Late stage review report

Student: Ryan McCrickerd | Supervisor: Dr Mikko Pakkanen

19th June 2019

Abstract

Following the LSR Guidelines, this document describes the research problems I intend to address in my thesis, and explains which of these are already, part or not completed. To identify the extent to which I consider research on particular topics completed, I have colour-coded all subsection headers: green means mostly, if not entirely, completed, orange means part completed, with no surprises expected towards completion, and red means not completed. I plan to have completed a draft of my thesis after half of the next academic year, to allow for some time to submit the three main chapters for publication (if not before) and to make improvements where possible or necessary. Although the *Future work* section here outlines a plan for roughly the next six months, a lot of the material here (mostly applied) was actually completed in the first year of the PhD programme, was included in the Early Stage Review report, and helped to motivate everything else. All of the material presented as such here is my own work (although some ideas in Section 2 originated from discussions with Eduardo Abi Jaber), completed under the supervision of Mikko Pakkanen, with also some guidance from Martin Rasmussen. For completeness, the working title for my thesis is currently:

The Heston model: reversionary regimes, pathwise limits and foreign exchange applications.

Contents

1 Introduction and overview

This section provides a general overview of the material which I intend for my thesis to contain. The very general goal (originating in part from personal experiences in risk management), is to draw upon state-of-the-art equity and foreign exchange (FX) models in order to develop a time-homogeneous multi-currency FX framework which is tractable (so clearly some form of 'optimal' compromise) and capable of capturing the behaviour of all inherent volatility surfaces. As is 'standard' to some extent, when I refer to a model as being 'tractable', I mean that it at least admits relatively fast and accurate procedures for both calibration (to e.g. vanilla options) and simulation (for e.g. exotics or xVA pricing).

In *Heston's reconciling reversionary regimes*, we first establish some basic relationships between diffusive, rough and jump models of volatility (all Heston-related), so that tractable foundations can be set which accommodate desirable characteristics of each. We suggest that these relationships are most conveniently appreciated by way of reversionary regimes (specific parameterisations) of the classical Heston model, or more precisely of the underlying Cox-Ingersoll-Ross (CIR) volatility process. In the rough case, an intuitive relationship manifests as a connection between roughness parameters and reversionary regimes. In contrast, we appear to find infinite-activity jumps arising in a very specific limit of the diffusive model, as the reversionary timescale becomes infinitesimally short. Specifically, reconciling and extending several results of the past, for example Keller-Ressel (2011), Forde & Jacquier (2011) and Mechkov (2015), we show that, once appropriately parameterised, the joint integrated variance and Heston log-price process converges marginally to IG-NIG Lévy motion, as defined in Barndorff-Nielsen & Shephard (2001). This apparent connection leads to the notion of a *consistent jump-diffusion* process as a tractable building block for volatility models, whereby (infinite-activity) jump components are consistent with their diffusive counterparts, understood to arise *strictly* as (practically-convenient) limits.

Since the connection established by this point is weak and marginal, and since the gap between marginal and functional relationships is practically significant, we follow the next chapter's theoretical path before returning to explore FX applications of Heston extensions. The primary goal of this chapter is to establish the strongest possible mode of convergence, in order to, again, ensure that we set stable foundations before leaving deeper theory behind. The unconventional pathwise approach we adopt here rewards us with a novel framework for wider model specification and analysis, which *always* exhibits unique, positive trajectories with jump (although not necessarily Lévy) limits.

In *A pathwise volatility framework with exit-time limits*, the focus becomes a theoretical exploration of the depths of this unusual connection between popular diffusion and jump processes. We explain how part of our motivation for the approach we take originates from Prokhorov's theorem (the characterisation of weak convergence by tightness and finite-dimensional distributions), the careless application of which may be considered dangerous when, as we will find, the limiting process turns out to not be naturally characterised by its finite-dimensional distributions—exhibiting *excursions*

in the sense of Whitt (2002, Chapter 15), which are invisible in stochastically-continuous, finite-dimensional settings. The approach we instead take is embodied by Skorokhod's representation theorem, which loosely states that we can always translate questions about weak convergence onto almost sure counterparts applicable to sequences on the same probability space. In fact, as in its Billingsley (1999, Theorem 6.7) formulation, we are always able to relieve ourselves of probability measures. The study of our problem on a pathwise basis, manifesting as an initial-value problem (IVP), provides a fruitful example of this.

After the thorough classical treatment of this CIR IVP, we find that the classical integrated CIR variance process converges weakly to the IG Lévy process on some (but not all) of Skorokhod's topologies. In contrast, it will become clear that the Heston log-price process *does not* converge weakly to the NIG process on any of Skorokhod's topologies. Our approach will, however, enable us to precisely characterise limits on wider topologies, which derive from spaces of higher-dimensional or set-valued processes. Besides remarking on deeper consequences, we make sure to consolidate the most practically-relevant and immediate consequences of these findings, for example by clarifying derivative types whose (theoretical) prices will or will not converge. This motivates definitions of Heston *excursion* and *jump-diffusion* processes, which become equivalent in finite-dimensional settings.

As an exciting by-product of this chapter's approach, a door to a general probability-free framework in which one may impose any probability measure on a space of continuous driving noise paths (inducing generalisations of Brownian motion), yet always retain pathwise unique volatility solutions and desirable properties such as non-negativity. To return to rough volatility briefly, we define a rough Heston variant connected to the original, but in contrast immediately admits pathwise unique solutions, which is (at the present time) considered a difficult open question for the original.

Finally, in *Future work: foreign exchange applications*, having developed a deep theoretical understanding for the Heston process's limiting behaviour, and practically-relevant consequences, we turn our attention to classical multi-asset problems and finite-dimensional applications, so henceforth proceed with a the Heston jump-diffusion framework. In the FX setting, this section can be understood as unifying the classical diffusive framework of De Col, Gnoatto, & Grasselli (2013) with the Lévy one of Ballotta, Deelstra, & Rayée (2017). To this end, we establish a natural correlation structure between jump components which proves tractable and derive the consequential measure-change transformations, making connections with Sato (1999) when possible. The resulting framework is, by its design, consistent in the usual FX sense, that all models (including implicit ones arising from inverses and ratios of others) exhibit numéraire invariance, despite the general delicacy of Lévy measure changes.

For those willing to forgo the framework's natural time-homogeity to improve calibration accuracy further (certainly valuable when the calibration time horizon spans large economical or political events), we also exhibit the ability of jump components to perform as a local volatility decorator

in the spirit of Dupire (1994). This has the attractive features of being relatively low-dimensional; collapsing the classical two-dimensional local volatility surface into three intuitive one-dimensional curves over time, and retaining essentially-exact simulation. This is achieved tractably with or without a classical stochastic volatility component, and enables near-perfect replication of volatility surfaces provided the relevant market does not exhibit rare behaviour such as multi-modality.

2 Heston's reconciling reversionary regimes

Section summary. We recall the 'mean-reverting mixture' representations, as defined in Keller-Ressel, Larsson, & Pulido (2018), of the rough and classical Heston models of volatility. A Markovian approximation of the rough Heston model is taken, leading to a rough-inspired parameterisation of the classical Heston model, where a quasi-roughness parameter controls its 'reversionary regime'. In a specific quasi-hyper-rough regime, jumps apparently arise in classical Heston's *time-integrated* variance and price processes as the model's reversionary timescale becomes arbitrarily small. Results of Gerhold, Gülüm, & Pinter (2016) are invoked to explain why the (single-factor) classical Heston model can, for all practical purposes, generate explosive implied volatility skews. This is verified numerically.

For the most part we work from a fixed probability space $(C, \mathscr{C}, \mathbb{W})$ where $C = C_0(\mathbb{R}_+, \mathbb{R})$ is the space of real continuous paths started from zero, $\mathscr{C}$ its sigma-algebra induced by the uniform norm, and $\mathbb{W}$ a Wiener measure making a canonical process defined by $W(\omega) = \omega$ for $\omega \in C$ Brownian motion. This notation will be extended into higher dimensions as required.

2.1 Motivation for approach: emerging volatility models and tractability

For a continuous log-price process H, consider the rough Heston model of El Euch & Rosenbaum (2019), which we define here using

$$V_t = \xi + v \int_0^t (t-s)^{-\zeta} \sqrt{V_s} dW_s^1, \quad H_t = \int_0^t \sqrt{V_s} d(\bar{\rho} W^2 + \rho W^1)_s - \frac{1}{2} \int_0^t V_s ds \tag{2.1}$$

where the roughness parameter ζ takes values in $[0, \frac{1}{2})$, and gives the instantaneous variance process V Hölder regularity $\frac{1}{2} - \zeta - \epsilon$ for any $\epsilon > 0$. See Abi Jaber, Larsson, & Pulido (2017, Theorem 7.1, Example 2.3) for (weak) existence and uniqueness results for this model. As described in Keller-Ressel et al. (2018, Section 5), this model is a specific case of a 'mean-reverting mixture' class of affine rough models, wherein the kernel $k(\tau) := \tau^{-\zeta}$ here admits in general a Laplace transform representation of a suitable 'reversionary measure' μ:

$$k(\tau) = \int_{\mathbb{R}_+} e^{-x\tau} \mu(dx).$$

In this class of models, the rough Heston case of Equation 2.1 and the classical Heston case (with reversion parameter κ) correspond respectively to the measures

$$\mu(dx) = \frac{x^{\zeta-1}}{\Gamma(\zeta)}dx, \quad \mu = \delta_\kappa.$$

For clarity when we select the classical reversionary case as in Heston (1993) we obtain the integral equation

$$V_t = \xi + v\int_0^t h(t-s)\sqrt{V_s}dW_s^1, \quad h(\tau) := e^{-\kappa\tau} \tag{2.2}$$

which of course is equivalent to the classical SDE of Cox, Ingersoll, & Ross (1985),

$$dV_t = v\sqrt{V_t}dW_t^1 + \kappa(\xi - V_t)dt, \quad V_0 = \xi, \tag{2.3}$$

which is popular in practice for interest and hazard rate modelling in addition to volatility. Since, in contrast to its rough counterpart, this (strong) SDE solution is both Markovian and affine, it will end up forming the core building block of our modelling framework. However, in the next section, we parameterise it in a 'quasi-rough' way, which leads to very surprising consequences.

2.2 Connecting volatility roughness with reversionary regimes

Clearly there are an infinitude of possible ways one could try to connect the rough and classical Heston models. See for example Abi Jaber (2018) and Abi Jaber & El Euch (2019). Since all will fall short in some way, we demonstrate only a simple and intuitive approach here.

Now let $k(\tau) = \tau^{-\zeta}$ denote rough Heston's kernel. Since this is clearly singular at the origin for any $\zeta \in (0, \frac{1}{2})$, we first make a small time-shift $\zeta\varepsilon \ll 1$, with the plan of later sending the time-scale ε back to zero. So we define

$$k^\varepsilon(\tau) := k(\tau + \zeta\varepsilon) = (\tau + \zeta\varepsilon)^{-\zeta} \tag{2.4}$$

and consider the approximating affine Volterra process (which actually defines a semimartingale)

$$V_t^\varepsilon = \xi + v\int_0^t k^\varepsilon(t-s)\sqrt{V_s^\varepsilon}dW_s^1. \tag{2.5}$$

We know, using the representation of Equation 2.2, the type of kernels which correspond to the classical Markovian case. Given the delicacy of the origin, we select an approximating kernel h^ε which matches the level and first derivative of k^ε at this point. This uniquely defines a quasi-rough kernel:

$$h^\varepsilon(\tau) := (\zeta\varepsilon)^{-\zeta}e^{-\varepsilon^{-1}\tau}$$

and we define here $0^{-0} := \lim_{\zeta\downarrow 0}\zeta^{-\zeta} = 1$. Notice that we may then *precisely* write Equation 2.5 as

$$V_t^\varepsilon = \xi + v\int_0^t h^\varepsilon(t-s)\sqrt{V_s^\varepsilon}dW_s^1 + vZ_t^\varepsilon \tag{2.6}$$

where Z^ε is built from a difference kernel d^ε according to

$$Z_t^\varepsilon := \int_0^t d^\varepsilon(t-s)\sqrt{V_s^\varepsilon}dW_s^1, \quad d^\varepsilon(\tau) := k^\varepsilon(\tau) - h^\varepsilon(\tau),$$

and, by its design, the difference kernel d^ε satisfies $d^\varepsilon(0) = (d^\varepsilon)'(0) = 0$. At this stage, notice that we have the uniform convergence $V^\varepsilon \to V^0$ as $\varepsilon \to 0$ where V^0 is rough Heston. This follows from the convergence of kernels $k^\varepsilon \to k^0 := k$ in L^2_{loc} and Abi Jaber (2018, Theorem 3.6). This non-Markovian-inducing term Z^ε in Equation 2.6 thus provides a sensible candidate to drop when considering a relationship with the classical case, especially since it is actually of finite variation ![1] *Omitted—plots of kernels k^ε, h^ε and d^ε for range of ε, ζ values.*

So we arrive at a rough-inspired Markovian version of Equation 2.5, specified by

$$V_t^\varepsilon = \xi + \int_0^t h^\varepsilon(t-s)\sqrt{V_s^\varepsilon}dW_s^1,$$

and being equivalent to the SDE of Equation 2.3 but with parameterisation

$$dV_t = v(\zeta\varepsilon)^{-\zeta}\sqrt{V_t}dW_t^1 + \varepsilon^{-1}(\xi - V_t)dt, \quad V_0 = \xi. \tag{2.7}$$

Notice that, due to the well-chosen shift of $\zeta\varepsilon$ in Equation 2.4, the timescale ε corresponds precisely with the reversionary time-scale for this resulting SDE (given generally by the reciprocal of its reversion speed).[2]

What we have obtained here is an elegant connection, whereby the 'quasi-roughness' parameter ζ now takes values in all of $\mathbb{R}_+$, and classifies this Markovian approximation into reversionary regimes (loosely defined by how drift and diffusion coefficients scale, relatively). For example, the three cases of $\zeta \in \{0, \frac{1}{2}, 1\}$ correspond to the regimes studied respectively in Heston (1993), Fouque, Papanicolaou, Sircar, & Sølna (2011) (most notably), and Mechkov (2015)—see McCrickerd (2019, Section 2.4) for more background details. Moreover, the component $\zeta^{-\zeta}$ is very peculiar in that it establishes a very natural parameterisation in all three of these historically-prevalent cases:

Heston (1993):	$dV_t^\varepsilon = \sqrt{V_t^\varepsilon}dW_t + \varepsilon^{-1}(1 - V_t^\varepsilon)dt$
Fouque et al. (2011):	$dV_t^\varepsilon = \varepsilon^{-\frac{1}{2}}\sqrt{2V_t^\varepsilon}dW_t + \varepsilon^{-1}(1 - V_t^\varepsilon)dt$
Mechkov (2015):	$\varepsilon dV_t^\varepsilon = \sqrt{V_t^\varepsilon}dW_t + (1 - V_t^\varepsilon)dt.$

We will see in the next section that, when we consider $\varepsilon \to 0$, the case of $\zeta = 1$, being 'quasi-hyper-rough' like the models of Jusselin & Rosenbaum (2018), is particularly special.

[1] To see this, expand $d^\varepsilon = d^\varepsilon(0) + 1 * (d^\varepsilon)'$ then stochastic Fubini can be employed since $(d^\varepsilon)' \in L^2_{\text{loc}}$, giving: $Z^\varepsilon := d^\varepsilon * dM = d^\varepsilon(0) * dM + (1 * (d^\varepsilon)') * dM = 1 * ((d^\varepsilon)' * dM)$ where $dM = \sqrt{V}dW$.
See Abi Jaber et al. (2017, Lemma 2.1) and preceding definitions.

[2] Another (less tractable) route can be taken to arrive at the same SDE. One can employ Abi Jaber et al. (2017, Lemma 2.6) to, after again dropping a non-Markovian finite variation term, arrive instead at

$$dV_t^\varepsilon = vk^\varepsilon(0)\sqrt{V_t^\varepsilon}dW_t - \frac{(k^\varepsilon)'(0)}{k^\varepsilon(0)}(\xi - V_t^\varepsilon)dt.$$

However, evaluating these coefficients (surprisingly) yields precisely Equation 2.7 !

2.3 Characteristic function limits and the apparent arrival of infinite-activity jumps

Because we started from the rough Heston model of Equation 2.1 with normalised expectations, for example satisfying $\mathbb{E}[(V,H)_t] = (\xi, -\frac{1}{2}\xi t)$ for all $t \geq 0$, we will actually find that *only* the quasi-hyper-rough regime defined by $\zeta = 1$ exhibits interesting behaviour as we send the reversionary time-scale ε back to zero. We find $\zeta \in [0,1)$ leads to trivial and $\zeta > 1$ degenerate results, although omit details here.[3]

There are several results in the literature relating the Heston log-price process H to the normal inverse-Gaussian (NIG) process of Barndorff-Nielsen (1997). Now we reconcile the most relevant ones before extending them. First assume the classical Heston set-up

$$dV_t = v\sqrt{V_t}W_t^1 + (\xi - V_t)dt, \quad dH_t = \sqrt{V_t}(\bar{\rho}W^2 + \rho W^1)_t - \frac{1}{2}V_t dt, \quad (V,H)_0 = (\xi, 0)$$

and notice this is consistent with setting $\zeta = \varepsilon = 1$ in Equation 2.7. The results of Keller-Ressel (2011) and Forde & Jacquier (2011) (connecting H here to the NIG distribution) are applicable to large times, but we would like to make a finite-time statement, like the main result of Mechkov (2015). For this, redefine time using $t \mapsto \varepsilon^{-1}t$, so that we can bring the effect of $t \to \infty$ into finite time through the equivalent limit $\varepsilon \to 0$. Defining suitably-scaled processes

$$(W,V,H)_t^\varepsilon := (\varepsilon^{\frac{1}{2}}W, \varepsilon^{-1}V, H)_{\varepsilon^{-1}t},$$

we then obtain (omitting the ε superscript on W)

$$\varepsilon dV_t^\varepsilon = v\sqrt{V_t^\varepsilon}W_t^1 + (\xi^\varepsilon - V_t^\varepsilon)dt, \quad dH_t^\varepsilon = \sqrt{V_t^\varepsilon}(\bar{\rho}W^2 + \rho W^1)_t - \frac{1}{2}V_t^\varepsilon dt, \quad (V,H)_0^\varepsilon = (\xi, 0) \tag{2.8}$$

where $\xi^\varepsilon := \varepsilon^{-1}\xi$. Now if we let $\varepsilon \to 0$, we *literally* bring the results of Keller-Ressel (2011) and Forde & Jacquier (2011) into finite time, leading to undesirable moment explosions since the reversion level ξ^ε goes to ∞. However, we can control this explosion by simply fixing the reversion level ξ^ε at $\xi^1 = \xi$. In direct analogy with Forde & Jacquier (2011, Theorem 2.1), we then obtain the finite time characteristic exponent limit

$$\lim_{\varepsilon \to 0} \frac{v^2}{\xi t} \log \mathbb{E}[e^{iuH_t^\varepsilon}] = 1 - \rho v i u - \sqrt{(1 - \rho v i u)^2 + v^2 u(i+u)}.$$

Moreover, under a straightforward parameter correspondence, this is precisely the result of Mechkov (2015, Equation 5), derived directly using characteristic functions of the re-scaled process H^ε with fixed $\xi^\varepsilon := \xi$. This provides the marginal convergence in distribution $H_t^\varepsilon \xrightarrow{\text{d}} H_t^0$ where H_t^0 is a NIG Lévy process, by Lévy's continuity theorem. We summarise this by $H_t^0 = \text{NIG}_t(W; \delta, \beta, \gamma, \mu)$, using the notation of Definition 2.1, where we have

$$\delta = \frac{\xi\bar{\rho}}{v}, \quad \beta = \frac{2\rho - v}{2\bar{\rho}^2 v}, \quad \gamma = \frac{1}{\bar{\rho}v}, \quad \mu = -\frac{\xi\rho}{v}. \tag{2.9}$$

[3] This is not necessarily the case if we, for example, allow ξ to depend on ε, but although interesting, this is not relevant in 'historically-normal' risk-neutral settings.

Since we appear to have found infinite-activity jumps arriving in the form of the NIG process,[4] this regime of $\zeta = 1$ appears to justify its 'quasi-hyper-rough' description. Now we extend these results. First we provide a slight generalisation of IG-NIG Lévy motion as compared with Barndorff-Nielsen & Shephard (2001, Example 4.2), which allows for full generality (e.g. non-zero skewness) in the NIG component.

Definition 2.1 (IG-NIG Lévy motion)**.** *We will call (Z, X) IG-NIG Lévy motion built from Brownian motion $B = (B^1, B^2)$ and write*

$$(Z, X) = \text{IG-NIG}(B; \delta, \beta, \gamma, \mu)$$

if for some $\delta, \gamma > 0$ and $\beta, \mu \in \mathbb{R}$, Z admits the IG exit-time representation

$$Z_t = \inf\left\{s > 0 : \gamma s - B^1_s > \delta t\right\} =: \text{IG}_t(B^1; \delta, \gamma)$$

and X admits the NIG 'normal variance-mean mixture' representation

$$X_t = B^2_{Z_t} + \beta Z_t + \mu t =: \text{NIG}_t(B; \delta, \beta, \gamma, \mu).$$

Whenever we drop reference to μ, this should be understood to take the value which ensures $\mathbb{E}[e^{X_t}] = 1$, equivalent to the constraint

$$t^{-1} \log \mathbb{E}[e^{X_t}] = \mu + \delta\left(\gamma - \sqrt{\gamma^2 - 2\beta - 1}\right) = 0.$$

For details on this IG representation see Applebaum (2009), and for more information on the NIG process Barndorff-Nielsen & Shephard (2001, Section 5.2), as well as Cont & Tankov (2003) specifically in our context. Notice that, using the parameter definitions of Equation 2.9, the martingality constraint in Definition 2.1 is naturally satisfied. Now we establish the following result.

Theorem 2.2 (Joint marginal convergence to IG-NIG Lévy motion)**.** *Let $\{(\tilde{V}, H)^\varepsilon\}_{\varepsilon > 0}$ be the family of processes defined, as in Equation 2.7 with $\zeta = 1$, by*

$$\varepsilon dV^\varepsilon_t = v\sqrt{V^\varepsilon_t} W^1_t + (\xi - V^\varepsilon_t)dt, \quad dH^\varepsilon_t = \sqrt{V^\varepsilon_t}(\bar{\rho} W^2 + \rho W^1)_t - \frac{1}{2} V^\varepsilon_t dt, \quad (V, H)^\varepsilon_0 = (\xi, 0)$$

where $\tilde{V}^\varepsilon$ are scaled time-integrals of each V^ε, verifying $\tilde{V}^\varepsilon_t = \bar{\rho}^2 \int_0^t V^\varepsilon_s ds$. Let $(\tilde{V}, H)^0$ be IG-NIG Lévy motion defined by

$$\tilde{V}^0_t = \inf\left\{s > 0 : \gamma s - W^1_s > \delta t\right\}, \quad H^0_t = W^2_{\tilde{V}^0_t} + \beta \tilde{V}^0_t + \mu t$$

where $\delta, \beta, \gamma, \mu$ are as in Equation 2.9. Then, as the reversionary timescale ε goes to zero, we have the convergence in distribution $(\tilde{V}, H)^\varepsilon_t \xrightarrow{\text{d}} (\tilde{V}, H)^0_t = \text{IG-NIG}_t(W; \delta, \beta, \gamma)$ for any $t > 0$.

[4] An infinite-activity Lévy process is one whose implicit Lévy measure explodes at zero, leading to an infinite number of (convergent) jumps over any time interval. See Cont & Tankov (2003, Section 3) for a basic introduction.

Proof. *Omitted—but follows convergence of characteristic functions and Lévy's continuity theorem. Graphically verified by plotting the convergence* $\mathbb{E}[e^{iu\cdot(\tilde{V},H)^{\varepsilon}_t}] \to \mathbb{E}[e^{iu\cdot(\tilde{V},H)^{0}_t}]$ *for* $t \in (0,1)$ *and a variety of* u *values.* □

It at first appears unusual that this result applies to the scaled processes $\tilde{V}^{\varepsilon}$ depending on $\bar{\rho}$, but this simply establishes the correct connection with the natural IG-NIG Lévy motion form in Definition 2.1. By the scaling properties of the IG distribution, it is straightforward to see that we indeed also have $\int_0^t V^{\varepsilon}_s ds =: \bar{V}^{\varepsilon}_t \xrightarrow{\mathrm{d}} \bar{V}^0_t = \mathrm{IG}_t(W^1; \delta, \gamma)$ with $\delta = \frac{\xi}{v}$, $\gamma = \frac{1}{v}$.

The practical relevance of this extension should be noted, given for example the implications on derivative prices and hedging strategies, referencing both integrated variance $\bar{V}^{\varepsilon}$ *and* spot $e^{H^{\varepsilon}}$ processes. Given some results of Alòs, García-Lorite, & Muguruza (2018), it would be interesting to look at, for example, VIX volatility surfaces generated by the family $\{\bar{V}^{\varepsilon}\}_{\varepsilon \geq 0}$ of processes as $\varepsilon \to 0$, but this is not our focus here.

2.4 Implied volatility surface and skew explosion demonstrations

Henceforth we consider the Heston model for $(V,H)^{\varepsilon}$ as in Theorem 2.2. For now we just focus on some surprising implied volatility implications of this result, which actually all follow from the combined results of Mechkov (2015) and Gerhold et al. (2016).

Given that implied volatilities $\bar{\sigma}$ accept representations like

$$\bar{\sigma}^{\varepsilon}(k,\tau) = f\big(\mathbb{E}[(e^k - e^{H^{\varepsilon}_{\tau}})_+]\big)$$

for an appropriately-defined continuous bijective positive function f, then it is a straightforward continuous-mapping exercise to obtain the point-wise convergence $\bar{\sigma}^{\varepsilon}(k,\tau) \to \bar{\sigma}^0(k,\tau)$ as $\varepsilon \to 0$, where σ^0 is the volatility surface generated by the NIG process H^0 in Theorem 2.2. This leads to the uniform convergence $\bar{\sigma}^{\varepsilon} \to \bar{\sigma}^0$ over any compact subset of $\mathbb{R} \times (0,\infty)$. This is demonstrated graphically in Figure 1, with the reversionary time-scale ε taking values in $\{1, 2^{-2}, 2^{-4}, 2^{-6}, 2^{-8}, 0\}$.

To the eye, the ATM skews $\psi^{\varepsilon}(0,\tau)$ implicit in Figure 1, defined generally by $\psi^{\varepsilon} := \frac{\partial \bar{\sigma}^{\varepsilon}}{\partial k}$, certainly appear to explode as $\tau \to 0$, when ε is low. In fact in was shown in Gerhold et al. (2016) that the limiting NIG process H^0 here will generate a skew ψ^0 which exhibits the power-law behaviour $\psi^0(0,\tau) \approx \psi_0 \tau^{-\frac{1}{2}}$ as $\tau \to 0$. We demonstrate the skews derived from Figure 1 more clearly in Figure 2, wherein we see that the Heston case with $\varepsilon = 2^{-8}$ is practically indistinguishable from the NIG case, both exhibiting log-log gradients of $-\frac{1}{2}$ as $\tau \to 0$.

These observations should of course be contrasted with the widely known result from Alòs, León, & Vives (2007), which provides that, in contrast to $\psi^0(0,\tau) \approx \psi_0 \tau^{-\frac{1}{2}}$, the Heston model's skew will instead verify $\psi^{\varepsilon}(0,\tau) \approx \psi_0$ as $\varepsilon \to 0$ for any $\varepsilon > 0$. Indeed, this can be clearly observed in Figure 2 for the cases of $\varepsilon \in \{1, 2^{-2}, 2^{-4}\}$ (red, orange, yellow), but the time-scale of this result's relevance is clearly so short in the remaining two cases that it is rendered practically irrelevant (only becoming visible before the maturity of the shortest traded option).

Since the Heston parameterisation of Theorem 2.2 originated from connections with rough volat-

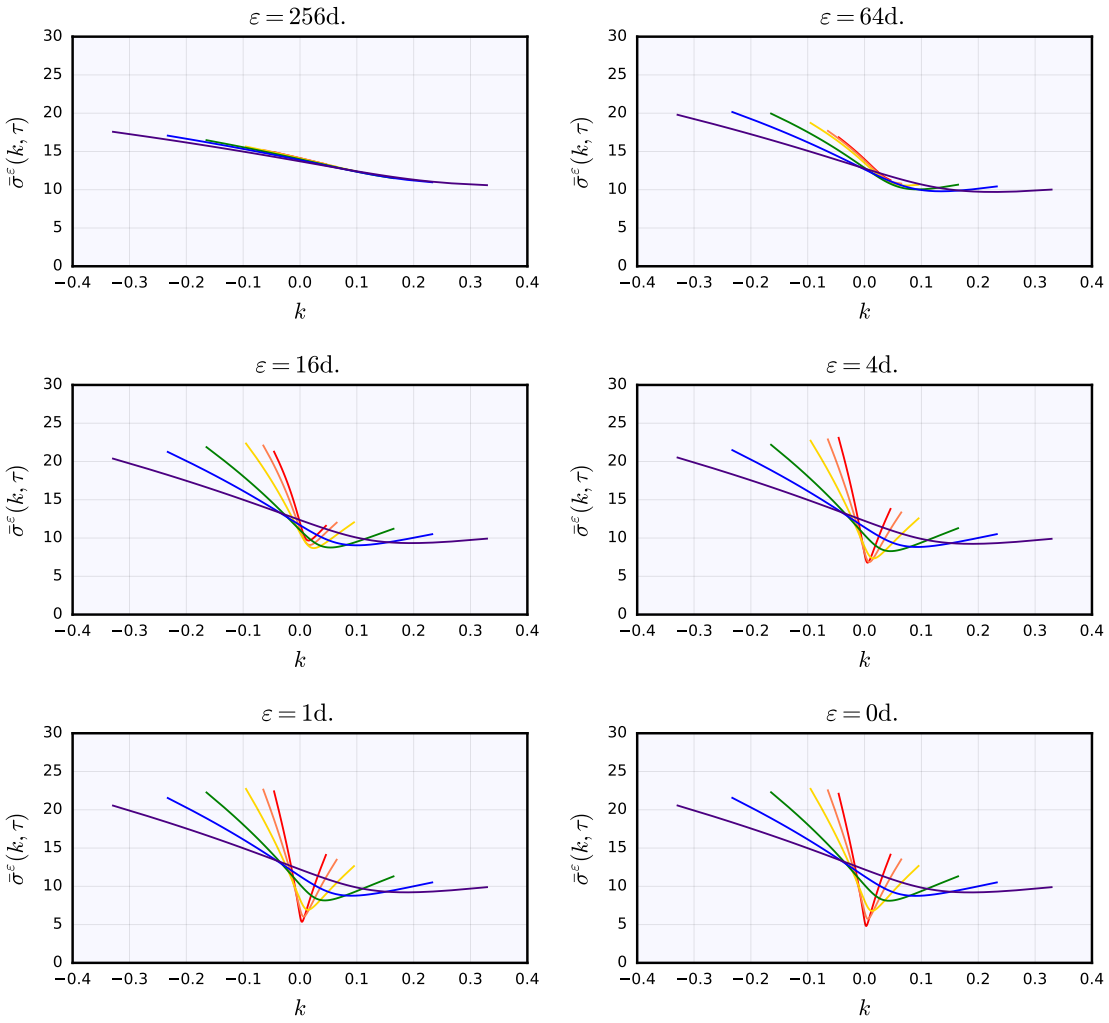


Figure 1: Implied volatility surfaces $\bar{\sigma}^{\varepsilon}$ from the Heston process H^{ε} as defined in Theorem 2.2, as the reversionary timescale ε varies from approximately one year (top left) to zero (bottom right). The last two plots are practically indistinguishable, suggesting that we might be able to make use of the tractable NIG process H^0 in a 'quasi-continuous' stochastic volatility framework. Other model parameter values are given by $\xi = 0.02$, $\rho = -0.7$, $v = 0.3$. Maturities shown range from one week (red) to one year (purple).

ility, we finally demonstrate in Figure 3 the 'quasi-hyper-rough' nature of the limiting surface $\bar{\sigma}^0$ generated from H^0, as compared with a rough Heston surface. We see that surfaces are very similar at the largest maturity of one year, but the quasi-hyper-rough surfaces justifies its description as such for shorter maturities.

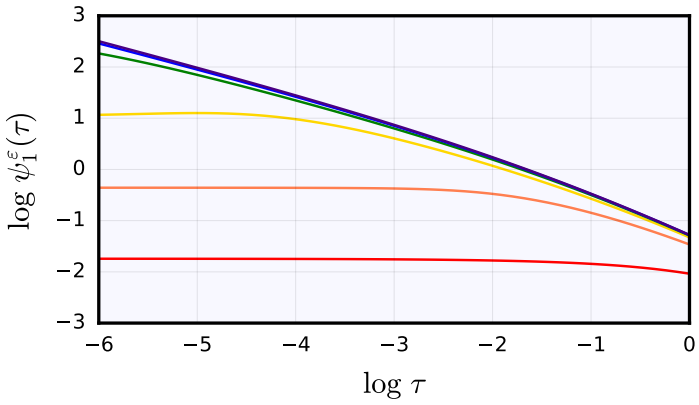


Figure 2: Implied volatility skews ψ_1^ε as the reversionary timescale ε varies from approximately one year (red) to zero (purple). The shortest maturity τ shown here verifies $\log \tau = -6$, so corresponds to roughly 0.5 trading days (shorter than the shortest observable in practice). This should be contrasted against popular plots like El Euch & Rosenbaum (2019, Figure 5.2), which do not present classical Heston in its best light.

3 A pathwise volatility framework with exit-time limits

Section summary. The CIR SDE from Section 2 is mapped onto an initial-value problem (IVP) driven by a continuous path ω started at zero, so it can be studied more generally, on a pathwise basis, as $\varepsilon \to 0$. Despite no constraints being placed on ω (e.g. Hölder continuity), existence, uniqueness and solution bounds are established for IVP solutions, which correspond to integrated-variance trajectories. Our bounds lead to exit-time limits for integrated variance trajectories which, as an example, maps back (weakly) to the IG Lévy process in the Wiener (SDE) setting. With the composition map not being continuous on the appropriate topologies, surprising limits are exposed for a price path or process driven by such IVP or SDE solutions. Much of this section is presently available in McCrickerd (2019), the next version of which I hope to extract from a near-complete thesis.

3.1 Motivation for approach: theorems of Prokhorov and Skorokhod

Summarising from Jacod & Shiryaev (2003), Prokhorov's theorem provides a characterisation of weak convergence, based on tightness, convergence of finite-dimensional distributions, and *characterisation* of a limit in terms of its finite-dimensional distributions. However, we are going to end up working with an unusual type of convergence, and in particular the finite-dimensional distributions (in time) of any limit that we can define does not naturally characterise it. Although Prokhorov's approach would allow us to establish convergence on a relatively weak topology, such as the excursion topology (E, M_2) of Whitt (2002, Chapter 15), we want to draw deeper functional

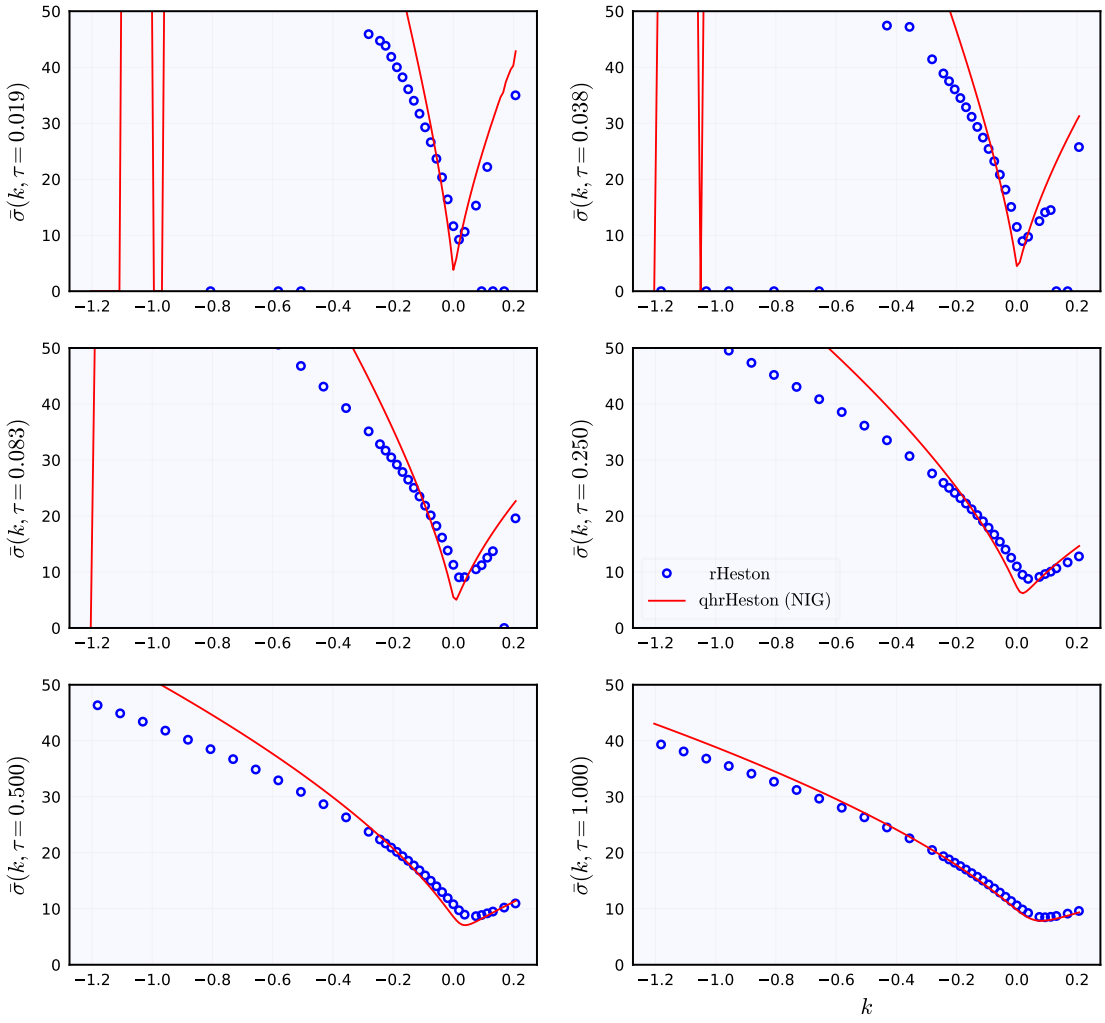


Figure 3: The rough Heston surface (blue) here was provided by E. Abi Jaber and is consistent with Equation 2.1 with parameters $\xi = 0.02$, $\rho = -0.7$, $v = 0.3$, $\zeta = 0.45$. The quasi-hyper-rough Heston surface (red) is generated by H^0 from Theorem 2.2 with the same parameters.

conclusions, ideally applicable to *any* exotic derivative we can define. The pathwise approach we take enables us to achieve this.

Although we will not *apply* Skorokhod's representation theorem, our approach provides a neat example of it. So, adapted from Billingsley (1999, Theorem 6.7) to coincide with our notation, we recall:

Theorem 3.1 (Skorokhod's representation theorem)**.** *Suppose that* $\{\bar{V}_\varepsilon\}_{\varepsilon \geq 0}$ *are processes verifying* $\bar{V}_\varepsilon \xrightarrow{\mathrm{d}} \bar{V}_0$ *on a separable topology* $(D, \mathscr{D})$*. Then there exist processes* $\{\bar{X}_\varepsilon\}_{\varepsilon \geq 0}$ *such that* $\bar{X}_\varepsilon \overset{\mathrm{d}}{=} \bar{V}_\varepsilon$*, yet pathwise convergence* $\bar{X}_\varepsilon(\omega) \to \bar{X}_0(\omega)$ *takes place for every* $\omega \in C$*.*

That is, we will *find* such processes $\{\bar{X}^\varepsilon\}_{\varepsilon\geq 0}$ verifying $\bar{X}_\varepsilon(\omega) \to \bar{X}_0(\omega)$ as $\varepsilon \to 0$ on a suitable topology for every $\omega \in C$. Here and throughout, $\xrightarrow{\text{d}}$ denotes weak convergence as defined in Billingsley (1999), and C the space of real continuous paths starting from zero.

3.2 Review of related problems and literature

In summary, there appears to be *very* few related pieces of work to this section, applicable to ODEs with no quasi-Lipschitz property in its spatial variable. See Kaper & Kwong (1988) for a seemingly-related IVP of type $x' = \phi(x) + p(t)$ but which makes very different assumptions. See also the theory of random ODEs in Han & Kloeden (2017), although this employs rather than contributes to ODE theory. See McCrickerd (2019, Section 4) for more details.

Invaluable reference texts for this section have been Coddington & Levinson (1955) and Agarwal & Lakshmikantham (1993). We will now refer to these as C&L and A&L. Yosie's simple and intuitive theorem of 1926 in particular embodies the approach to IVPs taken throughout, although is never applied directly.[5]

3.3 Topological primer: Skorokhod's non-standard and Whitt's topologies

Space prevents detailed coverage here, although the Preface of Whitt (2002), particularly pp. x–xii, is excellent in describing the issues we will face. $(C, \mathscr{C})$ will be used to denote the topology induced by the uniform norm on the space $C = C_0(\mathbb{R}_+, \mathbb{R})$ of paths. $(D, \mathscr{D})$ will denote the topology induced by Skorokhod's M_1 metric, as in Skorokhod (1956), onto the space $D = D_0(\mathbb{R}_+, \mathbb{R})$ of càdlàg paths.[6] This is the topology on which a result like $\bar{V}^\varepsilon \xrightarrow{\text{d}} \bar{V}^0$ from Section 2 will take place. As established in Skorokhod's original paper, convergence of a process on $(D, \mathscr{D})$ (as with all of his topologies) implies point-wise convergence on a dense set (Condition 2.4.1), which becomes the convergence of finite-dimensional distributions when the process exhibits stochastic continuity.

We will additionally require a Hausdorff topology on a space of set-valued processes, as in Whitt (2002, Chapter 15) (although do not cover this here), and a uniform product topology, to establish functional convergence results for families like $\{\varepsilon V^\varepsilon\}_{\varepsilon>0}$, $\{\int \sqrt{V^\varepsilon} dW\}_{\varepsilon>0}$ and finally $\{H^\varepsilon\}_{\varepsilon>0}$.

3.4 The Cox-Ingersoll-Ross initial-value problem

To treat the CIR SDE, and ultimately the Heston process, we first map the problem from an SDE to an time-change equation (TCE), and then to an initial-value problem (IVP), as in the following diagrammatic summary of problems as maps from spaces:

$$\text{SDE} : (C, \mathscr{C}, \mathbb{W}) \;\mapsto\; \text{TCE} : (C, \mathscr{C}, \mathbb{P} \neq \mathbb{W}) \;\mapsto\; \text{IVP} : (C, \mathscr{C}) = (C, \mathscr{C}, \mathbb{P} = \delta_\omega).$$

[5] Available from p.81 at worldscientific.com/doi/suppl/10.1142/1988/suppl_file/1988_chap01.pdf.
[6] Actually a very minor extension, as in Puhalskii & Whitt (1997).

$\mathbb{W}$ will always denote the Wiener measure on $(C, \mathscr{C})$, $\mathbb{P}$ an arbitrary measure, and δ_ω the Dirac measure at the particular path ω. This will lead us to the following:

Problem 3.2 (Cox-Ingersoll-Ross initial-value problem)**.** *For fixed reversionary timescale $\varepsilon > 0$ and noise trajectory $\omega \in C$, let $f_\varepsilon(\omega) = f : \mathbb{R}_+^2 \to \mathbb{R}$ be defined according to*

$$f(t,x) := \varepsilon^{-1}\Big(\omega(x) + t - x\Big) + 1.$$

Then, find a time-integrated volatility path $\varphi = \varphi_\varepsilon(\omega)$ (and so also volatility path φ'), such that for $t \in [0,T)$, with $T \in (0,\infty]$, we have

$$\varphi'(t) = f\big(t, \varphi(t)\big), \quad \varphi(0) = 0.$$

The IVP in Problem 3.2 identifies with the SDE for V^ε in Theorem 2.2 in the particular case of $\xi = v = 1$. We will deal with practically-necessary time-inhomogeneous generalisations in Section 3.6, which are relatively straightforward, and become applicable (for example) to an IVP with driving function depending on a positive and bijective path ϑ, of type

$$f(t,x) = \varepsilon^{-1}\Big(v\omega(x) + \vartheta(t) - x\Big) + \xi_0.$$

3.4.1 Problem derivation via the Dambis, Dubins-Schwarz theorem

Setting $\xi = v = 1$, we may write the CIR SDE of Theorem 2.2 in the unusual form

$$\varepsilon(V_t - 1) = \int_0^t \sqrt{V_u}\mathrm{d}W_u + t - \bar{V}_t. \tag{3.1}$$

We use $\bar{V}_t = \int_0^t V_s ds$ here and will adopt this notation more generally. We also drop reference to the reversionary time-scale ε until it is required again. Recall the DDS theorem:

Lemma 3.3 (Dambis, Dubins-Schwarz CIR time-change)**.** *Let W be the canonical process (Brownian motion) on $(C, \mathscr{C}, \mathbb{W})$ and V solve Equation 3.1. Define the exit-time $\hat{V}$ of the time-integrated process $\bar{V}$ by*

$$\hat{V}_t = \inf\Big\{s > 0 : \bar{V}_s > t\Big\}.$$

Then, the process B defined by $B_t := \int_0^{\hat{V}_t} \sqrt{V_u}\mathrm{d}W_u$ is another (non-canonical) Brownian motion on $(C, \mathscr{C}, \mathbb{W})$, and furthermore verifies $(B \circ \bar{V})_t = \int_0^t \sqrt{V_u}\mathrm{d}W_u$, where $(B \circ \bar{V})_t = B_{\bar{V}_t}$.

We will generally use $\hat{V}$ to denote the exit-time (actually bona fide inverse) of $\bar{V}$. In light of Lemma 3.3, consider a process X which solves the equation

$$\varepsilon(X_t - 1) = (W \circ \bar{X})_t + t - \bar{X}_t \tag{3.2}$$

which we will call the CIR Wiener TCE. Now we clarify the following relationship.

Lemma 3.4 (Weak equivalence of CIR SDE and Wiener TCE)**.** *Let W be the canonical process (Brownian motion) on $(C, \mathscr{C}, \mathbb{W})$ and V, X respectively solve Equation 3.1 and Equation 3.2. Then we have $X \overset{\mathrm{d}}{=} V$.*

Proof. See McCrickerd (2019, Lemma 3.0.2) if required. □

Of course we are free to consider such a TCE under any probability measure, unlike the SDE which is inseparable from the Wiener measure. This motivates the following which is clearly more general than the CIR Wiener TCE of Equation 3.2.

Definition 3.5 (CIR TCE)**.** *Let Z be the canonical process on the space $(C, \mathscr{C}, \mathbb{P})$ for any probability measure $\mathbb{P}$. Then we will say that the process X solves the CIR TCE if it verifies the following equation for $\varepsilon > 0$*

$$\varepsilon(X_t - 1) = (Z \circ \bar{X})_t + t - \bar{X}_t, \quad \bar{X}_t := \int_0^t X_u \mathrm{d}u. \tag{3.3}$$

Given that $\mathbb{P}$ in Definition 3.5 is arbitrary, we may consider a Dirac measure δ_ω. Then we are really dealing with an IVP. Under the correspondence

$$Z(\omega) = \omega, \quad \bar{X}(\omega) = \varphi, \quad X(\omega) = \varphi', \quad \bar{X}_0(\omega) = \varphi(0) = 0,$$

we therefore arrive from Equation 3.3 to the ordinary differential equation (ODE)

$$\varepsilon(\varphi'(t) - 1) = (\omega \circ \varphi)(t) + t - \varphi(t)$$

from the topological space $(C, \mathscr{C})$. Now clearly verification of this ODE with the initial condition $\varphi(0) = 0$ is equivalent to φ solving the CIR IVP of Problem 3.2.

* * *

We now consolidate results relating to the CIR IVP of Problem 3.2 into three main existence, limits and uniqueness theorems. Clearly it is important to keep the relationships $V \overset{\mathrm{d}}{=} X$ and $X(\omega) = \varphi'$ in mind in order to translate results applicable to the integrated volatility trajectory φ back onto the CIR SDE solution V (and TCE solution X). For example, when we prove that φ is bijective, this translates to a generalisation of the reflection principle of V at zero, and when we prove that φ is unique, this provides an alternative to Yamada-Watanabe uniqueness for V. Uniqueness in general is particularly surprising, given ω here is *any* element of C, e.g. $\omega(x) = x^\alpha$ for $\alpha \in (0, 1)$. However, we prove this last in order to clarify that other results do not actually depend on it.

We will use the following supremum S and exit-time E functionals

$$S(\varphi)(t) := \sup\Big\{\varphi(u) : u \in [0, t)\Big\}, \quad E(\varphi)(x) := \inf\Big\{u > 0 : \varphi(u) > x\Big\}, \tag{3.4}$$

which appear regularly in the limit theorems of Whitt (2002), and previous articles like Whitt (1971), Whitt (1980) and Puhalskii & Whitt (1997). We adopt the convention $\inf \varnothing := \infty$ in the definition of E, and let e denote the identity element of C, verifying $\mathrm{e}(t) = t$.

3.4.2 Existence and bijectivity of solutions

Solutions $\varphi : [0, T) \to \mathbb{R}^+$ of the CIR IVP turn out to exist and be bijective, regardless of the driving path $\omega \in C$. Notice that the bijectivity of φ means that φ' may attain zero (in general), but only countably-often over time.

Existence over some (explicit) interval $[0, \epsilon)$ is provided by the Cauchy-Peano existence theorem. See C&L (Theorem 1.1.2). Clearly we'd like to establish existence over a larger, practically helpful, interval. We find that actually, using the 'continuation' of IVP solutions, see C&L (Theorems 1.4.1 and 2.1.3), we can guarantee a solution exists over some $[0, T)$ where T satisfies

$$T > \sup\left\{x - \omega(x) : x \in \mathbb{R}_+\right\} - \varepsilon.$$

This is quite a strong statement for any desirable noise trajectory ω. For example, assuming ω be recurrent, defined here by $\omega \in R$ where

$$R := \left\{\omega \in C : \forall x > 0 \ \exists x' > x : \omega(x') = 0\right\}$$

ensures $T = \infty$; indefinite existence of a bijective solution $\varphi : \mathbb{R}_+ \to \mathbb{R}_+$. This is summarised in the following result.

Theorem 3.6 (Existence and bijectivity)**.** *Every CIR IVP admits a bijective solution $\varphi : [0, T) \to \mathbb{R}_+$. In addition to $T \in (0, \infty]$, we find $T + \varepsilon > S(\mathrm{e} - \omega)(\infty)$.*

Proof. See McCrickerd (2019, Section A.2). □

3.4.3 Temporal bounds and exit-time limits

Strangely, having established the bijectivity of solutions φ, it starts to become easier to work with their inverses $\hat{\varphi} = E(\varphi) = \varphi^{-1}$. It is most natural to establish bounds in space for $\hat{\varphi}$, which correspond to bounds in time for φ. This is achieved using geometrical arguments along the line of Yosie's theorem and the preceding A&L (Theorem 1.21.1), relating to *upper* and *lower solutions*. We end up obtaining the bounds $\hat{\varphi}(x) \in [\hat{\varphi}_-(x), \hat{\varphi}_+(x)]$ where

$$\hat{\varphi}_-(x) := S(\mathrm{e} - \omega)(x) - \varepsilon, \quad \hat{\varphi}_+(x) := S(\mathrm{e} - \omega)(x) + 2\sqrt{x\varepsilon}$$

for $x \geq 0$. So it becomes clear that the norm $\|\hat{\varphi} - S(\mathrm{e} - \omega)\|_x$ vanishes for any x as $\varepsilon \to 0$, which provides the uniform convergence of inverses $\hat{\varphi}_\varepsilon \to \hat{\varphi}_0 := S(\mathrm{e} - \omega)$ as $\varepsilon \to 0$. For clarity, this limit means

$$\hat{\varphi}_0(t) = \sup\left\{s - \omega(s) : s \in [0, t)\right\},$$

and mapping this back onto the CIR SDE yields the very surprising uniform convergence

$$E(\bar{V}^\varepsilon) =: \hat{V}^\varepsilon \xrightarrow{\mathrm{d}} \hat{V}^0, \quad \hat{V}^0_t := \sup\left\{s - W_s : s \in [0, t)\right\}.$$

There are numerous ways that we can now get to the limit of φ_ε, and $\bar{V}^\varepsilon$. The most straightforward, although a little sloppy, is to note that the map $E : (C, \mathscr{C}) \to (D, \mathscr{D})$ is continuous. This is established in Puhalskii & Whitt (1997). Then we simply have

$$\varphi_\varepsilon = E(\hat{\varphi}_\varepsilon) \to E(S(\mathrm{e} - \omega)) = E(\mathrm{e} - \omega) := \varphi_0$$

on $(D, \mathscr{D})$ as $\varepsilon \to 0$, which is demonstrated in Figure 4. For clarity, this limit means

$$\varphi_0(t) = \inf\left\{s > 0 : s - \omega(s) > t\right\}$$

and again mapping this back onto the CIR SDE yields the convergence

$$\bar{V}^\varepsilon \xrightarrow{\mathrm{d}} \bar{V}^0, \quad \bar{V}^0_t := \inf\left\{s > 0 : s - W_s > t\right\}$$

on $(D, \mathscr{D})$. We should recognise $\bar{V}^0$ as an IG Lévy process, and $\hat{V}^0$ its exit-time, studied in detatil in Vellaisamy & Kumar (2017). These results are summarised in the following.

Theorem 3.7 (Bounds in time and fast reversion limits)**.** *For each $x \in \mathbb{R}_+$, the exit-time $\hat{\varphi} := E(\varphi)$ verifies $\hat{\varphi}(x) \in [\hat{\varphi}_-(x), \hat{\varphi}_+(x)]$, where*

$$\hat{\varphi}_-(x) := S(\mathrm{e} - \omega)(x) - \varepsilon, \quad \hat{\varphi}_+(x) := S(\mathrm{e} - \omega)(x) + 2\sqrt{x\varepsilon}.$$

So as $\varepsilon \to 0$, we have convergence $\hat{\varphi}_\varepsilon \to \hat{\varphi}_0 := S(\mathrm{e} - \omega)$ on $(C, \mathscr{C})$. These bounds of $\hat{\varphi}$ in space are bounds of φ in time, so provide $\varphi_\varepsilon \to \varphi_0 := E(\mathrm{e} - \omega)$, but on $(D, \mathscr{D})$.

Proof. See McCrickerd (2019, Section A.3). □

3.4.4 Uniqueness and space of solutions

One can quickly develop intuition for why CIR IVP solutions are unique. Supposing $\varphi : [0, T) \to \mathbb{R}_+$ solves $x' = f(t, x)$, $x(0) = 0$, then $\hat{\varphi} : \mathbb{R}_+ \to [0, T)$ solves the inverse IVP given by

$$x' = \hat{f}(t, x) := f(x, t)^{-1}, \quad x(0) = 0.^{7}$$

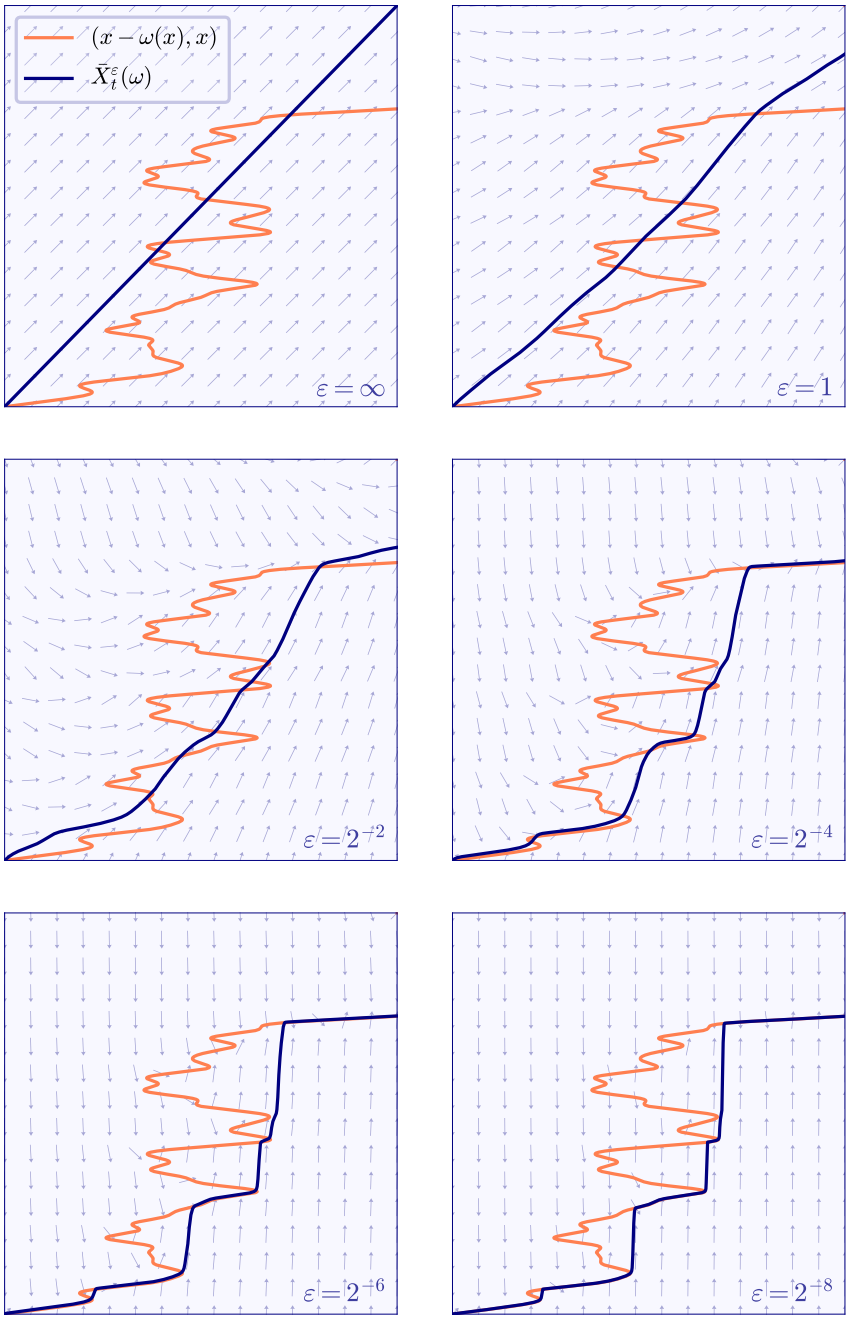
Although this is an unusual IVP in the CIR case (seems unlikely to arise naturally in practice), it is straightforward to check that it verifies the conditions of Peano's uniqueness theorem as given in A&L (Theorem 1.3.1), which means that $\hat{f}$ also verifies a one-sided Lipschitz condition:

$$\hat{f}(t, x') - \hat{f}(t, x) \leq L(x' - x), \quad x' \geq x, \quad L \geq 0.$$

This suggests inverse solutions $\hat{\varphi}$, and so also φ, are unique. However, actually completing a proof along these lines gets messy; we know that a solution φ can verify $\varphi' = 0$, which relates to an explosion of the inverse function $\hat{f}$. Although this explosion can be managed, a neater direct

[7] This follows elementary properties of calculus, like $\hat{\varphi}'(x) = \varphi'(\hat{\varphi}(x))^{-1}$. The reader should be convinced after verifying that if φ solves $x' = t$, then $\hat{\varphi}$ solves $x' = x^{-1}$.

Figure 4: Forward Euler approximations to the time-averaged volatility solution $\varphi_\varepsilon(t) = \bar{X}^\varepsilon_t(\omega)$ (blue) for $(t, x) \in [0, 1]^2$, and reversionary timescale ε varying down to approximately one day. Noise trajectory ω is given by a truncated (Brownian bridge) Karhunen-Loève series (orange). Specifically, we set $\omega(x) = \sum_{k=1}^{64} \frac{\xi_k}{k\pi} \sin(k\pi x)$ and use python's numpy.random module with seed(2) to generate $\xi = \{\xi_k\}_{k=1}^{64}$. Convergence to the limit $\varphi_\varepsilon \to \varphi_0 := E(\mathrm{e} - \omega)$ as $\varepsilon \to 0$ is clearly evident.



proof, which mirrors Peano's uniqueness (flipping space and time), is provided in McCrickerd (2019, Section A.4).

Given a driving path ω, we now have unique positive bijective solution $\varphi : [0,T) \to \mathbb{R}_+$. It is natural to therefore ask whether any such φ is a solution of the CIR IVP for some (unique?) $\omega \in C$. This turns out to be the case, so we see that the CIR IVP sets up another bijection (for each fixed $\varepsilon > 0$), from continuous paths ω to positive, differentiable and bijective paths φ.[8] These results are summarised in:

Theorem 3.8 (Uniqueness and space of solutions)**.** *There is precisely one solution φ of any CIR IVP, so also $\hat{\varphi}$. Moreover, for any $\varepsilon > 0$, any bijective and differentiable path $\varphi : [0,T) \to \mathbb{R}_+$ is a solution, when $\omega \in C$ is given by $\omega(t) := t - \hat{\varphi}(t) + \varepsilon\left(\hat{\varphi}'(t)^{-1} - 1\right)$.*

Proof. See McCrickerd (2019, Section A.4). That ω as defined here remains in C (e.g. that we don't find $\hat{\varphi}' \to 0$ in finite time) is not trivial, but is not difficult, to establish. □

3.4.5 Convergence of simulation schemes

The convergence of Euler-type simulation schemes is a relatively straightforward corollary of uniqueness, and the continuity of f on the entirety of $\mathbb{R}^2_+$. See McCrickerd (2019, Appendix A.4) if required.

3.5 Generality of the framework: from Black-Scholes to novel rough models

As mentioned, we have, somewhat unintentionally, arrived at a very general framework for specifying volatility models, with pathwise unique solutions being guaranteed. All one has to do is choose a desirable measure $\mathbb{P}$, which induces some (canonical) noise process Z on $(C, \mathscr{C}, \mathbb{P})$, and then the CIR IVP, or rather TCE, yields a volatility model as its solution.

Notice that selecting $\mathbb{P} = \delta_0$ yields the deterministic volatility process $X = \xi$, then we simply obtain the weakly equivalent Black-Scholes model in the form $H_t = W^2_{\xi t} - \frac{1}{2}\xi t$. Now consider again a rough Heston model for V which verifies

$$V_t = \xi + \int_0^t (t-s)^{-\zeta}\sqrt{V_s}dW_s, \quad \zeta \in (0, \tfrac{1}{2})$$

in contrast to the following alternative:

Definition 3.9 (Rough Heston 2.0)**.** *We will call H a rough Heston 2.0 process, built from Brownian motion (W^1, W^2), if $X_0 = \xi$ and thereafter*

$$\varepsilon dX_t = v dW^{\zeta}_{\int_0^t X_s ds} + (\xi - X_t)dt, \quad W^{\zeta}_t = \int_0^t (t-s)^{-\zeta} dW^1_s$$

[8] One could therefore make the grand claim that this IVP is sufficient to accommodate *all* stochastic volatility models, via a suitable measure $\mathbb{P}$ choice placed on $(C, \mathscr{C})$. Moreover, solutions are *always* $\mathbb{P}$-a.s. pathwise unique.

$$H_t = (\bar{\rho}W^2 + \rho W^1)_{\int_0^t X_s ds} - \frac{1}{2}\int_0^t X_s ds,$$

Notice that this is weakly equivalent to classical (and also rough) Heston when $\zeta = 0$.

Initial calculations by E. Abi Jaber appear to connect rough Heston 2.0 with the original through a generalisation of its kernel $(t-s)^{-\zeta}$ to a stochastic one of type:

$$(t-s)^{-\zeta} = \int_{\mathbb{R}_+} e^{-x(t-s)}\mu(dx) \mapsto \int_{\mathbb{R}_+} e^{-x\int_s^t X_u du}\mu(dx), \quad \mu(dx) := \frac{x^{\zeta-1}}{\Gamma(\zeta)}dx.$$

Notice that, regardless of the measure $\mathbb{P}$, our limiting results still apply. For example, in the rough Heston 2.0 case we obtain the unusual exit-time limit on $(D, \mathscr{D})$ given by

$$\bar{X}^\varepsilon \xrightarrow{\text{a.s.}} \bar{X}^0, \quad \bar{X}^0_t := \inf\left\{s > 0 : s - \int_0^s (s-u)^{-\zeta} dW_u > t\right\} \quad \text{as } \varepsilon \to 0.$$

Clearly it is difficult to imagine obtaining such a result, if instead taking Prokhorov's probabilistic approach as opposed to our pathwise one. Notice that, since W^ζ is only a Lévy process for the classical case of $\zeta = 0$, so is the limit $\bar{X}^0$.

3.6 Practically necessary time-inhomogeneous generalisations

In practice we need close control over $\mathbb{E}[V_t]$, whereas under the CIR SDE in Theorem 2.2 we simply have $\mathbb{E}[V_t] = \xi$. Likewise for the CIR IVP solution, in the Wiener setting. We will need to generate some positive 'forward variance' curve $\mathbb{E}[V_t] = \xi(t)$, equivalently $\mathbb{E}[\bar{V}_t] = \int_0^t \xi(s)ds$, obtained from traded derivatives (variance swaps directly or vanilla options indirectly).

There are two related ways of achieving this in practice. Either, one can work with a normalised CIR process X verifying $\mathbb{E}[X_t] = 1$ for all $t \geq 0$, and then exogenously scale $V_t := \xi(t)X_t$, or one can work with a generalised CIR process which for example solves

$$\varepsilon dV_t = v\sqrt{V_t}dW^1_t + (\vartheta'_\varepsilon(t) - V_t)dt, \quad V_0 = \xi_0, \quad \vartheta'_\varepsilon(t) := \xi(t) + \varepsilon\xi'(t). \tag{3.5}$$

Both approaches will be addressed, although the first will ultimately prove more useful in the multi-asset FX setting. The second, however, motivates the treatment of the following problem.

Problem 3.10 (Generalised Cox-Ingersoll-Ross initial-value problem)**.** *For fixed* $\varepsilon, v, \xi_0 > 0$*, positive and bijective* ϑ*, and* $\omega \in C$*, let* $f_\varepsilon(\omega) = f : \mathbb{R}^2_+ \to \mathbb{R}$ *be defined according to*

$$f(t,x) := \varepsilon^{-1}\Big(v\omega(x) + \vartheta(t) - x\Big) + \xi_0.$$

Then, find a solution $\varphi = \varphi_\varepsilon(\omega)$*, such that for* $t \in [0,T)$*, with* $T \in (0,\infty]$*, we have*

$$\varphi'(t) = f\big(t, \varphi(t)\big), \quad \varphi(0) = 0.$$

All results already derived in this section also apply to this generalisation in a natural way. The limits are particularly interesting, now admitting representations as exit-times with time-dependent

barriers. For example, for $V = V^\varepsilon$ which solves Equation 3.5, we find

$$\bar{V}^\varepsilon \to \bar{V}^0, \quad \bar{V}^0_t := \inf\left\{s > 0 : s - vW^1_s > \int_0^t \xi(u)du\right\} \text{ as } \varepsilon \to 0 \text{ on } (D, \mathscr{D})$$

and this limit naturally becomes

$$\bar{V}^0_t := \inf\left\{s > 0 : s - v\int_0^s (s-u)^{-\zeta} dW^1_u > \int_0^t \xi(u)du\right\}$$

in the rough Heston 2.0 case.

3.7 Price implications and the Heston excursion process

Before moving onto price trajectories, it is instructive to consider the behaviour of other families $\{\omega \circ \varphi_\varepsilon\}_{\varepsilon>0}$ and $\{\varepsilon\varphi'_\varepsilon\}_{\varepsilon>0}$, because these are implicitly present in the CIR IVP. It turns out that, like price processes, neither of these families find limits on $(D, \mathscr{D})$. The family $\{\omega \circ \varphi_\varepsilon + \mathrm{e}\}_{\varepsilon>0}$ is demonstrated in Figure 5, which finds the same limit as $\{\varphi_\varepsilon\}_{\varepsilon>0}$ almost everywhere, except for additional excursions. See also McCrickerd (2019, Figure 3) for the behaviour of $\{\varepsilon\varphi'_\varepsilon\}_{\varepsilon>0}$ which, except for excursions, becomes zero almost everywhere.

There are two ways of characterising the limits which all of these families find. We can work on Whitt's excursion topology $(E, \mathscr{E})$ of Whitt (2002, Chapter 15), where E is a space of set-valued functions which are a.e. singletons, or we can work parametrically in a space of space-time paths with the uniform product topology. The latter topology is finer (meaning convergence here implies convergence on $(E, \mathscr{E})$), so we only cover that here.

A log-price process H^ε defined by (for example)

$$H^\varepsilon_t = \int_0^t \sqrt{V^\varepsilon_s} d(W^2 + W^1)_s - \int_0^t V^\varepsilon_s ds \overset{\mathrm{d}}{=} (W^2 + W^1)_{\bar{V}^\varepsilon_t} - \bar{V}^\varepsilon_t$$

leads naturally to a log-price trajectory h_ε defined by

$$h_\varepsilon = \bar{\omega} \circ \varphi_\varepsilon + \omega \circ \varphi_\varepsilon - \varphi_\varepsilon = (\bar{\omega} + \omega - \mathrm{e}) \circ \varphi_\varepsilon$$

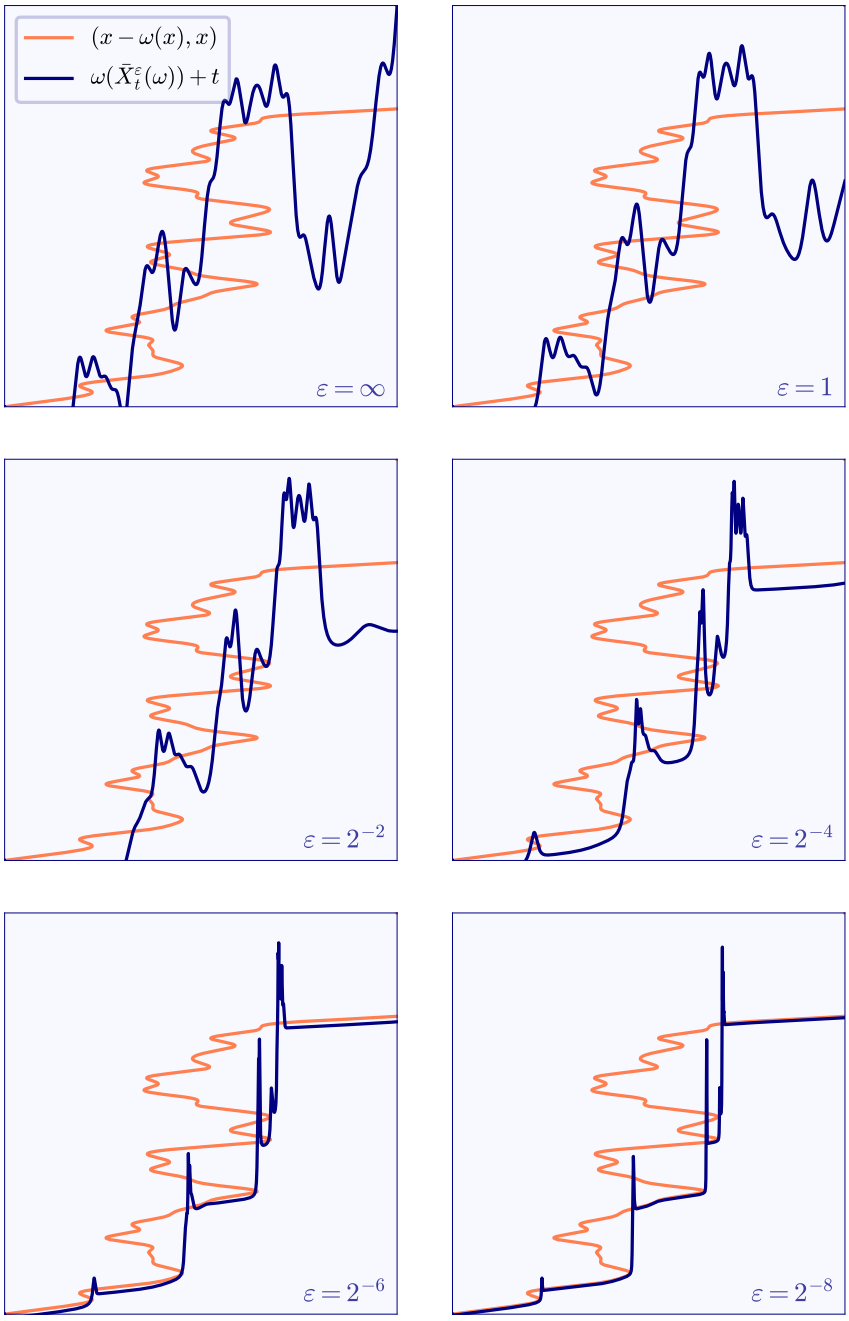
in the IVP setting, where $\bar{\omega} = W^2(\bar{\omega})$. Although we *do not* have $h_\varepsilon \to (\bar{\omega} + \omega + \mathrm{e}) \circ \varphi_0$ on $(D, \mathscr{D})$, consider the space-time parametric representation $(\hat{\varphi}_\varepsilon, \bar{\omega} + \omega + \mathrm{e})$ which traces out the same trajectory in time and space. We *do*, however, have

$$(\hat{\varphi}_\varepsilon, \bar{\omega} + \omega + \mathrm{e}) \to (\hat{\varphi}_0, \bar{\omega} + \omega + \mathrm{e})$$

on the uniform product topology, and this really tells us all we need to know about prices processes, explaining the excursions we see in Figure 5 via periods over which the limiting time trajectory $\hat{\varphi}_0$ stops.

This motivates the following three Heston extensions: the Heston *excursion, jump* and *jump-diffusion* processes, with the jump case coinciding with the excursionary one in finite-dimensional settings.

Figure 5: Here we repeat Figure 4, but show $\omega \circ \varphi + \mathrm{e} = \omega \circ \bar{X}_\varepsilon(\omega) + \mathrm{e}$ rather than $\bar{X}_\varepsilon(\omega)$, which may be considered related to a Heston price trajectory. The limit observed here as $\varepsilon \to 0$, which is almost everywhere equivalent to the exit-time limit $\bar{X}_0(\omega)$ of Figure 4, is best understood using parametric representations, and exists on Whitt's excursion topologies; see Whitt (2002, Chapter 15).



Definition 3.11 (Heston parametric excursion process)**.** *We will call the (space-time) parametric process $H = (H_\tau)_{\tau\geq 0}$ a Heston excursion process if it verifies*

$$\hat{V}_\tau = \sup\left\{s - vW_s^1 : s \in [0, \xi\tau)\right\}, \quad H_\tau = \left(\hat{V}_\tau,\ \bar{\rho}W_\tau^2 + \rho W_\tau^1 - \frac{1}{2}\tau\right).$$

Definition 3.12 (Heston jump process)**.** *We will call the NIG Lévy process H a Heston jump process if it verifies*

$$\bar{V}_t = \inf\left\{s > 0 : s - vW_s^1 > \xi t\right\}, \quad H_t = \bar{\rho}W^2_{\bar{V}_t} + \rho W^1_{\bar{V}_t} - \frac{1}{2}\bar{V}_t.$$

This equivalently admits the mixture representation

$$H_t = \bar{\rho}W^2_{\bar{V}_t^0} + \frac{2\rho - v}{2v}\bar{V}_t - \frac{\rho}{v}\xi t$$

and is the unique Lévy process with exponent $\psi(u) = t^{-1}\log\mathbb{E}[e^{iuH_t}]$ which verifies

$$v^2\xi^{-1}\psi(u) = 1 - \rho v i u - \sqrt{(1 - \rho v i u)^2 + v^2 u(i + u)}.$$

Definition 3.13 (Heston jump-diffusion process)**.** *We will call the process H a Heston jump-diffusion process if it verifies $H = H^\varepsilon + H^0$, where H^ε is a classical Heston process:*

$$\varepsilon dV_t^\varepsilon = v\sqrt{V_t^\varepsilon}dW_t^1 + ((1 - w)\xi - V_t^\varepsilon)dt, \quad dH_t^\varepsilon = \sqrt{V_t^\varepsilon}d(\bar{\rho}W^2 + \rho W^1)_t - \frac{1}{2}V_t^\varepsilon dt$$

with $V_0^\varepsilon = (1 - w)\xi$, and H^0 is a Heston jump process, independent from H^ε but consistent with its limit as $\varepsilon \to 0$, having characteristic exponent ψ which verifies

$$\frac{v^2}{w\xi}\psi(u) = 1 - \rho v i u - \sqrt{(1 - \rho v i u)^2 + v^2 u(i + u)}.$$

The full Heston jump-diffusion (and analgous excursion-diffusion) characteristic function is then given by

$$\mathbb{E}[e^{iuH_t}] = \exp\left(\left((1 - w)[\phi(u;t) + \varphi(u;t)] + w\psi(u;t)\right)\frac{\xi}{v^2}\right)$$

where

$$\varepsilon^{-1}\varphi(u;t) = \vartheta_1 - \vartheta_2\frac{\vartheta_1 + \vartheta_2\tanh(\vartheta_3 t)}{\vartheta_2 + \vartheta_1\tanh(\vartheta_3 t)},$$

$$\phi^\varepsilon(u;t) = \vartheta_1 t - \vartheta_2\vartheta_3^{-1}\log\left(\cosh(\vartheta_3 t) + \vartheta_1\vartheta_2^{-1}\sinh(\vartheta_3 t)\right),$$

$$\psi(u;t) = (\vartheta_1 - \vartheta_2)t,$$

$$\vartheta_1 = 1 - \rho v i u, \quad \vartheta_2^2 = \vartheta_1^2 + v^2 u(i + u), \quad \vartheta_3 = \tfrac{\vartheta_2}{2\varepsilon}.$$

Notice that for $w = 0$ we obtain pure diffusive Heston, and for $w = 1$ we obtain pure jump Heston, coinciding with the NIG process, parameterised in a similar way as in Mechkov (2015).

3.8 Consolidating practical consequences: pricing and hedging

As is the power of convergence results, it is now tempting to start applying mapping theorems, for example those presented in Skorokhod (1956, Theorem 3.2.3) and Billingsley (1999, Theorem 2.7). For example, the practitioner is likely interested in the kind of derivative prices which do and don't converge as $\varepsilon \to 0$. In this regard we provide these two consolidatory results.

Lemma 3.14 (Continuous mapping)**.** *Let $\bar{X}_\varepsilon$ solve the CIR TCE from $(C, \mathscr{C}, \mathbb{P})$ and $\varphi_\varepsilon := \bar{X}_\varepsilon(\omega)$. Let $h : (D, \mathscr{D}) \to (E, \mathscr{E})$ be a continuous function onto the topological space $(E, \mathscr{E})$. Then $h(\varphi_\varepsilon) \to h(\varphi_0)$ as $\varepsilon \to 0$ on $(E, \mathscr{E})$. Similarly, $h(\bar{X}_\varepsilon) \xrightarrow{\mathrm{d}} h(\bar{X}_0)$ on $(E, \mathscr{E})$, and for this h need only be continuous almost everywhere, with respect to the limit distribution $\mathbb{P}\bar{X}_0^{-1}$.*

Proof. See Skorokhod (1956, Theorem 3.2.3) or Billingsley (1999, Theorem 2.7). □

Although not that different to Lemma 3.14, in practice it is helpful to also observe the following result which is directly applicable to derivative pricing, Radon-Nikodym derivatives, etc.

Lemma 3.15 (Dominated mapping)**.** *Let h be continuous almost everywhere as in Lemma 3.14, map onto $(\mathbb{C}, |\cdot|)$ and be such that $\{h(\bar{X}_\varepsilon)\}_{\varepsilon>0}$ is dominated by a $\mathbb{P}$-integrable function. Then $\mathbb{E}[h(\bar{X}_\varepsilon)] \to \mathbb{E}[h(\bar{X}_0)]$.*

Proof. Since we have pathwise convergence $h(\varphi_\varepsilon) \to h(\varphi_0)$ as $\varepsilon \to 0$ on $(\mathbb{C}, |\cdot|)$, we may apply the dominated convergence theorem for $\mathbb{E}[h(\bar{X}_\varepsilon)] \to \mathbb{E}[h(\bar{X}_0)]$ provided $\{h(\bar{X}_\varepsilon)\}_{\varepsilon>0}$ is dominated by a $\mathbb{P}$-integrable function. □

Once pathwise limiting price trajectories are understood, it is fairly straightforward to see which derivative prices will converge to those generated by the NIG process, and which require instead the richer excursion process. For example, continuously-monitored Barrier option prices will not converge, but Asian and European options will.

4 Future work: foreign exchange applications

> **Proposed section summary.** The classical and jump Heston processes will be justified as building blocks for multi-asset frameworks, due to their tractability yet flexibility. The effect of foreign measure changes will be derived and verified so that one can ensure consistency with foreign-denominated derivatives. The *multi-Heston jump-diffusion* framework will be introduced. Volatility surface calibration and simulation procedures will be clarified. The ability of the framework to simultaneously replicate USD, EUR, JPY and GBP volatility surfaces over a range of dates will be demonstrated. Results will be contrasted with alternative (pure diffusion and jump) approaches provided by De Col et al. (2013) and Ballotta et al. (2017).

Much of the analysis for this section was completed in the first year of the PhD programme, prior to the preceding Section 3 and Section 2 here. Some of it was covered in the Early Stage Review

report. The quality of results largely motivated the choice to pursue time-homogeneous Heston-type dynamics for FX applications, in contrast to alternative solutions, for example offered by local or rough volatility options. Many of the important dependencies here (correlation structure, measure change transformations, implicit characteristic functions, simulation procedures) are verified in the jupyter notebooks available at bit.ly/2U2K2pm.

First we clarify some multi-Heston processes for FX applications, which are motivated in part by the findings in Section 2 and Section 3. The following process defines a basic multi-Heston diffusion which may be considered a particularly tractable (but not particularly realistic) form of the model proposed in De Col et al. (2013).

Definition 4.1 (Multi-Heston diffusion process)**.** *We will call a process $H = (H^i)_{i=1}^n$ a classical multi-Heston process of dimension $n \in \mathbb{N}$, built from $\mathbb{Q}$-Brownian motion $W = (W^i)_{i=0}^n$, if there exists a common CIR variance process satisfying*

$$\varepsilon dX_t = v\sqrt{X_t}dW_t^0 + (1 - X_t)dt, \quad X_0 = 1$$

and then each H^i follows Heston diffusion dynamics

$$dH_t^i = \sigma_i\sqrt{X_t}d(\bar{\rho}_i W^i + \rho_i W^0)_t - \frac{1}{2}\sigma_i^2 X_t dt, \quad [W^i, W^j]_t = \varrho_{ij}t.$$

The following multi-Heston process is motivated by results of Section 3. It characterises a weak limit of the n-dimensional multi-Heston process on the applicable $(n+1)$-dimensional uniform product topology.

Definition 4.2 (Multi-Heston excursion process)**.** *We will call the parametric process $H = (H^i)_{i=0}^n$ a multi-Heston excursion process of (spatial) dimension $n \in \mathbb{N}$, built from $\mathbb{Q}$-Brownian motion $W = (W^i)_{i=0}^n$, if there exists a common time process*

$$H_\tau^0 = \sup\left\{s - vW_s^0 : s \in [0, \tau)\right\}$$

and then each of $(H^i)_{i=1}^n$ follow Black-Scholes dynamics in the parametric index $\tau \in \mathbb{R}_+$:

$$H_\tau^i = \sigma_i\left(\bar{\rho}_i W^i + \rho_i W^0\right)_\tau - \frac{1}{2}\sigma_i^2\tau, \quad [W^i, W^j]_\tau = \varrho_{ij}\tau.$$

When projected into n-dimensional space, the Heston excursion process is finite-dimensionally equivalent to the following jump process, which fits into the framework of Ballotta et al. (2017).

Definition 4.3 (Multi-Heston jump process)**.** *We will call a process $H = (H^i)_{i=1}^n$ a multi-Heston jump process of dimension $n \in \mathbb{N}$, built from $\mathbb{Q}$-Brownian motion $W = (W^i)_{i=0}^n$, if there exists a common IG integrated-variance process*

$$\bar{X}_t = \inf\left\{s > 0 : s - vW_s^0 > t\right\}$$

and then each H^i follows Heston jump dynamics

$$H_t^i = \sigma_i\left(\bar{\rho}_i W_{\bar{X}_t}^i + \rho_i W_{\bar{X}_t}^0\right) - \frac{1}{2}\sigma_i^2\bar{X}_t, \quad [W^i, W^j]_t = \varrho_{ij}t.$$

Finally we bring the multi-Heston jump and diffusion processes together in the natural way, to define the core jump-diffusion process of the proposed framework.

Definition 4.4 (Multi-Heston jump-diffusion process)**.** *We will call a process* $H = (H^i)_{i=1}^n$ *a multi-Heston jump-diffusion process of dimension* $n \in \mathbb{N}$ *if it is additively-separable,* $H = H^0 + H^\varepsilon$*, into independent jump* H^0 *and diffusion* H^ε *multi-Heston processes each of dimension* n*.*

Notice that we have not constrained H^ε and H^0 to share the same parameters $\sigma_i, \rho_i, v, \varrho_{ij}$. Although this is clearly possible, it is not realistic in FX markets where, for example, risk-neutral distributions can be negatively skewed in short time, and positively skewed in long time.

The following headings outline a plan for this section, which in addition to completing the orange parts from Section 2 and Section 3, is expected to take roughly six months. Again, these headers are colour-coded to highlight the extent to which research on these topics is already completed.

Overview of foreign exchange problems and existing solutions. An overview of the typical problems which one faces in a multi-asset setting will be provided and pure diffusion and jump solutions of De Col et al. (2013) and Ballotta et al. (2017) covered. Focus will be placed on the effect of a 'foreign measure change', which enable us to derive implications (analytically) of our model, set up under one risk-neutral measure (identifying with our domestic bank account numéraire), on options prices which represent conditional expectations under another risk-neutral measure (identifying with a foreign bank account numéraire).

Motivation for approach: numéraire invariance and the jump volatility surface. Heston-type models for FX applications will be justified due to their invariance under foreign measure changes, which for example SABR-type models violate. See for example Graceffa, Brigo, & Pallavicini (2019) for a very recent exposition of this (also relevant more generally, given jump-diffusion focus). Further motivation for the use of the Heston jump process will be provided by the flexibility of NIG-based volatility surfaces which can be generated by such models. See Cont & Tankov (2003) for a comparison of jump-based models with local-volatility models. See also bit.ly/2U2K2pm for USD, EUR and JPY calibration examples, which verify that the Heston jump model replicates recent market data.

Effective measure changes and their Esscher transform connections

The transformative effect of foreign measure changes on the Heston jump process will be derived, and reconciled with results of Sato (1999). We will find that the IG integrated volatility process undergoes an Esscher transform, which identifies with a Girsanov transformation of the Brownian motion from which it is built. Results will contrasted with those from the classical Heston framework of De Col et al. (2013), before the reversionary time-scale ε is taken down to zero. Care will be taken to clarify the behaviour of implict Radon-Nikodym derivatives which, like the price processes these coincide with, can yield excursions.

A consistent jump-diffusion extension of the classical multi-Heston framework

The multi-Heston jump-diffusion framework will be defined and calibrated to implied volatility market data deriving from the USD, EUR, JPY and GBP currencies (six implicit volatility surfaces). Calibration procedures via characteristic functions will be provided which, as will prove necessary, will allow for near-arbitrary time-dependent forward variance curves. Results will be contrasted with those deriving purely from the jump and diffusion components, which respectively fall into the frameworks of Ballotta et al. (2017) and De Col et al. (2013). Depending on the quality of results here, the ability of the jump component to act like a local volatility decorator, through time-varying parameters, will be considered. This is known to yield near-exact replication of the liquid currency pairs in 'normal' market conditions. (One could see for example slide 18 of tiny.cc/gpmg8y for a demonstration showing 0.03% implied volatility RMSE over 65 GBPUSD quotes.)

Verification by and clarification of simulation procedures

Finally simple simulation procedures for the multi-Heston jump-diffusion process will be clarified, and used to verify the calibration results of the previous section. These are demonstrated in the pure jump case at bit.ly/2U2K2pm. Since we will typically find that, once a jump component is included (for which exact simulation is possible), the residual diffusion component verifies the Feller condition, simulation accuracy will be compared with other approaches which violate this.

References

Abi Jaber, E. (2018). Lifting the Heston model. *e-print arXiv:1810.04868*.

Abi Jaber, E., & El Euch, O. (2019). Multifactor Approximation of Rough Volatility Models. *SIAM Journal on Financial Mathematics*, *10*(2), 309–349.

Abi Jaber, E., Larsson, M., & Pulido, S. (2017). Affine Volterra processes. *e-print arXiv:1708.08796v2*.

Agarwal, R. P., & Lakshmikantham, V. (1993). *Uniqueness and Nonuniqueness Criteria for Ordinary Differential Equations*. World Scientific.

Alòs, E., León, J. A., & Vives, J. (2007). On the short-time behavior of the implied volatility for jump-diffusion models with stochastic volatility. *Finance and Stochastics*, *11*(4), 571–589.

Alòs, E., García-Lorite, D., & Muguruza, A. (2018). On smile properties of volatility derivatives and exotic products: understanding the VIX skew. *e-print arXiv:1808.03610v1*.

Applebaum, D. (2009). *Lévy Processes and Stochastic Calculus* (2nd ed.). Cambridge University Press.

Ballotta, L., Deelstra, G., & Rayée, G. (2017). Multivariate fx models with jumps: Triangles, quantos and implied correlation. *European Journal of Operational Research*, *260*(3), 1181 – 1199.

Barndorff-Nielsen, O. E. (1997). Normal Inverse Gaussian Distributions and Stochastic Volatility Modelling. *Scandinavian Journal of Statistics*, *24*(1), 1–13.

Barndorff-Nielsen, O. E., & Shephard, N. (2001). Modelling by Lévy Processess for Financial Econometrics. In *Lévy Processes*, (pp. 283–318). Birkhäuser Boston.

Billingsley, P. (1999). *Convergence of Probability Measures* (2nd ed.). John Wiley & Sons, Inc.

Coddington, A., & Levinson, N. (1955). *Theory of ordinary differential equations*. International series in pure and applied mathematics. McGraw-Hill.

Cont, R., & Tankov, P. (2003). *Financial Modelling with Jump Processes*. Chapman and Hall/CRC.

Cox, J. C., Ingersoll, J. E., & Ross, S. A. (1985). A Theory of the Term Structure of Interest Rates. *Econometrica*, *53*(2), 385–407.

De Col, A., Gnoatto, A., & Grasselli, M. (2013). Smiles all around: FX joint calibration in a multi-Heston model. *Journal of Banking & Finance*, *37*(10), 3799–3818.

Dupire, B. (1994). Pricing with a smile. *Risk Magazine*, (pp. 18–20).

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

Forde, M., & Jacquier, A. (2011). The large-maturity smile for the Heston model. *Finance and Stochastics*, *15*(4), 755–780.

Fouque, J.-P., Papanicolaou, G., Sircar, R., & Sølna, K. (2011). *Multiscale Stochastic Volatility for Equity, Interest Rate, and Credit Derivatives*. Cambridge University Press.

Gerhold, S., Gülüm, I. C., & Pinter, A. (2016). Small-Maturity Asymptotics for the At-The-Money Implied Volatility Slope in Lévy Models. *Applied Mathematical Finance*, *23*(2), 135–157.

Graceffa, F., Brigo, D., & Pallavicini, A. (2019). On the consistency of jump-diffusion dynamics for FX rates under inversion. *e-print arXiv:1905.05310*.

Han, X., & Kloeden, P. E. (2017). *Random Ordinary Differential Equations and Their Numerical Solution*. Springer Singapore.

Heston, S. L. (1993). A Closed-Form Solution for Options with Stochastic Volatility with Applications to Bond and Currency Options. *The Review of Financial Studies*, *6*(2), 327–343.

Jacod, J., & Shiryaev, A. N. (2003). *Limit Theorems for Stochastic Processes* (2nd ed.). Springer Berlin Heidelberg.

Jusselin, P., & Rosenbaum, M. (2018). No-arbitrage implies power-law market impact and rough volatility. *e-print arXiv:1805.07134*.

Kaper, H. G., & Kwong, M. K. (1988). Uniqueness results for some nonlinear initial and boundary value problems. *Archive for Rational Mechanics and Analysis*, *102*(1), 45–56.

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

Keller-Ressel, M., Larsson, M., & Pulido, S. (2018). Affine Rough Models. *e-print arXiv:1812.08486*.

McCrickerd, R. (2019). Pathwise volatility: Cox-Ingersoll-Ross initial-value problems and their fast reversion exit-time limits. *e-print arXiv:1902.01673v2*.

Mechkov, S. (2015). Fast-reversion limit of the Heston model. *e-print SSRN:2418631*.

Puhalskii, A. A., & Whitt, W. (1997). Functional large deviation principles for first-passage-time processes. *The Annals of Applied Probability*, *7*(2), 362–381.

Sato, K. (1999). *Lévy Processes and Infinitely Divisible Distributions*. Cambridge University Press.

Skorokhod, A. V. (1956). Limit Theorems for Stochastic Processes. *Theory of Probability & Its Applications*, *1*(3), 261–290.

Vellaisamy, P., & Kumar, A. (2017). First-exit times of an inverse Gaussian process. *Stochastics*, *90*(1), 29–48.

Whitt, W. (1971). Weak convergence of first passage time processes. *Journal of Applied Probability*, *8*(02), 417–422.

Whitt, W. (1980). Some Useful Functions for Functional Limit Theorems. *Mathematics of Operations Research*, *5*(1), 67–85.

Whitt, W. (2002). *Stochastic-Process Limits: An Introduction to Stochastic-Process Limits and Their Application to Queues*. Springer-Verlag New York.