20}$$

Thus we have established the following lemma:

**Lemma 3.1.**

$$\Lambda(p,\varepsilon t)=\varepsilon^{-2H}\Lambda(\varepsilon^{\alpha}p,t) \tag{21}$$

*in particular*

$$\Lambda(p,t)=t^{-2H}\Lambda(pt^{\alpha},1)\,. \tag{22}$$

## 3.2 | The small-time LDP

To simplify calculations, we make the following assumption throughout this section:

**Assumption 3.2.** $\lambda=0$.

*Remark* 3.3. The formal higher-order Laplace asymptotics in subsection 3.5 indicate that $\lambda$ will not affect the leading order small-time asymptotics, that is, $\lambda$ will not affect the rate function, as we would expect from previous works on small-time asymptotics for rough stochastic volatility models. The assumption that $\lambda=0$ is relaxed in the next section where we consider large-time asymptotics.

We now state the main small-time result in the paper $\left(\text{recallthat}\alpha=H+\frac{1}{2}\right)$:

**Theorem 3.4.** *For the Rough Heston model defined in (2), we have*

$$\lim_{t\to 0} t^{2H}\log\mathbb{E}\left(e^{\frac{p}{t^{\alpha}}X_t}\right)=\lim_{t\to 0} t^{2H}\log\mathbb{E}\left(e^{\frac{p}{t^{2H}}\frac{X_t}{t^{\frac{1}{2}-H}}}\right)=\left\{\begin{array}{ll}\bar{\Lambda}(p) & \text{if}\quad T^*(p)>1,\\ +\infty & \text{if}\quad T^*(p)\leq 1,\end{array}\right) \tag{23}$$

*where $\bar{\Lambda}(p):=V_0\Lambda(p)$, $\Lambda(p):=\Lambda(p,1)$, $\Lambda(p,t):=I^{1-\alpha}\psi(p,t)$, and $\psi(p,t)$ satisfies the Volterra differential equation*

$$D^{\alpha}\psi(p,t)=\frac{1}{2}p^2+p\rho\nu\psi(p,t)+\frac{1}{2}\nu^2\psi(p,t)^2 \tag{24}$$

*with initial condition $\psi(p,0)=0$, where $T^*(p)>0$ is the explosion time for $\psi(p,t)$ which is finite for all $p\neq 0$ (assuming $\nu>0$). Moreover, the scaling relation in the previous section show that $\Lambda(p)=|p|^{\frac{2H}{\alpha}}\Lambda(\mathrm{sgn}(p),|p|^{\frac{1}{\alpha}})$, so in fact we only need to solve (24) for $p=\pm 1$, and we can rewrite (23) in more familiar form as*

$$\lim_{t\to 0} t^{2H}\log\mathbb{E}\left(e^{\frac{p}{t^{\alpha}}X_t}\right)=\lim_{t\to 0} t^{2H}\log\mathbb{E}\left(e^{\frac{p}{t^{2H}}\frac{X_t}{t^{\frac{1}{2}-H}}}\right)=\left\{\begin{array}{ll}\bar{\Lambda}(p) & p\in(p_-,p_+),\\ +\infty & p\notin(p_-,p_+),\end{array}\right)$$

*where $p_{\pm}=\pm(T^*(\pm 1))^{\alpha}$, so $p_+>0$ and $p_-<0$. Then $X_t/t^{\frac{1}{2}-H}$ satisfies the LDP as $t\to 0$ with speed $t^{-2H}$ and good rate function $I(x)$ equal to the Fenchel–Legendre transform of $\bar{\Lambda}$.*

*Proof.* We first consider the following family of rescaled Rough Heston models:

$$dX_t^\varepsilon = -\frac{1}{2}\varepsilon V_t^\varepsilon dt + \sqrt{\varepsilon}\sqrt{V_t^\varepsilon}dB_t, \quad V_t^\varepsilon = V_0 + \frac{\varepsilon^\alpha}{\Gamma(\alpha)}\int_0^t (t-s)^{H-\frac{1}{2}}\lambda(\theta - V_s^\varepsilon)ds$$
$$+ \frac{\varepsilon^H}{\Gamma(\alpha)}\int_0^t (t-s)^{H-\frac{1}{2}}\nu\sqrt{V_s^\varepsilon}dW_s \tag{25}$$

with $X_t^\varepsilon = 0$, where $H = \alpha - \frac{1}{2} \in \left(0, \frac{1}{2}\right]$. Then, from Appendix B, we know that

$$\left(X_{(.)}^\varepsilon, V_{(.)}^\varepsilon\right) \overset{(d)}{=} (X_{\varepsilon(.)}, V_{\varepsilon(.)}) \tag{26}$$

(note this actually holds for all $\lambda > 0$, but we are only considering $\lambda = 0$ in this proof). Proceeding along similar lines to theorem 4.1 in Forde and Zhang (2017), we let $\tilde{X}_t^\varepsilon$ denote the solution to

$$d\tilde{X}_t^\varepsilon = \sqrt{\varepsilon}\sqrt{V_t^\varepsilon}dB_t \tag{27}$$

with $\tilde{X}_0^\varepsilon = 0$. From eq. 8 in El Euch and Rosenbaum (2018), we know that

$$\mathbb{E}(e^{p\tilde{X}_t}) = \mathbb{E}^{\mathbb{Q}_p}(e^{\frac{1}{2}p^2\int_0^t V_s ds}),$$

where $\tilde{X}_t := \tilde{X}_t^1$ and $\mathbb{Q}_p$ is defined as in El Euch and Rosenbaum (2018), but under $\mathbb{Q}_p$ the value of the mean reversion speed changes from zero to $\bar{\lambda} = \rho p \nu$, so

$$\mathbb{E}(e^{p\tilde{X}_t}) = e^{V_0 I^{1-\alpha}\psi(p,t)}$$

on some nonempty interval $[0, T^*(p))$, where

$$D^\alpha\psi(p,t) = \frac{1}{2}p^2 + p\rho\nu\psi(p,t) + \frac{1}{2}\nu^2\psi(p,t)^2$$

with $\psi(p,0) = 0$. Existence and uniqueness of solutions to these kind of fractional differential equations (FDEs) is standard, as is their equivalence to VIEs, see, for example, Gerhold et al. (2019) and chapter 12 of Gripenberg et al. (1990) for details.

From propositions 2 and 3 in Gerhold et al. (2019), we know that $\psi(p,t)$ blows up at some finite time $T^*(p) > 0$ (i.e., case A or B in the Gerhold et al., 2019, classification).Thus, we see that

$$\mathbb{E}(e^{\frac{p}{\varepsilon^\alpha}\tilde{X}_t^\varepsilon}) = \mathbb{E}(e^{\frac{p}{\varepsilon^\alpha}\tilde{X}_{\varepsilon t}}) = e^{V_0 I^{1-\alpha}\psi(\frac{p}{\varepsilon^\alpha},\varepsilon t)} = e^{\frac{1}{\varepsilon^{2H}}V_0 I^{1-\alpha}\psi(p,t)} \tag{28}$$

for all $t \in [0, T^*(p))$, which we can rewrite as $\mathbb{E}(e^{\frac{p}{t^\alpha}\tilde{X}_t}) = e^{\frac{\bar{\Lambda}(p)}{t^{2H}}}$. Thus, we see that

$$\lim_{t\to 0} t^{2H}\log\mathbb{E}(e^{\frac{p}{t^\alpha}\tilde{X}_t}) = \bar{\Lambda}(p)$$

and $\Lambda(p) := \Lambda(p,1) < \infty$ if and only if $T^*(p) > 1$. □

We now have the following obvious but important corollary of the $\Lambda$ scaling relation in (22):

**Corollary 3.5.**

$$\Lambda(q) = t^{2H}\Lambda(\frac{q}{t^\alpha}, t) = |q|^{\frac{2H}{\alpha}}\Lambda(\mathrm{sgn}(q), |q|^{\frac{1}{\alpha}}), \tag{29}$$

*where we have set* $p = 1 = \frac{|q|}{t^\alpha}$ *in (22), and* $t_q^* = |q|^{\frac{1}{\alpha}}$.

*Remark* 3.6. This implies that $\Lambda(p) \to \infty$ as $p \to p_\pm := \pm(T^*(\pm 1))^\alpha$. and more generally

$$pT^*(p)^\alpha = 1_{p>0}\, p_+ \, + \, 1_{p<0}\, p_- \, . \tag{30}$$

To prove the LDP, we first prove the corresponding LDP for $\tilde{X}_t$. From lemma 2.3.9 in Dembo and Zeitouni (1998), we know that

$$\lim_{t\to 0} t^{2H} \log \mathbb{E}(e^{\frac{p}{t^\alpha}\tilde{X}_t}) = \Lambda(p) = \Lambda(p,1) = I^{1-\alpha}\psi(p,t)|_{t=1}$$

is convex in $p$, and from (9) and (13) we know that

$$\frac{d}{dt}\Lambda(p,t) = \frac{1}{2}p^2 + p\rho\nu\psi(p,t) + \frac{1}{2}\nu^2\psi(p,t)^2$$

(where we have also used that $D^\alpha D^{1-\alpha} = D$), which shows that $\Lambda(p,t)$ is also differentiable in $t$, and thus from (29), we see that $\Lambda(p) = \Lambda(p,1)$ is differentiable in $p$ for $p > 0$. Moreover, the scaling relation easily yields that $\Lambda(p)$ is right differentiable at $p = 0$, since $\Lambda(p) = o(p)$. We also know that $\psi(p,t) \to \infty$ as $t \to T^*(p)$ (see propositions 2 and 3 in Gerhold et al., 2019), so $\Lambda(p,t) = I^{1-\alpha}\psi(p,t)$ also explodes at $T^*(p)$ by lemma 3 in Gerhold et al. (2019). Then from Corollary 3.5, we know that $\Lambda(p) = p^{\frac{2H}{\alpha}}\Lambda(\mathrm{sgn}(p), |p|^{\frac{1}{\alpha}})$, so $\Lambda(p) \to \infty$ as $p \to p_\pm = \pm(T^*(\pm 1))^\alpha$ and (by convexity and differentiability) $\Lambda$ is also essentially smooth, so by the Gärtner–Ellis theorem from large deviations theory (see theorem 2.3.6 in Dembo & Zeitouni, 1998), $\tilde{X}_1^\varepsilon/\varepsilon^{\frac{1}{2}-H}$ satisfies the LDP as $\varepsilon \to 0$ with speed $\varepsilon^{-2H}$ and rate function $I(x)$.

We now show that $X_1^\varepsilon/\varepsilon^{\frac{1}{2}-H}$ satisfies the same LDP, by showing that the nonzero drift of the log stock price can effectively be ignored at leading order in the limit as $\varepsilon \to 0$. Using that

$$\mathbb{E}(e^{\frac{p}{\varepsilon^{2\alpha}}\varepsilon\int_0^1 V_s^\varepsilon ds}) = \mathbb{E}(e^{\frac{p}{\varepsilon^{2H}}\int_0^1 V_s^\varepsilon ds}) = \mathbb{E}(e^{\frac{\sqrt{2p}}{\varepsilon^\alpha}\tilde{X}_1^\varepsilon}) = e^{\frac{1}{\varepsilon^{2H}}V_0\Lambda(\sqrt{2p})}$$

for $p \in (-\infty, \frac{1}{2}p_+)$ (and $+\infty$ otherwise) so

$$J(p) := \lim_{\varepsilon\to 0}\varepsilon^{2H}\log\mathbb{E}(e^{\frac{p}{\varepsilon^{2\alpha}}\varepsilon\int_0^1 V_s^\varepsilon ds}) = V_0\Lambda(\sqrt{2p})$$

so (again using part (a) of the Gärtner–Ellis theorem in Theorem 2.3.6 in Dembo & Zeitouni, 1998), $A_\varepsilon := \int_0^1 V_s^\varepsilon ds$ satisfies the upper bound LDP as $\varepsilon \to 0$ with speed $\varepsilon^{-2H}$ and good rate function $J^*$

equal to the FL transform of $J$. But we also know that

$$X_1^\varepsilon - \tilde{X}_1^\varepsilon = -\frac{1}{2}\varepsilon A_\varepsilon$$

and for any $a > 0$ and $\delta_1 > 0$

$$\mathbb{P}(|\frac{X_1^\varepsilon}{\varepsilon^{\frac{1}{2}-H}} - \frac{\tilde{X}_1^\varepsilon}{\varepsilon^{\frac{1}{2}-H}}| > \delta) = \mathbb{P}(\frac{1}{2}\varepsilon^{\frac{1}{2}+H}A_\varepsilon > \delta) = \mathbb{P}(A_\varepsilon > \frac{2\delta}{\varepsilon^{\frac{1}{2}+H}}) \leq \mathbb{P}(A_\varepsilon > a) \quad \leq \quad e^{-\frac{\inf_{a'\geq a} J^*(a')-\delta_1}{\varepsilon^{2H}}}$$

for any $\varepsilon$ sufficiently small, where we have use the upper bound LDP for $A_\varepsilon$ to obtain the final inequality. Thus

$$\limsup_{\varepsilon\to 0} \varepsilon^{2H} \log \mathbb{P}(|\frac{X_1^\varepsilon}{\varepsilon^{\frac{1}{2}-H}} - \frac{\tilde{X}_1^\varepsilon}{\varepsilon^{\frac{1}{2}-H}}| > \delta) \leq -\inf_{a'>a} J^*(a'),$$

but $a$ is arbitrary and (from lemma 2.3.9 in Dembo & Zeitouni, 1998), $J^*$ is a good rate function, so in fact

$$\limsup_{\varepsilon\to 0} \varepsilon^{2H} \log \mathbb{P}(|\frac{X_1^\varepsilon}{\varepsilon^{\frac{1}{2}-H}} - \frac{\tilde{X}_1^\varepsilon}{\varepsilon^{\frac{1}{2}-H}}| > \delta) = -\infty.$$

Thus, $\frac{X_1^\varepsilon}{\varepsilon^{\frac{1}{2}-H}}$ and $\frac{\tilde{X}_1^\varepsilon}{\varepsilon^{\frac{1}{2}-H}}$ are *exponentially equivalent* in the sense of definition 4.2.10 in Dembo and Zeitouni (1998), so (by theorem 4.2.13 in Dembo & Zeitouni, 1998) $\frac{X_1^\varepsilon}{\varepsilon^{\frac{1}{2}-H}}$ satisfies the same LDP as $\frac{\tilde{X}_1^\varepsilon}{\varepsilon^{\frac{1}{2}-H}}$.

## 3.3 | Asymptotics for call options and implied volatility

**Corollary 3.7.** *We have the following limiting behavior for out-of-the-money European put and call options with maturity t and log-strike* $t^{\frac{1}{2}-H}x$*, with* $x > 0$ *fixed:*

$$\lim_{t\to 0} t^{2H} \log \mathbb{E}((e^{X_t} - e^{xt^{\frac{1}{2}-H}})^+) = -I(x) \qquad (x > 0)$$

$$\lim_{t\to 0} t^{2H} \log \mathbb{E}((e^{xt^{\frac{1}{2}-H}} - e^{X_t})^+) = -I(x) \qquad (x < 0).$$

*Proof.* The lower estimate follows from the exact same argument used in appendix C in Forde and Zhang (2017) (see also theorem 6.3 in Friz, Gerhold, & Pinter, 2018). The proof of the upper estimate is the same as in theorem 6.3 in Friz, Gerhold, et al. (2018). □

**Corollary 3.8.** *Let* $\hat{\sigma}_t(x)$ *denote the implied volatility of a European put/call option with log-moneyness x under the Rough Heston model in (2) for* $\lambda = 0$*. Then for* $x \neq 0$ *fixed, the implied* *volatility satisfies*

$$\hat{\sigma}(x) := \lim_{t \to 0} \hat{\sigma}_t(t^{\frac{1}{2}-H} x) = \frac{|x|}{\sqrt{2I(x)}}. \tag{31}$$

*Proof.* Follows from corollary 7.2 in Gao and Lee (2014). See also the proof of corollary 4.1 in Friz, Gerhold, et al. (2018) for details on this, but the present situation is simpler, as we only require the leading order term here. □

## 3.4 | Series expansion for the asymptotic smile and calibration

Proceeding as in lemma 12 in Gerhold et al. (2019), we can compute a fractional power series for $\psi(p, t)$ (and hence $\Lambda(p, t)$) and then using (29), we find that

$$\bar{\Lambda}(p) = \frac{2V_0}{\nu^2} \sum_{n=1}^{\infty} a_n(1) p^{1+n} \frac{\Gamma(\alpha n + 1)}{\Gamma(2 + (n-1)\alpha)},$$

where the $a_n = a_n(u)$ coefficients are defined (recursively) as in Gerhold et al. (2019) except for our application here (based on (13)) we have to set $\lambda = 0$, and $c_1 = \frac{1}{2}u^2$ instead of $\frac{1}{2}u(u-1)$ (note this series will have a finite radius of convergence). Using the Lagrange inversion theorem, we can then derive a power series for $I(x)$ which takes the form

$$\hat{\sigma}(x) = \sqrt{V_0} + \frac{\rho\nu}{2\Gamma(2+\alpha)\sqrt{V_0}} x + \nu^2 \frac{\Gamma(1+2\alpha) + 2\rho^2\Gamma(1+\alpha)^2(2 - 3\frac{\Gamma(2+2\alpha)}{\Gamma(2+\alpha)^2})}{8V_0^{\frac{3}{2}}\Gamma(1+\alpha)^2\Gamma(2+2\alpha)} x^2 + O(x^3) \tag{32}$$

(compare this to theorem 3.6 in Bayer et al. (2018) for a general class of rough models and theorem 4.1 in Forde and Jacquier (2011b) for a Markovian local-stochastic volatility model). We can rewrite this expansion more concisely in dimensionless form as

$$\hat{\sigma}(x) = \sqrt{V_0}\,[1 + \frac{\rho}{2\Gamma(2+\alpha)} z + \frac{\Gamma(1+2\alpha) + 2\rho^2\Gamma(1+\alpha)^2(2 - 3\frac{\Gamma(2+2\alpha)}{\Gamma(2+\alpha)^2})}{8\Gamma(1+\alpha)^2\Gamma(2+2\alpha)} z^2 + O(z^3)],$$

where the dimensionless quantity $z = \frac{\nu x}{V_0}$.

*Remark* 3.9. In principle, one can use (32) to calibrate $V_0$, $\rho$, and $\nu$ to observed/estimated values of $\hat{\sigma}(0)$, $\hat{\sigma}'(0)$, and $\hat{\sigma}''(0)$ (i.e., the short-end implied vol level, skew, and convexity, respectively).

### 3.4.1 | Wing behavior of the rate function

From eq (3.2) in Roberts and Olmstead (1996), we expect that $\psi(p, t) \sim \frac{const.}{(T^*(p)-t)^{\alpha}}$ as $t \to T^*(p)$ and thus $\Lambda(p, t) = I^{1-\alpha}\psi(p, t) \sim \frac{const.}{(T^*(p)-t)^{2\alpha-1}}$ as $t \to T^*(p)$. Assuming this is consistent with the

$p$-asymptotics, then (by (30)) we have

$$\Lambda(p) = \Lambda(p,1) \sim \frac{const.}{(T^*(p)-1)^{2\alpha-1}} = \frac{const.}{((\frac{p_+}{p})^{1/\alpha}-1)^{2\alpha-1}} \sim \frac{const.}{(p_+-p)^{2\alpha-1}} \qquad (p \to p_+)$$

so $p^*(x)$ in $I(x) = \sup_p(px - V_0\Lambda(p))$ satisfies $p^*(x) = p_+ - const.\cdot x^{-1/2\alpha}(1+o(1))$, so $I(x) = p_+x + const.\cdot x^{1-\frac{1}{2\alpha}}(1+o(1))$ as $x \to \infty$.

## 3.5 | Higher-order Laplace asymptotics

If we now relax the assumption that $\lambda = 0$, and work with the original $X^\varepsilon$ process in (25) (as opposed to the driftless $\tilde{X}^\varepsilon$ process in (27)), then we know that

$$\mathbb{E}(e^{pX_t^\varepsilon}) = \mathbb{E}(e^{pX_{\varepsilon t}}) = e^{V_0 I^{1-\alpha} g_\varepsilon(p,t) + \varepsilon^\alpha \lambda\theta I^1 g_\varepsilon(p,t)}$$

for $t$ in some nonempty interval $[0, T_\varepsilon^*(p))$, where

$$g_\varepsilon(\frac{p}{\varepsilon^\alpha}, t) = \frac{\psi(p,t)}{\varepsilon^{2H}}, \tag{33}$$

which satisfies

$$D^\alpha g_\varepsilon(p,t) = \frac{1}{2}\varepsilon(p^2-p) + (p\rho\nu - \lambda)\varepsilon^\alpha g_\varepsilon(p,t) + \frac{1}{2}\varepsilon^{2H}\nu^2 g_\varepsilon(p,t)^2 \tag{34}$$

with initial condition $g_\varepsilon(p,0) = 0$. Setting

$$g_\varepsilon(\frac{p}{\varepsilon^\alpha}, t) = \frac{\psi_\varepsilon(p,t)}{\varepsilon^{2H}} \tag{35}$$

and setting $p \mapsto \frac{p}{\varepsilon^\alpha}$, and substituting for $g_\varepsilon(\frac{p}{\varepsilon^\alpha}, t)$ in (34) and multiplying by $\varepsilon^{2H}$ as before, we find that

$$D^\alpha \psi_\varepsilon(p,t) = \frac{1}{2}p^2 + p\rho\nu\psi_\varepsilon(p,t) + \frac{1}{2}\nu^2\psi_\varepsilon(p,t)^2 - \varepsilon^\alpha(\frac{1}{2}p + \lambda\psi_\varepsilon(p,t))$$

with $\psi_\varepsilon(p,0) = 0$. If we now formally try a higher-order series approximation of the form $\psi_\varepsilon(p,t) := \psi(p,t) + \varepsilon^{\frac{1}{2}+H}\psi_1(p,t)$, we find that $\psi_1(p,t)$ must satisfy

$$D^\alpha \psi_1(p,t) = -\frac{1}{2}p - \lambda\psi(p,t) + p\rho\nu\psi_1(p,t) + \nu^2\psi(p,t)\psi_1(p,t)$$

with $\psi_1(p,0) = 0$, which is a linear VIE for $\psi_1(p,t)$.

*Remark* 3.10. Let $\Delta_\varepsilon(p,t) = \psi_\varepsilon(p,t) - \psi(p,t) - \varepsilon^{\frac{1}{2}+H}\psi_1(p,t)$ denote the error term. Then $\Delta_\varepsilon(p,t)$ satisfies

$$\begin{aligned} D^\alpha \Delta_\varepsilon(p,t) &= p\nu\rho\Delta_\varepsilon(p,t) + \frac{1}{2}\nu^2\Delta_\varepsilon(p,t)^2 + \nu^2\Delta_\varepsilon(p,t)\psi(p,t) \\ &\quad + \varepsilon^{\frac{1}{2}+H}\Delta_\varepsilon(p,t)(-\lambda + \nu^2\psi_1(p,t)) \\ &\quad + \varepsilon^{2H+1}(-\lambda\psi_1(p,t) + \frac{1}{2}\nu^2\psi_1(p,t)^2) \end{aligned}$$

and the rescaled error $\bar{\Delta}_\varepsilon(p,t) := \Delta_\varepsilon(p,t)/\varepsilon^{\frac{1}{2}+H}$ satisfies

$$\begin{aligned} D^\alpha \bar{\Delta}_\varepsilon(p,t) &= \bar{\Delta}_\varepsilon(p,t)(p\nu\rho + \nu^2\psi(p,t)) + \\ &\quad + \varepsilon^{\frac{1}{2}+H}(-\lambda\psi_1(p,t) + \frac{1}{2}\nu^2\psi_1(p,t)^2 + (-\lambda + \nu^2\psi_1(p,t))\bar{\Delta}_\varepsilon(p,t) + \frac{1}{2}\nu^2\bar{\Delta}_\varepsilon(p,t)^2). \end{aligned}$$

We know that $\psi(p,t)$ is continuous on $[0,T^*(p))$. In order to make this rigorous, one would need to apply Gripenberg et al. (1990) to this, noting that the leading order solution is zero, then replace $p$ with $ik$ for $k$ real, then show this convergence is uniform on compact sets, and then argue away the tails as in Forde, Jacquier, and Lee (2012).

*Remark* 3.11. Setting $\psi_1(p,t) = \sum_{n=1}^\infty b_n(p)t^{\alpha n}$ we see that

$$\begin{aligned} \sum_{n=1}^\infty \frac{n\alpha\Gamma(n\alpha)}{\Gamma(1+(n-1)\alpha)} b_n(p)t^{(n-1)\alpha} &= -\frac{1}{2}p - \lambda\sum_{n=1}^\infty \bar{a}_n(p)t^{\alpha n} \\ &\quad + p\rho\nu\sum_{n=1}^\infty b_n(p)t^{\alpha n} + \nu^2\sum_{n=1}^\infty \bar{a}_n(p)t^{\alpha n}\sum_{m=1}^\infty b_m(p)t^{\alpha m}, \end{aligned}$$

where $\bar{a}_n(p) = \frac{2}{\nu^2}a_n(p)$, and we have set $\lambda = 0$ and $c_1 = \frac{1}{2}p^2$ in computing the $a_n(p)$ coefficients, so

$$\begin{aligned} \alpha\Gamma(\alpha)b_1(p) = -\frac{1}{2}p, \quad \frac{(n+1)\alpha\Gamma((n+1)\alpha)}{\Gamma(1+n\alpha)}b_{n+1}(p) &= -\lambda\bar{a}_n(p) + \rho p\nu b_n(p) \\ + \nu^2\sum_{k=1}^{n-1}a_k(p)b_{n-k}(p) \end{aligned}$$

so we have fractional power series for $\psi_1(p,t)$ on some finite radius of convergence.

Returning now to the main calculation, we see that if $p_\varepsilon(x)$ denotes the density of $\frac{X_1^\varepsilon}{\varepsilon^\alpha}$, then

$$p_\varepsilon(\frac{x}{\varepsilon^{2H}}) = \frac{1}{2\pi}\int_{-\infty}^\infty e^{-\frac{ikx}{\varepsilon^{2H}}} e^{\frac{1}{\varepsilon^{2H}}(F(k)+\varepsilon^{\frac{1}{2}+H}G(k))+\frac{\varepsilon^\alpha}{\varepsilon^{2H}}\lambda\theta(F_1(k)+\varepsilon^{\frac{1}{2}+H}G_1(k))}dk,$$

where $F(k) := V_0 I^{1-\alpha}\psi(ik,1)$, $G(k) := V_0 I^{1-\alpha}\psi_1(ik,1)$, $F_1 := I^1\psi(ik,1)$, and $G_1 := I^1\psi_1(ik,1)$. The saddle point $k^* = k^*(x) = ip^*(x)$ of $\bar{F}(k) = -ikx + F(k)$ satisfies $\bar{F}'(k^*) = 0$ which always falls on the imaginary axis (and in our case $p^*(x) \in (0, p_+)$ when $x > 0$ and $p^*(x) < 0 \in (p_-, 0)$ when $x < 0$), and

$$\begin{aligned}\bar{F}(k) &= \bar{F}(k^*) + \frac{1}{2}F''(k^*)(k-k^*)^2 + O((k-k^*)^3)\\ &= \bar{F}(k^*) - \frac{1}{2}\bar{\Lambda}''(p^*)(k-k^*)^2 + O((k-k^*)^3)\end{aligned}$$

(recall that $\bar{\Lambda}(p) = F(-ip)$) and $p^* = ik^* \in (p_-, p_+)$. Then proceeding along similar lines to Forde et al. (2012) and using Laplace's method we have for all $x \in \mathbb{R}$

$$p_\varepsilon\left(\frac{x}{\varepsilon^{2H}}\right) = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{\frac{1}{\varepsilon^{2H}}(\bar{F}(k)+\varepsilon^{\frac{1}{2}+H}G(k))+\varepsilon^{\frac{1}{2}-H}\lambda\theta(F_1(k)+\varepsilon^{\frac{1}{2}+H}G_1(k))}dk, \tag{36}$$

$$\approx \frac{1}{2\pi}e^{\varepsilon^{\frac{1}{2}-H}(G(k^*)+\lambda\theta F_1(k^*))}\int_{-\infty}^{\infty} e^{\frac{1}{\varepsilon^{2H}}(\bar{F}(k^*)-\frac{1}{2}\bar{\Lambda}''(p^*)(k-k^*)^2)}dk, \tag{37}$$

$$\begin{aligned}&\approx \frac{1}{2\pi}e^{\varepsilon^{\frac{1}{2}-H}(G(k^*)+\lambda\theta F_1(k^*))}e^{-\frac{I(x)}{\varepsilon^{2H}}}\int_{-\infty}^{\infty} e^{-\frac{1}{\varepsilon^{2H}}\frac{1}{2}\bar{\Lambda}''(p^*)(k-k^*)^2}dk\\ &= \frac{\varepsilon^H e^{-\frac{I(x)}{\varepsilon^{2H}}}}{\sqrt{2\pi\bar{\Lambda}''(p^*)}}[1 + \varepsilon^{\frac{1}{2}-H}(G(k^*) + \lambda\theta F_1(k^*)) + O(\varepsilon^{(1-2H)\wedge 2H})],\end{aligned} \tag{38}$$

where the $O(\varepsilon^{2H})$ part of the error terms comes from the next order term in theorem 7.1 in chapter 4 in Olver (1974), and the $\varepsilon^{(1-2H)}$ term comes from the second-order term in expanding the exponential. The meaning of $\approx$ in the above estimates is as follows: we expect to have asymptotic equality with a relative error term that does not interfere with the error term in (38), but since we did not carry out the tail estimate of the saddle point approximation, we do not know its size. Then letting $z = \frac{k}{\varepsilon^\alpha}$, we see that

$$\begin{aligned}C_\varepsilon(x) = \mathbb{E}((e^{X_1^\varepsilon} - e^{x\varepsilon^{\frac{1}{2}-H}})^+) &= \frac{1}{2\pi}e^{x\varepsilon^{\frac{1}{2}-H}}\int_{-ip^*-\infty}^{-ip^*+\infty}\mathrm{Re}\left(\frac{e^{-izx\varepsilon^{\frac{1}{2}-H}}}{-iz-z^2}\mathbb{E}(e^{izX_1^\varepsilon})\right)dz\\ &= \frac{1}{2\pi}e^{x\varepsilon^{\frac{1}{2}-H}}\int_{-ip^*-\infty}^{-ip^*+\infty}\mathrm{Re}\left(\frac{e^{-i\frac{k}{\varepsilon^{2H}}x}}{-i\frac{k}{\varepsilon^\alpha}-(\frac{k}{\varepsilon^\alpha})^2}\mathbb{E}(e^{i\frac{k}{\varepsilon^\alpha}X_1^\varepsilon})\right)d\frac{k}{\varepsilon^\alpha},\end{aligned} \tag{39}$$

$$= \frac{\varepsilon^{-\alpha}}{2\pi}e^{x\varepsilon^{\frac{1}{2}-H}}\int_{-ip^*-\infty}^{-ip^*+\infty}\mathrm{Re}(e^{i\frac{k}{\varepsilon^{2H}}x}(-\frac{\varepsilon^{2\alpha}}{k^2} - i\frac{\varepsilon^{3\alpha}}{k^3} + O(\varepsilon^{4\alpha}))\,\mathbb{E}(e^{i\frac{k}{\varepsilon^\alpha}X_1^\varepsilon}))dk, \tag{40}$$

$$= \frac{\varepsilon^{\frac{1}{2}+2H} e^{-\frac{I(x)}{\varepsilon^{2H}}}}{(p^*)^2\sqrt{2\pi\bar{\Lambda}''(p^*)}}[1+\varepsilon^{\frac{1}{2}-H}(x+G(k^*)+\lambda\theta F_1(k^*))+O(\varepsilon^{(1-2H)\wedge 2H})]$$

$$= \frac{A(x)\varepsilon^{\frac{1}{2}+2H} e^{-\frac{I(x)}{\varepsilon^{2H}}}}{\sqrt{2\pi}}[1+\varepsilon^{\frac{1}{2}-H}(x+G(k^*)+\lambda\theta F_1(k^*))+O(\varepsilon^{(1-2H)\wedge 2H})], \tag{41}$$

where

$$A(x) = \frac{1}{(p^*)^2\sqrt{\bar{\Lambda}''(p^*)}}. \tag{42}$$

The $\varepsilon$-dependence of the leading order term here is exactly the same as in corollary 7.1 in the recent article of Friz, Gassiat, et al. (2018) (in Friz, Gassiat, et al. (2018) $\varepsilon^2 = t$ whereas here $\varepsilon = t$) which deals with a general class of rough stochastic volatility models (which excludes Rough Heston). The difficulty in making the expansions (38) and (41) rigorous is the step from (40) to (41), or, more explicitly, from (36) to (37). The expansion of $\bar{F}$ used in (37) is valid locally, close to the saddle point. An estimate for the integrand in (36) is needed to argue that this is good enough, that is, that the asymptotic behavior of (36) is captured by integrating over an appropriate neighborhood of the saddle point. This is usually done by establishing monotonicity of the integrand, but seems nontrivial here; cf. lemma 6.4 in Forde et al. (2012), which uses the explicit characteristic function of the classical Heston model.

More generally, we can formally substitute a fractional power series of the form $\psi_\varepsilon(p,t) = \sum_{n=0}^{\infty}\psi_n(p,t)\varepsilon^{(n+1)\alpha}$ (where $\psi_0(p,t) := \psi(p,t)$), and we find that $(\psi_n)_{n\geq 1}$ satisfies a nested sequence of linear FDEs:

$$D^{\alpha}\psi_1(p,t) = -\frac{1}{2}p - \lambda\psi_0(p,t) + p\rho\nu\psi_1(p,t) + \nu^2\psi_0(p,t)\psi_1(p,t)$$

$$D^{2\alpha}\psi_2(p,t) = -\lambda\psi_1(p,t) + p\rho\nu\psi_2(p,t) + \nu^2\psi_0(p,t)\psi_2(p,t) + \frac{1}{2}\nu^2\psi_1(p,t)^2$$

$$\ldots$$

$$D^{n\alpha}\psi_n(p,t) = -\lambda\psi_{n-1}(p,t) + p\rho\nu\psi_n(p,t) + \frac{1}{2}\nu^2\Big[\sum_{k=0}^{n}\psi_k(p,t)\psi_{n-k}(p,t) + 1_{\frac{1}{2}n\in\mathbb{N}}\cdot\psi_{\frac{1}{2}n}(p,t)^2\Big] \tag{43}$$

with $\psi_n(p,0) = 0$, and in principle we can then compute fractional power series expansions for each $\psi_n(p,t)$ of the form $\psi_n(p,t) = \sum_{m=1}^{\infty} a_{m,n}(p)t^{\alpha m}$, as in Remark 3.11.

### 3.5.1 | Higher-order expansion for implied volatility

**Formal corollary of** (41): Let $\hat{\sigma}_t(x)$ denote the implied volatility of a European put/call option with log-moneyness $x$ under the Rough Heston model in (2) for $\lambda \geq 0$. Then for $x \neq 0$ fixed, the implied volatility satisfies

$$\hat{\sigma}_t(t^{\frac{1}{2}-H}x)^2 = \frac{|x|}{\sqrt{2I(x)}} + t^{2H}\Sigma_1(x) + o(t^{2H}), \tag{44}$$

where

$$\Sigma_1(x) = \frac{x^2 \log A_1(x)}{2I(x)^2},$$

and where $A_1(x) = 2A(x)I(x)^{\frac{3}{2}}/x^{1}$.

*Proof.* Let $L_t = -\log C_t(x)$, where $C_t(x)$ is defined as in (41). Then using corollary 7.1 and eq (7.2) in Gao and Lee (2014) we see that

$$|\frac{1}{t}G_-^2(k_t, L_t - \frac{3}{2}\log L_t + \log\frac{k_t}{4\sqrt{\pi}}, u) - \hat{\sigma}_t^2(k_t)| = o(\frac{k_t^2}{L_t^2 t}),$$

where $G_-(k,u) := \sqrt{2}(\sqrt{u+k} - \sqrt{u})$. Then

$$L_t - \frac{3}{2}\log L_t + \log\frac{k_t}{4\sqrt{\pi}} = \frac{I(x)}{t^{2H}} - (\frac{1}{2} + 2H)\log t - \log\frac{A(x)}{\sqrt{2\pi}} - \frac{3}{2}\log(\frac{I(x)}{t^{2H}}(1 - \log A(x)\frac{t^{2H}}{I(x)}))$$
$$+ \log\frac{x}{4\sqrt{\pi}} + (\frac{1}{2} - H)\log t,$$

where $A(x)$ is defined as in (42). Collecting $\log t$ terms we find that their sum vanishes, so

$$L_t - \frac{3}{2}\log L_t + \log\frac{k_t}{4\sqrt{\pi}} = \frac{I(x)}{t^{2H}} - \log\frac{A(x)}{\sqrt{2\pi}} - \frac{3}{2}\log I(x) + \log\frac{x}{4\sqrt{\pi}} + o(1)$$
$$= \frac{I(x)}{t^{2H}} - \log A_1(x) + o(1).$$

Then using that

$$G_-^2(k,u) = \frac{k^2}{2u} - \frac{k^3}{4u^2} + O(\frac{k^4}{u^3})$$

as $k/u \to 0$, we obtain the result. □

### 3.5.2 | Using these approximations in practice

Equation (41) is of little use in practice, since the leading order Laplace approximation ignores the variation of the function $\frac{1}{k^2}$ in the integrand, and even if we partially take account of this effect by going to next order with Laplace's method using the formula in theorem 7.1 in chapter 4 in Olver (1974) (which we have checked and tried), it still frequently gives a worse estimate that the leading order estimate $\hat{\sigma}(x)$ because the higher-order error terms being ignored are too large, and since $H$ is usually very small in practice, $t^H$ converges very slowly to zero. If we instead compute an approximate call price using the Fourier integral along the horizontal contour going through the saddle point in (39) (using, e.g., the NIntegrate" command in Mathematica) and use our higher-order asymptotic estimate $\psi(ik,t)+\varepsilon^{\frac{1}{2}+H}\psi_1(ik,t)$ for $\log\mathbb{E}(e^{i\frac{k}{\varepsilon^\alpha}X^\varepsilon}))$, and then compute the *exact* implied volatility associated with this price (which avoids the problems with the Laplace approximation), then (for the parameters we considered) we found this approximation to be an order of magnitude closer to the Monte Carlo value than the leading order approximation $\hat{\sigma}(x)$ (see graphs and table below). See Lord and Kahl (2007) for more on computing the optimal contour of integration for such problems.

## 3.6 | Small-time moderate deviations

Inspired by Bayer et al. (2018), if we replace (35) with

$$g_\varepsilon(\frac{p}{\varepsilon^q},t)=\frac{\psi_\varepsilon(p,t)}{\varepsilon^{2H-2\beta}},$$

where $q=\frac{1}{2}-H+\beta$, then we find that

$$D^\alpha\psi_\varepsilon(p,t)=\frac{1}{2}p^2-\frac{1}{2}p\varepsilon^{\frac{1}{2}-H+\beta}+p\varepsilon^{-2H+3\beta}\rho\nu\psi_\varepsilon(p,t)-\varepsilon^{\frac{1}{2}-3H+4\beta}\lambda\psi_\varepsilon(p,t)$$
$$+\frac{1}{2}\varepsilon^{-4H+6\beta}\nu^2\psi_\varepsilon(p,t)^2$$

and we see that all nonconstant terms on the right-hand side are $o(1)$ as $\varepsilon\to 0$ if $\beta\in(\frac{2}{3}H,H)$ and $H\in(0,\frac{1}{2})$. Following similar calculations as above, we formally obtain that $\lim_{t\to 0}t^{2H-2\beta}\log\mathbb{E}(e^{\frac{p}{t^{2H-2\beta}}\frac{X_t}{t^q}})=V_0I^{1-\alpha}I^\alpha(\frac{1}{2}p^2)=\frac{1}{2}V_0p^2$ for *all* $p\in\mathbb{R}$, which (modulo some rigour) implies that $X_t/t^q$ satisfies the LDP with speed $\frac{1}{t^{2H-2\beta}}$ and Gaussian rate function $I(x)=\frac{1}{2}x^2/V_0$. Note that $\beta=H$ corresponds to the central limit or Edgeworth regime, see Forde et al. (2019) for details.

| $x$ | $\hat{\sigma}(x)$ | Higher order $T=0.00005$ | Monte Carlo $T=0.00005$ | Higher order $T=0.005$ | Monte Carlo $T=0.005$ |
|---|---|---|---|---|---|
| **−0.10** | **20.2068%** | **20.2023%** | **20.2020%** | **20.1615%** | **20.1589 %** |
| −0.08 | 20.141% | 20.1364% | 20.1363% | 20.0953% | 20.0931% |
| −0.06 | 20.0869% | 20.0822% | 20.0824% | 20.0407% | 20.0388% |
| −0.04 | 20.045% | 20.0404% | 20.0407% | 19.9986% | 19.9968% |
| −0.02 | 20.016% | 20.0113% | 20.0119% | 19.9693% | 19.9676% |
| 0.00 | 20.0000% | - | 19.9942% | - | 19.9513% |
| 0.02 | 19.9973% | 19.9926% | 19.9921% | 19.9503% | 19.9509% |
| 0.04 | 20.0079% | 20.0033% | 20.0029% | 19.9610% | 19.9613 % |
| 0.06 | 20.0316% | 20.0270% | 20.0266% | 19.9850% | 19.9850% |
| 0.08 | 20.068% | 20.0634% | 20.0629% | 20.0218% | 20.0213% |
| 0.10 | 20.1166% | 20.1120% | 20.1114% | 20.0709% | 20.0699% |

*Notes*. Table of numerical results corresponding to the right plot in Figure 3 and the left plot in Figure 4.

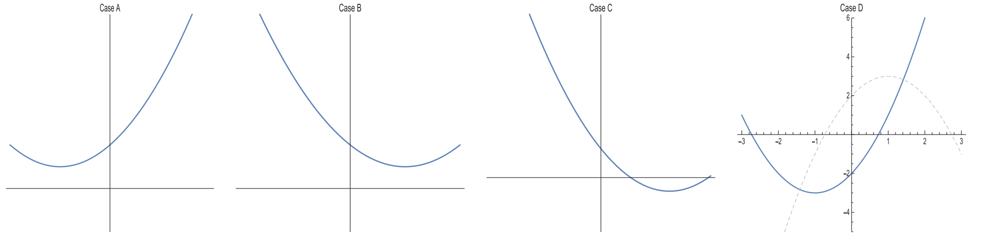


**FIGURE 1** Here we have plotted the quadratic function $G(p, w)$ as a function of $w$ for the four cases described in Gerhold et al. (2019) [Color figure can be viewed at wileyonlinelibrary.com]
*Note*. In cases A and B, there are no roots and the solution $\psi(p, t)$ to (13) increases without bound whereas in cases C and D we have a stable fixed point (the lesser of the two roots) and an unstable root, so a solution starting at the origin increases (decreases) until it reaches the stable fixed point. For Case D, we have also drawn the curve arising from the reflection transformation used in the proof in Appendix C.

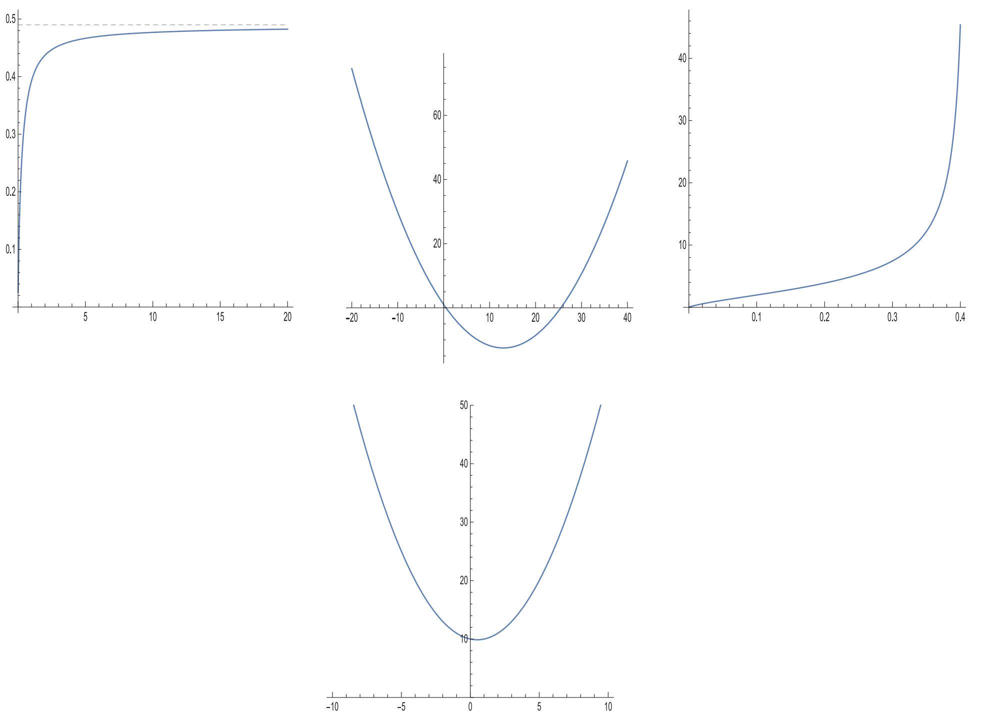

**FIGURE 2** Here we have solved for the solution $f(p, t)$ to Equation (10) numerically by discretizing the VIE with 2,000 time steps, and plotted $f(p, t)$ a function of $t$ and the corresponding quadratic function $G(p, w)$ as a function of $w$ with $p$ fixed [Color figure can be viewed at wileyonlinelibrary.com]
*Note*. In the first case $\alpha = .75$, $\lambda = 2$, $\rho = -0.1$, $\nu = .4$, and $p = 2$ and $f(p, t)$ tends to a finite constant, and in the second case $\alpha = .75$, $\lambda = 1$, $\rho = 0.1$, $\nu = 1$, and $p = 5$ and we see that $f(p, t)$ has an explosion time at some $T^*(p) \approx 0.4$. VIE, Volterra integral equation.

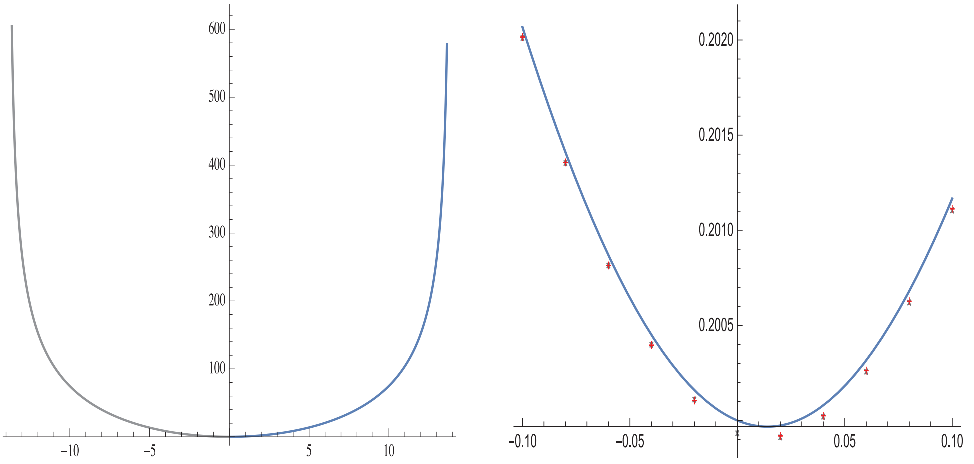


**FIGURE 3** On the left, we have plotted $\Lambda(p)$ using an Adams scheme to numerically solve the VIE in (13) with 2,000 time steps combined with Corollary 3.5, for $\alpha = .75$, $V_0 = .04$, $\nu = .15$, $\rho = -0.02$, and we find that $p_+ = T^*(1) \approx 34.5$ and $p_- = T^*(-1) \approx 33.25$. On the right, we have plotted the corresponding asymptotic small-maturity smile $\hat{\sigma}(x)$ (in blue) versus the higher-order approximation using Equation (39) (red "+" signs), and the smile points obtained from a simple Euler-type Monte Carlo scheme with maturity $T = 0.00005$, $10^5$ simulations, and 1,000 time steps in Matlab (gray crosses), Matlab and Mathematica code available on request [Color figure can be viewed at wileyonlinelibrary.com]
*Note*. We did not use the Adams scheme to compute $\hat{\sigma}(x)$; rather have used the first 15 terms in the series expansion for $\bar{\Lambda}(p)$ in Subsection 3.4 and then numerically computed its Fenchel–Legendre transform and used this to compute $I(x)$ and hence $\hat{\sigma}(x)$. We see that the Monte Carlo and higher-order smile points can barely be distinguished by the naked eye. For $|x|$ small, we have found this method of computing $\hat{\sigma}(x)$ to be far superior to using an Adams scheme, since the numerical computation of the fractional integral $I^{1-\alpha} f(p,t)$ for $|t| \ll 1$ can lead to numerical artifacts when computing the FL transform of $\bar{\Lambda}(p,1)$ close to $x = 0$.

## 4 | LARGE-TIME ASYMPTOTICS

In this section, we derive large-time large deviation asymptotics for the Rough Heston model, and we begin making the following assumption throughout this section:

**Assumption 4.1.** $\lambda > 0, \rho \le 0$.

Recall that $f(p,t)$ in (9) satisfies

$$D^{\alpha} f(p,t) = H(p, f(p,t)) \tag{45}$$

subject to $f(p,0) = 0$, where $H(p,w) := \frac{1}{2}p^2 - \frac{1}{2}p + (p\rho\nu - \lambda)w + \frac{1}{2}\nu^2 w^2$. We write

$$U_1(p) := \frac{1}{\nu^2}[\lambda - p\rho\nu - \sqrt{\lambda^2 - 2\lambda\rho\nu p + \nu^2 p(1 - p\bar{\rho}^2)}]$$

for the smallest root of $H(p,.)$, and note that $U_1(p)$ is real if and only if $p \in [\underline{p}, \bar{p}]$, where

$$\underline{p} := \frac{\nu - 2\lambda\rho - \sqrt{4\lambda^2 + \nu^2 - 4\lambda\rho\nu}}{2\nu(1-\rho^2)}, \quad \bar{p} := \frac{\nu - 2\lambda\rho + \sqrt{4\lambda^2 + \nu^2 - 4\lambda\rho\nu}}{2\nu(1-\rho^2)}.$$

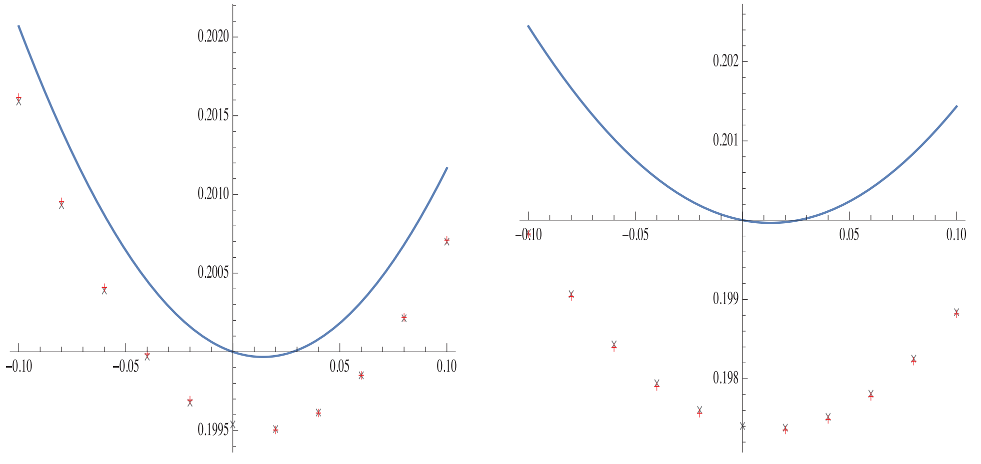


**FIGURE 4** On the left here we have the same plot as above but with $T = 0.005$ and for the right plot $T = 0.005$ and $\alpha = .6$ (i.e., $H = 0.1$), and again we see that the higher-order approximation makes a significant improvement over the leading order smile [Color figure can be viewed at wileyonlinelibrary.com]
*Note*. Of course we would not expect such close agreement for smaller values of $\alpha$, or larger values of $T$, $|x|$, or $|\rho|$, for example, $\rho = -0.65$ reported in, for example, El Euch, Gatheral & Rosenbaum (2018), but the point here is really just to verify the correctness of the asymptotic formula in (31), and give a starting point for other authors/practitioners who wish to test refinements/variants of our formula. We have not repeated numerical results for the large-time case at the current time, since it is intuitively fairly clear that our large maturity formula is correct (since it just boils down to computing the stable fixed point of the VIE) and for maturities ≈30 years with a small step size, the code would take a prohibitively long time to give good results given that each simulation takes $O(N^2)$ for a rough model (where $N$ is the number of time steps), and it is difficult to verify the formula numerically even for the standard Heston model.

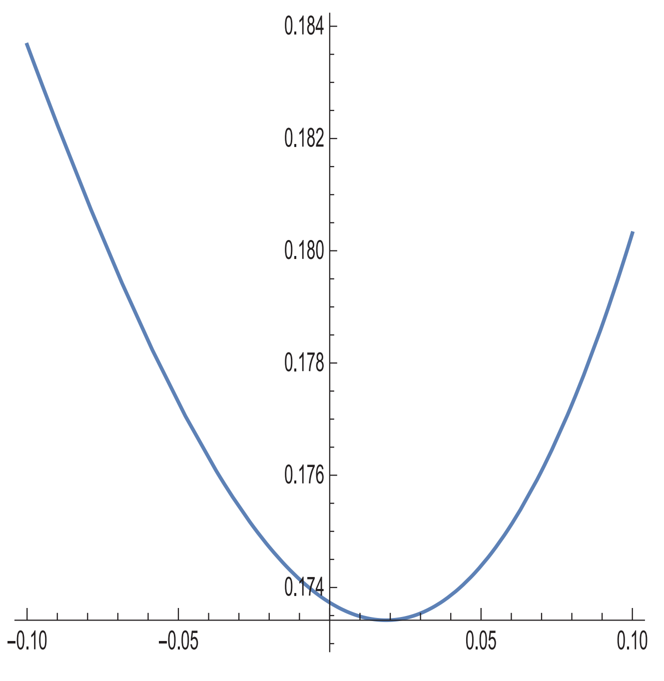


**FIGURE 5** Here we have plotted the $H = 0$ asymptotic short-maturity smile (i.e., $\hat{\sigma}_0(x)$ in (51)), for $\nu = .2$, $\rho = -0.1$, and $V_0 = .04$ [Color figure can be viewed at wileyonlinelibrary.com]
*Note*. We have used a 10-term small-$t$ series approximation to the solution to (48) combined with the scaling property in (49), and the Alan Lewis Fourier inversion formula for call options given in, for example, eq (1.4) in El Euch, Gatheral & Rosenbaum (2018) using Gauss–Legendre quadrature for the inverse Fourier transform with 1,600 points over a range of [0,40].

**Proposition 4.2.**

$$V(p) := \lim_{t\to\infty}\frac{1}{t}\log\mathbb{E}(e^{pX_t}) = \begin{cases}\lambda\theta U_1(p) & p\in[\underline{p},\bar{p}],\\ +\infty & p\notin[\underline{p},\bar{p}].\end{cases}$$

*Proof.* Gerhold et al. (2019) show that the explosion time for the Rough Heston model $T^*(p)<\infty$ if and only if $T^*(p)<\infty$ for the corresponding standard Heston model (i.e., the case $\alpha=1$).

From the usual quadratic solution formula $\frac{-b\pm\sqrt{b^2-4ac}}{2a}$, we know that $H(p,.)$ has two distinct real roots (or a single root) if and only if

$$(\lambda-\rho p\nu)^2 \geq (p^2-p)\nu^2, \tag{46}$$

which is the same as the condition $e_1(p)\geq 0$ in condition (C) in Gerhold et al. (2019). We note that $\bar{p}$, $\underline{p}$ are the zeros of $e_1(p)$.

We now have to verify that under our assumptions that $\lambda>0$ and $\rho\leq 0$, $T^*(p)<\infty$ if and only $e_1(p)<0$. We have two cases to consider to verify this claim:

- Suppose $e_1(p)\geq 0$. Then case B in Gerhold et al. (2019) is impossible by definition, and $p\in[\underline{p},\bar{p}]$, and eq. (3.5) in Forde and Jacquier (2011a) is satisfied. Eq. (3.4) in Forde and Jacquier (2011a) is

$$\lambda>\rho\nu p$$

  in our current notation, and by the assertion on p. 769 in Forde and Jacquier (2011a) that "(3.4) is implied by (3.5)", we see that it holds, which is equivalent to $e_0(p)<0$. Therefore, case A is impossible. So we are in the nonexplosive case C or D of the Gerhold et al. (2019) classification. We note that case C is by definition equivalent now to $c_1(p)>0$.
- Suppose $e_1(p)<0$. By definition we are not in case C. And we have $p\notin[\underline{p},\bar{p}]$, but from p. 769 in Forde and Jacquier (2011a), we know the interval [0,1] is strictly contained in $[\underline{p},\bar{p}]$. Hence, case D is also impossible, and we are in the explosive cases A or B.

Hence our claim is verified. We can now rewrite (45) in integral form as

$$f(p,t)=\frac{1}{\Gamma(\alpha)}\int_0^t (t-s)^{\alpha-1}H(p,f(p,s))ds.$$

Clearly, we have $H(p,w)\searrow 0$ as $w\nearrow U_1(p)$. Assume to begin with that $U_1(p)>0$ (by an easy calculation, this is exactly case C in the Gerhold et al. (2019) classification). Then from the proof of proposition 4 in Gerhold et al. (2019), we know that $0\leq f(p,t)\leq U_1(p)$.

Moreover, $w^*=U_1(p)$ is the smallest root of $H(p,w)$, so $H(p,w)\geq H_\delta := H(p,U_1(p)-\delta)$ for $w\leq U_1(p)-\delta$ and $\delta\in(0,U_1(p))$; hence we must have

$$\frac{H_\delta}{\Gamma(\alpha)}\int_0^t (t-s)^{\alpha-1}1_{f(p,s)\leq U_1(p)-\delta}\,ds < U_1(p)$$

for all $t > 0$. This implies that $\frac{H_\delta}{\Gamma(\alpha)}(t-1)^{\alpha-1}\int_1^t 1_{f(p,s)\le U_1(p)-\delta}ds < U_1(p)$, or equivalently

$$t - 1 - \int_1^t 1_{f(p,s)>U_1(p)-\delta}\,ds \le \frac{\Gamma(\alpha)}{H_\delta}U_1(p)(t-1)^{1-\alpha}.$$

Then we see that

$$\begin{aligned}\frac{1}{t}\int_0^t f(p,s)ds \ge \frac{1}{t}\int_1^t f(p,s)ds \ge \frac{1}{t}\int_1^t f(p,s)1_{f(p,s)>U_1(p)-\delta}ds\\ &\ge \frac{1}{t}(U_1(p)-\delta)\left(t-1-\frac{\Gamma(\alpha)}{H_\delta}U_1(p)(t-1)^{1-\alpha}\right)\\ &\ge U_1(p)-2\delta\end{aligned}$$

for $t$ sufficiently large. Thus, $U_1(p)-2\delta \le \frac{1}{t}\int_0^t f(p,s)ds \le U_1(p)$, so $\frac{1}{t}\int_0^t f(p,s)ds \to U_1(p)$ as $t \to \infty$. Then using that

$$\log \mathbb{E}(e^{pX_t}) = V_0 I^{1-\alpha}f(p,t) + \lambda\theta I f(p,t)$$

and that $f(p,t)$ is bounded, the result follows. We proceed similarly for the case $U_1(p) < 0$ (i.e., case D in the Gerhold et al. (2019) classification, see also Lemma 4.5). □

**Corollary 4.3.** *$X_t/t$ satisfies the LDP as $t \to \infty$ with speed $t$ and rate function $V^*(x)$ equal to the Fenchel–Legendre transform of $V(p)$, as for the standard Heston model.*

*Proof.* Since $U_1'(p) \to +\infty$ as $p \to \bar{p}$ and $U_1'(p) \to -\infty$ as $p \to \underline{p}$, the function $\lambda\theta U_1(p)$ is essentially smooth; so the stated LDP follows from the Gärtner–Ellis theorem in large deviations theory. □

*Remark* 4.4. We can easily add stochastic interest rates into this model by modeling the short rate $r_t$ by an independent Rough Heston process, and proceeding as in Forde and Kumar (2016) (we omit the details), see also Forde (2011).

Note that we have not proved that $f(p,t) \to U_1(p)$, but to establish the leading order behavior in Proposition 4.2, this is not necessary, rather we only needed to show that $I^1 f(p,t) \sim tU_1(p)$. Nevertheless, this convergence would be required to go to higher order, so for completeness we prove this property as well, as a special case of the following general result:

**Lemma 4.5.** *Consider functions $G(y)$ and $K(z)$ which satisfy the following:*

- *$G(y)$ is analytic and increasing on $[0, y_0]$ and decreasing on $[y_0, \infty)$ where $y_0 \ge 0$;*
- *$G(0) \ge 0$;*
- *$K(z)$ is positive, continuous, and strictly decreasing for $z > 0$;*
- *$\int_0^t K(z)dz$ is finite for each $t > 0$ and diverges as $t \to \infty$;*
- *$K(z+\alpha)/K(z)$ is strictly increasing in $z$ for each fixed $\alpha$ greater than zero.*

*Then the solution to* $y(t) = \int_0^t K(t-s)G(y(s))ds$ *is monotonically increasing, and if G has at least one positive root then* $y(t)$ *converges to the smallest positive root of G as* $t \to \infty$.

*Proof.* See Appendix C. □

This lemma can be applied to both cases C and D. As shown in Gerhold et al. (2019), the solution in case C is bounded between zero and the smallest positive root of $G$ (denoted $a$ in that paper) so $G$ need only satisfy the conditions of the above lemma on the interval $[0, a]$ which it does with $y_0 = 0$. For case D, multiplying the defining integral equation by $-1$ and applying the transformations $-y(t) \to y(t)$ and $-G(-y(t)) \to G(y(t))$ (see final plot in Figure 3) we recover an integral equation of the desired form (again $G$ need only satisfy the conditions of the lemma over the corresponding interval $[0, a]$).

## 4.1 | Asymptotics for call options and implied volatility

**Corollary 4.6.** *We have the following large-time asymptotic behavior for European put/call options in the large-time, large log-moneyness regime:*

$$-\lim_{t\to\infty}\frac{1}{t}\log\mathbb{E}(S_t - S_0e^{xt})^+ = V^*(x) - x \qquad (x \geq \frac{1}{2}\bar{\theta}),$$
$$-\lim_{t\to\infty}\frac{1}{t}\log(S_0 - \mathbb{E}(S_t - S_0e^{xt})^+) = V^*(x) - x \qquad (-\frac{1}{2}\theta \leq x \leq \frac{1}{2}\bar{\theta}),$$
$$-\lim_{t\to\infty}\frac{1}{t}\log(\mathbb{E}(S_0e^{xt} - S_t)^+) = V^*(x) - x \qquad (x \leq -\frac{1}{2}\theta),$$

*where* $\bar{\theta} = \frac{\lambda\theta}{\lambda-\rho\nu}$.

*Proof.* See corollary 2.4 in Forde and Jacquier (2011a). □

**Corollary 4.7.** *We have the following asymptotic behavior in the large-time, large log-moneyness regime, where* $\hat{\sigma}_t(kt)$ *is the implied volatility of a European put/call option with strike* $S_0e^{xt}$*:*

$$\hat{\sigma}_\infty(x)^2 = \lim_{t\to\infty}\hat{\sigma}_t^2(xt) = \frac{\omega_1}{2}(1 + \omega_2\rho x + \sqrt{(\omega_2 x + \rho)^2 + \bar{\rho}^2}),$$

*where*

$$\omega_1 = \frac{4\lambda\theta}{\nu^2\bar{\rho}^2}[\sqrt{(2\lambda-\rho\nu)^2 + \nu^2\bar{\rho}^2} - (2\lambda - \rho\nu)], \quad \omega_2 = \frac{\nu}{\lambda\theta}.$$

*Proof.* See proposition 1 in Gatheral and Jacquier (2011) (note that for the Rough Heston model $\lambda$ has to be replaced with $\frac{\lambda}{\Gamma(\alpha)}$ and $\nu$ replaced with $\frac{\nu}{\Gamma(\alpha)}$, but the effect of the $\alpha$ here cancels out in the final formula for $\hat{\sigma}_\infty(k)$. □

### 4.2 | Higher-order large-time behavior

We can formally try going to higher order; indeed, using the ansatz $f(p,t) = U_1(p)t + U_2(p)t^{-\alpha}(1+o(1))$ for $p \in [\underline{p}, \bar{p}]$, and we find that

$$U_2(p) = -\frac{U_1(p)}{(\lambda - U_1(p)\nu^2 - p\rho\nu)\Gamma(1-\alpha)}$$

but if we try and go higher order again, the fractional derivative on the left-hand side of (10) does not exist. Using the same approach as in Forde et al. (2011), one should be able to use this to compute a higher-order large-time saddle point approximation for call options. For the sake of brevity, we defer the details of this for future work.

## 5 | ASYMPTOTICS IN THE $H \to 0$ LIMIT

In this section, we will show that for fixed $t$, the log stock price $X_t^{(\alpha)} := X_t$ converges as $\alpha \to \frac{1}{2}$ that is, as $H \to 0$ in an appropriate sense. To match the assumptions of theorem 13.1.1 on p. 384 of Gripenberg et al. (1990) (on the continuity of the solutions to a parameterized family of VIEs), we define $h(\alpha, w) := G(p, w)$ for $\alpha \geq \frac{1}{2}$ (which is independent of $\alpha$). The kernel $a(t,s,\alpha) := (t-s)^{\alpha-1}/\Gamma(\alpha)$ is of continuous type; see definition 9.5.2 in Gripenberg et al. (1990), and the remark to theorem 12.1.1 in Gripenberg et al. (1990), which states local integrability of $k$ as a sufficient condition for this property, and we can easily verify that

$$\sup_{t\in[0,T]} \left| \int_0^t (a(t,s,\alpha) - a(t,s,\tfrac{1}{2}))ds \right| \to 0$$

as $\alpha \to \frac{1}{2}$, so the uniform continuity assumption in theorem 13.1.1 of Gripenberg et al. (1990) is satisfied. Moreover, the solution to the VIE is unique for $\alpha \in (0,1)$, see theorem 3.1.4 in Brunner (2017), or Satz 1 in Dinghas (1958). Note that the Lipschitz condition (3.1) in Dinghas (1958) has a fixed Lipschitz constant $\Gamma(\alpha+1)$, but since the function $H$ defining our VIE (see (45)) does not depend on time, the factor $t^\alpha$ on the left-hand side of condition (3.1) in Dinghas (1958) (using our notation) allows for an arbitrary Lipschitz constant, on a sufficiently small time interval. Moreover, once uniqueness on a small time interval is established, there is a unique continuation (if any) by a standard extension procedure described on p. 107 of Brunner (2017).

Then from theorem 13.1.1(ii) in Gripenberg et al. (1990), $f(p,t;\alpha)$ is continuous in $\alpha$ and $t$ on $\{(\alpha,t) : \alpha \in [\frac{1}{2}, 1), 0 \leq t < \hat{T}_\alpha(p)\}$, where $[0, \hat{T}_\alpha(p))$ denotes the maximal interval on which a continuous solution of the VIE exists. Moreover, since theorem 13.1.1 of Gripenberg et al. (1990) is multidimensional, we can apply it to $(\mathrm{Re}(f), \mathrm{Im}(f))$ to conclude that $f(i\theta,t;\alpha) \to f(i\theta,t;\frac{1}{2})$ for $\theta \in \mathbb{R}$. Using the analyticity of $f(.,t,0)$, for example, from lemma 7 in Gerhold et al. (2019), we have that $f(i\theta,t;\frac{1}{2})$ is continuous at $\theta = 0$, so we can apply Lévy's convergence theorem and verify that $X_t^{(\alpha)}$ tends weakly to some random variable $X_t^{(\frac{1}{2})}$ as $\alpha \to \frac{1}{2}$, for which

$$\mathbb{E}(e^{pX_t^{(\frac{1}{2})}}) = e^{V_0 I^{\frac{1}{2}} f(p,t) + \lambda\theta I^1 f(p,t)}$$

for $p$ in some open interval $I = (p_-(t), p_+(t)) \supset [0,1]$, where $f(p,t)$ satisfies

$$D^{\frac{1}{2}} f(p,t) = \frac{1}{2}(p^2 - p) + (p\rho\nu - \lambda) f(p,t) + \frac{1}{2}\nu^2 f(p,t)^2$$

with initial condition $f(p,0) = 0$.

Thus we have a $H = 0$ "model," or more precisely a family of marginals for $X_t^{(\frac{1}{2})}$ for all $t \in [0,T]$), with nonzero skewness. This is in contrast to the Rough Bergomi model, which for the vol-of-vol $\gamma \in (0,1)$ tends to a model with zero skew in the limit as $H \to 0$ (see Forde et al., 2020, for details).

Then using similar scaling arguments to Section 3, we know that

$$\mathbb{E}(e^{pX_{\varepsilon t}^{(\frac{1}{2})}}) = e^{V_0 I^{\frac{1}{2}} f_\varepsilon(p,t) + \varepsilon^{\frac{1}{2}} \lambda \theta I^1 f_\varepsilon(p,t)}$$

for $p \in (p_-(\varepsilon t), p_+(\varepsilon t)) \supset [0,1]$, where $f_\varepsilon(p,t)$ satisfies

$$D^{\frac{1}{2}} f_\varepsilon(p,t) = \frac{1}{2}\varepsilon(p^2 - p) + \varepsilon^{\frac{1}{2}}(p\rho\nu - \lambda) f_\varepsilon(p,t) + \frac{1}{2}\nu^2 f_\varepsilon(p,t)^2$$

with initial condition $f_\varepsilon(p,0) = 0$. Then setting $f_\varepsilon(\frac{p}{\sqrt{\varepsilon}}, t) = \phi_\varepsilon(p,t)$ as in eq. 49 in Forde et al. (2019), we find that $\phi_\varepsilon(p,t)$ satisfies

$$D^{\frac{1}{2}} \phi_\varepsilon(p,t) = \frac{1}{2}p^2 - \frac{1}{2}p\sqrt{\varepsilon} + p\rho\nu\phi_\varepsilon(p,t) + \frac{1}{2}\nu^2\phi_\varepsilon(p,t)^2 - \lambda\varepsilon^{\frac{1}{2}}\phi_\varepsilon(p,t) \tag{47}$$

with $\phi_\varepsilon(p,0) = 0$, for $p \in (\frac{p_-(\varepsilon t)}{\sqrt{\varepsilon}}, \frac{p_+(\varepsilon t)}{\sqrt{\varepsilon}})$. We can then apply theorem 13.1.1 in Gripenberg et al. (1990) as above to show that $\phi_\varepsilon(p,t)$ tends to the solution $\phi$ of

$$D^{\frac{1}{2}} \phi(p,t) = \frac{1}{2}p^2 + p\rho\nu\phi(p,t) + \frac{1}{2}\nu^2\phi(p,t)^2 \tag{48}$$

as $\varepsilon \to 0$ for $p \in (p_-^0, p_+^0)$ where $p_\pm^0 := \lim_{\varepsilon \to 0} \frac{p_\pm(\varepsilon t)}{\sqrt{\varepsilon}}$. Thus setting $t = 1$, we see (again using Lévy's convergence theorem) that $X_\varepsilon^{(\frac{1}{2})}/\sqrt{\varepsilon}$ tends weakly to a (non-Gaussian) random variable $Z$ as $t \to 0$ for which $\mathbb{E}(e^{pZ}) = e^{V_0 I^{\frac{1}{2}}\phi(p,.)(1)}$. Two interesting and difficult open questions now arise: (a) is this property *time-consistent*, that is, does it remain true at a future time $t$ when we condition on the history of $V$ up to $t$, and (b) is $V$ itself a well-defined process in the $\alpha \to \frac{1}{2}$ limit, or does it, for example, tend to a non-Gaussian field which is not pointwise defined. We answer the second question in Subsections 5.2 and 5.3.

*Remark* 5.1. Note that the scaling property in this case simplifies to

$$\Lambda(p,t) = \Lambda(pt^{\frac{1}{2}}, 1), \tag{49}$$

where $\Lambda(p,t) := I^{1-\alpha}\phi(p,t)$ with $\alpha = \frac{1}{2}$.

### 5.1 | Implied vol asymptotics in the $H = 0, t \to 0$ limit—full smile effect for the Edgeworth FX options regime

Following a similar argument to lemma 5 in Mijatovic & Tankov (2016) one can establish the following small-time behavior for European put options in the Edgeworth regime:

$$\frac{1}{\sqrt{t}}\mathbb{E}((e^{x\sqrt{t}} - e^{X_t})^+) \quad \sim \quad e^{x\sqrt{t}}\,\mathbb{E}((x - \frac{X_t}{\sqrt{t}})^+) \quad \sim \quad \mathbb{E}((x - \frac{X_t}{\sqrt{t}})^+) \sim P(x) := \mathbb{E}((x - Z)^+)$$

as $t \to 0$, where $Z$ is the non-Gaussian random variable defined in the previous subsection, and $f \sim g$ here means that $f/g \to 1$. From, for example, Fukasawa (2017) or lemma 3.3 in Forde et al. (2019), we know that for the Black–Scholes model with volatility $\sigma$

$$\frac{1}{\sqrt{t}}\mathbb{E}((e^{x\sqrt{t}} - e^{X_t})^+) \sim P_B(x,\sigma) \quad := \quad \mathbb{E}((x - \sigma W_1)^+), \tag{50}$$

where $W$ is a standard Brownian motion. From this, we can easily deduce that

$$\hat{\sigma}_0(x) := \lim_{t\to 0}\hat{\sigma}_t(x\sqrt{t}, t) = P_B(x,.)^{-1}(P(x)) \tag{51}$$

for $x > 0$, where $\hat{\sigma}_t(x,t)$ denotes the implied volatility of a European put option with strike $e^{x\sqrt{t}}$, maturity $t$, and $S_0 = 1$, and $P_B(x,\sigma)$ is the Bachelier model put price formula. Hence, we see the full smile effect in the small-time FX options Edgeworth regime unlike the $H > 0$ case discussed in, for example, Fukasawa (2017), El Euch et al. (2019), and Forde et al. (2019), where the leading order term is just Black–Scholes, followed by a next order skew term, followed by an even higher-order term.

### 5.2 | A closed-form expression for the skew, the $H \to 0$ limit, and calibrating a time-dependent correlation function

We now consider a driftless version of the model where $dX_t = \sqrt{V_t}dB_t$ and $V_t = V_0 + \frac{1}{\Gamma(\alpha)}\int_0^t (t - s)^{\alpha-1}\nu\sqrt{V_s}dW_s$. Then

$$\mathbb{E}(X_T^3) = 3\mathbb{E}(X_T\langle X\rangle_T) = 3\mathbb{E}(\int_0^T \sqrt{V}_s(\rho dW_s + \bar{\rho}dB_s)\int_0^T V_t dt) = 3\rho\,\mathbb{E}(\int_0^T \sqrt{V}_s dW_s \int_0^T V_t dt)$$

so formally we need to compute

$$\begin{aligned}\mathbb{E}(\sqrt{V}_s V_t dW_s) &= \mathbb{E}(\sqrt{V}_s(V_0 + \frac{1}{\Gamma(\alpha)}\int_0^t (t-u)^{\alpha-1}\nu\sqrt{V_u}dW_u)dW_s)\\ &= \mathbb{E}(\sqrt{V}_s\frac{\nu}{\Gamma(\alpha)}(t-s)^{\alpha-1}\sqrt{V_s}ds\,1_{s<t})\\ &= \frac{\nu}{\Gamma(\alpha)}(t-s)^{\alpha-1}\,1_{s<t}\,\mathbb{E}(V_s)ds = \frac{\nu}{\Gamma(\alpha)}(t-s)^{\alpha-1}\,1_{s<t}\,V_0 ds.\end{aligned}$$

Thus

$$\mathbb{E}(X_T^3) = 3\rho \int_0^T \int_0^t \mathbb{E}(\sqrt{V}_s V_t dW_s) = \frac{3V_0 \rho \nu T^{1+\alpha}}{\Gamma(\alpha)\alpha(1+\alpha)}. \tag{52}$$

If we now relax the assumption that $V$ is driftless and assume a given initial variance curve $\xi_0(t)$ and a general $L^2$ kernel $\kappa$ then

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

(where $\kappa$ is computed in Proposition 2.2). Then

$$\mathbb{E}(\sqrt{V}_s V_t dW_s) = \mathbb{E}(\sqrt{V}_s(\xi_0(t) + \int_0^t \kappa(t-u)\sqrt{V_u} dW_u) dW_s) = \kappa(t-s)\, 1_{s<t} \mathbb{E}(V_s)$$

and

$$\mathbb{E}(X_T^3) = 3\rho \int_0^T \int_0^t \mathbb{E}(\sqrt{V}_s V_t dW_s) = 3\rho \int_0^T \int_0^t \kappa(t-s)\, \xi_0(s) ds dt.$$

*Remark* 5.2. If we allow $\rho$ to be time-dependent, then $\mathbb{E}(X_t^3) = 3\rho(t) \int_0^T \int_0^t \kappa(t-s)\, \xi_0(s) ds dt$ and we can use this equation to calibrate $\rho(t)$ to the observed *skew term structure*, that is, the value of $\mathbb{E}(X_t^3)$ at each $t$ in some interval $[0, T]$ implied by European option prices via the Breeden–Litzenberger formula. Note we have ignored the drift terms of $X$ to simplify the computations here but in the small-time limit these drift terms will be higher order.

## 5.3 | Weak convergence of the $V$ process on pathspace to a tempered distribution, and the hyper-rough Heston model

From theorem 4.3 in Abi Jaber et al. (2019) with $\alpha \in (\frac{1}{2}, 1)$, $a(v) = \nu^2 v$, $\sigma(v) = \nu\sqrt{v}$, $b(v) = \lambda(\theta - v)$, $A(v) = \nu^2 v$, and $f \in L^1([0, T])$, we know that

$$\mathbb{E}(e^{\int_0^T f(T-t) V_t dt}) = e^{V_0 \int_0^T f(t) dt + \frac{1}{2}\nu^2 V_0 \int_0^T \psi_\alpha(t)^2 dt},$$

where $\psi_\alpha$ satisfies the Ricatti–Volterra equation:

$$\psi_\alpha(t) = \int_0^t c_\alpha (t-s)^{\alpha-1} (f(s) + \frac{1}{2}\nu^2 \psi_\alpha(s)^2) ds \tag{53}$$

and $c_\alpha = \frac{1}{\Gamma(\alpha)}$.

**Proposition 5.3.** *$V$ tends to a random tempered distribution $V^{(\frac{1}{2})}$ in distribution as $\alpha \to \frac{1}{2}$ with respect to the strong and weak topologies (see p. 2 in Bierme, Durieu, and Wang, 2017, for definitions),* *where* $V^{(\frac{1}{2})}$ *is a random tempered distribution*[2] *and for all* $f$ *in the Schwartz space* $\mathcal{S}$ *we have*

$$\mathbb{E}(e^{\int_0^T f(T-t)V_t^{(\frac{1}{2})}dt}) = e^{V_0 \int_0^T f(t)dt + \frac{1}{2}\nu^2 V_0 \int_0^T \psi(t)^2 dt},$$

*where* $\psi$ *satisfies the following VIE:*

$$\psi(t) = \int_0^t c_{\frac{1}{2}}(t-s)^{-\frac{1}{2}}(f(s) + \frac{1}{2}\nu^2\psi(s)^2)ds.$$

*Proof.* See Appendix D. □

Let $A_t$ satisfy $A_t = V_0 t + \frac{\nu}{\Gamma(\frac{1}{2})}\int_0^t (t-s)^{-\frac{1}{2}} W_{A_s} ds$. Then $A_t$ is of the same form as $X_t$ in Abi Jaber (2019), with their $dG_0(t) = V_0 dt$. Then from theorem 2.5 in Abi Jaber (2019) (with $a = b = 0$ and $c = \nu^2$) we know that

$$\begin{aligned}\mathbb{E}(e^{\int_0^T f(T-t)dA_t}) = e^{\int_0^T F(T-s,\psi(T-s))dG_0(s)} &= e^{V_0 \int_0^T (f(T-s)+\frac{1}{2}\nu^2\psi(T-s)^2)ds} \\ &= e^{V_0 \int_0^T (f(s)+\frac{1}{2}\nu^2\psi(s)^2)ds},\end{aligned} \tag{54}$$

where $F(s,u) = f(u) + \frac{1}{2}cu^2$, and $\psi$ satisfies

$$\psi(t) = \int_0^t K(t-s)F(s,\psi(s))ds = \int_0^t c_{\frac{1}{2}}(t-s)^{-\frac{1}{2}}(f(s) + \frac{1}{2}\nu^2\psi(s)^2)ds.$$

The process $A_t$ here is the driftless *hyper-rough* Heston model for $H = 0$ discussed in the next subsection, and note that $\psi$ satisfies the same VIE as (53) (and by, e.g., theorem 3.1.4 in Brunner, 2017, we know the solution is unique), so the limiting field $V^{(\frac{1}{2})}$ has the same law as the random measure $dA_t$. Moreover, from proposition 4.6 in Jusselin and Rosenbaum (2020) (which uses the law of the iterated logarithm for $B$) $A$ is a.s. not continuously differentiable but is only known to be $2\alpha - \varepsilon$ Hölder continuous for all $\varepsilon > 0$. Hence $A$ exhibits (non-Gaussian) "field"-type behavior.

## 5.4 | The hyper-rough Heston model for $H = 0$—driftless and general cases

If $\lambda = 0$ and $\alpha \in (\frac{1}{2}, 1)$ and we set $A_t := \int_0^t V_s ds$, then using the stochastic Fubini theorem, we see that

$$\begin{aligned}A_t - V_0 t &= \frac{1}{\Gamma(\alpha)}\int_0^t \int_0^s (s-u)^{\alpha-1}\nu\sqrt{V_u}dW_u ds = \frac{1}{\Gamma(\alpha)}\int_0^t \nu\sqrt{V_u}dW_u \int_u^t (s-u)^{\alpha-1}ds \\ &= \frac{\nu}{\alpha\Gamma(\alpha)}\int_0^t (t-u)^{\alpha}\sqrt{V_u}dW_u\end{aligned}$$

(using Dambis-Dubins-Schwarz time change)

$$= \frac{\nu}{\alpha\Gamma(\alpha)} \int_0^t (t-u)^{\alpha} dB_{A_u}$$

(where $B_t := X_{T_t}, T_t = \inf\{s : A_s > t\}$) so $B$ is a Brownian motion)

$$= \frac{\nu}{\alpha\Gamma(\alpha)} B_{A_u}(t-u)^{\alpha}|_{u=0}^{t} + \frac{\nu}{\Gamma(\alpha)} \int_0^t (t-u)^{\alpha-1} B_{A_u} du$$

$$= \nu I^{\alpha} B_{A_t}.$$

We can now take

$$A_t = V_0 t + \nu I^{\alpha} B_{A_t} \tag{55}$$

as the *definition* of the Rough Heston model for $\alpha \in [\frac{1}{2}, 1)$ (i.e., allowing for the possibility that $\alpha = \frac{1}{2}$), where $B$ is now a *given* Brownian motion (this is the so-called *hyper-rough Heston* model introduced in Jusselin and Rosenbaum (2020) for the case of zero drift. Note that for a given sample path $B_t(\omega)$, we can regard (55) as a (random) fractional ordinary differential equation (ODE) of the form:

$$A(t) = V_0 t \, + I^{\alpha} f(A(t)), \tag{56}$$

where $f(t) = B_t(\omega)$.

### 5.4.1 | The case $\lambda > 0$

For the case when $\lambda > 0$, using (8) we see that

$$A_t - \int_0^t \xi_0(s) ds = \int_0^t \int_0^s \kappa(s-u)\sqrt{V_u} dW_u ds = \int_0^t \sqrt{V_u} \int_u^t \kappa(s-u) ds \, dW_u$$

$$= \int_0^t F(t-u)\sqrt{V_u} dW_u \text{ (where } F(t-u) = \int_u^t \kappa(s-u) ds)$$

$$= \int_0^t F(t-u) dM_u$$

(where $dM_t = \sqrt{V_t} dW_t$)

$$= \int_0^t F(t-u) dB_{A_u}$$

(where $B_t := M_{T_t}, T_t = \inf\{s : A_s > t\}$) so $B$ is a Brownian motion)

$$= B_{A_u} F(t-u)|_{u=0}^{t} + \int_0^t \kappa(t-u) B_{A_u} du$$

$$= \int_0^t \kappa(t-u) B_{A_u} du,$$

where we have used (7) to verify that $F(t-u) \to 0$ as $u \to t$.

## ACKNOWLEDGMENTS

We thank Peter Friz, Eduardo Abi Jaber, Martin Larsson, and Martin Keller-Ressel for helpful discussions. S. Gerhold acknowledges financial support from the Austrian Science Fund (FWF) under grant P 30750.

## ORCID

*Martin Forde* https://orcid.org/0000-0002-1606-9921
*Stefan Gerhold* https://orcid.org/0000-0002-4172-3956

## ENDNOTES

[1] We thank Peter Friz and Paolo Pigato for clarifying the main steps in this result.
[2] see, for example, Duplantier, Rhodes, Sheffield, and Vargas (2017) for more details on tempered distributions.

## REFERENCES

Abi Jaber, E. (2019). Weak existence and uniqueness for affine stochastic Volterra equations with $L^1$-kernels (preprint).

Abi Jaber, E., Larsson, M., & Pulido, S. (2019). Affine Volterra processes. *Annals of Applied Probability*, *29*(5), 3155–3200.

Bayer, C., Friz, P. K., Gulisashvili, A., Horvath, B., & Stemper, B. (2018). Short-time near-the-money skew in rough fractional volatility models. *Quantitative Finance*, *19*(5), 779–798.

Bierme, H., Durieu, O., & Wang, Y. (2017). Generalized random fields and Lévy's continuity theorem on the space of tempered distributions. Preprint, 2017.

Brunner, H. (2017). *Volterra integral equations*. Cambridge: Cambridge University Press.

Dandapani, A., Jusselin, P., & Rosenbaum, M. (2019). From quadratic Hawkes processes to super-Heston rough volatility models with Zumbach effect. Arxiv preprint.

Dembo, A., & Zeitouni, O. (1998). *Large deviations techniques and applications*. Boston: Jones and Bartlet Publishers.

Dinghas, A. (1958). Zur existenz von fixpunkten bei abbildungen vom abel-liouvilleschen typus. *Mathematische Zeitschrift*, *70*, 174–189.

Duplantier, D., Rhodes, R., Sheffield, S., & Vargas, V. (2017). Log-correlated Gaussian fields: An overview. *Geometry, Analysis and Probability*, *310*, 191–216. https://doi.org/10.1007/978-3-319-49638-2_9

El Euch, O., Fukasawa, M., Gatheral, J., & Rosenbaum, M. (2019). Short-term at-the-money asymptotics under stochastic volatility models. *SIAM Journal on Financial Mathmatics*, *10*(2), 491–511.

El Euch, O., Fukasawa, M., & Rosenbaum, M. (2018). The microstructural foundations of leverage effect and rough volatility. *Finance and Stochastics*, *12*(6), 241–280.

El Euch, O., Gatheral, J., & Rosenbaum, M. (2018). Roughening Heston. *Risk*, pp. 84–89, May 2019.

El Euch, O., & Rosenbaum, M. (2018). Perfect hedging in Rough Heston models. *Annals of Applied Probability*, *28*(6), 3813–3856.

El Euch, O., & Rosenbaum, M. (2019). The characteristic function of Rough Heston models. *Mathematical Finance*, *29*(1), 3–38.

Forde, M. (2011). Large-time asymptotics for an uncorrelated stochastic volatility model. *Statistics & Probability Letters*, *81*(8), 1230–1232.

Forde, M., Fukasawa, M., Gerhold, S., & Smith, B. (2020). The Rough Bergomi model as $H \to 0$—Skew flattening/blow up and non-Gaussian Rough volatility. Preprint.

Forde, M., & Jacquier, A. (2011a). The large-maturity smile for the Heston model. *Finance and Stochastics*, *15*, 755–780.

Forde, M., & Jacquier, A. (2011b). Small-time asymptotics for an uncorrelated local-stochastic volatility model. *Applied Mathematical Finance*, *18*, 517–535.

Forde, M., Jacquier, A., & Lee, R. (2012). The small-time smile and term structure of implied volatility under the Heston model. *SIAM Journal on Financial Mathematics*, *3*, 690–708.

Forde, M., Jacquier, A., & Mijatovic, A. (2011). A note on essential smoothness in the Heston model. *Finance and Stochastics*, *15*, 781–784.

Forde, M., & Kumar, R. (2016). Large-time option pricing using the Donsker-Varadhan LDP—Correlated stochastic volatility with stochastic interest rates and jumps. *Annals of Applied Probability*, *6*, 3699–3726.

Forde, M., Smith, B., & Viitasaari, L. (2019). Rough volatility and CGMY jumps with a finite history and the Rough Heston model—Small-time asymptotics in the $k\sqrt{t}$ regime. Preprint.

Forde, M., & Zhang, H. (2017). Asymptotics for Rough stochastic volatility models. *SIAM Journal on Financial Mathematics*, *8*, 114–145.

Friz, P., Gerhold, S., & Pinter, A. (2018). Option pricing in the moderate deviations regime. *Mathematical Finance*, *28*(3), 962–988.

Friz, P. K., Gassiat, P., & Pigato, P. (2018). Precise asymptotics: Robust stochastic volatility models. Preprint.

Fukasawa, M. (2017). Short-time at-the-money skew and Rough fractional volatility. *Quantitative Finance*, *17*(2), 189–198.

Gao, K., & Lee, R. (2014). Asymptotics of implied volatility to arbitrary order. *Finance and Stochastics*, *18*, 349–392.

Gatheral, J., & Jacquier, A. (2011). Convergence of Heston to SVI. *Quantitative Finance*, *11*(8), 1129–1132.

Gatheral, J., & Keller-Ressel, M. (2019). Affine forward variance models. *Finance and Stochastics*, *23*(3), 501–533. https://doi.org/10.2139/ssrn.3105387.

Gerhold, S., Gerstenecker, C., & Pinter, A. (2019). Moment explosions in the rough heston model. *Decisions in Economics and Finance*, *42*(2), 575–608.

Gripenberg, G., Londen, S. O., & Staffans, O. (1990). *Volterra integral and functional equations*. Cambridge: Cambridge University Press.

Jacquier, A., & Pannier, A. (2020). Large and moderate deviations for stochastic Volterra systems. Preprint.

Jaisson, T., & Rosenbaum, M. (2016). Rough fractional diffusions as scaling limits of nearly unstable heavy tailed Hawkes processes. *Annals of Applied Probability*, *26*(5), 2860–2882.

Jusselin, P., & Rosenbaum, M. (2020). No-arbitrage implies power-law market impact and Rough volatility. *Mathematical Finance*, *30*(4), 1309–1336.

Lord, R., & Kahl, C. (2007). *Optimal Fourier inversion in semi-analytical option pricing*. Tinbergen Institute Discussion Paper No. 2006-066/2.

Mann, W. R., & Wolf, F. (1951). Heat transfer between solids and gases under nonlinear boundary conditions. *Quarterly of Applied Mathematics*, *9*(2), 163–184.

Mijatovic, A. & Tankov, P. (2016). A new look at short-term implied volatility in asset price models with jumps. *Mathematical Finance*, *26*(1), 149–183.

Miller, R. K., & Feldstein, A. (1971). Smoothness of solutions of volterra integral equations with weakly singular kernels. *SIAM Journal on Mathematical Analysis*, *2*(2), 242–258.

Olver, F. W. (1974). *Asymptotics and special functions*. New York: Academic Press.

Roberts, C. A., & Olmstead, W. E. (1996). Growth rates for blow-up solutions of nonlinear Volterra equations. *Quarterly of Applied Mathematics*, *54*(1), 153–159.

**How to cite this article:** Forde M, Gerhold S, Smith B. Small-time, large-time, and $H \to 0$ asymptotics for the Rough Heston model. *Mathematical Finance*. 2021;31:203–241. https://doi.org/10.1111/mafi.12290

## APPENDIX A: COMPUTING THE KERNEL FOR THE ROUGH HESTON VARIANCE CURVE

Let $Z_t = \int_0^t \sqrt{V_s} dW_s$, and we recall that

$$V_t = V_0 + \frac{1}{\Gamma(\alpha)} \int_0^t (t-s)^{\alpha-1} \lambda(\theta - V_s) ds + \frac{1}{\Gamma(\alpha)} \int_0^t (t-s)^{\alpha-1} \nu \sqrt{V_s} dW_s$$
$$= \tilde{\xi}_0(t) - \frac{\lambda}{\nu}(\varphi * V) + \varphi * dZ,$$

where $*$ denotes the convolution of two functions, $\varphi * dZ = \int_0^t \varphi(t-s) dZ_s$ and $\tilde{\xi}_0(t) = V_0 + \frac{1}{\Gamma(\alpha)} \int_0^t (t-s)^{\alpha-1} \lambda\theta ds = V_0 + \frac{\lambda\theta}{\alpha\Gamma(\alpha)} t^\alpha$, and $\varphi(t) = \frac{\nu}{\Gamma(\alpha)} t^\alpha$. Now define $\kappa$ to be the unique function which satisfies

$$\kappa = \varphi - \frac{\lambda}{\nu}(\varphi * \kappa). \tag{A.1}$$

Such a $\kappa$ exists and is known as the *resolvent* of $\varphi$. Then we see that

$$V_t - \frac{\lambda}{\nu} \kappa * V_t = \tilde{\xi}_0(t) - \frac{\lambda}{\nu} \varphi * V + \varphi * dZ - \frac{\lambda}{\nu} \kappa * [\tilde{\xi}_0(t) - \frac{\lambda}{\nu} \varphi * V + \varphi * dZ]$$
$$= \xi_0(t) - \frac{\lambda}{\nu}(\varphi - \frac{\lambda}{\nu} \kappa * \varphi) * V + (\varphi - \frac{\lambda}{\nu} \kappa * \varphi) * dZ$$
$$= \xi_0(t) - \frac{\lambda}{\nu} \kappa * V + \kappa * dZ,$$

where $\xi_0(t) = \tilde{\xi}_0(t) - \frac{\lambda}{\nu} \kappa * \tilde{\xi}_0(t)$, and we have used (A.1) in the final line. Canceling the $-\frac{\lambda}{\nu} \kappa * V$ terms, we see that

$$V_t = \xi_0(t) + \kappa * dZ = \xi_0(t) + \int_0^t \kappa(t-s) \sqrt{V_s} dW_s$$
$$\Rightarrow \qquad \xi_t(u) = \mathbb{E}(V_u | \mathcal{F}_t) = \xi_0(u) + \int_0^t \kappa(u-s) \sqrt{V_s} dW_s$$

and thus

$$d\xi_t(u) = \kappa(u-t) \sqrt{V_t} dW_t,$$

that is, the correct $\kappa$ function is the solution to (A.1). If we take the Laplace transform of (A.1), we get

$$\hat{\kappa}(z) = \hat{\varphi}(z) - \frac{\lambda}{\nu} \hat{\varphi}(z) \hat{\kappa}(z), \tag{A.2}$$

and (A.2) is just an algebraic equation now, which we can solve explicitly to get $\hat{\kappa}(z) = \frac{\hat{\varphi}(z)}{1+\frac{\lambda}{\nu}\hat{\varphi}(z)}$. But we know that $\varphi(t) = \frac{\nu}{\Gamma(\alpha)}t^{\alpha}$ whose Laplace transform is $\hat{\varphi}(z) = \nu z^{-\alpha}$, so $\hat{\kappa}(z)$ evaluates to

$$\hat{\kappa}(z) = \frac{\nu z^{-\alpha}}{1 + \lambda z^{-\alpha}}.$$

Then the inverse Laplace transform of $\hat{\kappa}(z)$ is given by

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

## APPENDIX B: THE RESCALED MODEL

We first let

$$\begin{aligned}
dX_t^{\varepsilon} &= -\frac{1}{2}\varepsilon V_t^{\varepsilon} dt + \sqrt{\varepsilon}\sqrt{V_t^{\varepsilon}} dW_t \\
V_t^{\varepsilon} - V_0 &= \frac{\varepsilon^{\gamma}}{\Gamma(\alpha)} \int_0^t (t-s)^{H-\frac{1}{2}} \lambda(\theta - V_s^{\varepsilon}) ds + \frac{\varepsilon^{H}}{\Gamma(\alpha)} \int_0^t (t-s)^{H-\frac{1}{2}} \nu \sqrt{V_s^{\varepsilon}} dW_s \\
&\overset{(d)}{=} \frac{\varepsilon^{\gamma}}{\Gamma(\alpha)} \int_0^t (t-s)^{H-\frac{1}{2}} \lambda(\theta - V_s^{\varepsilon}) ds + \frac{\varepsilon^{H-\frac{1}{2}}}{\Gamma(\alpha)} \int_0^t (t-s)^{H-\frac{1}{2}} \nu \sqrt{V_s^{\varepsilon}} dW_{\varepsilon s} \\
&= \frac{\varepsilon^{\gamma}}{\Gamma(\alpha)} \int_0^{\varepsilon t} (t - \frac{u}{\varepsilon})^{H-\frac{1}{2}} \lambda(\theta - V_{u/\varepsilon}^{\varepsilon}) \frac{1}{\varepsilon} du + \frac{\varepsilon^{H-\frac{1}{2}}}{\Gamma(\alpha)} \int_0^{\varepsilon t} (t - \frac{u}{\varepsilon})^{H-\frac{1}{2}} \nu \sqrt{V_{u/\varepsilon}^{\varepsilon}} dW_u ,
\end{aligned}$$

where we have set $u = \varepsilon s$. Now set $V'_{\varepsilon t} = V_t^{\varepsilon}$. Then

$$\begin{aligned}
V'_{\varepsilon t} - V_0 &= \frac{\varepsilon^{\gamma-1}}{\Gamma(\alpha)} \int_0^{\varepsilon t} (t - \frac{u}{\varepsilon})^{H-\frac{1}{2}} \lambda(\theta - V'_u) du + \frac{\varepsilon^{H-\frac{1}{2}}}{\Gamma(\alpha)} \int_0^{\varepsilon t} (t - \frac{u}{\varepsilon})^{H-\frac{1}{2}} \nu \sqrt{V'_u}\, dW_u \\
&= \frac{\varepsilon^{\gamma-1}}{\varepsilon^{H-\frac{1}{2}}\Gamma(\alpha)} \int_0^{\varepsilon t} (\varepsilon t - u)^{H-\frac{1}{2}} \lambda(\theta - V'_u) du + \frac{\varepsilon^{H-\frac{1}{2}}}{\varepsilon^{H-\frac{1}{2}}\Gamma(\alpha)} \int_0^{\varepsilon t} (\varepsilon t - u)^{H-\frac{1}{2}} \nu \sqrt{V'_u}\, dW_u \\
&= \frac{1}{\Gamma(\alpha)} \int_0^{\varepsilon t} (\varepsilon t - u)^{H-\frac{1}{2}} \lambda(\theta - V'_u)\, du + \frac{1}{\Gamma(\alpha)} \int_0^{\varepsilon t} (\varepsilon t - u)^{H-\frac{1}{2}} \nu \sqrt{V'_u}\, dW_u,
\end{aligned}$$

where the last line follows on setting $\gamma - 1 = H - \frac{1}{2}$, that is, $\gamma = \alpha$. Thus for this choice of $\gamma$, $V_{\varepsilon(.)} \overset{(d)}{=} V_{(.)}^{\varepsilon}$.

## APPENDIX C: PROOF OF MONOTONICITY OF THE SOLUTION FOR A GENERAL CLASS OF VOLTERRA INTEGRAL EQUATIONS

Recall that $y(t)$ satisfies

$$y(t) = \int_0^t K(t-s)G(y(s))ds.$$

One can easily verify that the kernel used for the Rough Heston model satisfies the stated properties in Lemma 4.5.

In the classical case $K(t) \equiv 1$, the integral equation clearly reduces to an ODE, and it is well known that the solution of this is at least continuously differentiable on the domain of existence. In the following, it will be assumed that the solution $y(t)$ is analytic for $t > 0$. This is proved for the kernel relevant to the Rough Heston model in Miller and Feldstein (1971) (theorem 6), see also the end of p. 14 in Gerhold et al. (2019).

What follows is a natural extension of the technique used in Mann and Wolf (1951) (theorem 8). Using the properties of convolution and differentiating under the integral sign, we have

$$y(t) = \int_0^t K(t-s)G(y(s))ds = \int_0^t K(s)G(y(t-s))ds, \tag{C.1}$$

$$y'(t) = K(t)G(0) + \int_0^t K(s)G'(y(t-s))y'(t-s)ds, \tag{C.2}$$

$$= K(t)G(0) + \int_0^t K(t-s)G'(y(s))y'(s)ds \tag{C.3}$$

$G(0) > 0$ so $y'(t) \to +\infty$ as $t \to 0^+$ and since $G(y)$ is increasing for $y \le y_0$ we have that $y'(t) > 0$ until $y(t)$ reaches $y_0$, that is, the solution increases. For $y \ge y_0$, $G(y)$ is decreasing and suppose that $y(t)$ ceases to be increasing at some point. This implies (assuming a continuous derivative) the existence of a $t_0$ and an interval $I = [t_0, t_1]$ such that $y'(t_0) = 0$ and $y'(t_1) < 0$ for all $t_1 \in I$ (if $y(t)$ and hence $y'(t)$ is analytic then the zeros of the derivative are isolated and a sufficiently small interval $I$ exists). Using the integral equation for $y'(t)$:

$$y'(t_0) = K(t_0)G(0) + \int_0^{t_0} K(t_0-s)G'(y(s))y'(s)ds = 0, \tag{C.4}$$

$$y'(t_1) = K(t_1)G(0) + \int_0^{t_0} K(t_1-s)G'(y(s))y'(s)ds + \int_{t_0}^{t_1} K(t_1-s)G'(y(s))y'(s)ds.$$

We can rewrite the kernels in the first and second terms of the expression for $y'(t_1)$ as

$$K(t_1) = \frac{K(t_1)}{K(t_0)}K(t_0), \quad K(t_1-s) = \frac{K(t_1-s)}{K(t_0-s)}K(t_0-s)$$

and we can easily check that the quotient in the second expression here decreases monotonically from $K(t_1)/K(t_0)$ to zero.

By the mean value theorem for definite integrals there exists a $\tau \in (0, t_0)$ such that:

$$\int_0^{t_0} \frac{K(t_1 - s)}{K(t_0 - s)} K(t_0 - s) G'(y(s)) y'(s) ds = \frac{K(t_1 - \tau)}{K(t_0 - \tau)} \int_0^{t_0} K(t_0 - s) G'(y(s)) y'(s) ds$$
$$= -\frac{K(t_1 - \tau)}{K(t_0 - \tau)} K(t_0) G(0), \tag{C.5}$$

where the second equality follows from (C.4). Substituting this into our expression for $y'(t_1)$:

$$y'(t_1) = \frac{K(t_1)}{K(t_0)} K(t_0) G(0) + \frac{K(t_1 - \tau)}{K(t_0 - \tau)} \int_0^{t_0} K(t_0 - s) G'(y(s)) y'(s) ds$$
$$+ \int_{t_0}^{t_1} K(t_1 - s) G'(y(s)) y'(s) ds$$
$$= K(t_0) G(0) \underbrace{\left(\frac{K(t_1)}{K(t_0)} - \frac{K(t_1 - \tau)}{K(t_0 - \tau)}\right)}_{>0} + \int_{t_0}^{t_1} K(t_1 - s) \underbrace{G'(y(s)) y'(s)}_{>0} ds > 0 \tag{C.6}$$

and we have used (C.4) in the second line. But this is a contradiction so the solution remains increasing.

As discussed elsewhere in this paper, when studying the Rough Heston model, the nonlinearity in the integral equation has the generic form $G(y) = (y - \theta_1)^2 + \theta_2$, that is, a quadratic with positive leading coefficient (for simplicity set to 1 here) and minimum of $\theta_2$ obtained at $y = \theta_1$. Depending on the values of $\{\theta_1, \theta_2\}$, the following cases due to Gerhold et al. (2019) are distinguished:

- (C) $G(0) > 0$, $\theta_1 > 0$ and $\theta_2 < 0$
- (D) $G(0) \leq 0$

Case C is already in the form considered here with $y_0 = 0$. In case D, applying the transformation $y(t) \to -y(t)$ and $-G(-y(t)) \to G(y(t))$ (reflecting in the $x$ and then $y$-axis) yields a function $G(y)$ which is a quadratic with negative leading coefficient and thus increases until it reaches it's maximum after which it decreases which is of the type considered here.

## APPENDIX D

From theorem 13.1.1(ii) in Gripenberg et al. (1990), the unique solution $\psi^{(\alpha)}$ to

$$\psi^{(\alpha)}(t) = \int_0^t c_\alpha (t - s)^{\alpha - 1} (f(s) + \frac{1}{2} \nu^2 \psi^{(\alpha)}(s)^2) ds$$

tends pointwise to the solution of

$$\psi_{\frac{1}{2}}(t) = \int_0^t c_{\frac{1}{2}}(t-s)^{-\frac{1}{2}}(f(s) + \frac{1}{2}\nu^2\psi_{\frac{1}{2}}(s)^2)ds,$$

which is also unique by, for example, theorem 3.1.4 in Brunner (2017). Now consider any sequence $f_\varepsilon \in \mathcal{S}$ with $\|f_\varepsilon\|_{m,j} \to 0$ as $\varepsilon \to 0$ for all $m, j \in \mathbb{N}_0^n$ for any $n \in \mathbb{N}$ (i.e., under the Schwartz space seminorm defined in eq. 1 in Bierme, Durieu, and Wang, 2017). Then the convergence here implies in particular that $f_\varepsilon$ tends to $f$ pointwise. Then from theorem 13.1.1. in Gripenberg et al. (1990), the unique solution $\psi_\varepsilon$ to

$$\psi_\varepsilon(t) = \int_0^t c_{\frac{1}{2}}(t-s)^{-\frac{1}{2}}(f_\varepsilon(s) + \frac{1}{2}\nu^2\psi_\varepsilon(s)^2)ds$$

tends pointwise to the solution to

$$\psi_0(t) = \int_0^t c_{\frac{1}{2}}(t-s)^{-\frac{1}{2}}\frac{1}{2}\nu^2\psi_0(s)^2 ds$$

which is zero. Then from Lévy's continuity theorem for generalized random fields in the space of tempered distributions (see theorem 2.3 and corollary 2.4 in Bierme, Durieu, & Wang, 2017), we obtain the stated result.