

This electronic thesis or dissertation has been downloaded from the King's Research Portal at <https://kclpure.kcl.ac.uk/portal/>



Small-time & $H \downarrow 0$ limits of Rough Volatility Models

Smith, Ben

Awarding institution:
King's College London

The copyright of this thesis rests with the author and no quotation from it or information derived from it may be published without proper acknowledgement.

END USER LICENCE AGREEMENT



Unless another licence is stated on the immediately following page this work is licensed

under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International

licence. <https://creativecommons.org/licenses/by-nc-nd/4.0/>

You are free to copy, distribute and transmit the work

Under the following conditions:

- Attribution: You must attribute the work in the manner specified by the author (but not in any way that suggests that they endorse you or your use of the work).
- Non Commercial: You may not use this work for commercial purposes.
- No Derivative Works - You may not alter, transform, or build upon this work.

Any of these conditions can be waived if you receive permission from the author. Your fair dealings and other rights are in no way affected by the above.

Take down policy

If you believe that this document breaches copyright please contact librarypure@kcl.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.

Small-time & $H \downarrow 0$ limits of Rough Volatility Models



Benjamin Francisco Smith

Supervisor: Dr. M. Forde

Department of Mathematics
King's College London

This dissertation is submitted for the degree of
Doctor of Philosophy

February 2023

I would like to dedicate this thesis to my parents, Christian and Elizabeth Smith.

Declaration

I hereby declare that except where specific reference is made to the work of others, the contents of this dissertation are original and have not been submitted in whole or in part for consideration for any other degree or qualification in this, or any other university. This dissertation is my own work and contains nothing which is the outcome of work done in collaboration with others, except as specified in the text and Acknowledgements. This dissertation contains fewer than 65,000 words including appendices, bibliography, footnotes, tables and equations and has fewer than 150 figures.

Benjamin Francisco Smith

February 2023

Acknowledgements

First and foremost I'd like to thank my supervisor Martin Forde for giving me the opportunity to study for a PhD and for introducing me to many new and interesting areas of Mathematics. Without Martin's help the completion of this PhD would have been impossible.

Secondly I'd like to thank King's College London and the EPSRC for their support and funding.

Finally, many thanks to my fellow PhD students Andrei, Zach and Axel whose friendship proved invaluable especially during the pandemic.

Abstract

This thesis consists of 4 chapters each of which discusses the properties of various models from Mathematical Finance in particular Rough Volatility models.

In Chapter 1 we discuss the origins of Rough Volatility and cover various theoretical topics needed for the later chapters in the thesis. Other competing approaches are briefly discussed.

Chapter 2 (whose contents can also be found in the paper [FGS21]) introduces the Rough Heston model and demonstrates it's affine structure. In the absence of the semimartingale property of the variance, the affine structure is what we shall use to determine both small-time and large-time asymptotics for this model as well as the $H \downarrow 0$ limit.

Chapter 3 (based on the preprint titled "Small-time VIX smile and the stationary distribution for the Rough Heston model" found at <https://nms.kcl.ac.uk/martin.forde/>) stays with the Rough Heston Model but here we examine the small-time asymptotic behaviour of the VIX. We see that the model produces skewness and convexity features similar to those seen in the market.

In Chapter 4 we introduce the Gaussian Multiplicative Chaos (GMC) of the (re-scaled) Riemann-Liouville (RL) process and prove various forms of convergence as H tends to zero by comparing with the Multifractal Random Walk. We show the GMC emerges in the limit from the Rough Bergomi model as H tends to zero. We derive an approximation for the skew. In addition to the original paper on which this chapter is based [FFGS20] we prove convergence in L^1 using the abstract Shamov framework and derive a Karhunen-Loeve type expansion for the $H = 0$ field.

Table of contents

1	Introduction	9
1.1	Rough Volatility	9
1.1.1	Volatility is Rough	9
1.1.2	Rough Bergomi Model	10
1.1.3	Rough Heston Model	11
1.1.4	Skew	13
1.1.5	Aside: Hedging in Rough Models	15
1.1.6	Monte-Carlo Simulation	16
1.2	Gaussian Measures	18
1.2.1	The Cameron-Martin Space	19
1.2.2	The Theorems of Girsanov and Cameron-Martin	19
1.3	Random Measures & Multifractality	21
1.4	Other Approaches	23
1.4.1	Calibration	23
1.4.2	Machine Learning	25
2	Small-time, large-time and $H \rightarrow 0$ asymptotics for the Rough Heston model	26
2.1	Introduction	26
2.2	Rough Heston and other variance curve models - basic properties	29
2.2.1	Computing $\mathbb{E}(V_t)$	29
2.2.2	Computing $\mathbb{E}(V_u \mathcal{F}_t)$	30
2.2.3	Evolving the variance curve	31
2.2.4	The characteristic function of the log stock price	31
2.2.5	The generalized time-dependent Rough Heston model and fitting the initial variance curve	32
2.2.6	Other affine and non-affine variance curve models	32
2.3	Small-time asymptotics	32
2.3.1	Scaling relations	32
2.3.2	The small-time LDP	34

2.3.3	Asymptotics for call options and implied volatility	38
2.3.4	Series expansion for the asymptotic smile and calibration	38
2.3.5	Higher order Laplace asymptotics	39
2.4	Large-time asymptotics	46
2.4.1	Asymptotics for call options and implied volatility	49
2.4.2	Higher order large-time behaviour	49
2.5	Asymptotics in the $H \rightarrow 0$ limit	50
2.5.1	Implied vol asymptotics in the $H = 0, t \rightarrow 0$ limit - full smile effect for the Edgeworth FX options regime	52
2.5.2	A closed-form expression for the skewness, the $H \rightarrow 0$ limit and calibrating a time-dependent correlation function	52
2.5.3	Weak convergence of the V process on pathspace to a tempered distribution, and the hyper-rough Heston model	53
2.5.4	The hyper-rough Heston model for $H = 0$ - driftless and general cases	55
2.6	Appendix	57
2.6.1	Appendix A: Computing the kernel for the Rough Heston variance curve	57
2.6.2	Appendix B: The re-scaled model	58
2.6.3	Appendix C: Monotonicity property	58
2.6.4	Appendix D: Limit of Volterra equations	60
3	Small-time VIX smile for the Rough Heston model	62
3.1	Introduction	62
3.2	The Model	64
3.2.1	The small-time LDP for $(\text{VIX}_T^2 - \text{VIX}_0^2)/T^{\frac{1}{2}-H}$	66
3.2.2	VIX call option asymptotics	72
3.2.3	VIX future and implied volatility asymptotics	72
3.2.4	Small log-moneyness expansions	76
3.2.5	The Edgeworth regime	77
3.2.6	Fourier inversion formula for VIX calls for $T > 0$	78
3.3	Appendix	78
3.3.1	Appendix A: Uniqueness of solutions to fractional Riccati VIEs	78
3.3.2	Appendix B: Derivation of the VIE	80
3.3.3	Appendix C: Uniform moment bound	85
3.3.4	Appendix D: Asymptotics for VIX call options	85

4	The Riemann-Liouville field and its GMC as $H \rightarrow 0$, and skew flattening for the rough Bergomi model	88
4.1	Introduction	88
4.2	The Riemann-Liouville process and its GMC as $H \rightarrow 0$	90
4.2.1	Constructing a Gaussian multiplicative chaos from Z^H as $H \rightarrow 0$	91
4.2.2	Construction and properties of the usual Bacry-Muzy multifractal random measure (MRM) via Gaussian white noise on triangles	94
4.3	ξ_γ for the full sub-critical range $\gamma \in (0, \sqrt{2})$	95
4.3.1	The Sandwich lemma	95
4.3.2	Existence of a limiting law for ξ_γ for $\gamma \in (0, \sqrt{2})$	98
4.3.3	Existence of the GMC measure for $\gamma \in (0, \sqrt{2})$ using the Shamov approximation theorem	100
4.3.4	Local multifractality	103
4.4	Application to the Rough Bergomi model - skew flattening/blowup as $H \rightarrow 0$	104
4.4.1	$H \rightarrow 0$ behaviour for the usual rough Bergomi model	106
4.4.2	A closed-form expression for $\mathbb{E}((\tilde{X}_t^H)^3)$	106
4.4.3	Convergence of the skew to zero	107
4.4.4	Speed of convergence of the skew to zero	108
4.4.5	A $H = 0$ model - pros and cons	108
4.5	Explicit spectral expansions for Z^H for $H \geq 0$	110
4.5.1	The Cameron-Martin space of a log-correlated Gaussian field	111
4.5.2	Characterizing H_μ when $C = AA^*$	112
4.5.3	Karhunen-Loève type expansions	113
4.5.4	Choice of basis and explicit computation of terms	114
4.5.5	Using the spectral expansion to sample the GMC mass $\xi_\gamma^H([0, T])$ for $H \ll 1$ and $H = 0$	115
4.6	Appendix	118
4.6.1	Appendix A: Definition and properties of $F_H(k)$ and $G_H(k)$ for the Sandwich lemma	118
4.6.2	Appendix B: Monotonicity properties of $g_H(s, t)$	119
4.6.3	Appendix C: Proof of Proposition 4.4.3	120
4.6.4	Appendix D: Proof of skew formula	121
References		122

Chapter 1

Introduction

1.1 Rough Volatility

1.1.1 Volatility is Rough

The origins of modern mathematical finance as it is currently known can be traced back to the introduction of the Black-Scholes model which even today is considered the canonical model for financial assets in continuous time:

$$dS_t/S_t = \mu dt + \sigma dW_t \quad (1.1)$$

where W is a Brownian Motion.

This model was revolutionary in that one could price all European options (and some exotics) with corresponding unique replication strategies i.e. perfect hedging strategies. Such a model was ubiquitous in industry and contributed to the boom in the derivatives market during the end of the twentieth century.

Whilst simple this model has some substantial drawbacks. Given market prices for vanilla options we can infer the value of σ required to give the observed market price, the so-called *implied volatility*. When the implied volatility is computed from market observed equity options prices across strikes we typically see what has come to be known as a *volatility smile* i.e. a non-constant implied volatility which typically increases as we move sufficiently far away from the "at-the-money" (ATM) strike. This is in contrast to the "flat" implied volatility characteristic of the Black-Scholes model. The gradient of the smile at-the-money (where the strike is the same as the spot value) is known as the skew and it has been empirically observed to diverge as the maturity tends to zero.

Faced with the task of building more realistic models for asset prices we can divide

the various proposed approaches into two broad types. The first, which we shall not address in this thesis, consists of models that incorporate jumps into the sample paths of the asset prices thus allowing for the framework of (potentially infinite-activity) Lévy processes amongst other things.

The second type known as Stochastic Volatility models are models constructed by replacing the constant σ in the Black-Scholes model with an adapted stochastic process σ_t . A historically important class of Stochastic Volatility models are the Local Volatility models introduced by Dupire [Dup93] where $\sigma_t = \sigma(t, S_t)$ for some deterministic function $\sigma(t, S)$. Dupire shows that, by integrating the forward Kolmogorov equation, one can calibrate exactly to given market prices for vanilla prices at all strikes and maturities. In particular, given a stochastic volatility model with its associated call option prices one can find the associated local stochastic volatility (see also Gyongy [G86]). Whilst very flexible this class of models fails to capture the dynamic skew structure (i.e. power-law ATM skew *for all t*, see section 1.1.4) observed in the market for short dated options (such options are of great interest to practitioners, they tend to be the most liquid markets) and thus present great difficulties in controlling the error of hedging portfolios, see Fukasawa [Fuk17].

Another popular approach to Stochastic Volatility is to model σ_t as a diffusion which leads to many widely used models such as the Heston Model (where σ_t is a CIR process) or the SABR model. Whilst able to generate smiles similar to those observed in the market, the desired skew structure remains elusive.

Gatheral et al. [GJR18] posit that the volatility process could be modelled by a fractional Brownian motion with Hurst parameter $H < 1/2$ and showed that this choice was consistent with the scaling of realised volatility time series observed in the market with H values typically of order 0.1 (this analysis was refined by Fukasawa [FTW19] who found the H values to be even smaller). This was then followed up by another paper [BFG16] which showed that a rough model (termed the rBergomi model) could reproduce the power-law skew structure seen in options markets. Rough models had thus been shown to reproduce features of both time series data for financial assets ("under the P-measure") and also the features of option prices ("ie under Q-measure").

1.1.2 Rough Bergomi Model

In his 2005 paper [B05] Bergomi introduced the so called Bergomi model which, rather than specifying the dynamics of the spot volatility V_t , modeled the dynamics of the *forward*

variances $\xi_t^T := \mathbb{E}_t[V_T]$. The added flexibility of the model allowed for separate calibration to variance swaps which can be used as hedges. The one-factor model can be written as follows:

$$\frac{dS_t}{S_t} = \sqrt{\xi_t^T} dW_t \quad (1.2)$$

$$\frac{d\xi_t^T}{\xi_t^T} = \alpha e^{-k(T-t)} dB_t \quad (1.3)$$

where W and B are correlated Brownian motions.

This model being a diffusion makes it amenable to approximations of the smile found in the literature such as the Guyon-Bergomi expansion [Guy21b] which Friz, Bayer and Gatheral [BFG16] use to derive an expression for the ATM skew showing that it converges to a constant (which is to be expected for all classical diffusion models). The initially ad-hoc suggestion of changing the exponential kernel to a power kernel is made so as to achieve power-law behaviour of the ATM skew. Furthermore, the authors obtain such a pricing model by taking the so called RFSV model (in the statistical measure) and making a change of measure resulting in the rBergomi model:

$$\frac{dS_t}{S_t} = \sqrt{\xi_t(t)} dW_t \quad (1.4)$$

$$V_t = \xi_0(t) e^{\nu \int_0^t (t-s)^{H-\frac{1}{2}} ds - \frac{1}{2} \nu^2 \frac{t^{2H}}{2H}} \quad (1.5)$$

The aforementioned papers ([BFG16], [Fuk17]) show that indeed this reproduces the correct skew term structure and there is even a pathwise large deviation principle established in the paper by Jaquier, Pakkanen and Stone [JPS18] and [GJPSW19].

1.1.3 Rough Heston Model

One of the most famous and heretofore widely employed stochastic volatility models is the (classical) Heston model [H93]

$$dS_t = \sqrt{V_t} S_t dW_t \quad (1.6)$$

$$dV_t = \lambda(V - \theta) dt + \nu \sqrt{V_t} dB_t \quad (1.7)$$

for which the characteristic function of the log stock price can be found in semi-closed form:

$$\log \mathbb{E}(e^{pX_t}) = V_0 f(p, t) + \lambda \theta \int_0^t f(p, s) ds \quad (1.8)$$

where $f(p, t)$ solves the non-linear Riccati equation with initial condition $f(p, 0) = 0$:

$$\partial_t f(p, t) = \frac{1}{2}(p^2 - p) + (p\rho v - \lambda)f(p, t) + \frac{1}{2}v^2 f(p, t)^2 \quad (1.9)$$

The Riccati equation can be linearised and solved efficiently yielding a closed-form expression for the characteristic function. This allows us to price any European option using Fourier inversion (see Lee [L04] and [LK07]).

This feature of the model is due to its affine structure. More specifically, an affine (Volterra) process X_t is a model of the form (see Abi-Jaber [ALP19] for definitions of the relevant parameters):

$$X_t = X_0 + \int_0^t K(t-s)b(X_s)ds + \int_0^t K(t-s)\sigma(X_s)dW_s \quad (1.10)$$

where $a(x) := \sigma(x)\sigma(x)^T$ and $b(x)$ are affine functions.

Choosing a constant kernel with appropriate σ and drift b gives rise to the classical Heston model. By changing the volatility kernel to the power law kernel $t^{H-1/2}/\Gamma(H+1/2)$ gives rise to the Rough Heston model:

$$\begin{aligned} dS_t &= \sqrt{V_t}S_t dW_t \\ V_t &= V_0 + \frac{\lambda}{\Gamma(H+\frac{1}{2})} \int_0^t (t-s)^{H-\frac{1}{2}}(V_s - \theta)ds + \frac{v}{\Gamma(H+\frac{1}{2})} \int_0^t (t-s)^{H-\frac{1}{2}}\sqrt{V_t}dB_t \end{aligned} \quad (1.11)$$

This is a rough model in the sense discussed previously (the volatility sample paths have Hölder continuity $H - \varepsilon$ paths and the desired ATM skew term structure is observed as can be shown using large deviations results, see Chapter 2).

This model remains affine and we have a similar expression for the characteristic function (here I^α and D^α denote the Riemann-Liouville integral and derivative respectively):

$$\log \mathbb{E}[e^{pX_t}] = V_0 I^{\frac{1}{2}-H} f(p, t) + \lambda \theta \int_0^t f(p, s)ds \quad (1.12)$$

where now we have a fractional Riccati equation:

$$D^\alpha f(p, t) = \frac{1}{2}(p^2 - p) + (p\rho v - \lambda)f(p, t) + \frac{1}{2}v^2 f(p, t)^2 \quad (1.13)$$

such equations can be solved efficiently using an Adams scheme, so (akin to the classical case), we have a method to price any European option using Fourier inversion with e.g. gaussian quadrature.

This affine structure will be exploited in chapters 2 and 3 to derive a large deviation principle for the re-scaled log stock price and the associated VIX index.

As detailed by Rosenbaum and co-authors [ER19], one can model prices over short time scales as a Hawkes process which is a generalized Poisson process where the intensity itself depends on the history of the process via a kernel. In order to capture observed market stylised facts (namely: Market endogeneity, lack of statistical arbitrage, buy-sell asymmetry and metaorder splitting) one has to pick certain values for the parameters of the process. In a certain high-frequency limit with where the jump sizes tend to zero and the frequency of jumps tends to infinity one obtains weak convergence to the rough Heston model using C-tightness arguments from [JS13]. Hence there is a market microstructure justification for Rough Volatility models, although the original Hawkes model has the disadvantage that the stock can only go up or down by 1, and this simplicity is exploited when they appeal to C-tightness arguments.

1.1.4 Skew

Driving a volatility process with a Gaussian Volterra process significantly increases the analytical complexity of the model. In particular, the process becomes non-Markovian (due to the power kernel) and is no longer a semi-martingale. Generating a single Monte-Carlo sample path is now $O(N^2)$ (compared to $O(N)$ for a traditional model) where N is the number of time steps, and an option price can no longer be represented via the solution to a finite-dimensional PDE.

To analyse the implied volatilities generated by such models the literature has mostly focused on asymptotic estimates for the implied volatility smile in particular the short term at-the-money skew (this being a feature not explained by the previous generation of models).

In the paper by Alos et al. [ALV07], starting from a general unspecified Stochastic volatility model (with jumps), an expression for the at-the-money skew is derived in terms of Malliavin derivatives of the volatility process σ . In the case where σ is a mean-reverting process driven by a Riemann-Liouville process with $H \in (0, 1/2)$ they show that (see the paper for the definitions of the relevant quantities):

$$\lim_{T \rightarrow t} (T - t)^{H-1/2} \frac{\partial}{\partial x} I_t(x) = -c\sqrt{2\alpha} \frac{\rho}{\sigma_t} f'(Y_t) \quad (1.14)$$

where $I_t(x)$ is the implied volatility with log-price x , $\sigma_t = f(Y_t)$ and Y_t is a "rough" OU process with c as the vol-of-vol parameter and α the mean reversion parameter. We see the so-called power-law skew phenomenon (i.e. the ATM skew blows up like a power-law as the maturity tends to zero). This Malliavin calculus approach was extended to more complicated products such as VIX options in the paper by Alos, Lorite and Muguruza [AGM18]. The recent article of Jaquier et al. [JMP21] also derives the at-the-money convexity term for a two-factor rough Bergomi model with three correlation coefficients to allow for more realistic positive sloping VIX skew using similar Malliavian methods, which has implications for making smart initial guesses for calibrating such models.

The above results concern the implied volatility precisely at-the-money. A different approach is taken by Fukasawa [Fuk17] who considers the so-called Edgeworth regime i.e. the asymptotics of call option prices $c(t, k_t)$ where log-strike $k_t = k\sqrt{t}$. By expanding around the Black-Scholes price, Fukasawa shows that a classical Markovian local-stochastic volatility model cannot dynamically generate power-law skew. As discussed in [Fuk17], one can calibrate a local volatility model to a volatility surface with exploding skew (see [BDFP22]) but as time evolves the ATM short-end skew will flatten requiring a re-calibration of the model. If the same (Edgeworth) expansion is conducted for a stochastic volatility model driven by a fractional Brownian motion one obtains:

$$\frac{\sigma_t(\sqrt{\theta}z, \theta) - \sigma_t(\sqrt{\theta}\xi, \theta)}{\sqrt{\theta}(z - \xi)} \sim \theta^{H-1/2} \quad (1.15)$$

as $\theta \rightarrow 0$ where $\sigma_t(k, \theta)$ is the implied vol at time t with log-strike k and maturity T . A similar Edgeworth expansion is derived in the paper by El Euch, Fukasawa et al. [EFGR19] by expanding the log stock density at maturity using Fourier methods. This expansion provides a skew term, a curvature term and an at-the-money correction term.

Forde and Zhang [FZ17] take a new approach and characterise the asymptotic behaviour of European call option prices in the Large deviations regime when the log-moneyness $k_t = xt^{1/2-H}$. Specifically, they consider a simple correlated rough model of the form:

$$dS_t = S_t \sigma(Y_t) (\bar{\rho} dW_t + \rho dB_t) \quad (1.16)$$

$$Y_t = B_t^H \quad (1.17)$$

where B_t^H is a fractional Brownian motion and σ satisfies a certain mild continuity and growth conditions. A large deviation principle is established for $t^{H-1/2}X_t$ (X_t being the log stock price in this setup), which yields the following expression for the asymptotic implied

volatility at log-moneyness x :

$$\hat{\sigma}_0(x) = \lim_{t \rightarrow 0} \hat{\sigma}(xt^{1/2-H}, t) = \frac{|x|}{\sqrt{2\Lambda^*(x)}} \quad (1.18)$$

where $\Lambda^*(x) = \inf_{y>x} I(y)$ and $I(y)$ is the rate function which has a variational representation which can be computed numerically using the Ritz method by optimizing over a finite number of Fourier coefficients although this is quite cumbersome to compute in practice (the rough Heston short-maturity smile is much easier by comparison).

In a later paper, Friz and co-authors [BFGHS18] provide an explicit expansion for the Forde-Zhang rate function which allows for computations of the smile (in particular the atm skew). Using a stochastic Taylor series expansion combined with a Laplace approximation in the spirit of Ben Arous [Ben88], they establish the asymptotic behaviour of the smile in the moderate deviations regime which sits between the Edgeworth regime discussed by Fukasawa and the large deviations regime of Forde and Zhang:

$$\frac{\sigma_t(\theta^{1/2-H+\beta}z, \theta) - \sigma_t(\theta^{1/2-H+\beta}\xi, \theta)}{\theta^{1/2-H+\beta}(z - \xi)} \sim \rho \frac{\sigma'_0}{\sigma_0} \langle K1, 1 \rangle \theta^{H-1/2} \quad (1.19)$$

for $0 < H < \frac{1}{2}, 0 < \beta < \frac{2}{3}H$ and as $\theta \rightarrow 0$. These results are placed in the context of Rough Paths (actually, in this case, regularity structures) by Friz, Gassiat and Pigato [FGP18a] though their setup does not a priori contain the Rough Heston model.

1.1.5 Aside: Hedging in Rough Models

This thesis mainly focuses on pricing and so we only briefly discuss the issue of hedging.

In classical stochastic volatility models, the task of hedging a derivative is greatly simplified by the Markov property. An application of Ito's formula readily yields the classic PDE formulation familiar from Black-Scholes. Clearly (as previously mentioned) in rough volatility models we do not have either Markovianity or Ito's lemma (in the conventional sense) and so the classical approach will fail.

One approach found in the literature is that of Fukasawa, Horvath and Tankov [FHT21] where they show that, in a certain class of models, Markovianity is recovered in the enlarged space consisting of the spot and the forward variance curve. Using this they show that, formally, an option can be hedged by the stock and a variance swap (an instance of the Clark-Ocone formula is needed). A similar approach is adopted by Euch and Rosenbaum [ER18] where again they show a similar Markov property involving the forward variance

curve and explain how to hedge an option using the spot and the forward variance curve.

These two approaches are analytically rather complex and once any form of transaction costs are considered the problem becomes intractable. An alternative more data driven approach inspired by the Deep Hedging framework of Buehler [Bue19] was discussed in the paper [HTZ21] where a hedging portfolio was modelled as a neural network and machine learning techniques were applied to find the optimal hedge.

1.1.6 Monte-Carlo Simulation

As previously mentioned, Rough models display features observed in the markets that conventional stochastic volatility models cannot capture. This extra sophistication in the model, however, adds much difficulty to the task of pricing options in particular if one hopes to price more exotic path dependent contracts such as Asian, Barrier, forward-starting or VIX options. In the absence of analytical formulae one must turn to Monte-Carlo methods.

We first consider a class of models sometimes referred to as *simple* Rough Volatility models where the variance process is a deterministic function of the underlying fractional process. Following the notation of [Gas22]:

$$\sigma_t = f(t, \hat{W}_t), \quad \hat{W}_t = \int_0^t (t-s)^{H-1/2} dW_s \quad (1.20)$$

with the standard price dynamics $dS_t/S_t = \sigma_t(\rho dW_t + \sqrt{1-\rho^2} d\bar{W}_t)$. Pricing options will ultimately boil down to simulating random variables of the form:

$$I = \int_0^T f(t, \hat{W}_t) dW_t \quad (1.21)$$

(the integrated variance is also needed but that will be a Lebesgue integral and hence much less problematic).

Denote by I_n a Monte-Carlo estimate of I . To quantify how good the approximation we consider the weak error associated to a function Φ :

$$\mathcal{E}_\Phi = \mathbb{E}[\Phi(I)] - \mathbb{E}[\Phi(I_n)] \quad (1.22)$$

The (power) rate at which this tends to zero is the *weak error rate*, which essentially tells us how many extra simulations are required to reduce the weak error to a given tolerance.

The first task in a Monte-Carlo is to simulate \hat{W}_t . One can utilise the Cholesky decomposition (since fBm is a Gaussian process) however this becomes very computationally expensive. Alternative much faster schemes include the much used Hybrid scheme [BLP17] and the rDonsker scheme [HJM17]. Having simulated $f(t, \hat{W}_t)$ on some grid one may simulate the Ito integral I using a standard method e.g. Euler discretisation.

The simple case of $f(t, x) = x$ is known as the *rough Stein-Stein* model. Using the Cholesky method, [BHT20] show that for general $\Phi \in C_b^{\lceil 1/H \rceil}$ the weak error rate is $H + 1/2$ and if Φ is a quadratic the rate is 1. [Gas22] shows that for a $\Phi \in C_b^{2\lceil 1/4H \rceil + 3}$ one has a weak rate of $(3H + 1/2) \wedge 1$ and in the case of the Hybrid scheme a weak rate of $H + 1/2$ is obtained (a similarly result can be found in [BFN22]). It should be remarked that Monte-Carlo pricing is not the only approach for the rough Stein-Stein model and that European options can be priced using the analytic results of [A20].

For the classical and rough Heston model the square root coefficient in the vol-of-vol term is not Lipschitz at 0 placing it outside the framework of simple rough volatility models. This is a problematic feature of the model since for many commonly used parameter values the variance actually does hit zero, a property that complicates simulation and weak/strong uniqueness results (discussed in [JP20]).

Andersen addresses this issue in his paper [A07] by introducing the so-called QE scheme. The transition density for the variance process is a multiple of a non-central chi-square distribution. Andersen considers two different regimes: High vol where the chi-square can be approximated as a squared Gaussian (this corresponds to the Q in QE i.e. quadratic) and the low vol regime where the chi-square is approximated by the sum of an exponential (hence E) and a Dirac mass at zero to reflect the possibility of hitting zero. Andersen reports favourable performance of the scheme relative to previous proposals though these results are only empirical and theoretical justification remains an open question.

To simulate the Rough Heston variance process, in [G22], Gatheral describes the so-called HQE scheme (see algorithm 6.1 in the paper). Using the L^2 averaging technique introduced in [HJM17] and matching moments, one approximates the evolution of the one step forward variance curve denoted $\hat{\xi}_n$ and then simulates two independent QE variables with parameters depending on $\hat{\xi}_n$. The variance is then a linear combination of these two QE random variables (by construction guaranteed to be non-negative).

Similar to the QE scheme, the HQE scheme has not been analysed on a theoretical level. One benchmark for performance are the smiles generated by the characteristic

function methods previously discussed and Gatheral reports order 1 weak convergence (justifying his use of Richardson extrapolation though the order of weak convergence remains unproven). With regards to bias, Gatheral reports good agreement with the smile obtained via characteristic function methods. Another benchmark is the formula derived in Chapter 2 for the third moment of the *driftless* log stock price:

$$\mathbb{E}(X_T^3) = 3\rho \int_0^T \int_0^t \mathbb{E}(\sqrt{V_s} V_t dW_s dt) = \frac{3V_0 \rho \nu T^{1+\alpha}}{\Gamma(\alpha)\alpha(1+\alpha)} \quad (1.23)$$

and numerical experiments conducted by the author indicate that agreement can be achieved though this becomes harder to establish for higher values of ν and lower values of H due to the slow convergence of Monte-Carlo.

An alternative approximation for Rough models is the *Markovian Approximation* [AE19a],[A19a], [BB21] which is based on the following representation for the kernel $K(t) = t^{H-1/2}$:

$$K(t) = \int_0^\infty e^{-\gamma t} \mu(d\gamma), \quad t > 0 \quad (1.24)$$

for some Borel measure μ . Approximating μ by a positive linear combination of Dirac masses approximates K by a linear combination of exponentials. This realises the volatility process as a combination of OU processes driven by a single Brownian Motion which are standard to simulate. [Rom22] combines the Markovian approximation with the Hybrid approach [BLP17] where the kernel $K(t)$ is replaced by a linear combination of exponentials for $t > \kappa\Delta$ (unchanged elsewhere) for some low integer κ where Δ is the discretisation scale. The numerical results in that paper indicate that the Hybrid Markovian scheme outperforms the Markovian approximation and the original Hybrid scheme [BLP17].

1.2 Gaussian Measures

In order to make quantitative statements about prices we typically have to resort to asymptotics. In classical mathematical finance the spot and volatility are semimartingales which allows us to make use of well known tools such as Ito's lemma, Girsanov's theorem etc...

As mentioned previously, the rough volatility framework, by construction, lies outside of the classical semi-martingale paradigm and thus the associated prices are not amenable to the classical set of approaches (eg Feynman-Kac PDEs) and other techniques are required.

One option is to assume the market follows an affine (Volterra) process whose struc-

ture allows for the computation of characteristic functions. Other approaches include the Freidlin-Wentzell approach (Forde-Zhang), Malliavin calculus approach (Fuksawa, Alos) and the Rough Paths framework (Friz) all of which, to varying extents, utilise the theory of Gaussian measures [Bog91].

1.2.1 The Cameron-Martin Space

Given a (linear space) X with it's dual X^* a Gaussian measure is a (Borel) probability measure γ with the property that for all $\theta \in X^*$, $\theta(x)$ has a (without loss of generality, centred) Gaussian distribution with respect to γ .

This naturally induces a Hilbert space structure on X^* via the covariance structure

$$R_\gamma(\theta_1)(\theta_2) := \mathbb{E}[\theta_1(x)\theta_2(x)] \quad (1.25)$$

The closure of X^* under this inner product (in the language of Bogachev) is known as the reproducing kernel Hilbert space of the measure γ denoted X_γ^* . Note that the limit points obtained under this closure will not in general be elements of X^* , they are what are known as measurable linear functionals.

This construction can be dualised by considering the following norm with associated subset of X :

$$|h|_{H(\gamma)} := \sup\{l(h) : l \in X^*, R_\gamma(l)(l) \leq 1\} \quad (1.26)$$

$$H(\gamma) := \{h \in X : |h|_{H(\gamma)} < \infty\} \quad (1.27)$$

where one can see that the above norm is simply the operator norm acting on the Hilbert space X_γ^* and so $H(\gamma)$ is the dual of X_γ^* in the conventional sense. This space $H(\gamma)$ is known as the Cameron-Martin space of the measure γ . Naturally, being a Hilbert space, X_γ^* is isometrically isomorphic to it's dual (namely $H(\gamma)$) and this Riesz isomorphism is simply:

$$l \rightarrow R_\gamma(l, \cdot) \quad (1.28)$$

1.2.2 The Theorems of Girsanov and Cameron-Martin

The Cameron-Martin space encodes all the information of the (centred) Gaussian measure γ and in fact there is a one-to-one correspondence between candidate Cameron-Martin spaces and Gaussian measures on a space (subject to regularity conditions).

The Cameron-Martin space has the property that for any element $h = R_\gamma(g) \in H_\gamma$ the measure $\gamma_h := \gamma(\cdot - h)$ (i.e. the translation of the original measure by the element h) is equivalent to γ with density:

$$\rho_h = \exp(g(x) - \frac{1}{2}|h|_{H(\gamma)}^2) \quad (1.29)$$

indeed this property is an alternative characterisation of H_γ . This is the Cameron-Martin theorem and it underpins many of the topics covered in this thesis.

The canonical example of a Gaussian process is the Brownian motion W_t on the unit interval. It has mean zero and its covariance function is $\mathbb{E}[W_t W_s] = t \wedge s$. The corresponding Cameron-Martin space is the Sobolev space:

$$H^1 := \{f \in C^1[0, 1] : \int_0^1 \dot{f}(t)^2 dt < \infty, f(0) = 0\} \quad (1.30)$$

Most introductory accounts of continuous-time finance introduce a Black-Scholes price process and use Girsanov's theorem to add a drift to the underlying Brownian motion rendering the process risk neutral. This classic result in semi-martingale theory bears a remarkable resemblance to the above Cameron-Martin result and indeed both these results coincide when the underlying change of measure is deterministic (as a result of this these theorems are often conflated). However whereas Girsanov's theorem relies on a martingale decomposition in time (i.e. a Brownian motion W_t is a martingale in time) the Cameron-Martin theorem corresponds to a martingale decomposition in a different basis namely the reproducing kernel Hilbert space (this decomposition is known as the Karhunen-Loeve expansion [Ber17b]).

Whilst agreeing for deterministic shifts, these results nonetheless apply in different contexts. Girsanov's theorem in its classical form applies to semi-martingales but has no Gaussian assumption whereas the Cameron-Martin theorem can be applied to Gaussian processes that are not semi-martingales (e.g. fractional Brownian motion) and to more exotic Gaussian objects such as random fields which don't even exist pointwise (we shall discuss this in the case of the $H = 0$ rough Bergomi model).

That the Cameron-Martin space emerges when considering asymptotics of stochastic processes is to be expected in part due to its link to equivalent measures as detailed by the eponymous theorem. Even in the most elementary case of one dimensional Cramer's theorem, the proof of the large deviations upper bound involves changing the measure

such that the large deviation is no longer atypical and then applying elementary estimates [DZ98]. This measure change technique is the standard technique in proving large deviation principles even in much more complex situations e.g. models driven by rough paths (see [FGP18a] and references therein).

As described by Forde-Zhang [FZ17], if the dynamics of the model are driven exclusively by Gaussian (possibly fractional) noise then by applying various contraction principles the large deviation asymptotics of the price are reduced to the large deviations of the Gaussian noise. The characteristic function has the form:

$$\Lambda(\theta) = \log \mathbb{E}[e^{\theta(x)}] = \frac{1}{2} R_\gamma(\theta)(\theta) \quad (1.31)$$

The Gärtner-Ellis theorem (in this setting) states that (under certain conditions, see Dembo and Zeitouni [DZ98]) the rate function of a sequence of Gaussian measures is given by the Legendre transform of limiting logarithmic moment generating function. We can formally compute this as

$$\Lambda^*(h) = \sup_{\theta \in R_\gamma} (\theta(h) - \frac{1}{2} \langle \theta, \theta \rangle). \quad (1.32)$$

Taking the derivative with respect to θ gives $h(\cdot) - \langle \theta, \cdot \rangle = 0$. Substituting back in yields

$$\Lambda^*(h) = \frac{1}{2} |h|_{H(\gamma)}^2 \quad (1.33)$$

and hence we (formally) see the emergence of the Cameron-Martin space in the context of large-deviations.

1.3 Random Measures & Multifractality

A natural concern when studying the dynamics of a market is that of scale. Individual trades occur on short time scales which are then aggregated and drive longer term movements. Conversely long term positions are taken by investors and then dynamically hedged in the short term. Inspired by Kolmogorov's theory of turbulence in fluids, the notion of multifractality is introduced as a means of connecting the behaviour of the model over different time-scales. Bacry and Muzy [BM03], consider the q th moment of the length l increment of a random process $X(t)$:

$$m(q, l) = \mathbb{E}(|\delta_l X(t)|^q) \quad (1.34)$$

where $\delta_l X(t) = X(t+l) - X(t)$. We say this process exhibits multifractality if:

$$m(q, l) \sim C_q l^{\xi(q)} \quad (1.35)$$

for some concave function $\xi(q)$ as $l \rightarrow 0$. In a series of papers by Bacry and Muzy [BDM01] [BDM01b] [BBM13], a family of stationary models were proposed of the form:

$$X(t) = B(M([0, t])) \quad (1.36)$$

where $M(t)$ is a random measure and B is a self-similar process (typically a Brownian motion independent of M). The multifractality of the process is thus reduced to the multifractality of the random measure M .

By considering a multiplicative cascade model, multifractal random measures have been constructed on discrete scales (see the discussion in the introduction to [BM03]). To achieve this on a continuous scale Bacry & Muzy consider a general infinitely divisible random measure in the upper half plane. The integral of this over a conical domain gives a field ω_t and $M(A)$ is the integral of $e^{\gamma\omega_t}$ over the region A where γ is known as the intermittency parameter. Such measures can be shown to be exactly multifractal (not just asymptotically). See section 4.2.2 for an explicit construction.

When the underlying infinitely-divisible random measure is Gaussian, this model is referred to as the Multifractal Random Walk (MRW). The field ω_t is a Gaussian field with covariance structure

$$C(s, t) = \log\left(\frac{T}{|t-s|}\right)^+ \quad (1.37)$$

and the associated multifractal exponent:

$$\xi(q) = q - \frac{1}{2}\gamma^2(q^2 - q). \quad (1.38)$$

Random measures of this type (integrals of exponentials of Gaussian fields) are known as Gaussian Multiplicative Chaos (GMC) measures:

$$M(A) = \int_A e^{\gamma X(t) - \frac{1}{2}\gamma^2 \text{Var}(X(t))} dt \quad (1.39)$$

When X_t is a well-defined Gaussian process there is no issue defining the above measure. When X is only a field (i.e. not defined pointwise) then it is not immediately obvious if the above measure even exists let alone what properties it may have. For this one must establish an approximation $X_n(t)$ which is defined pointwise that converges in some sense to the orig-

inal Gaussian X as $n \rightarrow \infty$ and consider the limit of the corresponding random measures M_n .

For mollifier approximations to log-correlated Gaussian fields, Rhodes and Vargas [RV10] establish the convergence in law of the random measures. Berestycki [Ber17b] proves the stronger convergence in probability and in L^1 for suitable circle-average approximations. The much more general framework introduced by Shamov [Sha16] establishes the conditions under which a general approximation converges in probability and in L^1 .

These limiting measures are substantially different to their approximations. They are singular with respect to the Lebesgue measure and in fact all their mass is concentrated on the so-called set of thick points (see Berestycki [Ber17],[Ber17b]) of the Gaussian field (this set has Lebesgue measure 0). Furthermore, for γ greater than $\sqrt{2d}$ (d being the dimension of the measure space in question) the limiting measure is in fact 0 (since the set of thick points is empty).

Such measures have been used extensively in Quantum Field Theory. When the covariance of the Gaussian field is the Green's function of the Laplacian, X is the Gaussian Free Field. The associated random measure is the Liouville measure which was used by Vargas et al. to construct a rigorous mathematical model of Liouville Quantum Gravity (see the lecture notes by Berestycki [Ber17]).

1.4 Other Approaches

In this section we briefly discuss approaches other than Rough Volatility to the tasks of pricing and calibration. This is included to put Rough Volatility in context but will not be the focus of this thesis.

1.4.1 Calibration

The rough volatility approach is inherently parametric. As previously discussed it reproduces many stylised features observed in both financial time series and market prices which is clearly of theoretical interest. A financial institution is concerned with the more everyday task of pricing and hedging exotic derivatives in the presence of market data. For this one requires a market model which is consistent with market prices so as to avoid arbitrage and correctly prices any hedging instruments in the market.

The traditional approach would be to pick a particular model (e.g. a Heston, SABR or rough volatility model) and then calibrate the parameters to fit market data (for example

SPX and VIX futures and options) also known as model calibration. This is naturally a non-trivial task, the map from model parameters to prices is not available in closed form.

One approach explored in the paper by Guo, Loeper and Wang [GLW22] is to re-cast the calibration problem as a semi-martingale optimal transport problem. More explicitly, we are looking for a (continuous) semi-martingale which returns the observed market prices (a finite number in this case) and minimises some objective functional. If such a functional is chosen to be strictly convex then our search can be restricted to the set of Markov semi-martingales.

This problem can be dualised resulting in a dual problem consisting of an optimisation over a single real variable with a constraint taking the form of a non-linear Hamilton-Jacobi equation with a penalty term. This leads to a max-min problem of the form:

$$\inf_{a \in \mathcal{A}} \sup_{w_i} \left(- \sum_{i=1}^n w_i c_i + \sum_{i=1}^n w_i \mathbb{E}((X_T - K_i)^+) + \text{cost term} \right) \quad (1.40)$$

The infimum and supremum are exchanged by means of the Fenchel-Rockafellar theorem.

This approach can be adapted to more complex products and can be used give a partial solution to the SPX-VIX calibration problem as shown by Obloj and co-authors[GLWO20].

A similar approach (specialised here to the SPX-VIX problem) is that employed by Guyon [Guy21]. Instead of focusing on continuous-time models Guyon considers a discrete-time model. More specifically he answers the following question: given SPX smiles at maturities $\{T, T + \Delta\}$ and a T -maturity VIX smile can we find a market model that perfectly calibrates to these smiles in an arbitrage free way. To solve this Guyon introduces the so called "VIX-constrained martingale Schrödinger problem":

$$D_{\bar{\mu}} = \inf_{\mu \in P(\mu_1, \mu_V, \mu_2)} H(\mu | \hat{\mu}) \quad (1.41)$$

where $H(\mu | \hat{\mu})$ is the relative entropy between μ and a chosen sensible reference measure $\bar{\mu}$.

This problem can be dualised as follows

$$D_{\bar{\mu}} = \sup_{u \in U} J_{\bar{\mu}}(u) \quad (1.42)$$

where

$$J_{\bar{\mu}}(u) = \mathbb{E}^1[u_1(S_1)] + \mathbb{E}^V[u_V(V)] + \mathbb{E}^2[u_1(S_2)] - \log \mathbb{E}^{\bar{\mu}}[e^{u_1 + u_V + u_2 + \Delta_S + \Delta_L}] \quad (1.43)$$

and U is the set of admissible portfolios consisting of European options on S_1, S_2, V , delta hedged positions in the stock and forward-starting log contract.

If this supremum is attained at an optimal portfolio u^* then (due to the particular choice of the relative entropy as the relevant cost function) the optimal measure μ^* can be expressed as:

$$\frac{d\mu^*}{d\bar{\mu}} = e^{u_1^* + u_V^* + u_2^* + \Delta_S^* + \Delta_L^*} / Z \quad (1.44)$$

where Z is a normalisation constant so that the new measure has unit mass.

This optimal portfolio can be approximated using the Sinkhorn fixed-point algorithm. Once this portfolio has been found (to the desired accuracy) the Radon-Nikodym derivative above yields the optimal arbitrage-free calibrating model.

Both these approaches have the advantage that they work completely generally and can fit smiles or prices arbitrarily well but have the downside of returning models that are (in the first case) diffusions and so will have the standard problems associated to such processes (insufficient skew on the short end) and in the second case the resulting model will be in discrete time and so unsuitable for pricing exotic options. Furthermore it is a non trivial task to check if a given triple (μ_1, μ_V, μ_2) is arbitrage free.

1.4.2 Machine Learning

Another approach to calibration is to learn the pricing map by approximating the map with a neural network as discussed by Horvath et al. [HMT19]. Given a set of training data (a set of option prices for a large range of different (random) model parameter values) one can closely approximate the model pricing map with a neural network after optimizing the weights and bias of the NN. The training is performed as a one-off task (which can be lengthy), but once this is done, calibration can be very quick. This separation of the calibration into two different steps allows for a significant speed up in the computations which is useful for practitioners.

Chapter 2

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

2.1 Introduction

[JR16] 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. [ER19] show that the characteristic function of the log stock price for the Rough Heston model is the solution to a fractional Riccati equation which is non-linear (see also [EFR18] and [ER18]), 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 2.2.2 below for a proof of this). Theorem 2.1 in [ER18] 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, [ER18] 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 [DJR19] 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 VIE techniques in this paper).

[GK19] 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 [ER19], 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 Riccati equation, or equivalently is a non-linear function of the solution to a VIE.

[GGP19] 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. [GGP19] 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 M.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 [GGP19].

[JP20] compute a small-time 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

$$Y_t = Y_0 + \int_0^t K_2(t-s)\zeta(Y_s)dW_s \quad (2.1)$$

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 $K_2(t) = \text{const.}t^{H-\frac{1}{2}}$ the right term is proportional to the α -th fractional integral of ζv (where $\alpha = H + \frac{1}{2}$), and in this case [JP20] show that $Y_{\varepsilon(\cdot)}$ satisfies

an LDP as $\varepsilon \rightarrow 0$ with rate function

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

(see Proposition 4.3 in [JP20]) 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 (2.1) 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 Volterra integral equations.

Corollary 7.1 in [FGP18a] provides a sharp small-time expansion in the [FZ17] 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 behaviour obtained in [FZ17], and some earlier intermediate results in Bayer et al. [BFGHS18]. [FSV19] 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 [EFGR19] and [FSV19] 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 chapter, we establish small-time 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 $\hat{\sigma}(x)$. 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 [FJ11].

In the final section, using Lévy's convergence theorem and result from [GLS90] 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 $\alpha \rightarrow \frac{1}{2}$ 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 non-symmetric when $\rho \neq 0$. From this we show that $X_t^{(\frac{1}{2})} / \sqrt{t}$ tends weakly to a non-symmetric random variable as $t \rightarrow 0$, which leads to a non-trivial asymptotic smile in the Edgeworth (or central limit theorem) regime. where the log-moneyness scale as $z = k\sqrt{t}$ as $t \rightarrow 0$. 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 $H \rightarrow 0$ (in contrast to the Rough Bergomi model discussed in Chapter 4 where the skew flattens or blows up depending on the vol-of-vol parameter γ) and using the expression in [ALP19] 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 non-linear VIE and (from Theorem 2.5 in [A19b]) this tempered distribution has the same law as the $H = 0$ hyper rough Heston model.

2.2 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 non-affine forward variance models. Most of the results in this section are given in various locations in [ER18],[ER19] and [GK19], 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 [JR16]:

$$\begin{aligned} dX_t &= -\frac{1}{2}V_t dt + \sqrt{V_t} dB_t \\ V_t &= V_0 + \frac{1}{\Gamma(\alpha)} \int_0^t (t-s)^{\alpha-1} \lambda(\theta - V_s) ds + \frac{1}{\Gamma(\alpha)} \int_0^t (t-s)^{\alpha-1} \nu \sqrt{V_s} dW_s \end{aligned} \quad (2.2)$$

for $\alpha \in (\frac{1}{2}, 1)$, $\theta > 0$, $\lambda \geq 0$ 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 .

2.2.1 Computing $\mathbb{E}(V_t)$

Proposition 2.2.1

$$\mathbb{E}(V_t) = V_0 - (V_0 - \theta) \int_0^t f^{\alpha, \lambda}(s) ds \quad (2.3)$$

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 page 7 in [GK19]), and Proposition 3.1 in [ER18] for an alternate proof). Let $r(t) = f^{\alpha, \lambda}(t)$. Taking expectations of (3.1) and using that the expectation of the stochastic integral term is zero, we see that

$$\mathbb{E}(V_t) = V_0 + \frac{1}{\Gamma(\alpha)} \int_0^t (t-s)^{\alpha-1} \lambda(\theta - \mathbb{E}(V_s)) dt. \quad (2.4)$$

2.2 Rough Heston and other variance curve models - basic properties

Let $k(t) := \frac{\lambda t^{\alpha-1}}{\Gamma(\alpha)}$ and $f(t) := \mathbb{E}(V_t) - \theta$. Then we can re-write (2.4) as

$$f(t) = (V_0 - \theta) - k * f(t). \quad (2.5)$$

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 (2.5) 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 re-arrange 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 $r(t) = \lambda t^{\alpha-1} E_{\alpha, \alpha}(-\lambda t^\alpha)$. ■

2.2.2 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

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

If $\lambda = 0$, we can re-write this expression as

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

Proposition 2.2.2 (see [ER19]). For $\lambda > 0$

$$d\xi_t(u) = \kappa(u-t) \sqrt{V_t} dW_t = \kappa(u-t) \sqrt{\xi_t(t)} dW_t \quad (2.6)$$

2.2 Rough Heston and other variance curve models - basic properties

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

$$\kappa(x) = vx^{\alpha-1}E_{\alpha,\alpha}(-\lambda x^\alpha) \sim \frac{1}{\Gamma(\alpha)}vx^{\alpha-1} \quad (2.7)$$

as $x \rightarrow 0$ (see also page 6 in [GK19] and page 29 in [ER18]).

Proof. See Appendix Appendix A. ■

Remark 2.2.1 Integrating (2.6) and setting $u = t$ we see that

$$V_t = \xi_0(t) + \int_0^t \kappa(t-s)\sqrt{V_s}dW_s. \quad (2.8)$$

Remark 2.2.2 From (2.6), we see that $\xi_t(\cdot)$ is Markov in $\xi_t(\cdot)$. However V is not Markov in itself.

2.2.3 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 (2.3) and the expression for $\kappa(t)$ in Proposition 2.2.2 (which are both expressed in terms of the Mittag-Leffler function).

2.2.4 The characteristic function of the log stock price

From Corollary 3.1 in [ER19] (see also Theorem 6 in [GGP19]), we know that for all $t \geq 0$

$$\mathbb{E}(e^{pX_t}) = e^{V_0 t^{1-\alpha} f(p,t) + \lambda \theta t^1 f(p,t)} \quad (2.9)$$

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

$$D^\alpha f(p, t) = \frac{1}{2}(p^2 - p) + (p\rho v - \lambda)f(p, t) + \frac{1}{2}v^2 f(p, t)^2 \quad (2.10)$$

with initial condition $f(p, 0) = 0$, where $I^\alpha f$ denotes the fractional integral operator of order α (see e.g. page 16 in [ER19] for definition) and D^α denotes the fractional derivative operator of order α (see page 17 in [ER19] for definition).

2.2.5 The generalized time-dependent Rough Heston model and fitting the initial variance curve

If we now replace the constant θ 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 re-arrange 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 [ER18]).

2.2.6 Other affine and non-affine variance curve models

Another well known (and non-affine) 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 $d\xi_t(u) = \eta e^{-\lambda(u-t)} \xi_t(u) dW_t$.

2.3 Small-time asymptotics

2.3.1 Scaling relations

Let

$$d\tilde{X}_t^\varepsilon = \sqrt{\varepsilon} \sqrt{V_t^\varepsilon} dB_t \tag{2.11}$$

where V_t^ε is defined in Appendix B. This is simply the small-noise Rough Heston variance process (for now we set $\lambda = 0$).

This satisfies (see Appendix B):

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

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

$$\mathbb{E}(e^{p\tilde{X}_t}) = e^{V_0 t^{1-\alpha} \psi(p,t)} \tag{2.12}$$

where $\psi(p, t)$ satisfies:

$$D^\alpha \psi(p, t) = \frac{1}{2} p^2 + p\rho v \psi(p, t) + \frac{1}{2} v^2 \psi(p, t)^2 \quad (2.13)$$

with $\psi(p, 0) = 0$.

This expression for the characteristic function is simply the [ER19] formula with all the drift terms removed. A more direct proof of this result can be found in Appendix B of the next chapter which is a minor variant of [ALP19].

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

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

Combining this with (2.13) we see that

$$\varepsilon^{-\alpha} D^\alpha (\psi(p, \varepsilon(\cdot)))(t) = \frac{1}{2} p^2 + p\rho v \psi(p, \varepsilon t) + \frac{1}{2} v^2 \psi(p, \varepsilon t)^2. \quad (2.14)$$

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

$$\varepsilon^{-\alpha-2\gamma} D^\alpha (\psi(\varepsilon^\gamma q, \varepsilon(\cdot)))(t) = \frac{1}{2} q^2 + q\rho v \varepsilon^{-\gamma} \psi(\varepsilon^\gamma q, \varepsilon t) + \frac{1}{2} v^2 \varepsilon^{-2\lambda} \psi(\varepsilon^\gamma q, \varepsilon t)^2 \quad (2.15)$$

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

$$D^\alpha (\varepsilon^\alpha \psi(\varepsilon^{-\alpha} q, \varepsilon(\cdot)))(t) = \frac{1}{2} q^2 + q\rho v \varepsilon^\alpha \psi(\varepsilon^{-\alpha} q, \varepsilon t) + \frac{1}{2} v^2 \varepsilon^{2\alpha} \psi(\varepsilon^{-\alpha} q, \varepsilon t)^2 \quad (2.16)$$

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

$$\psi(p, t) = \varepsilon^\alpha \psi(\varepsilon^{-\alpha} p, \varepsilon t) \quad (2.17)$$

From the form of the characteristic function in (2.12), the function $\Lambda(p, t) := I^{1-\alpha}\psi(p, t)$ is clearly of interest too. Using the scaling relation on $\psi(p, t)$:

$$I^{1-\alpha}\psi(p, \varepsilon t) = \frac{1}{\Gamma(1-\alpha)} \int_0^{\varepsilon t} (\varepsilon t - s)^{-\alpha} \psi(p, s) ds \quad (2.18)$$

$$= \frac{\varepsilon}{\Gamma(1-\alpha)} \int_0^t (\varepsilon t - \varepsilon u)^{-\alpha} \psi(p, \varepsilon u) du \quad (2.19)$$

$$= \frac{\varepsilon^{1-\alpha}}{\Gamma(1-\alpha)} \int_0^t (t - u)^{-\alpha} \varepsilon^{-\alpha} \psi(\varepsilon^\alpha p, u) du \quad (2.20)$$

$$= \varepsilon^{-2H} I^{1-\alpha} \psi(\varepsilon^\alpha p, t) \quad (2.21)$$

Thus we have established the following lemma:

Lemma 2.3.1

$$\Lambda(p, \varepsilon t) = \varepsilon^{-2H} \Lambda(\varepsilon^\alpha p, t) \quad (2.22)$$

in particular

$$\Lambda(p, t) = t^{-2H} \Lambda(pt^\alpha, 1). \quad (2.23)$$

2.3.2 The small-time LDP

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

Assumption 2.3.2 $\lambda = 0$.

Remark 2.3.1 The formal higher order Laplace asymptotics indicate that λ will not affect the leading order small-time asymptotics, i.e. λ 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 chapter (recall that $\alpha = H + \frac{1}{2}$):

Theorem 2.3.3 *For the Rough Heston model defined in (2.2), we have*

$$\lim_{t \rightarrow 0} t^{2H} \log \mathbb{E}(e^{\frac{p}{t^\alpha} X_t}) = \lim_{t \rightarrow 0} t^{2H} \log \mathbb{E}(e^{\frac{p}{t^{2H}} \frac{X_t}{t^{1-H}}}) = \begin{cases} \bar{\Lambda}(p) & \text{if } T^*(p) > 1 \\ +\infty & \text{if } T^*(p) \leq 1 \end{cases} \quad (2.24)$$

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

$$D^\alpha \psi(p, t) = \frac{1}{2}p^2 + p\rho v\psi(p, t) + \frac{1}{2}v^2\psi(p, t)^2 \quad (2.25)$$

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 $v > 0$). Moreover, the scaling relation in the previous section show that $\Lambda(p) = |p|^{\frac{2H}{\alpha}}\Lambda(\text{sgn}(p), |p|^{\frac{1}{\alpha}})$, so in fact we only need to solve (2.25) for $p = \pm 1$, and we can re-write (2.24) in more familiar form as

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

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 $t \rightarrow 0$ 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 re-scaled Rough Heston models:

$$dX_t^\varepsilon = -\frac{1}{2}\varepsilon V_t^\varepsilon dt + \sqrt{\varepsilon}\sqrt{V_t^\varepsilon}dB_t \quad (2.26)$$

$$V_t^\varepsilon = V_0 + \frac{\varepsilon^\alpha}{\Gamma(\alpha)} \int_0^t (t-s)^{H-\frac{1}{2}} \lambda(\theta - V_s^\varepsilon) ds + \frac{\varepsilon^H}{\Gamma(\alpha)} \int_0^t (t-s)^{H-\frac{1}{2}} v \sqrt{V_s^\varepsilon} dW_s \quad (2.27)$$

with $X_0^\varepsilon = 0$, where $H = \alpha - \frac{1}{2} \in (0, \frac{1}{2}]$. Then from Appendix B we know that

$$(X_{(\cdot)}^\varepsilon, V_{(\cdot)}^\varepsilon) \stackrel{(d)}{=} (X_{\varepsilon(\cdot)}, V_{\varepsilon(\cdot)}) \quad (2.28)$$

(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 [FZ17], we let \tilde{X}_t^ε denote the solution to

$$d\tilde{X}_t^\varepsilon = \sqrt{\varepsilon}\sqrt{V_t^\varepsilon}dB_t \quad (2.29)$$

with $\tilde{X}_0^\varepsilon = 0$. From the previous section we know that

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

on some non-empty interval $[0, T^*(p))$, where

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

with $\psi(p, 0) = 0$. Existence and uniqueness of solutions to these kind of fractional differential equations (FDE) is standard, as is their equivalence to VIEs, see e.g. [GGP19] and chapter 12 of [GLS90] for details.

From Propositions 2 and 3 in [GGP19], we know that $\psi(p, t)$ blows up at some finite time $T^*(p) > 0$ (i.e. case A or B in the [GGP19] classification). Thus we see that

$$\mathbb{E}(e^{\frac{p}{\varepsilon^\alpha} \tilde{X}_t^\varepsilon}) = \mathbb{E}(e^{\frac{p}{\varepsilon^\alpha} \tilde{X}_{\varepsilon t}}) = e^{V_0 I^{1-\alpha} \psi(\frac{p}{\varepsilon^\alpha}, \varepsilon t)} = e^{\frac{1}{\varepsilon^{2H}} V_0 I^{1-\alpha} \psi(p, t)} \quad (2.30)$$

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

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

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 Λ scaling relation in (2.23):

Corollary 2.3.4

$$\Lambda(q) = t^{2H} \Lambda(\frac{q}{t^\alpha}, t) = |q|^{\frac{2H}{\alpha}} \Lambda(\text{sgn}(q), |q|^{\frac{1}{\alpha}}) \quad (2.31)$$

where we have set $p = 1 = \frac{|q|}{t^\alpha}$ in (2.23), and $t_q^* = |q|^{\frac{1}{\alpha}}$.

Remark 2.3.2 This implies that $\Lambda(p) \rightarrow \infty$ as $p \rightarrow p_\pm := \pm(T^*(\pm 1))^\alpha$. and more generally

$$p T^*(p)^\alpha = 1_{p>0} p_+ + 1_{p<0} p_- . \quad (2.32)$$

To prove the LDP, we first prove the corresponding LDP for \tilde{X}_t . From Lemma 2.3.9 in [DZ98], we know that

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

is convex in p , and from (2.9) and (2.13) we know that

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

(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 (2.31), 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 $\psi(p, t) \rightarrow \infty$ as $t \rightarrow T^*(p)$ (see Propositions 2 and 3 in [GGP19]), so $\Lambda(p, t) = I^{1-\alpha} \psi(p, t)$ also explodes at $T^*(p)$ by Lemma 3 in [GGP19].

Then from Corollary 2.3.4, we know that $\Lambda(p) = p^{\frac{2H}{\alpha}} \Lambda(\text{sgn}(p), |p|^{\frac{1}{\alpha}})$, so $\Lambda(p) \rightarrow \infty$ as $p \rightarrow p_{\pm} = \pm(T^*(\pm 1))^{\alpha}$ and (by convexity and differentiability) Λ is also essentially smooth, so by the Gärtner-Ellis theorem from large deviations theory (see Theorem 2.3.6 in [DZ98]), $\tilde{X}_1^{\varepsilon}/\varepsilon^{\frac{1}{2}-H}$ satisfies the LDP as $\varepsilon \rightarrow 0$ with speed ε^{-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 non-zero drift of the log stock price can effectively be ignored at leading order in the limit as $\varepsilon \rightarrow 0$. First, note that

$$\mathbb{E}(e^{\frac{p}{\varepsilon^{2\alpha}} \varepsilon \int_0^1 V_s^{\varepsilon} ds}) = \mathbb{E}(e^{\frac{p}{\varepsilon^{2H}} \int_0^1 V_s^{\varepsilon} ds}) = \mathbb{E}(e^{\frac{\sqrt{2p}}{\varepsilon^{\alpha}} \hat{X}_1^{\varepsilon}}) = e^{\frac{1}{\varepsilon^{2H}} V_0 \Lambda_0(\sqrt{2p})}$$

for $p \in (-\infty, \frac{1}{2}p_+)$ (and $+\infty$ otherwise). The \hat{X} process after the second equals sign corresponds to the \tilde{X} process when $\rho = 0$ as reflected in the subscript of Λ after the third equals sign. So

$$J(p) := \lim_{\varepsilon \rightarrow 0} \varepsilon^{2H} \log \mathbb{E}(e^{\frac{p}{\varepsilon^{2\alpha}} \varepsilon \int_0^1 V_s^{\varepsilon} ds}) = V_0 \Lambda_0(\sqrt{2p})$$

so (again using part a) of the Gärtner-Ellis theorem in Theorem 2.3.6 in [DZ98]), $A_{\varepsilon} := \int_0^1 V_s^{\varepsilon} ds$ satisfies the upper bound LDP as $\varepsilon \rightarrow 0$ with speed ε^{-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 $\delta_1 > 0$

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

for any ε sufficiently small, where we have used the upper bound LDP for A_{ε} to obtain the final inequality. Thus

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

but a is arbitrary and (from Lemma 2.3.9 in [DZ98]), J^* is a good rate function, so in fact

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

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 [DZ98], so (by Theorem 4.2.13 in [DZ98]) $\frac{X_1^{\varepsilon}}{\varepsilon^{\frac{1}{2}-H}}$ satisfies the same LDP as $\frac{\tilde{X}_1^{\varepsilon}}{\varepsilon^{\frac{1}{2}-H}}$. ■

2.3.3 Asymptotics for call options and implied volatility

Corollary 2.3.5 *We have the following limiting behaviour for out-of-the-money European put and call options with maturity t and log-strike $t^{\frac{1}{2}-H}x$, with $x \in \mathbb{R}$ fixed:*

$$\begin{aligned} \lim_{t \rightarrow 0} t^{2H} \log \mathbb{E}((e^{X_t} - e^{xt^{\frac{1}{2}-H}})^+) &= -I(x) & (x > 0) \\ \lim_{t \rightarrow 0} t^{2H} \log \mathbb{E}((e^{xt^{\frac{1}{2}-H}} - e^{X_t})^+) &= -I(x) & (x < 0). \end{aligned}$$

Proof. The lower estimate follows from the exact same argument used in Appendix C in [FZ17] (see also Theorem 6.3 in [FGP18b]). The proof of the upper estimate is the same as in Theorem 6.3 in [FGP18b]. ■

Corollary 2.3.6 *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 (3.1) for $\lambda = 0$. Then for $x \neq 0$ fixed, the implied volatility satisfies*

$$\hat{\sigma}(x) := \lim_{t \rightarrow 0} \hat{\sigma}_t(t^{\frac{1}{2}-H}x) = \frac{|x|}{\sqrt{2I(x)}}. \quad (2.33)$$

Proof. Follows from Corollary 7.2 in [GL14]. See also the proof of Corollary 4.1 in [FGP18b] for details on this, but the present situation is simpler, as we only require the leading order term here. ■

2.3.4 Series expansion for the asymptotic smile and calibration

Proceeding as in Lemma 12 in [GGP19], we can compute a fractional power series for $\psi(p, t)$ (and hence $\Lambda(p, t)$) and then using (2.31), we find that

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

where the $a_n = a_n(u)$ coefficients are defined (recursively) as in [GGP19] except for our application here (based on (2.13)) 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

$$\hat{\sigma}(x) = \sqrt{V_0} + \frac{\rho v}{2\Gamma(2+\alpha)\sqrt{V_0}}x + v^2 \frac{\Gamma(1+2\alpha) + 2\rho^2\Gamma(1+\alpha)^2(2 - 3\frac{\Gamma(2+2\alpha)}{\Gamma(2+\alpha)^2})}{8V_0^{\frac{3}{2}}\Gamma(1+\alpha)^2\Gamma(2+2\alpha)}x^2 + O(x^3). \quad (2.34)$$

(compare this to Theorem 3.6 in [BFGHS18] for a general class of rough models and Theorem 4.1 in [FJ11b] for a Markovian local-stochastic volatility model). We can re-write this expansion more concisely in dimensionless form as

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

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

Remark 2.3.3 In principle one can use (2.34) to calibrate V_0 , ρ and v 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).

Wing behaviour of the rate function

From Eq 3.2 in [RO96], we expect that $\psi(p, t) \sim \frac{\text{const.}}{(T^*(p)-t)^\alpha}$ as $t \rightarrow T^*(p)$ and thus $\Lambda(p, t) = I^{1-\alpha} \psi(p, t) \sim \frac{\text{const.}}{(T^*(p)-t)^{2\alpha-1}}$ as $t \rightarrow T^*(p)$. Assuming this is consistent with the p -asymptotics, then (by (2.32)) we have

$$\Lambda(p) = \Lambda(p, 1) \sim \frac{\text{const.}}{(T^*(p)-1)^{2\alpha-1}} = \frac{\text{const.}}{\left(\left(\frac{p_+}{p}\right)^{1/\alpha} - 1\right)^{2\alpha-1}} \sim \frac{\text{const.}}{(p_+ - p)^{2\alpha-1}} \quad (p \rightarrow p_+)$$

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

2.3.5 Higher order Laplace asymptotics

The following is a shortened version of section 3.5 in [FGS21]. If we now relax the assumption that $\lambda = 0$, and work with the original X^ε process in (2.27) (as opposed to the driftless \tilde{X}^ε process), 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 non-empty interval $[0, T_\varepsilon^*(p))$, where

$$g_\varepsilon\left(\frac{p}{\varepsilon^\alpha}, t\right) = \frac{\Psi(p, t)}{\varepsilon^{2H}} \quad (2.35)$$

which satisfies

$$D^\alpha g_\varepsilon(p, t) = \frac{1}{2}\varepsilon(p^2 - p) + (p\rho v - \lambda)\varepsilon^\alpha g_\varepsilon(p, t) + \frac{1}{2}\varepsilon^{2H} v^2 g_\varepsilon(p, t)^2 \quad (2.36)$$

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

$$g_\varepsilon\left(\frac{p}{\varepsilon^\alpha}, t\right) = \frac{\Psi_\varepsilon(p, t)}{\varepsilon^{2H}} \quad (2.37)$$

and setting $p \mapsto \frac{p}{\varepsilon^\alpha}$, and substituting for $g_\varepsilon\left(\frac{p}{\varepsilon^\alpha}, t\right)$ in (2.36) and multiplying by ε^{2H} as before, we find that

$$D^\alpha \Psi_\varepsilon(p, t) = \frac{1}{2}p^2 + p\rho v \Psi_\varepsilon(p, t) + \frac{1}{2}v^2 \Psi_\varepsilon(p, t)^2 - \varepsilon^\alpha \left(\frac{1}{2}p + \lambda \Psi_\varepsilon(p, t)\right)$$

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

$$D^\alpha \Psi_1(p, t) = -\frac{1}{2}p - \lambda \Psi(p, t) + p\rho v \Psi_1(p, t) + v^2 \Psi(p, t) \Psi_1(p, t)$$

with $\Psi_1(p, 0) = 0$, which is a linear VIE for $\Psi_1(p, t)$.

This gives us an approximation to the characteristic function and, using Laplace and Saddle point techniques as in [FJL12], (formally) gives the following expression (see [FGS21] for definitions):

$$\begin{aligned} C_\varepsilon(x) &= \mathbb{E}((e^{X_1^\varepsilon} - e^{x\varepsilon^{\frac{1}{2}-H}})^+) \\ &= \frac{1}{2\pi} e^{x\varepsilon^{\frac{1}{2}-H}} \int_{-ip^*-\infty}^{-ip^*+\infty} \operatorname{Re}\left(\frac{e^{-izx\varepsilon^{\frac{1}{2}-H}}}{-iz - z^2} \mathbb{E}(e^{izX_1^\varepsilon})\right) dz \end{aligned} \quad (2.38)$$

$$= \frac{A(x)\varepsilon^{\frac{1}{2}+2H} e^{-\frac{I(x)}{\varepsilon^{2H}}}}{\sqrt{2\pi}} [1 + \varepsilon^{\frac{1}{2}-H}(x + G(k^*) + \lambda \theta F_1(k^*)) + O(\varepsilon^{(1-2H)\wedge 2H})] \quad (2.39)$$

where

$$A(x) = \frac{1}{(p^*)^2 \sqrt{\bar{\Lambda}''(p^*)}} \quad (2.40)$$

The ε -dependence of the leading order term here is exactly the same as in Corollary 7.1 in the article of Friz et al. [FGP18a] (in [FGP18a] $\varepsilon^2 = t$ whereas here $\varepsilon = t$) which deals with a general class of rough stochastic volatility models (which excludes Rough Heston).

(2.39) 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 [Olv74] (which we have checked and tried), it still frequently

gives a worse estimate than 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 saddlepoint (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 graph and tables below). See [LK07] for more on computing the optimal contour of integration for such problems.

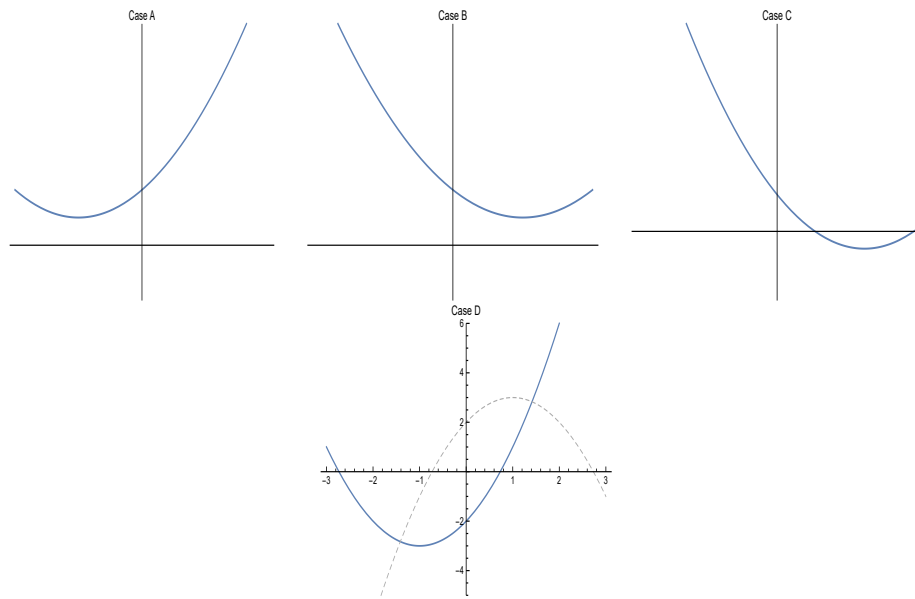


Fig. 2.1 Here we have plotted the quadratic function $G(p, w)$ as a function of w for the four cases described in [GGP19]. In cases A and B there are no roots and the solution $\psi(p, t)$ to (2.13) increases without bound whereas in cases C and D we have a stable fixed point (the lesser of the two roots) 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.

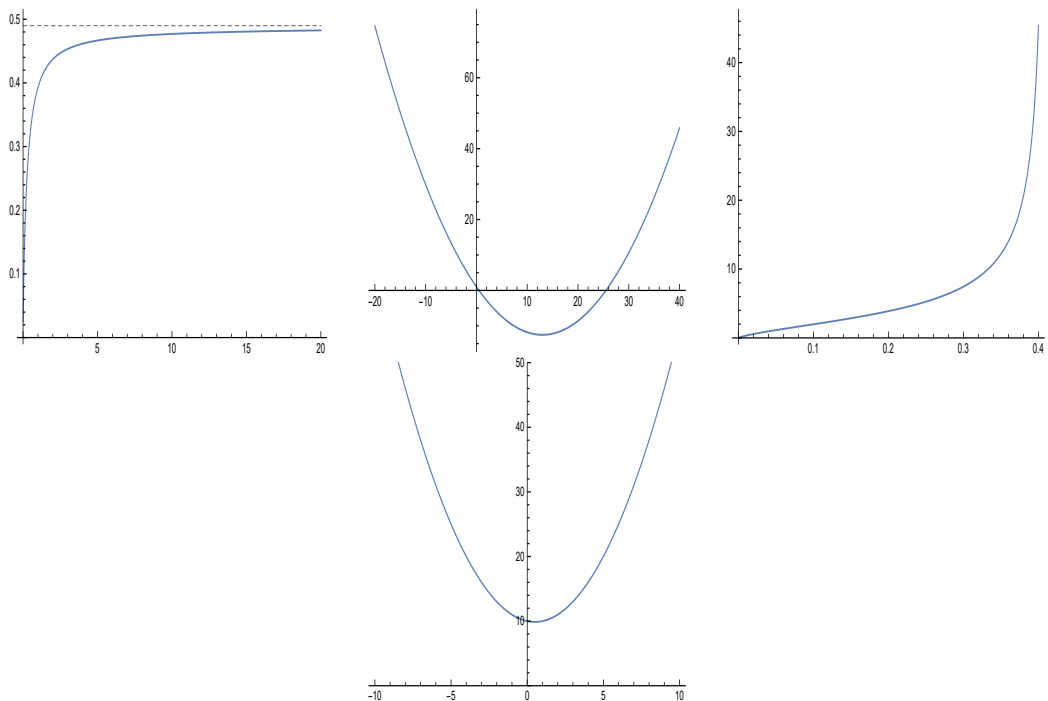


Fig. 2.2 Here we have solved for the solution $f(p, t)$ to (2.10) numerically by discretizing the VIE with 2000 time steps, and plotted $f(p, t)$ as a function of t and the corresponding quadratic function $G(p, w)$ as a function of w with p fixed. In the first case $\alpha = .75$, $\lambda = 2$, $\rho = -0.1$, $v = .4$ and $p = 2$ and $f(p, t)$ tends to a finite constant, and in the second case $\alpha = .75$, $\lambda = 1$, $\rho = 0.1$, $v = 1$ and $p = 5$ and we see that $f(p, t)$ has an explosion time at some $T^*(p) \approx 0.4$.

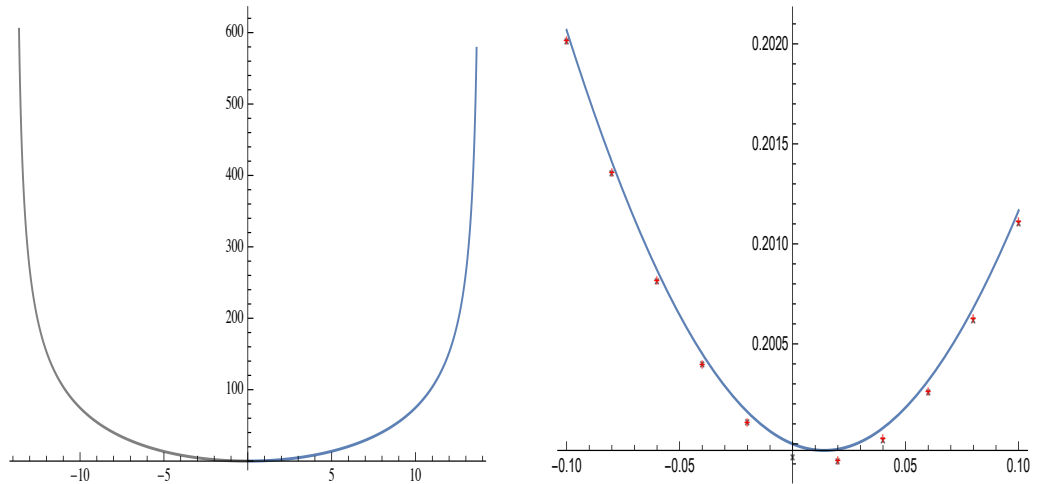


Fig. 2.3 On the left we have plotted $\Lambda(p)$ using an Adams scheme to numerically solve the VIE in (2.13) with 2000 time steps combined with Corollary 2.3.4, for $\alpha = 0.75$, $V_0 = .04$, $\nu = .15$, $\rho = -0.02$, and we find that $p_+ = T^*(1) \approx 34.5$ and $p_- = T^*(-1) \approx 33.25$. On the right we have plotted the corresponding asymptotic small-maturity smile $\hat{\sigma}(x)$ (in blue) versus the higher order approximation using Eq (2.38) (red “+” signs), and the smile points obtained from a simple Euler-type Monte Carlo scheme with maturity $T = .00005$, 10^5 simulations and 1000 time steps in Matlab (grey crosses), Matlab and Mathematica code available on request. 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 2.3.4 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 artefacts when computing the FL transform of $\bar{\Lambda}(p, 1)$ close to $x = 0$.

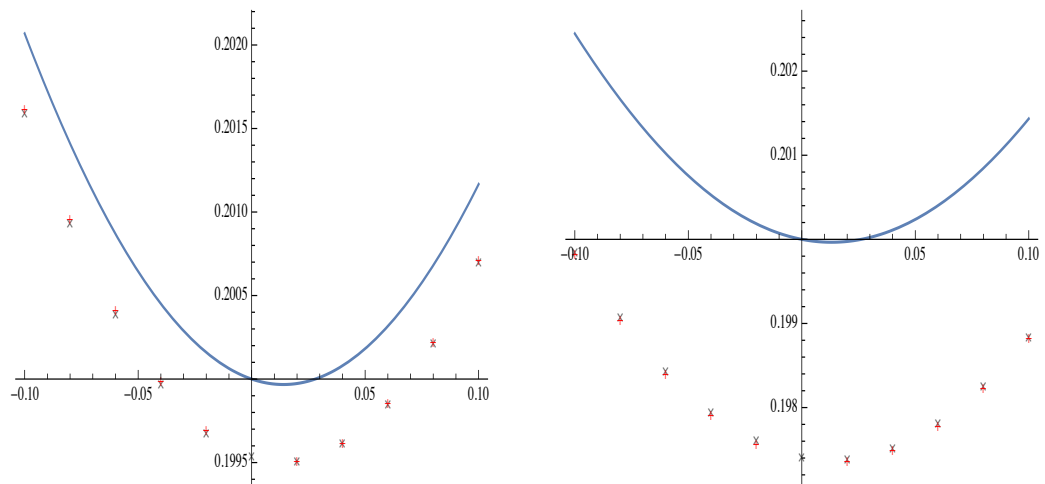


Fig. 2.4 On the left here we have the same plot as above but with $T = .005$ and for the right plot $T = .005$ and $\alpha = .6$ (i.e. $H = 0.1$), and again we see that the higher order approximation makes a significant improvement over the leading order smile. Of course we would not expect such close agreement for smaller values of α , or larger values of T , $|x|$ or $|\rho|$, e.g. $\rho = -0.65$ reported in e.g. [EGR18], but the point here is really just to verify the correctness of the asymptotic formula in (2.33), 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.

x	$\hat{\sigma}(x)$	Higher order $T = .00005$	Monte Carlo $T = .00005$	Higher order $T = .005$	Monte Carlo $T = .005$
-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%

Table of numerical results corresponding to the right plot in Figure 2.3 and the left plot in Figure 2.4.

2.4 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 2.4.1 $\lambda > 0$, $\rho \leq 0$.

Recall that $f(p, t)$ in (2.9) satisfies

$$D^\alpha f(p, t) = H(p, f(p, t)) \quad (2.41)$$

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

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

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

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

Proposition 2.4.2

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

Proof. [GGP19] 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 $H(p, \cdot)$ has two distinct real roots (or a single root) if and only if

$$(\lambda - \rho p v)^2 \geq (p^2 - p)v^2 \quad (2.42)$$

which is the same as the condition $e_1(p) \geq 0$ in condition C) in [GGP19]. 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 if $e_1(p) < 0$. We have two cases to consider to verify this claim:

- Suppose $e_1(p) \geq 0$. Then case B in [GGP19] is impossible by definition, and $p \in [\underline{p}, \bar{p}]$, and Eq (3.5) in [FJ11] is satisfied. Eq (3.4) in [FJ11] is

$$\lambda > \rho v p$$

in our current notation, and by the assertion on p.769 in [FJ11] 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 non-explosive cases C or D of the [GGP19] 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 [FJ11], 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 re-write (2.41) 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 [GGP19] classification). Then from the proof of Proposition 4 in [GGP19], 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 $\frac{H_\delta}{\Gamma(\alpha)} (t-1)^{\alpha-1} \int_1^t 1_{f(p, s) \leq U_1(p) - \delta} ds < U_1(p)$, or equivalently if we flip the inequality inside the indicator function

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

Then we see that

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

for t sufficiently large. Thus $U_1(p) - 2\delta \leq \frac{1}{t} \int_0^t f(p,s)ds \leq U_1(p)$, so $\frac{1}{t} \int_0^t f(p,s)ds \rightarrow U_1(p)$ as $t \rightarrow \infty$. 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 [GGP19] classification, see also Lemma 2.4.4). ■

Corollary 2.4.3 X_t/t satisfies the LDP as $t \rightarrow \infty$ 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 $U_1'(p) \rightarrow +\infty$ as $p \rightarrow \bar{p}$ and $U_1'(p) \rightarrow -\infty$ as $p \rightarrow \underline{p}$, 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 2.4.1 We can easily add stochastic interest rates into this model by modelling the short rate r_t by an independent Rough Heston process, and proceeding as in [FK16] (we omit the details), see also [F11].

Note that we have not proved that $f(p,t) \rightarrow U_1(p)$, but to establish the leading order behaviour in Proposition 2.4.2, 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 2.4.4 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 $y_0 \geq 0$;
- $G(0) \geq 0$;
- $K(z)$ is positive, continuous and strictly decreasing for $z > 0$;
- $\int_0^t K(z)dz$ is finite for each $t > 0$ and diverges as $t \rightarrow \infty$;
- $K(z + \alpha)/K(z)$ is strictly increasing in z for each fixed α 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 $t \rightarrow \infty$.

Proof. See Appendix C. ■

This lemma can be applied to both cases C and D. As shown in [GGP19], 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 $-y(t) \rightarrow y(t)$ and $-G(-y(t)) \rightarrow G(y(t))$ (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 $[0, a]$).

2.4.1 Asymptotics for call options and implied volatility

Corollary 2.4.5 *We have the following large-time asymptotic behaviour for European put/call options in the large-time, large log-moneyness regime:*

$$\begin{aligned} -\lim_{t \rightarrow \infty} \frac{1}{t} \log \mathbb{E}(S_t - S_0 e^{xt})^+ &= V^*(x) - x & (x \geq \frac{1}{2} \bar{\theta}), \\ -\lim_{t \rightarrow \infty} \frac{1}{t} \log(S_0 - \mathbb{E}(S_t - S_0 e^{xt})^+) &= V^*(x) - x & (-\frac{1}{2} \theta \leq x \leq \frac{1}{2} \bar{\theta}), \\ -\lim_{t \rightarrow \infty} \frac{1}{t} \log(\mathbb{E}(S_0 e^{xt} - S_t)^+) &= V^*(x) - x & (x \leq -\frac{1}{2} \theta), \end{aligned}$$

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

Proof. See Corollary 2.4 in [FJ11]. ■

Corollary 2.4.6 *We have the following asymptotic behaviour 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 $S_0 e^{xt}$:*

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

where

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

Proof. See Proposition 1 in [GJ11] (note that for the Rough Heston model λ has to be replaced with $\frac{\lambda}{\Gamma(\alpha)}$ and ν replaced with $\frac{\nu}{\Gamma(\alpha)}$, but the effect of the α here cancels out in the final formula for $\hat{\sigma}_\infty(k)$). ■

2.4.2 Higher order large-time behaviour

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 $p \in [p, \bar{p}]$, 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 (2.10) does not exist. Using the same approach as in [FJM11], one should be able to use this to compute a higher order large-time saddlepoint approximation for call options. For the sake of brevity, we defer the details of this for future work.

2.5 Asymptotics in the $H \rightarrow 0$ limit

In this section, we will show that for fixed t , the log stock price $X_t^{(\alpha)} := X_t$ converges as $\alpha \rightarrow \frac{1}{2}$ i.e. as $H \rightarrow 0$ in an appropriate sense. To match the assumptions of Theorem 13.1.1 on p.384 of [GLS90] (on the continuity of the solutions to a parametrized family of VIEs), we define $h(\alpha, w) := G(p, w)$ for $\alpha \geq \frac{1}{2}$ (which is independent of α). The kernel $a(t, s, \alpha) := (t-s)^{\alpha-1}/\Gamma(\alpha)$ is of continuous type; see Definition 9.5.2 in [GLS90], and the remark to Theorem 12.1.1 in [GLS90], which states local integrability of k as a sufficient condition for this property, and we can easily verify that

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

as $\alpha \rightarrow \frac{1}{2}$, so the uniform continuity assumption in Theorem 13.1.1 of [GLS90] is satisfied. Moreover the solution to the VIE is unique for $\alpha \in (0, 1)$, see Theorem 3.1.4 in [Brun17], or Satz 1 in [Di58]. Note that the Lipschitz condition (3.1) in [Di58] has a fixed Lipschitz constant $\Gamma(\alpha + 1)$, but since the function H defining our VIE (see (2.41)) does not depend on time, the factor t^α on the left hand side of condition (3.1) in [Di58] (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 [Brun17].

Then from Theorem 13.1.1 ii) in [GLS90], $f(p, t; \alpha)$ is continuous in α 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 [GLS90] is multi-dimensional, we can apply it to $(\text{Re}(f), \text{Im}(f))$ to conclude that $f(i\theta, t; \alpha) \rightarrow f(i\theta, t; \frac{1}{2})$ for $\theta \in \mathbb{R}$. Using the analyticity of $f(\cdot, t, 0)$, e.g. from Lemma 7 in [GGP19], 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 $\alpha \rightarrow \frac{1}{2}$, 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

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

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 non-zero 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 $H \rightarrow 0$, see Chapter 4 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

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

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 [FSV19], we find that $\phi_{\varepsilon}(p, t)$ satisfies

$$D^{\frac{1}{2}}\phi_{\varepsilon}(p, t) = \frac{1}{2}p^2 - \frac{1}{2}p\sqrt{\varepsilon} + p\rho v\phi_{\varepsilon}(p, t) + \frac{1}{2}v^2\phi_{\varepsilon}(p, t)^2 - \lambda\varepsilon^{\frac{1}{2}}\phi_{\varepsilon}(p, t) \quad (2.43)$$

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 [GLS90] as above to show that $\phi_{\varepsilon}(p, t)$ tends to the solution ϕ of

$$D^{\frac{1}{2}}\phi(p, t) = \frac{1}{2}p^2 + p\rho v\phi(p, t) + \frac{1}{2}v^2\phi(p, t)^2 \quad (2.44)$$

as $\varepsilon \rightarrow 0$ for $p \in (p_-^0, p_+^0)$ where $p_{\pm}^0 := \lim_{\varepsilon \rightarrow 0} \frac{p_{\pm}(\varepsilon t)}{\sqrt{\varepsilon}}$. 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 $t \rightarrow 0$ for which $\mathbb{E}(e^{pZ}) = e^{V_0 I^{\frac{1}{2}}\phi(p, \cdot)(1)}$. Two interesting and difficult open questions now arise: is this property *time-consistent*, i.e. does it remain true at a future time t when we condition on the history of V up to t , and ii) is V itself a well defined process in the $\alpha \rightarrow \frac{1}{2}$ limit, or does it e.g. tend to a non-Gaussian field which is not pointwise defined. We answer the second question in section 2.5.

Remark 2.5.1 Note that the scaling property in this case simplifies to

$$\Lambda(p, t) = \Lambda(pt^{\frac{1}{2}}, 1) \quad (2.45)$$

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

2.5.1 Implied vol asymptotics in the $H = 0, t \rightarrow 0$ limit - full smile effect for the Edgeworth FX options regime

Following a similar argument to Lemma 5 in [MT16] one can establish the following small-time behaviour for European put options in the Edgeworth regime:

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

as $t \rightarrow 0$, where Z is the non-Gaussian random variable defined in the previous subsection, and $f \sim g$ here means that $f/g \rightarrow 1$. From e.g. [Fuk17] or Lemma 3.3 in [FSV19], we know that for the Black-Scholes model with volatility σ

$$\frac{1}{\sqrt{t}} \mathbb{E}((e^{x\sqrt{t}} - e^{X_t})^+) \sim P_B(x, \sigma) := \mathbb{E}((x - \sigma W_1)^+) \quad (2.46)$$

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

$$\hat{\sigma}_0(x) := \lim_{t \rightarrow 0} \hat{\sigma}_t(x\sqrt{t}, t) = P_B(x, \cdot)^{-1}(P(x)) \quad (2.47)$$

for $x > 0$, where $\hat{\sigma}_t(x, t)$ denotes the implied volatility of a European put option with strike e^x , 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 e.g. [Fuk17], [EFGR19], [FSV19], where the leading order term is just Black-Scholes, followed by a next order skew term, followed by an even higher order term.

2.5.2 A closed-form expression for the skewness, the $H \rightarrow 0$ 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 $V_t = V_0 + \frac{1}{\Gamma(\alpha)} \int_0^t (t-s)^{\alpha-1} v \sqrt{V_s} dW_s$. 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

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

Thus

$$\mathbb{E}(X_T^3) = 3\rho \int_0^T \int_0^t \mathbb{E}(\sqrt{V_s}V_t dW_s dt) = \frac{3V_0\rho v T^{1+\alpha}}{\Gamma(\alpha)\alpha(1+\alpha)}. \quad (2.48)$$

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 κ then

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

(where κ is computed in Proposition 2.2.2). 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 2.5.2 If we allow ρ 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 *skewness term structure*, i.e. 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.

2.5.3 Weak convergence of the V process on pathspace to a tempered distribution, and the hyper-rough Heston model

From Theorem 4.3 in [ALP19] we know that

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

for $f \in L^1([0, T])$, where ψ_α satisfies the Riccati-Volterra equation:

$$\psi_\alpha(t) = \int_0^t c_\alpha(t-s)^{\alpha-1} (f(s) + \frac{1}{2}v^2\psi_\alpha(s)^2) ds \quad (2.49)$$

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

Proposition 2.5.1 *V tends to a random tempered distribution $V^{(\frac{1}{2})}$ in distribution as $\alpha \rightarrow \frac{1}{2}$ with respect to the strong and weak topologies (see page 2 in [BDW17] for definitions), where $V^{(\frac{1}{2})}$ is a random tempered distribution¹ and for all f in the Schwartz space \mathcal{S} we have*

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

where ψ satisfies the following VIE:

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

Proof. See Appendix D. ■

Let A_t satisfy $A_t = V_0 t + \frac{v}{\Gamma(\frac{1}{2})} \int_0^t (t-s)^{-\frac{1}{2}} W_{A_s} ds$. Then A_t is of the same form as X_t in [A19b], with their $dG_0(t) = V_0 dt$. Then from Theorem 2.5 in [A19b] (with $a = b = 0$ and $c = v^2$) we know that

$$\mathbb{E}(e^{\int_0^T f(T-t)dA_t}) = e^{\int_0^T F(T-s, \psi(T-s)) dG_0(s)} = e^{V_0 \int_0^T (f(T-s) + \frac{1}{2}v^2\psi(T-s)^2) ds} = e^{V_0 \int_0^T (f(s) + \frac{1}{2}v^2\psi(s)^2) ds} \quad (2.50)$$

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

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

The process A_t here is the driftless *hyper-rough* Heston model for $H = 0$ discussed in the next subsection, and e note that ψ satisfies the same VIE as (2.49) (and by e.g. Theorem 3.1.4 in [Brun17] 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 [JR18] (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 behaviour.

¹see e.g. [DRSV17] for more details on tempered distributions

2.5.4 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

$$\begin{aligned}
 A_t - V_0 t &= \frac{1}{\Gamma(\alpha)} \int_0^t \int_0^s (s-u)^{\alpha-1} v \sqrt{V_u} dW_u ds \\
 &= \frac{1}{\Gamma(\alpha)} \int_0^t v \sqrt{V_u} dW_u \int_u^t (s-u)^{\alpha-1} ds \\
 &= \frac{v}{\alpha \Gamma(\alpha)} \int_0^t (t-u)^\alpha \sqrt{V_u} dW_u \\
 &\text{(using Dambis-Dubins-Schwarz time change)} \\
 &= \frac{v}{\alpha \Gamma(\alpha)} \int_0^t (t-u)^\alpha dB_{A_u} \\
 &\text{(where } B_t := X_{T_t}, T_t = \inf\{s : A_s > t\} \text{) so } B \text{ is a Brownian motion)} \\
 &= \frac{v}{\alpha \Gamma(\alpha)} B_{A_u} (t-u)^\alpha \Big|_{u=0}^t + \frac{v}{\Gamma(\alpha)} \int_0^t (t-u)^{\alpha-1} B_{A_u} du \\
 &= v I^\alpha B_{A_t}.
 \end{aligned} \tag{2.51}$$

We can now take

$$A_t = V_0 t + v I^\alpha B_{A_t} \tag{2.52}$$

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 [JR18] for the case of zero drift. Note that for a given sample path $B_t(\omega)$, we can regard (2.52) as a (random) fractional ODE of the form:

$$A(t) = V_0 t + I^\alpha f(A(t)) \tag{2.53}$$

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

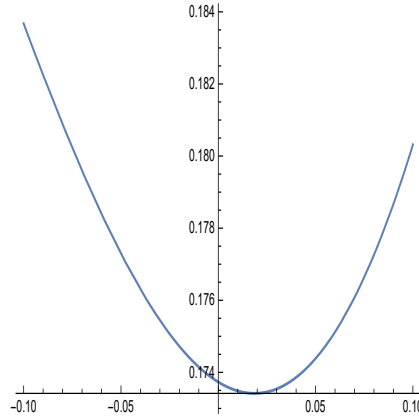


Fig. 2.5 Here we have plotted the $H = 0$ asymptotic short-maturity smile (i.e. $\hat{\sigma}_0(x)$ in (2.47)), for $\nu = .2$, $\rho = -.1$ and $V_0 = .04$. We have used a 10-term small- t series approximation to the solution to (2.44) combined with the scaling property in (2.45), and the Alan Lewis Fourier inversion formula for call options given in e.g. Eq 1.4 in [EGR18] using Gauss-Legendre quadrature for the inverse Fourier transform with 1600 points over a range of $[0, 40]$.

The case $\lambda > 0$

For the case when $\lambda > 0$, using (2.8) we see that

$$\begin{aligned}
 A_t - \int_0^t \xi_0(s) ds &= \int_0^t \int_0^s \kappa(s-u) \sqrt{V_u} dW_u ds & (2.54) \\
 &= \int_0^t \sqrt{V_u} \int_u^t \kappa(s-u) ds dW_u \\
 &= \int_0^t F(t-u) \sqrt{V_u} dW_u \quad (\text{where } F(t-u) = \int_u^t \kappa(s-u) ds) \\
 &= \int_0^t F(t-u) dM_u \\
 &\quad (\text{where } dM_t = \sqrt{V_t} dW_t) \\
 &= \int_0^t F(t-u) dB_{A_u} \\
 &\quad (\text{where } B_t := M_{T_t}, T_t = \inf\{s : A_s > t\}) \text{ so } B \text{ is a Brownian motion}) \\
 &= B_{A_u} F(t-u) \Big|_{u=0}^t + \int_0^t \kappa(t-u) B_{A_u} du \\
 &= \int_0^t \kappa(t-u) B_{A_u} du
 \end{aligned}$$

where we have used (2.7) to verify that $F(t-u) \rightarrow 0$ as $u \rightarrow t$.

2.6 Appendix

2.6.1 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

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

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 κ to be the unique function which satisfies

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

Such a κ exists and is known as the *resolvent* of φ . Then we see that

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

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. Cancelling the $-\frac{\lambda}{\nu} \kappa * V$ terms, we see that

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

and thus

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

i.e. the correct κ function is the solution to (A.1). If we take the Laplace transform of (A.1), we get

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

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

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

2.6.2 Appendix B: The re-scaled model

We first let

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

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

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

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

2.6.3 Appendix C: Monotonicity property

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 2.4.4.

In the classical case $K(t) \equiv 1$ the integral eq 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 [MF71] (Theorem 6), see also the end of page 14 in [GGP19].

What follows is a natural extension of the technique used in [MW51] (Theorem 8). Using the properties of convolution and differentiating under the integral sign, we have:

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

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

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

$G(0) > 0$ so $y'(t) \rightarrow +\infty$ as $t \rightarrow 0^+$ and since $G(y)$ is increasing for $y \leq y_0$ we have that $y'(t) > 0$ until $y(t)$ reaches y_0 i.e. the solution increases. For $y \geq y_0$, $G(y)$ 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 $y'(t)$:

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

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

We can re-write the kernels in the first and second terms of the expression for $y'(t_1)$ as:

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

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:

$$\begin{aligned} \int_0^{t_0} \frac{K(t_1 - s)}{K(t_0 - s)} K(t_0 - s) G'(y(s)) y'(s) ds &= \frac{K(t_1 - \tau)}{K(t_0 - \tau)} \int_0^{t_0} K(t_0 - s) G'(y(s)) y'(s) ds \\ &= -\frac{K(t_1 - \tau)}{K(t_0 - \tau)} K(t_0) G(0) \end{aligned} \quad (\text{C.5})$$

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

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

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 non-linearity in the integral equation has the generic form $G(y) = (y - \theta_1)^2 + \theta_2$ i.e. a quadratic with positive leading coefficient (for simplicity set to 1 here) and minimum of θ_2 obtained at $y = \theta_1$. Depending on the values of $\{\theta_1, \theta_2\}$ the following cases due to [GGP19] 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 $y(t) \rightarrow -y(t)$ and $-G(-y(t)) \rightarrow G(y(t))$ (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 its maximum after which it decreases which is of the type considered here.

2.6.4 Appendix D: Limit of Volterra equations

From Theorem 13.1.1 ii) in [GLS90], the unique solution $\psi^{(\alpha)}$ to

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

tends pointwise to the solution of

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

which is also unique by e.g. Theorem 3.1.4 in [Brun17]. Now consider any sequence $f_\varepsilon \in \mathcal{S}$ with $\|f_\varepsilon\|_{m,j} \rightarrow 0$ as $\varepsilon \rightarrow 0$ for all $m, j \in \mathbb{N}_0^n$ for any $n \in \mathbb{N}$ (i.e. under the Schwartz space semi-norm defined in Eq 1 in [BDW17]). Then the convergence here implies in particular that f_ε tends to f pointwise. Then from Theorem 13.1.1. in [GLS90], the unique solution ψ_ε to

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

tends pointwise to the solution to

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

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 [BDW17]), we obtain the stated result.

Chapter 3

Small-time VIX smile for the Rough Heston model

3.1 Introduction

The Rough Heston stochastic volatility model was introduced in Jaisson&Rosenbaum[JR16], and (using C -tightness arguments from Jacod&Shiryayev[JS13]) they show that the model arises naturally as a weak large-time limit of a high-frequency market microstructure model driven by two nearly unstable Hawkes process. [ER19] show that the characteristic function of the log stock price for the Rough Heston model admits a quasi-closed form solution via the solution to a non-linear Volterra integral equation (VIE) (see also [EFR18] and [ER18]), 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 usual fractional kernel $t^{H-\frac{1}{2}}$ for the V process multiplied by a *Mittag-Leffler* function. The instantaneous variance process V for the model is $(H - \varepsilon)$ -Hölder continuous like fractional Brownian motion (see e.g. Theorem 3.2 in [JR16]) and the model exhibits power law skew in the small-time limit (see the previous chapter and Corollary 3.4 in [FSV21]). [DJR19] introduce an extension of this model known as the *super Rough Heston model* which incorporates the empirically observed *strong Zumbach effect* as a weak limit of a market microstructure model driven by a quadratic Hawkes process (also using C -tightness arguments) but this model is no longer affine and thus not directly amenable to VIE techniques or Edgeworth and large deviation asymptotics, so it is difficult to prove anything about the qualitative behaviour or dynamics of the smile (and the Zumbach term is a drift term and hence very unlikely to affect leading order large deviation asymptotics). A variant of this model is used in [GJR20], which attains a better fit to SPX and VIX options in practice than conventional rough volatility models, but Guyon[Guy20b] remarks if we calibrate this model to the VIX smile, the short-maturity

at-the-money SPX skew is still too small compared to what is observed in practice (see below for discussion on the addition of jumps in [FS21]).

The theoretical value of the VIX index at time t is $\text{VIX}_t = \sqrt{-\frac{2}{\Delta} \mathbb{E}(\log \frac{S_{t+\Delta}}{S_t} | \mathcal{F}_t)}$, where S_t is the S&P 500 index value at time t , $\Delta = 30$ days and \mathcal{F}_t is the market filtration, so VIX_t^2 is effectively a rolling 30-day Variance swap rate. A VIX option is a European call or put option on VIX_T for some maturity T , and if we replace the spot value S_0 in the Black-Scholes formula with the VIX future price $\mathbb{E}(\text{VIX}_T)$, we can define the implied volatility of a VIX call or put in the usual way by inverting the Black-Scholes formula. VIX options are very liquid in practice (although their bid/offer spreads are still comparatively high), and empirical VIX smiles typically exhibit positive skew with negative convexity (see plots in [GJR20],[Guy20],[DeM18],[HJT20] et al.), although e.g. Markovian diffusion models like the standard Heston model can give rise to negative VIX implied vol skews..

In this chapter, we work with a generalized version of the Rough Heston model as used in [GR19] with initial variance curve $\xi_0(t)$ and the corresponding dynamics of the forward variance $\xi_t(u)$. We first derive an explicit formula for simulating VIX_T in Eq (3.4) (note that no such formula exists for e.g. the quadratic rough Heston model in [GJR20] though there are approximations see [Rom22]) and we then perform a formal small T -expansion of $\xi_T(u)$ which suggests that $(\text{VIX}_T^2 - \text{VIX}_0^2)/T^{\frac{1}{2}-H} \sim c\tilde{X}_T/T^{\frac{1}{2}-H}$ in some sense as $T \rightarrow 0$ for some constant $c > 0$, where $\tilde{X}_t = \int_0^t \sqrt{V_s} dW_s$ is the martingale component of the log stock price for the driftless rough Heston model when the correlation $\rho = 1$. This leads us to guess that $(\text{VIX}_T^2 - \text{VIX}_0^2)/T^{\frac{1}{2}-H}$ satisfies the same small-time LDP as $c\tilde{X}_T/T^{\frac{1}{2}-H}$, for which we can readily compute a small-time LDP with minor amendments to the main arguments in Theorem 2.3.3 to allow for non-flat $\xi_0(t)$. We then make this rigorous by showing that $(\text{VIX}_T^2 - \text{VIX}_0^2)/T^{\frac{1}{2}-H}$ and $c\tilde{X}_T/T^{\frac{1}{2}-H}$ are *exponentially equivalent* as $T \rightarrow 0$ and hence satisfy the same LDP, and this is proved using a minor variant/extension of Theorem 7.1 in Abi Jaber et al.[ALP19] for the exponential-affine formula for $\mathbb{E}(e^{u\tilde{X}_T + (f*\tilde{X})_T} | \mathcal{F}_t^W)$ for a general function f and $u \in \mathbb{R}$ such that T is less than the explosion time $T^*(u)$. Specifically we show that $\varepsilon^{2H} \log \mathbb{E}(e^{\varepsilon^{-\alpha} p(\text{VIX}_\varepsilon^2 - \text{VIX}_0^2 - c\tilde{X}_\varepsilon)}) = V_0 I^{1-\alpha} \phi_\varepsilon(p, 1)$ where $\alpha = H + \frac{1}{2}$, and $\phi_\varepsilon(p, t)$ satisfies a family of VIEs whose solution tends uniformly to zero on $[0, 1]$ as $\varepsilon \rightarrow 0$ for all $p \in \mathbb{R}$ and I^r denotes the r -th order fractional integral operator. We later translate this LDP into VIX call option and implied volatility asymptotics, and we compute a small log-moneyness expansion for the asymptotic VIX smile using expansions previously derived in chapter 2 which yields tractable expressions for the overall level, skew and convexity of the short-end VIX smile. We also mention Proposition 18 in [AGM18] which shows that the derivative of the VIX implied volatility with respect to log-moneyness at-the-money tends to a finite constant as $T \rightarrow 0$ (as opposed to exploding power-law behaviour $\propto T^{H-\frac{1}{2}}$) for a standard (and mixed) rough Bergomi-type model, and we have

verified this behaviour numerically. However numerical computations suggest that for the rough Heston model, the [AGM18] measure of at-the-money skew does indeed appear to be $O(T^{H-\frac{1}{2}})$ as $T \rightarrow 0$, as one would guess from (3.24) below.

Unfortunately, since the limiting VIX smile only depends on the factor $v/\sqrt{V_0}$ and not on ρ , we cannot simultaneously fit the overall level and skew of observed limiting VIX smile using the standard rough Heston model. To circumvent this issue, the companion article [FS21] enriches the model with an additional independent CGMY (a.k.a. KoBoL)-type Lévy process L as in [FSV21] with $Y \in (1, 2)$, and using a simple modification of the main result in [FSV21] for the Edgeworth regime where log-moneyness scales like $x\sqrt{T}$, we show that one can simultaneously use the rough Heston parameters to fit the at-the-money VIX level and skew as $T \rightarrow 0$, and the CGMY parameters to fit the observed level, at-the-money correction and at-the-money skew of SPX options as $T \rightarrow 0$ (using the main Theorem in [FSV21] adapted for our rough Heston V process), and the drift of the V process can be made to be fully consistent with the initial observed variance curve structure.

3.2 The Model

We consider a generalized Rough Heston model for a log stock price process $X_t = \log S_t$ of the same form in Gatheral&Radoičić[GR19]:

$$\begin{aligned} dX_t &= -\frac{1}{2}V_t dt + \sqrt{V_t}(\rho dW_t + \bar{\rho} dB_t) \\ V_t &= \xi_0(t) + c_\alpha \int_0^t (t-s)^{\alpha-1} v \sqrt{V_s} dW_s \end{aligned} \quad (3.1)$$

for $H \in (0, \frac{1}{2})$, $\alpha = H + \frac{1}{2}$, $c_\alpha = \frac{1}{\Gamma(\alpha)}$ and $v > 0$, with some initial variance curve $\xi_0(t)$ with $\xi_0(\cdot)$ continuous, where W, B are two independent Brownian motions, $\bar{\rho} = \sqrt{1 - \rho^2}$ with $|\rho| \leq 1$, and we assume $X_0 = 0$ and zero interest rate without loss of generality. Note we do not have a mean reversion term λ in (3.1) since such a term will not materially affect the asymptotics at the leading order large deviations level that we consider here once we re-calibrate to the observed initial variance curve $\xi_0(t)$, but would add further headache to our already lengthy analysis in e.g. Appendix B.

It is not known whether we have pathwise uniqueness for (3.1) even when $\xi_0(t)$ is constant because \sqrt{v} is not Lipschitz at zero (see section 4.2.3 in [JP20] for more on this), but we do have weak uniqueness (see Theorem 3.4 in [ALP19]) and uniqueness in law for V on $C([0, T])$, since we can explicitly compute an exponential-affine formula for the Fourier transform of V on pathspace in terms of a Volterra integral equation with a unique

solution, see Appendix B (which is based on Theorem 7.1 in [ALP19]) (see also Theorem 6.1 in [ALP19]).

When $\xi_0(t)$ is a non-constant function of t , conditions were derived in [AE19b] (see Theorem 2.1) guaranteeing the existence of a non-negative weak solution for V . An example of such an admissible initial variance curve would be a smooth non-decreasing function such that the spot variance $V_0 = \xi_0(0) \geq 0$. We shall assume henceforth that our chosen ξ_0 satisfies the [AE19b] conditions.

To clarify these points further, if we assume we have two solutions U and V to (3.1), then

$$\begin{aligned} \mathbb{E}((V_t - U_t)^2) &= \frac{1}{\Gamma(\alpha)^2} \mathbb{E} \left(\int_0^t (t-s)^{2H-1} \mathbf{v} (\sqrt{V_s} - \sqrt{U_s})^2 ds \right) \quad (3.2) \\ &\leq \frac{1}{\Gamma(\alpha)^2} \mathbb{E} \left(\int_0^t (t-s)^{2H-1} \mathbf{v} |V_s - U_s| ds \right) \\ &\leq \frac{1}{\Gamma(\alpha)^2} \int_0^t (t-s)^{2H-1} \mathbf{v} \mathbb{E}((V_s - U_s)^2)^{\frac{1}{2}} ds \end{aligned}$$

so $f(t) := \mathbb{E}((V_t - U_t)^2)$ satisfies

$$f(t) \leq \frac{1}{\Gamma(\alpha)^2} \int_0^t (t-s)^{2H-1} \mathbf{v} \sqrt{f(s)} ds \quad (3.3)$$

but unfortunately there is a non-zero solution to $f(t) = \int_0^t (t-s)^{2H-1} \mathbf{v} \sqrt{f(s)} ds$ in addition to the trivial zero solution (see Example 3.1.18 in [Brun17] for general $H \in (0, 1)$ and for $H = \frac{1}{2}$, $f(t) = \frac{1}{4} \mathbf{v}^2 t^2$), so we cannot directly use a comparison principle in e.g. Appendix A.2 in [ACLP19] to assert that $f(t) \leq 0$. If however we replace the $\sqrt{\cdot}$ coefficient in (3.1) with a Lipschitz function $\sigma(\cdot)$ which agrees with $\sqrt{\cdot}$ for $\cdot \geq \delta > 0$, this comparison theorem approach does show that we have pathwise uniqueness for V up to the hitting time of V to δ for any $\delta > 0$ (see also [JP20]). One can also adapt Lemma 4.10 in [JP20] to show that $V_t > 0$ Lebesgue a.e. even if V hits zero, and it is currently an open problem for what parameter combinations this is possible.

We let $\mathcal{F}_t = \mathcal{F}_t^{W,B}$. Then we know that $\xi_t(u) := \mathbb{E}(V_u | \mathcal{F}_t)$ is given by

$$\xi_t(u) = \xi_0(u) + \frac{\mathbf{v}}{\Gamma(\alpha)} \int_0^t (u-s)^{H-\frac{1}{2}} \sqrt{V_s} dW_s$$

so

$$d\xi_t(u) = \frac{\mathbf{v}}{\Gamma(\alpha)}(u-t)^{H-\frac{1}{2}}\sqrt{V_t}dW_t$$

$$\text{and } \text{VIX}_T^2 := -\frac{2}{\Delta}\mathbb{E}(\log \frac{S_{T+\Delta}}{S_T} | \mathcal{F}_T) = \frac{1}{\Delta}\mathbb{E}(\int_T^{T+\Delta} V_u du | \mathcal{F}_T) = \frac{1}{\Delta} \int_T^{T+\Delta} \xi_T(u) du.$$

3.2.1 The small-time LDP for $(\text{VIX}_T^2 - \text{VIX}_0^2)/T^{\frac{1}{2}-H}$

Using the stochastic Fubini theorem and Taylor's remainder theorem, we see that

$$\begin{aligned} \text{VIX}_T^2 &= \frac{1}{\Delta} \int_T^{T+\Delta} \xi_T(u) du \\ &= \frac{1}{\Delta} \int_T^{T+\Delta} (\xi_0(u) + \int_0^T \frac{\mathbf{v}}{\Gamma(\alpha)}(u-s)^{H-\frac{1}{2}}\sqrt{V_s}dW_s) du \\ &= \frac{1}{\Delta} \int_T^{T+\Delta} \xi_0(u) du + c_1 \int_0^T ((T+\Delta-s)^{\frac{1}{2}+H} - (T-s)^{\frac{1}{2}+H})\sqrt{V_s}dW_s \\ &= \frac{1}{\Delta} \int_T^{T+\Delta} \xi_0(u) du + c_1 \int_0^T ((T+\Delta-s)^{\frac{1}{2}+H} - (T-s)^{\frac{1}{2}+H})d\tilde{X}_s \end{aligned} \quad (3.4)$$

where $c_1 = \frac{\mathbf{v}}{\Delta\Gamma(\alpha)(\frac{1}{2}+H)}$ and

$$\tilde{X}_t = \int_0^t \sqrt{V_s}dW_s$$

is the martingale component of the log stock price process X when $\rho = 1$.

Then using that

$$\frac{1}{\Delta} \int_T^{T+\Delta} \xi_0(u) du = \frac{1}{\Delta} \int_0^\Delta \xi_0(u) du + \frac{1}{\Delta} T(\xi_0(\Delta) - \xi_0(0)) + o(T) = \text{VIX}_0^2 + O(T)$$

we (formally) expect that

$$\frac{\text{VIX}_T^2 - \text{VIX}_0^2}{T^{\frac{1}{2}-H}} \sim \frac{c_1\Delta^{\frac{1}{2}+H}}{T^{\frac{1}{2}-H}} \int_0^T \sqrt{V_s}dW_s = \frac{c\tilde{X}_T}{T^{\frac{1}{2}-H}} \quad (3.5)$$

as $T \rightarrow 0$, where

$$c := c_1\Delta^{\frac{1}{2}+H} = \frac{\mathbf{v}}{\Gamma(\alpha)\alpha}\Delta^{H-\frac{1}{2}}.$$

From Theorem 2.3.3 we know that $\tilde{X}_T/T^{\frac{1}{2}-H}$ satisfies an LDP with some rate function $I^{\rho=1}(x)$ and speed T^{-2H} as $T \rightarrow 0$, so based on above we conjecture the following result, for which the full proof is given below.

Theorem 3.2.1 $(\text{VIX}_T^2 - \text{VIX}_0^2)/T^{\frac{1}{2}-H}$ satisfies the LDP as $T \rightarrow 0$ with speed T^{-2H} and rate function $J(x) := I^{\rho=1}(\frac{x}{c})$ where $I^{\rho=1}(x)$ is the same as $I(x)$ in Theorem 2.3.3 for the special case when $\rho = 1$, and $J(x)$ is the Fenchel-Legendre transform of

$$\bar{\Lambda}^{\rho=1}(cp) := \lim_{T \rightarrow 0} T^{2H} \log \mathbb{E}(e^{\frac{p}{T^\alpha}(\text{VIX}_T^2 - \text{VIX}_0^2)})$$

for $p \in (-\infty, \frac{p_+}{c})$ and $\bar{\Lambda}^{\rho=1}(cp) = +\infty$ otherwise, where $\bar{\Lambda}^{\rho=1}$ and p_+ are the same as $\bar{\Lambda}$ and p_+ in Theorem 2.3.3 for the special case where the correlation ρ is $+1$, and c is defined above.

The following Proposition extends and streamlines the proof of main Theorem 2.3.3 to the case of the generalized rough Heston model in (3.1).

Proposition 3.2.2 $(X_T + \frac{1}{2}\langle X \rangle_T)/T^{\frac{1}{2}-H}$ and $X_T/T^{\frac{1}{2}-H}$ satisfies the same LDP as $T \rightarrow 0$ as in Theorem 2.3.3.

Proof. Recall that $X_t + \frac{1}{2}\langle X \rangle_t$ is just the martingale component of the log stock price X_t . Then from Theorem B.1 in Appendix B, we know that

$$\mathbb{E}(e^{p(X_t + \frac{1}{2}\langle X \rangle_t)}) = e^{\int_0^t \xi_0(t-s)(\frac{1}{2}p^2 + p\rho v\psi(p,s) + \frac{1}{2}v^2\psi(p,s)^2)ds} = e^{\int_0^t \xi_0(t-s)D^\alpha\psi(p,s)ds}$$

for $t \in [0, T_\psi^*(p))$, where $\psi(p, \cdot)$ satisfies the fractional Riccati VIE:

$$\psi(p, t) = \int_0^t c_\alpha(t-s)^{\alpha-1} \left(\frac{1}{2}p^2 + p\rho v\psi(p, s) + \frac{1}{2}v^2\psi(p, s)^2 \right) ds \quad (3.6)$$

and $T_\psi^*(p) > 0$ is the explosion time for ψ , and (by e.g. Appendix A) this solution is unique. Then

$$\mathbb{E}(e^{\frac{p}{\varepsilon^\alpha}(\frac{1}{2}\langle X \rangle_{\varepsilon t} + X_{\varepsilon t})}) = e^{\int_0^{\varepsilon t} \xi_0(\varepsilon t-s)(\frac{1}{2}\frac{p^2}{\varepsilon^{2\alpha}} + \frac{p}{\varepsilon^\alpha}\rho v\psi(\frac{p}{\varepsilon^\alpha}, s) + \frac{1}{2}v^2\psi(\frac{p}{\varepsilon^\alpha}, s)^2)ds}$$

and

$$\begin{aligned} \psi(\frac{p}{\varepsilon^\alpha}, \varepsilon t) &= \int_0^{\varepsilon t} c_\alpha(\varepsilon t-s)^{\alpha-1} \left(\frac{1}{2}\frac{p^2}{\varepsilon^{2\alpha}} + \frac{p}{\varepsilon^\alpha}\rho v\psi(\frac{p}{\varepsilon^\alpha}, s) + \frac{1}{2}v^2\psi(\frac{p}{\varepsilon^\alpha}, s)^2 \right) ds \\ &= \varepsilon \int_0^t c_\alpha(\varepsilon t - \varepsilon s)^{\alpha-1} \left(\frac{1}{2}\frac{p^2}{\varepsilon^{2\alpha}} + \frac{p}{\varepsilon^\alpha}\rho v\psi(\frac{p}{\varepsilon^\alpha}, \varepsilon s) + \frac{1}{2}v^2\psi(\frac{p}{\varepsilon^\alpha}, \varepsilon s)^2 \right) ds \end{aligned}$$

for $t \in [0, \frac{1}{\varepsilon}T_\psi^*(\frac{p}{\varepsilon^\alpha}))$. Then multiplying both sides by ε^α , we see that $\psi^\varepsilon(p, t) := \varepsilon^\alpha\psi(\frac{p}{\varepsilon^\alpha}, \varepsilon t)$ satisfies

$$\psi^\varepsilon(p, t) = \int_0^t c_\alpha(t-s)^{\alpha-1} \left(\frac{1}{2}p^2 + p\rho v\psi^\varepsilon(p, s) + \frac{1}{2}v^2\psi^\varepsilon(p, s)^2 \right) ds$$

for $t \in [0, \frac{1}{\varepsilon} T_\psi^*(\frac{p}{\varepsilon\alpha})]$, so we see that $\psi^\varepsilon(p, \cdot)$ and $\psi(p, \cdot)$ satisfy the same VIE, and hence are equal. Then

$$\begin{aligned}
 \varepsilon^{2H} \log \mathbb{E}(e^{\frac{p}{\varepsilon\alpha}(\frac{1}{2}\langle X \rangle_{\varepsilon t} + X_{\varepsilon t})}) &= \varepsilon^{2H} \log e^{\int_0^{\varepsilon t} \xi_0(\varepsilon t - s) (\frac{1}{2} \frac{p^2}{\varepsilon^2 \alpha} + \frac{p}{\varepsilon\alpha} \rho v \psi(\frac{p}{\varepsilon\alpha}, s) + \frac{1}{2} v^2 \psi(\frac{p}{\varepsilon\alpha}, s)^2) ds} \\
 &= \varepsilon^{2H} \log e^{\varepsilon \int_0^t \xi_0(\varepsilon t - \varepsilon s) (\frac{1}{2} \frac{p^2}{\varepsilon^2 \alpha} + \frac{p}{\varepsilon\alpha} \rho v \psi(\frac{p}{\varepsilon\alpha}, \varepsilon s) + \frac{1}{2} v^2 \psi(\frac{p}{\varepsilon\alpha}, \varepsilon s)^2) ds} \\
 &= \int_0^t \xi_0(\varepsilon t - \varepsilon s) (\frac{1}{2} p^2 + p \rho v \psi(p, s) + \frac{1}{2} v^2 \psi(p, s)^2) ds \\
 &\rightarrow V_0 \int_0^t (\frac{1}{2} p^2 + p \rho v \psi(p, s) + \frac{1}{2} v^2 \psi(p, s)^2) ds \quad (3.7) \\
 &= \bar{\Lambda}(p, t) \quad (3.8)
 \end{aligned}$$

as $\varepsilon \rightarrow 0$ if $t < T_\psi^*(p)$ using the bounded convergence theorem, since $\xi_0(\cdot)$ is continuous at zero and ψ is bounded on $[0, t]$ for if $t < T_\psi^*(p)$. From Lemma 2.3.9 in [DZ98], we know that $\bar{\Lambda}(p, t)$ is convex in p , and from (3.6) we also know that

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

so $\Lambda(p, t)$ is differentiable in t .

Using the scaling relation in Corollary 2.3.4 we also know that $\Lambda(p, 1) = p^{\frac{2H}{\alpha}} \Lambda(\text{sgn}(p), |p|^{\frac{1}{\alpha}})$, so setting $\Lambda(p) := \Lambda(p, 1)$ we know that $\Lambda(p)$ is differentiable in p . Moreover, the quadratic $Q(w) := \frac{1}{2} p^2 + \rho p v w^2 + \frac{1}{2} w^2$ has no real roots so we are in Case A or B in [GGP19] where the VIE for $\psi(p, \cdot)$ has no fixed point, so T_ψ^* is finite and explodes at rate $\text{const.}/(T_\psi^*(p) - t)^\alpha$ (see Lemma 3 in [GGP19]).

From the integral in (3.8) and the aforementioned known explosion rate and the scaling relation, we see that $\bar{\Lambda}(p, t)$ also tends to $+\infty$ as $p \nearrow p_+ = T_\psi^*(+1)^\alpha$ or as $p \searrow p_- = -T_\psi^*(-1)^\alpha$, and (by convexity and differentiability) Λ is also essentially smooth. Moreover, from the monotonicity of the L^p -norm, we know that $\bar{\Lambda}(p, t) = \infty$ for $p \notin (p_-, p_+)$ as well. Hence by the Gärtner-Ellis theorem from large deviations theory (see Theorem 2.3.6 in [DZ98]), $(X_\varepsilon + \frac{1}{2}\langle X \rangle_\varepsilon)/\varepsilon^{\frac{1}{2}-H}$ satisfies the LDP as $\varepsilon \rightarrow 0$ with speed ε^{-2H} and rate function $I(x)$. Finally the LDP for $X_\varepsilon/\varepsilon^{\frac{1}{2}-H}$ is obtained using exponential equivalence as in the proof of Theorem 2.3.3. ■

Proof. (of Theorem 3.2.1). Setting $T = \varepsilon$ to make the notation consistent with Chapter 2 and integrating (3.4) by parts, we see that

$$\text{VIX}_\varepsilon^2 - \frac{1}{\Delta} \int_\varepsilon^{\varepsilon+\Delta} \xi_0(u) du = \text{VIX}_\varepsilon^2 - \text{VIX}_0^2 - \zeta \quad (3.9)$$

$$= c_1 \int_0^\varepsilon ((\varepsilon + \Delta - s)^{\frac{1}{2}+H} - (\varepsilon - s)^{\frac{1}{2}+H}) d\tilde{X}_s \quad (3.10)$$

$$= c_1 \Delta^{\frac{1}{2}+H} \tilde{X}_\varepsilon - c_1 \alpha \int_0^\varepsilon ((\varepsilon - s)^{H-\frac{1}{2}} - (\varepsilon + \Delta - s)^{H-\frac{1}{2}}) \tilde{X}_s ds$$

$$= c_1 \Delta^{\frac{1}{2}+H} \tilde{X}_\varepsilon + \int_0^\varepsilon f(\varepsilon - s) \tilde{X}_s ds$$

$$= c \tilde{X}_\varepsilon + (f * \tilde{X})_\varepsilon \quad (3.11)$$

where $\zeta := \frac{1}{\Delta} \int_\varepsilon^{\varepsilon+\Delta} \xi_0(u) du - \text{VIX}_0^2 = O(\varepsilon)$ and $f(s) := c_1 \alpha ((s + \Delta)^{H-\frac{1}{2}} - s^{H-\frac{1}{2}})$ and we note that $f \in L^2$. As discussed above, from Theorem 2.3.3 we know that the leading order term $c \tilde{X}_T / T^{\frac{1}{2}-H}$ satisfies the stated LDP as $T \rightarrow 0$, so the issue is just to argue away the remainder term, using exponential equivalence as in Chapter 2.

From Theorem B.1 (which is adapted from Eqs 2.8-2.10 in [AE19b] and Lemma 7.3 in [ALP19]), we know that

$$\mathbb{E}(e^{p \int_0^\varepsilon f(\varepsilon-s) \tilde{X}_s ds}) = e^{\int_0^\varepsilon \xi_0(\varepsilon-s) D^\alpha \psi_2(p,s) ds} = e^{\int_0^\varepsilon \xi_0(\varepsilon-s) g(p,s) ds}$$

where

$$\psi_1(p,t) = p \int_0^t f(s) ds \quad (3.12)$$

$$\psi_2(p,t) = \int_0^t c \alpha (t-s)^{\alpha-1} \left(\frac{1}{2} \psi_1(p,s)^2 + \psi_1(p,s) \nu \psi_2(p,s) + \frac{1}{2} \nu^2 \psi_2(p,s)^2 \right) ds$$

for $\varepsilon \leq T_{\psi_2}(p)$ where $T_{\psi_2}(p)$ is the explosion time for ψ_2 (note that $g(p,t) := D^\alpha \psi_2(p,t) = \frac{1}{2} \psi_1(p,t)^2 + \psi_1(p,t) \nu \psi_2(p,t) + \frac{1}{2} \nu^2 \psi_2(p,t)^2$), and recall that $f(s) := c_1 \alpha ((s + \Delta)^{H-\frac{1}{2}} - s^{H-\frac{1}{2}})$. We first note that

$$\psi_1(\varepsilon^{-\alpha} p, \varepsilon t) = \frac{p}{\varepsilon^{H+\frac{1}{2}}} c_1 \alpha \int_0^{\varepsilon t} ((s + \Delta)^{H-\frac{1}{2}} - s^{H-\frac{1}{2}}) ds = p c \varepsilon^{-\alpha} h_\varepsilon(t) \quad (3.13)$$

where h_ε denotes the bounded, continuous function

$$h_\varepsilon(t) := \Delta^{-\frac{1}{2}-H} \alpha \int_0^{\varepsilon t} ((s + \Delta)^{H-\frac{1}{2}} - s^{H-\frac{1}{2}}) ds \leq 0$$

defined for $t \in [0, 1]$, which tends to zero pointwise as $\varepsilon \rightarrow 0$ (this will be needed below).

Then

$$\begin{aligned}
 \psi_2(p\varepsilon^{-\alpha}, \varepsilon t) &= \int_0^{\varepsilon t} c_\alpha(\varepsilon t - s)^{\alpha-1} \left(\frac{1}{2} \psi_1(p\varepsilon^{-\alpha}, s)^2 + \psi_1(p\varepsilon^{-\alpha}, s) \nu \psi_2(p\varepsilon^{-\alpha}, s) \right. \\
 &\quad \left. + \frac{1}{2} \nu^2 \psi_2(p\varepsilon^{-\alpha}, s)^2 \right) ds \\
 &= \varepsilon \int_0^t c_\alpha(\varepsilon t - \varepsilon s)^{\alpha-1} \left(\frac{1}{2} \psi_1(p\varepsilon^{-\alpha}, \varepsilon s)^2 + \psi_1(p\varepsilon^{-\alpha}, \varepsilon s) \nu \psi_2(p\varepsilon^{-\alpha}, \varepsilon s) \right. \\
 &\quad \left. + \frac{1}{2} \nu^2 \psi_2(p\varepsilon^{-\alpha}, \varepsilon s)^2 \right) ds \\
 &= \varepsilon^\alpha \int_0^t c_\alpha(t - s)^{\alpha-1} \left(\frac{1}{2} (p c h_\varepsilon(s))^2 + p c h_\varepsilon(s) \nu \psi_2(p\varepsilon^{-\alpha}, \varepsilon s) \right. \\
 &\quad \left. + \frac{1}{2} \nu^2 \psi_2(p\varepsilon^{-\alpha}, \varepsilon s)^2 \right) ds
 \end{aligned}$$

for $t \in [0, \frac{1}{\varepsilon} T_{\psi_2}(\varepsilon^{-\alpha} p)]$. Multiplying by ε^α and cancelling powers of ε , we see that

$$\begin{aligned}
 \psi_2^\varepsilon(p, t) &:= \varepsilon^\alpha \psi_2(p\varepsilon^{-\alpha}, \varepsilon t) \\
 &= \int_0^t c_\alpha(t - s)^{\alpha-1} \left(\frac{1}{2} (p c h_\varepsilon(s))^2 + p c h_\varepsilon(s) \nu \psi_2^\varepsilon(p, s) + \frac{1}{2} \nu^2 \psi_2^\varepsilon(p, s)^2 \right) ds
 \end{aligned} \tag{3.14}$$

i.e. $\psi_2^\varepsilon(p, t)$ satisfies

$$D^\alpha \psi_2^\varepsilon(p, t) = \frac{1}{2} (p c h_\varepsilon(t) + \nu \psi_2^\varepsilon(p, t))^2. \tag{3.15}$$

Then we see that

$$\begin{aligned}
 \varepsilon^{2H} \log \mathbb{E}(e^{p\varepsilon^{-\alpha}(\text{VIX}_\varepsilon^2 - \zeta - \text{VIX}_0^2 - c\tilde{X}_\varepsilon)}) &= \varepsilon^{2H} \log \mathbb{E}(e^{p\varepsilon^{-\alpha} \int_0^\varepsilon f(\varepsilon - s) \tilde{X}_s ds}) \\
 &= \varepsilon^{2H} \log e^{\int_0^\varepsilon \xi_0(\varepsilon - s) g(p\varepsilon^{-\alpha}, s) ds} \\
 &= \varepsilon^{2H} \log e^{\varepsilon \int_0^1 \xi_0(\varepsilon - \varepsilon u) g(p\varepsilon^{-\alpha}, \varepsilon u) du} \\
 &= \varepsilon^{2\alpha} \int_0^1 \xi_0(\varepsilon - \varepsilon u) g(p\varepsilon^{-\alpha}, \varepsilon u) du \\
 &= \int_0^1 \xi_0(\varepsilon - \varepsilon u) g_\varepsilon(p, u) du \\
 &= \int_0^1 \xi_0(\varepsilon - \varepsilon u) \frac{1}{2} (p c h_\varepsilon(s) + \nu \psi_2^\varepsilon(p, s))^2 ds.
 \end{aligned}$$

where $g_\varepsilon(p, u) := \varepsilon^{2\alpha} g(p\varepsilon^{-\alpha}, \varepsilon u)$.

Recall that $h_\varepsilon(t) \rightarrow 0$ as $\varepsilon \rightarrow 0$, so we expect $\psi^\varepsilon(t)$ to tend to zero as well. To use Theorem 13.1.1 in [GLS90] to prove this, we first need to verify uniqueness for the solution ψ_ε , which we can do using the general argument given in Appendix A.

Since (3.15) with h_ε replaced by zero has a unique solution equal to the zero function, from Theorem 13.1.1 i) in [GLS90] we know there is a subsequence ε_n such that $\psi_{\varepsilon_n}(p, t)$

converges uniformly to zero on $[0, 1]$. Now suppose $\psi^\varepsilon(p, \cdot)$ does not converge uniformly to zero. Then we can find a subsequence ε_k such that $\psi_{\varepsilon_k}(p, \cdot)$ stays uniformly far from zero for all $k \in \mathbb{N}$. This subsequence has no subsequence that converges to zero, which contradicts Theorem 13.1.1 i) in [GLS90].

Then using (3.15) and the bounded convergence theorem and the continuity of $\xi_0(t)$ at $t = 0$ we see that

$$\begin{aligned} \lim_{\varepsilon \rightarrow 0} \varepsilon^{2H} \log \mathbb{E}(e^{p\varepsilon^{-\alpha}(\text{VIX}_\varepsilon^2 - \zeta - \text{VIX}_0^2 - c\tilde{X}_\varepsilon)}) &= \lim_{\varepsilon \rightarrow 0} \int_0^1 \xi_0(\varepsilon - \varepsilon u) D^\alpha g_\varepsilon(p, u) du \\ &= 0 \end{aligned} \quad (3.16)$$

for all $p \in \mathbb{R}$, and since $\zeta = O(\varepsilon)$ the limit is unchanged if we remove ζ here. Finally, setting $R_T := \text{VIX}_T^2 - \text{VIX}_0^2 - c\tilde{X}_T$, and using (3.16) we see that for $x > 0$ and $p > 0$

$$\lim_{T \rightarrow 0} T^{2H} \log \mathbb{P}\left(\frac{R_T}{T^{\frac{1}{2}-H}} > x\right) \leq \lim_{T \rightarrow 0} T^{2H} \log \mathbb{E}\left(e^{\frac{p}{T^{2H}}\left(\frac{R_T}{T^{\frac{1}{2}-H}} - x\right)}\right) = 0 - xp$$

and taking the inf over $p \geq 0$ we see that the left hand side is $-\infty$. Similarly for $x < 0$ and $p < 0$

$$\lim_{T \rightarrow 0} T^{2H} \log \mathbb{P}\left(\frac{R_T}{T^{\frac{1}{2}-H}} < x\right) \leq \lim_{T \rightarrow 0} T^{2H} \log \mathbb{E}\left(e^{\frac{p}{T^{2H}}\left(\frac{R_T}{T^{\frac{1}{2}-H}} - x\right)}\right) = -xp$$

and again we can take the inf over $p \leq 0$. Combining these observations, we see that

$$\lim_{T \rightarrow 0} T^{2H} \log \mathbb{P}\left(\left|\frac{R_T}{T^{\frac{1}{2}-H}}\right| > x\right) = -\infty$$

which shows that $(\text{VIX}_T^2 - \text{VIX}_0^2 - \zeta)/T^{\frac{1}{2}-H}$ and $c\tilde{X}_T/T^{\frac{1}{2}-H}$ are exponentially equivalent as $T \rightarrow 0$ (where ζ is defined in (3.10)) (see Definition 4.2.10 in [DZ98]), so the LDP follows from Theorem 4.2.13 in [DZ98] (as used in Theorem 2.3.3). Finally we can remove the ζ term here since ζ is deterministic and $o(T^{\frac{1}{2}-H})$ so $(\text{VIX}_T^2 - \text{VIX}_0^2)/T^{\frac{1}{2}-H}$ and $c\tilde{X}_T/T^{\frac{1}{2}-H}$ are also exponentially equivalent. ■

Remark 3.2.1 The lower bound for p in Theorem 3.2.1 is $-\infty$ (as opposed to some finite negative constant p_-) because for $\rho = 1$ and $p < 0$, $\bar{\Lambda}^{\rho=1}(\cdot)$ falls under case C for the ABCD classification used in [GGP19] (for our case we have to use driftless versions of the quantities defined in Eq 7 and 8 in [GGP19]; specifically $c_1(u) = \frac{1}{2}u^2$, $e_0(u) = \frac{1}{2}\rho v u$ and $e_1(u) = e_0(u)^2 - \frac{1}{4}v^2 u^2$, because we are working with \tilde{X} not the true log stock price process X). But since the ρ value associated with the \tilde{X} process is 1, we are in the special double root case for Eq 10 in [GGP19] where $e_1(u) = 0$ (borderline between C and B), but

since we still fall in Case C, there is no explosion for the VIE in Eq 2.25 for any negative p -value.

Remark 3.2.2 For the driftless case where $\xi_0(t) \equiv V_0$, using a simple ansatz and local martingale arguments, Proposition 4.6 in [GK19] derives the following exponential-affine formula for the mgf of $\int_T^{T+\Delta} \xi_T(u) du$:

$$\mathbb{E}(e^{h \int_T^{T+\Delta} (\xi_T(u) - V_0) du}) = e^{V_0 \int_0^T g(T+\Delta-u) du} = e^{V_0 \int_\Delta^{\Delta+T} g(s) ds}$$

for h in a certain interval, where g satisfies the non-standard VIE $g(t) = \frac{1}{2} (\int_0^t \frac{v}{\Gamma(\alpha)} (t-v)^{H-\frac{1}{2}} g(v) dv)^2$ for $t \geq \Delta$ and $g(t) = h$ for $t \in [0, \Delta]$ (note g is discontinuous at $t = \Delta$ or else we have a contradiction). This is clearly very relevant for pricing VIX options at non-zero maturities using Fourier inversion methods (see Subsection 3.2.6 for more details), but we will not need to use this VIE here.

3.2.2 VIX call option asymptotics

We now translate the LDP in Theorem 3.2.1 into small-time asymptotics for VIX call options for the same large deviations regime used in Chapter 2:

Corollary 3.2.3 *For $x > 0$ we have the following asymptotic behaviour for close-to-the money VIX call option prices:*

$$\lim_{T \rightarrow 0} T^{2H} \log \mathbb{E}((\text{VIX}_T - \text{VIX}_0 e^{xT^{\frac{1}{2}-H}})_+) = -J(2\text{VIX}_0^2 x)$$

where J is the rate function defined in the main Theorem 3.2.1.

Proof. See Appendix D. ■

Remark 3.2.3 For $x < 0$ (using very similar arguments), we obtain the following small-time behaviour for close-to-the-money VIX put options:

$$\lim_{T \rightarrow 0} T^{2H} \log \mathbb{E}((\text{VIX}_0 e^{xT^{\frac{1}{2}-H}} - \text{VIX}_T)_+) = -J(2\text{VIX}_0^2 x).$$

3.2.3 VIX future and implied volatility asymptotics

Lemma 3.2.4 $\frac{\text{VIX}_T^2 - \text{VIX}_0^2}{\sqrt{T}}$ tends weakly to $c\sqrt{V_0}Z$, where Z is a standard Normal.

Proof. From Theorem 3.3.2 (which is adapted from Eqs 2.8-2.10 in [AE19b] and Lemma 7.3 in [ALP19]), we know that

$$\mathbb{E}(e^{pc\tilde{X}_\varepsilon + p \int_0^\varepsilon f(\varepsilon-s)\tilde{X}_s ds}) = e^{\int_0^\varepsilon \xi_0(\varepsilon-s)D^\alpha \psi_2(p,s) ds} = e^{\int_0^\varepsilon \xi_0(\varepsilon-s)g(p,s) ds}$$

where (not to be confused with the ψ functions of the previous subsection)

$$\begin{aligned} \psi_1(p,t) &= p(c + \int_0^t f(s) ds) \\ \psi_2(p,t) &= \int_0^t c_\alpha(t-s)^{\alpha-1} \left(\frac{1}{2} \psi_1(p,s)^2 + \psi_1(p,s) \nu \psi_2(p,s) + \frac{1}{2} \nu^2 \psi_2(p,s)^2 \right) ds \end{aligned} \quad (3.17)$$

for $\varepsilon \leq T_{\psi_2}^*(p)$ where $T_{\psi_2}^*(p)$ is the explosion time for ψ_2 (where $g(p,t) := D^\alpha \psi_2(p,t) = \frac{1}{2} \psi_1(p,t)^2 + \psi_1(p,t) \rho \nu \psi_2(p,t) + \frac{1}{2} \nu^2 \psi_2(p,t)^2$), and recall that $f(s) := c_1 \alpha((s+\Delta)^{H-\frac{1}{2}} - s^{H-\frac{1}{2}})$. We first note that

$$\begin{aligned} \psi_1(\varepsilon^{-\frac{1}{2}} p, \varepsilon t) &= p \varepsilon^{-\frac{1}{2}} c + \frac{p}{\varepsilon^{\frac{1}{2}}} c_1 \alpha \int_0^{\varepsilon t} ((s+\Delta)^{H-\frac{1}{2}} - s^{H-\frac{1}{2}}) ds \\ &= p c \varepsilon^{-\frac{1}{2}} (1 + h_\varepsilon(t)) \end{aligned} \quad (3.18)$$

where h_ε is defined as above. Then

$$\begin{aligned} \psi_2(p \varepsilon^{-\frac{1}{2}}, \varepsilon t) &= \int_0^{\varepsilon t} c_\alpha(\varepsilon t - s)^{\alpha-1} \left(\frac{1}{2} \psi_1(p \varepsilon^{-\frac{1}{2}}, s)^2 + \psi_1(p \varepsilon^{-\frac{1}{2}}, s) \nu \psi_2(p \varepsilon^{-\frac{1}{2}}, s) \right. \\ &\quad \left. + \frac{1}{2} \nu^2 \psi_2(p \varepsilon^{-\frac{1}{2}}, s)^2 \right) ds \\ &= \varepsilon \int_0^t c_\alpha(\varepsilon t - \varepsilon s)^{\alpha-1} \left(\frac{1}{2} \psi_1(p \varepsilon^{-\frac{1}{2}}, \varepsilon s)^2 + \psi_1(p \varepsilon^{-\frac{1}{2}}, \varepsilon s) \nu \psi_2(p \varepsilon^{-\frac{1}{2}}, \varepsilon s) \right. \\ &\quad \left. + \frac{1}{2} \nu^2 \psi_2(p \varepsilon^{-\frac{1}{2}}, \varepsilon s)^2 \right) ds \\ &= \varepsilon^\alpha \int_0^t c_\alpha(t-s)^{\alpha-1} \left(\frac{1}{2} (p c \varepsilon^{-\frac{1}{2}} (1 + h_\varepsilon(s)))^2 + p c \varepsilon^{-\frac{1}{2}} (1 + h_\varepsilon(s)) \nu \psi_2(p \varepsilon^{-\frac{1}{2}}, \varepsilon s) \right. \\ &\quad \left. + \frac{1}{2} \nu^2 \psi_2(p \varepsilon^{-\frac{1}{2}}, \varepsilon s)^2 \right) ds \end{aligned} \quad (3.19)$$

for $t \in [0, \frac{1}{\varepsilon} T_{\psi_2}(\varepsilon^{-\frac{1}{2}} p)]$. Multiplying by $\sqrt{\varepsilon}$ and cancelling powers of ε , we see that $\psi_2^\varepsilon(p,t) := \sqrt{\varepsilon} \psi_2(p \varepsilon^{-\frac{1}{2}}, t)$ satisfies

$$\psi_2^\varepsilon(p,t) := \varepsilon^H \int_0^t c_\alpha(t-s)^{\alpha-1} \left(\frac{1}{2} (p c (1 + h_\varepsilon(s)))^2 + p c (1 + h_\varepsilon(s)) \nu \psi_2^\varepsilon(p,s) + \frac{1}{2} \nu^2 \psi_2^\varepsilon(p,s)^2 \right) ds$$

i.e. $\psi_2^\varepsilon(p, t)$ satisfies $D^\alpha \psi_2^\varepsilon(p, t) = \frac{1}{2} \varepsilon^H (pc(1 + h_\varepsilon(t)) + v\psi_2^\varepsilon(p, t))^2$. Then we see that

$$\begin{aligned} \mathbb{E}(e^{p\varepsilon^{-\frac{1}{2}}(\text{VIX}_\varepsilon^2 - \zeta - \text{VIX}_0^2)}) &= e^{\int_0^\varepsilon \xi_0(\varepsilon-s)\frac{1}{2}(\psi_1(\frac{p}{\sqrt{\varepsilon}}, s) + v\psi_2(\frac{p}{\sqrt{\varepsilon}}, s))^2 ds} \\ &= e^{\varepsilon \int_0^1 \xi_0(\varepsilon-\varepsilon s)\frac{1}{2}(\psi_1(\frac{p}{\sqrt{\varepsilon}}, \varepsilon s) + v\psi_2(\frac{p}{\sqrt{\varepsilon}}, \varepsilon s))^2 ds} \\ &= e^{\varepsilon \int_0^1 \xi_0(\varepsilon-\varepsilon s)\frac{1}{2}(pc\varepsilon^{-\frac{1}{2}}(1+h_\varepsilon(t)) + v\psi_2(\frac{p}{\sqrt{\varepsilon}}, \varepsilon s))^2 ds} \\ &= e^{\int_0^1 \xi_0(\varepsilon-\varepsilon s)\frac{1}{2}(pc(1+h_\varepsilon(t)) + v\psi_2^\varepsilon(p, s))^2 ds} \end{aligned}$$

ε^H and $h_\varepsilon(t) \rightarrow 0$ as $\varepsilon \rightarrow 0$, so we expect $\psi_2^\varepsilon(t)$ to tend to zero as well. To use Theorem 13.1.1 in [GLS90] to prove this, we first need to verify uniqueness for the solution ψ^ε , which we can do using the general argument given in Appendix A.

From Theorem 13.1.1 i) in [GLS90] (as above) we know that $\psi_2^\varepsilon(p, \cdot)$ converges uniformly to zero on any compact interval, and $\psi_2^\varepsilon(p, t)$ is continuous in ε and t on $\{(\varepsilon, t) : \varepsilon \in [0, 1], 0 \leq t < T_\varepsilon^*(p)\}$. Then using the above equations, the bounded convergence theorem and the continuity of $\xi_0(t)$ at $t = 0$ we see that

$$\lim_{\varepsilon \rightarrow 0} \mathbb{E}(e^{p\varepsilon^{-\frac{1}{2}}(\text{VIX}_\varepsilon^2 - \zeta - \text{VIX}_0^2)}) = e^{\frac{1}{2}p^2 V_0 c^2}$$

for all $p \in \mathbb{R}$, and since $\zeta = O(\varepsilon)$ the limit is unchanged if we remove ζ here. Finally, since Theorem 13.1.1 in [GLS90] is multi-dimensional, we can apply it to $(\text{Re}(\psi), \text{Im}(\psi))$ with p replaced by ik with $k \in \mathbb{R}$ as we discuss in Section 2.5. The result then follows from Lévy's convergence theorem. ■

We also have the following asymptotic estimate for the small-time behaviour of VIX futures which will be needed for the implied volatility asymptotics below.

Lemma 3.2.5 $\mathbb{E}(\text{VIX}_T - \text{VIX}_0) = O(\sqrt{T})$ as $T \rightarrow 0$.

Proof. Recall that $\text{VIX}_\varepsilon^2 - \text{VIX}_0^2 - \zeta = c\tilde{X}_\varepsilon + (f * \tilde{X})_\varepsilon$ from (3.11) where $\zeta = O(\varepsilon)$ and $f(s) := c_1 \alpha((s + \Delta)^{H-\frac{1}{2}} - s^{H-\frac{1}{2}})$ and setting $\hat{X}_t := \sqrt{V_0} \int_0^t (\rho dW_t + \bar{\rho} dB_t)$ and $\varepsilon = t$, we see that

$$\begin{aligned} \mathbb{E}((\text{VIX}_t^2 - \text{VIX}_0^2 - \zeta)^2)^{\frac{1}{2}} &= \mathbb{E}((c\tilde{X}_t + (f * \tilde{X})_t)^2)^{\frac{1}{2}} \\ &\leq \mathbb{E}([c(\tilde{X}_t - \hat{X}_t) + (f * (\tilde{X} - \hat{X}))_t]^2)^{\frac{1}{2}} + \mathbb{E}((c\hat{X}_t + (f * \hat{X})_t)^2)^{\frac{1}{2}} \\ &\leq c\mathbb{E}((\tilde{X}_t - \hat{X}_t)^2)^{\frac{1}{2}} + \mathbb{E}((f * (\tilde{X} - \hat{X}))_t^2)^{\frac{1}{2}} + c\sqrt{V_0}\sqrt{t} \\ &+ \mathbb{E}((\int_0^t f(t-s) \int_0^s \sqrt{V_0} dW_u ds)^2)^{\frac{1}{2}}. \end{aligned}$$

Using stochastic Fubini we can re-write the final term as $\mathbb{E}((\sqrt{V_0} \int_0^t \int_u^t f(t-s) ds dW_u)^2)^{\frac{1}{2}} = O(t)$. We also note that

$$\begin{aligned} \mathbb{E}((\tilde{X}_t - \hat{X}_t)^2) &= \int_0^t \mathbb{E}((\sqrt{V_s} - \sqrt{V_0})^2) ds = \int_0^t (\xi_0(s) + 2\sqrt{V_0}\mathbb{E}(\sqrt{V_s}) + V_0) ds \\ &\leq \int_0^t (\xi_0(s) + 2\sqrt{V_0}\mathbb{E}(V_s)^{\frac{1}{2}} + V_0) ds \\ &\leq \int_0^t (\xi_0(s) + 2\sqrt{V_0}\xi_0(s)^{\frac{1}{2}} + V_0) ds \sim 4V_0t \quad (3.20) \end{aligned}$$

as $t \rightarrow 0$, and for the convolution term (from Jensen) we see that

$$\begin{aligned} \mathbb{E}((f * (\tilde{X} - \hat{X}))_t^2)^{\frac{1}{2}} &= \mathbb{E}((t \cdot \frac{1}{t} \int_0^t f(t-s)(\tilde{X}_s - \hat{X}_s) ds)^2)^{\frac{1}{2}} \quad (3.21) \\ &\leq t \int_0^t f(t-s)^2 \mathbb{E}((\tilde{X}_s - \hat{X}_s)^2) ds = O(t). \end{aligned}$$

using (3.20) and the fact that $f \in L^2$.

Putting all this together, we see that $\frac{1}{\sqrt{t}} \mathbb{E}((\text{VIX}_t^2 - \text{VIX}_0^2 - \zeta)^2)^{\frac{1}{2}} \leq \bar{c}$ for some constant $\bar{c} > 0$ and t sufficiently small, and since $\zeta = O(t)$ we can remove the ζ term and the result still holds. Thus $\Upsilon_T := \frac{\text{VIX}_T^2 - \text{VIX}_0^2}{\sqrt{T}}$ is U.I., and (from Lemma 3.2.4) we know that $\Upsilon_T \xrightarrow{w} c\sqrt{V_0}Z$ as $\varepsilon \rightarrow 0$, where Z is a standard Normal.

Then from (3.10) and the Ito isometry we know that

$$\begin{aligned} \mathbb{E}((\text{VIX}_T^2 - \text{VIX}_0^2)^2)^{\frac{1}{2}} &\leq \mathbb{E}((\text{VIX}_T^2 - \frac{1}{\Delta} \int_T^{T+\Delta} \xi_0(u) du)^2)^{\frac{1}{2}} + |\frac{1}{\Delta} \int_T^{T+\Delta} \xi_0(u) du - \text{VIX}_0^2| \\ &= c_1 (\int_0^T ((T+\Delta-s)^{\frac{1}{2}+H} - (T-s)^{\frac{1}{2}+H})^2 \xi_0(s) ds)^{\frac{1}{2}} + |\zeta| \rightarrow 0 \quad (3.22) \end{aligned}$$

as $T \rightarrow 0$, so $\text{VIX}_T^2 \rightarrow \text{VIX}_0^2$ in L^2 and hence also in probability.

Now define $Y_T := \frac{1}{\sqrt{\text{VIX}_T + \text{VIX}_0}}$. Then Y_T is a continuous function of VIX_T^2 so (by the continuous mapping theorem) $Y_T \rightarrow Y_0$ (a constant) in probability, and clearly $Y_T \leq \frac{1}{\sqrt{\text{VIX}_0}}$. Note that $\frac{\text{VIX}_T - \text{VIX}_0}{\sqrt{T}} = \Upsilon_T Y_T$, and from above we know that $\Upsilon_T \xrightarrow{w} c\sqrt{V_0}Z$. From the general standard result that if $X_n \xrightarrow{w} X$ and $Y_n \rightarrow c$ (a constant) in probability, then $(X_n, Y_n) \xrightarrow{w} (X, c)$, we see that (Υ_T, Y_T) tends weakly to $(c\sqrt{V_0}Z, Y_0)$, and from the continuous mapping theorem $\Upsilon_T Y_T$ tends weakly to $Y_0 Z$. Moreover, Y_T is uniformly bounded so $\Upsilon_T Y_T$ is also U.I. Then by Theorem 3.5 in Billingsley[Bil99], $\mathbb{E}(\Upsilon_T Y_T) \rightarrow Y_0 \mathbb{E}(Z) = 0$. ■

Corollary 3.2.6 *If $\hat{\sigma}_{\text{VIX}}(K, T)$ denotes the implied volatility of a VIX call or put option with strike K , we see that*

$$\hat{\sigma}_{\text{VIX}}(x) := \lim_{T \rightarrow 0} \hat{\sigma}_{\text{VIX}}(\text{VIX}_0 e^{xT^{H-\frac{1}{2}}}, T) = \frac{|x|}{\sqrt{2J(2\text{VIX}_0^2 x)}} \quad (3.23)$$

for $x \in \mathbb{R}$, where J is the rate function introduced in the main Theorem 3.2.1.

Proof. Let $C^{\text{BS}}(S, K, \sigma, T)$ denote the usual Black-Scholes call option formula with zero interest rate and dividend. Then can easily verify that for any $b \in \mathbb{R}$

$$\lim_{T \rightarrow 0} T^{2H} \log C^{\text{BS}}(\text{VIX}_0 + b\sqrt{T}, \text{VIX}_0 e^{xT^{\frac{1}{2}-H}}, \sigma, T) = -\frac{x^2}{2\sigma^2}$$

so from Lemma 3.2.5 (and using that C^{BS} is monotonic in its first argument) we see that

$$\lim_{T \rightarrow 0} T^{2H} \log C^{\text{BS}}(\mathbb{E}(\text{VIX}_T), \text{VIX}_0 e^{xT^{\frac{1}{2}-H}}, \sigma, T) = -\frac{x^2}{2\sigma^2}.$$

For any $\delta \in (0, J(2\text{VIX}_0^2 x))$, we can then choose σ so that $-J(2\text{VIX}_0^2 x) = -\frac{x^2}{2\sigma^2} - \delta$. Then from Corollary 3.2.3

$$\begin{aligned} -J(2\text{VIX}_0^2 x) &= \limsup_{T \rightarrow 0} T^{2H} \log \mathbb{E}((\text{VIX}_T - \text{VIX}_0 e^{xT^{\frac{1}{2}-H}})^+) \\ &= \limsup_{T \rightarrow 0} T^{2H} \log C^{\text{BS}}(\mathbb{E}(\text{VIX}_T), \text{VIX}_0 e^{xT^{\frac{1}{2}-H}}, \hat{\sigma}_{\text{VIX}}(x, T), T) \\ &< \lim_{T \rightarrow 0} T^{2H} \log C^{\text{BS}}(\mathbb{E}(\text{VIX}_T), \text{VIX}_0 e^{xT^{\frac{1}{2}-H}}, \sigma, T) = -\frac{x^2}{2\sigma^2}. \end{aligned}$$

Since $C^{\text{BS}}(\cdot)$ is monotonically increasing in the σ argument, we see that $\limsup_{T \rightarrow 0} \hat{\sigma}_{\text{VIX}}(x, T) \leq \sigma$. Finally we let $\delta \rightarrow 0$, and we proceed similarly for the lower bound. ■

3.2.4 Small log-moneyness expansions

Using section 2.3.4, we obtain the following small-moneyness expansion

$$\begin{aligned} \bar{\Lambda}^{\rho=1}(p) &= \frac{1}{2} V_0 p^2 + \frac{V_0 v}{2\Gamma(2+\alpha)} p^3 + O(p^3) \\ (\bar{\Lambda}^{\rho=1})^*(x) &= \frac{1}{2} \frac{x^2}{V_0} - \frac{vx^3}{2V_0^2\Gamma(2+\alpha)} + O(x^4) \end{aligned}$$

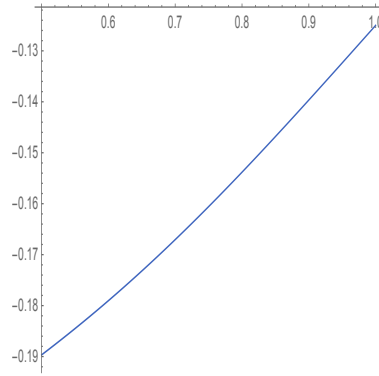


Fig. 3.1 Here we have plotted the factor $\frac{\Gamma(2+\alpha)^2\Gamma(1+2\alpha)+\Gamma(1+\alpha)^2(4\Gamma(2+\alpha)^2-6\Gamma(2+2\alpha))}{4\Gamma(1+\alpha)^2\Gamma(2+\alpha)^2\Gamma(2+2\alpha)}$ which appears in the convexity term in (3.24) as a function of α , and we see that this factor is strictly negative for all admissible α values.

and combining this with (3.23) we find that

$$\begin{aligned} \hat{\sigma}_{\text{VIX}}(x) &= \frac{\bar{c}v\sqrt{V_0}}{2\text{VIX}_0^2} + \frac{vx}{2\sqrt{V_0}\Gamma(2+\alpha)} \\ &+ \frac{\text{VIX}_0^2 v \alpha \Delta^{1-\alpha} \Gamma(\alpha)}{V_0^{\frac{3}{2}}} \frac{\Gamma(2+\alpha)^2\Gamma(1+2\alpha)+\Gamma(1+\alpha)^2(4\Gamma(2+\alpha)^2-6\Gamma(2+2\alpha))}{4\Gamma(1+\alpha)^2\Gamma(2+\alpha)^2\Gamma(2+2\alpha)} x^2 \\ &+ O(x^3) \end{aligned} \tag{3.24}$$

where $\bar{c} := \frac{1}{\Gamma(\alpha)\alpha}\Delta^{\alpha-1}$ and we see that the linear skew term is positive, and note there is no VIX smile if $v = 0$, since in this case V_t is constant. Moreover, since the fraction in front of the x^2 term only depends on α , we can readily verify from a graph that the $O(x^2)$ convexity term is strictly negative (see Figure 3.1 below), which is consistent with what is observed in practice, see e.g. [JMP21] and plots in [GJR20],[Guy20],[DeM18],[HJT20] et al. for more on this point. Since V_0 is already fixed from $\xi_0(\cdot)$ i.e. $V_0 = \xi_0(0)$ we see that we cannot independently fit the overall level and the skew of the VIX smile in the small- T limit. This issue is addressed in the companion article [FS21] by the addition of an independent CGMY-jump component to the model which allows the SPX and VIX short-maturity smiles to decouple in some sense.

3.2.5 The Edgeworth regime

Proceeding as in [FSV21], we have also formally verified the following asymptotic behaviour for VIX options in the Edgeworth regime under driftless rough Heston model

:

$$\hat{\sigma}_{\text{VIX}}(\sqrt{V_0}e^{x\sqrt{V_0}}, T) = \frac{v}{\sqrt{V_0}} \left(\frac{\Delta^{\alpha-1}}{2\alpha\Gamma(\alpha)} + \frac{1}{2\Gamma(2+\alpha)} xT^H + o(T^H) \right)$$

and we see that the at-the-money and skew terms are essentially the same as in the large deviations regime in (3.24). Since the answer is not surprising, we omit the details of the proof. To make this rigorous would require very fiddly tail estimates with Fourier arguments as in [EFGR19], which is beyond the scope of the discussion here.

3.2.6 Fourier inversion formula for VIX calls for $T > 0$

Note we have the following Fourier inversion formula for exact pricing of VIX call options, where we have used Cauchy and Fubini's theorem in the first and second lines respectively:

$$\begin{aligned} \mathbb{E}((\text{VIX}_T - K)^+) &= \frac{1}{2\pi} \int_0^\infty (v^{\frac{1}{2}} - K)^+ \int_{-\infty}^\infty e^{-i(u-ia)v} \phi(u-ia, T) dudv \\ &= \frac{1}{2\pi} \int_{-\infty}^\infty \phi(u-ia, T) \int_0^\infty (v^{\frac{1}{2}} - K)^+ e^{-i(u-ia)v} dv du \\ &= \frac{1}{2\sqrt{\pi}} \int_{-\infty}^\infty \phi(u-ia, T) \frac{\text{Erfc}(K\sqrt{a+iu})}{2(a+iu)^{\frac{3}{2}}} du \end{aligned}$$

where $\phi(u, T) := e^{V_0 T^{1-\alpha} \psi_2(iu, T)}$ is the characteristic function of VIX_T^2 , Erfc is the complementary error function and $a > 0$ such that ia is inside the strip of analyticity of $\phi(\cdot, T)$ (the condition $T\psi_2(\text{Re}(a)) > T$ is sufficient for this, by the same reasoning as in the proof of Theorem 7 of [GGP19]).

3.3 Appendix

3.3.1 Appendix A: Uniqueness of solutions to fractional Riccati VIEs

Following Theorem 3.1.2 and 3.1.4 in Chapter 3 in [Brun17], we consider a general non-linear VIE of the form

$$u(t) = \int_0^t \frac{1}{\Gamma(\alpha)} (t-s)^{H-\frac{1}{2}} \left(\frac{1}{2}(p+h(s))^2 + v(p+h(s))u(s) + \frac{1}{2}v^2u(s)^2 \right) ds \quad (\text{A.1})$$

where h is bounded and continuous, and suppose we have two continuous solutions u and u_2 to (A.1) on some interval $[0, T]$. Then

$$|u_2(t) - u(t)| \leq \int_0^t (t-s)^{H-\frac{1}{2}} L |u_2(s) - u(s)| ds$$

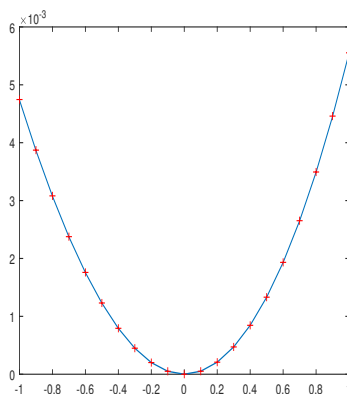


Fig. 3.2 Here we have plotted $\bar{\Lambda}^{\rho=1}(pc)$ in blue using an Adams scheme with 2000 time steps applied to the driftless rough Heston VIE in Eq 2.25 with $\rho = 1$ versus $T^{2H} \log \mathbb{E}(e^{\frac{p}{T^\alpha}(\text{VIX}_T^2 - V_0)})$ (red) obtained using Monte Carlo for $T = .0001$, $\xi_0(t) = V_0 = .04$, $H = .25$, $\nu = .25$, and $\Delta = 1/12$ with 100,000 simulations, and we see both quantities are in close agreement.

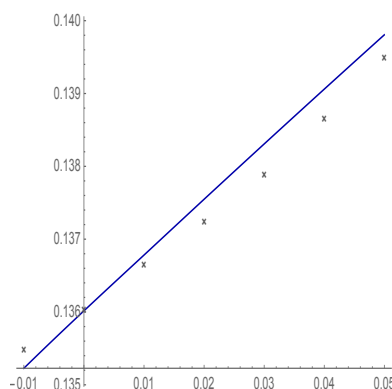


Fig. 3.3 Here we have computed $\hat{\sigma}(x)$ using the same method as for Figure 2.3 with 15 terms (blue) versus the VIX implied volatility computed using Monte Carlo (crosses) for $T = .0001$, $V_0 = 1$, $H = .25$, $\nu = .25$ and $\Delta = 1$ with 10,000,000 simulations and 200 time steps. It is difficult to verify exact agreement here since we can no longer exploit the usual Romano-Touzi/Willard conditioning trick for the Monte Carlo since ρ is effectively 1 here, and because of this the MC results for the left portion of the smile are less accurate since there is a significantly lower exponentially small probability that these (put) options expire in-the-money.

for some local Lipschitz constant L (since the function $\frac{1}{2}(p+h(s))^2 + v(p+h(s))u + \frac{1}{2}v^2u^2$ is locally Lipschitz in u), and we write this more succinctly as $\Delta \leq -k * \Delta$, where $k(t) := -Lt^{H-\frac{1}{2}}$ and $\Delta = |u_2 - u_1|$. The Laplace transform of k is $\hat{k}(\lambda) = -c\lambda^{-\alpha}$ where $c = L\Gamma(\alpha)$, and (from the definition of Eq 2.11 in [ALP19]) the resolvent r of k satisfies

$$k * r = k - r$$

which implies that

$$\hat{k}\hat{r} = \hat{k} - \hat{r}$$

and hence

$$\hat{r}(\lambda) = 1 - (1 + \hat{k}(\lambda))^{-1} = \frac{c}{c - \lambda^\alpha}.$$

Then \hat{r} is the Laplace transform of $r(t) = -ct^{\alpha-1}E_{\alpha,\alpha}(ct^\alpha)$ which is non-positive (see e.g. Table 1 in [ALP19] with $c \mapsto -c$ and end of proof of Proposition 2.2.1). Then using the following Lemma (taken from Appendix A.2 in [ACLP19]), we see that in fact $\Delta \equiv 0$, so we have uniqueness.

Lemma 3.3.1 (See Appendix A.2 in [ACLP19]. Suppose $f, g, k \in L^1([0, T])$. Assume k has non-positive resolvent r . Then if $f \leq g - k * f$, then $f \leq g - r * g$.

Proof. Write $f + k * f = g - h$ for $h \geq 0$, so $\hat{f} + \hat{k}\hat{f} = \hat{g} - \hat{h}$. Then from the definition of the resolvent: $\hat{k}\hat{r} = \hat{k} - \hat{r}$ we find that

$$\begin{aligned} \hat{f} + \frac{\hat{r}}{1 - \hat{r}}\hat{f} &= \hat{g} - \hat{h} \\ \Rightarrow \hat{f}(1 - \hat{r}) + \hat{r}\hat{f} &= \hat{f} = \hat{g} - \hat{h} - \hat{r}(\hat{g} - \hat{h}) \end{aligned}$$

so $f = g - h - r * (g - h) \leq g - r * g$. ■

3.3.2 Appendix B: Derivation of the VIE

Theorem 3.3.2 (minor variant of Theorem 7.1 in [ALP19] without their restriction that $\text{Re}(\psi_1) \in [0, 1]$ and no drift term). Consider the d -dimensional stochastic convolution equation:

$$X_t = X^0(t) + \int_0^t \tilde{K}(t-s)\sigma(X_s)dW_s$$

so $\mathbb{E}(X_t) = X^0(t)$, where $\tilde{K} \in L^2([0, T]; \mathbb{R}^{d \times d})$, $a(x) = \sigma^T(x)\sigma(x) = x_1A^1 + \dots + x_dA^d$, A^i is a (symmetric) $d \times d$ matrix for each $i = 1..d$, $A(u) = (uA^1u^T, uA^2u^T, \dots, uA^du^T)$, W is a d -dimensional Brownian motion, and

$$\psi = u\tilde{K} + (f + \frac{1}{2}A(\psi)) * \tilde{K} \quad (\text{B.1})$$

for $f \in L^1$ and let Y satisfy $dY_t = -\frac{1}{2}\psi(T-t)^2\sigma(X_t)^2dt + \psi(T-t)\sigma(X_t)dW_t$ with

$$Y_0 = uX^0(T) + (f * X_0)_T + \frac{1}{2} \int_0^T \psi(T-s)a(X^0(t))\psi(T-s)^T ds.$$

Then if ψ is bounded on $[0, T]$, then e^Y is an \mathcal{F}_t^W -martingale on $[0, T]$ and we have the exponential-affine formula:

$$\mathbb{E}(e^{uX_T + (f * X)_T} | \mathcal{F}_t^W) = e^{Y_t}.$$

For our specific case of interest for the Rough Heston model, $f_2 = 0$, $u_2 = 0$ and $X^0(t) = (0, \xi_0(t))$ so we can re-write (B.1) in component form as

$$\begin{aligned} \psi_1 &= u_1 + f_1 * 1 \\ \psi_2 &= u_2K + \frac{1}{2}(\psi_1^2 + 2\nu\psi_1\psi_2 + \nu^2\psi_2^2) * K = I^\alpha F(\psi_1, \psi_2) \end{aligned}$$

where $K(t) = \frac{t^{\alpha-1}}{\Gamma(\alpha)}$ and $F(\psi_1, \psi_2) = \frac{1}{2}(\psi_1 + \nu\psi_2)^2$, and $Y_0 = uX^0(T) + (f * X_0)_T + \frac{1}{2} \int_0^T \psi(T-s)a(X^0(T))\psi(T-s)^T ds = I^1(\xi_0(T - (\cdot))D^\alpha\psi_2)(T)$, which further simplifies to the familiar expression $V_0I^{1-\alpha}\psi_2(T)$ if $\xi_0(\cdot)$ is flat.

Proof. We let $\mathcal{F}_t := \mathcal{F}_t^W$ throughout, and we first note that

$$\mathbb{E}(X_s | \mathcal{F}_t) = X_0(s) + \int_0^{s \wedge t} \tilde{K}(s-v)dM_v \quad (\text{B.2})$$

where $dM_t = \sigma(X_t)dW_t$. Now let

$$Y_t = \mathbb{E}(uX_T + \int_0^T f(T-s)X_s ds | \mathcal{F}_t) + \frac{1}{2} \left(\int_0^T - \int_0^t \right) \psi(T-s)a(\mathbb{E}(X_s | \mathcal{F}_t))\psi(T-s)^T ds$$

for $t \leq T$. Using (B.2) and the affine property of $a(\cdot)$ we can re-write Y_t in the form $Y_0 + (\dots)$ as

$$\begin{aligned}
 Y_t &= uX^0(T) + (f * X_0)(T) + \frac{1}{2} \int_0^T \psi(T-s)a(X_0(s))\psi(T-s)^T ds \\
 &+ u \int_0^t \tilde{K}(T-s)dM_s + \int_0^T f(T-s) \int_0^{s \wedge t} \tilde{K}(s-v)dM_v ds \\
 &+ \frac{1}{2} \int_0^T \psi(T-s)a\left(\int_0^{s \wedge t} \tilde{K}(s-v)dM_v\right)\psi(T-s)^T ds - \frac{1}{2} \int_0^t \psi(T-s)a(X_s)\psi(T-s)^T ds
 \end{aligned} \tag{B.3}$$

(the sum of the three terms on the right hand side in the first line here is Y_0). From Fubini we see that the fifth term here can be re-written as

$$\int_0^T f(T-s) \int_0^{s \wedge t} \tilde{K}(s-v)dM_v ds = \int_0^t \int_v^T f(T-s)\tilde{K}(s-v)dsdM_v.$$

Similarly

$$\begin{aligned}
 &\int_0^T \psi(T-s)a\left(\int_0^{s \wedge t} \tilde{K}(s-v)dM_v\right)\psi(T-s)^T ds \\
 &= \int_0^T \psi(T-s)\left(\sum_{i=1}^d A^i \int_0^{s \wedge t} \tilde{K}(s-v)dM_v^i\right)\psi(T-s)^T ds \\
 &= \int_0^t \int_v^T A(\psi(T-s))\tilde{K}(s-v)dsdM_v
 \end{aligned} \tag{B.4}$$

and recall that $A(u) = (uA^1(u)u^T, uA^2(u)u^T, \dots, uA(u)^d u^T)$. Thus

$$\begin{aligned}
 Y_t &= Y_0 + u \int_0^t \tilde{K}(T-v)dM_v + \int_0^t \int_v^T f(T-s)\tilde{K}(s-v)dsdM_v \\
 &+ \frac{1}{2} \int_0^t \int_v^T A(\psi(T-s))\tilde{K}(s-v)dsdM_v - \frac{1}{2} \int_0^t \psi(T-s)a(X_s)\psi(T-s)^T ds \\
 &= Y_0 + \int_0^t (u\tilde{K}(T-v) + \int_v^T f(T-s)\tilde{K}(s-v)ds) + \frac{1}{2} \int_v^T A(\psi(T-s))\tilde{K}(s-v)ds)dM_v \\
 &- \frac{1}{2} \int_0^t \psi(T-s)a(X_s)\psi(T-s)^T ds
 \end{aligned}$$

and we note that

$$\begin{aligned}
 \int_v^T f(T-s)\tilde{K}(s-v)ds &= \int_0^{T-v} f(T-(s+v))\tilde{K}(s)ds \\
 &= (\tilde{K} * f)(T-v)
 \end{aligned}$$

$$\text{and similarly } \int_v^T A(\psi(T-s))\tilde{K}(s-v)ds = (A(\psi) * \tilde{K})(T-v)$$

so

$$Y_t = Y_0 + \int_0^t (u\tilde{K}(T-v) + (f * \tilde{K})(T-v) + \frac{1}{2}(A(\psi) * \tilde{K})(T-v))dM_v - \frac{1}{2} \int_0^t \psi(T-s)a(X_s)\psi(T-s)^T ds.$$

Comparing this expression to the ‘‘Driftless’’ Riccati eq:

$$\psi = u\tilde{K} + (f + \frac{1}{2}A(\psi)) * \tilde{K}$$

we see that

$$Y_t = Y_0 + \int_0^t \psi(T-s)\sigma(X_s)dW_s - \frac{1}{2} \int_0^t \psi(T-s)a(X_s)\psi(T-s)^T ds$$

so e^{Y_t} is a local martingale. If e^Y is a true martingale on $[0, T]$ (see the end of the proof for clarification on this point), then $\mathbb{E}(e^{Y_T} | \mathcal{F}_t) = e^{Y_t}$ and in particular

$$\mathbb{E}(e^{Y_T}) = \mathbb{E}(e^{uX_T + (f * X)_T}) = e^{Y_0} = e^{uX^0(T) + (f * X_0)(T) + \frac{1}{2} \int_0^T \psi(T-s)a(X_0(s))\psi(T-s)^T ds}. \quad (\text{B.5})$$

In our specific case $X_t = \begin{pmatrix} \tilde{X}_t \\ V_t \end{pmatrix}$ with kernel $\tilde{K} = \begin{pmatrix} 1 & 0 \\ 0 & K \end{pmatrix}$ and $X^0(t) = (0, \xi_0(t))$. Then

$$\sigma(X_t) = \sqrt{V_t} \begin{pmatrix} 0 & 1 \\ 0 & v \end{pmatrix} \text{ so}$$

$$a(X_t) := \sigma(X_t)\sigma^T(X_t) = V_t \begin{pmatrix} 0 & 1 \\ 0 & v \end{pmatrix} \begin{pmatrix} 0 & 0 \\ 1 & v \end{pmatrix} = V_t \begin{pmatrix} 1 & v \\ v & v^2 \end{pmatrix}$$

which implies that $A^1 = 0$ and

$$A^2(\psi) = \begin{pmatrix} \psi_1 & \psi_2 \end{pmatrix} \begin{pmatrix} 1 & v \\ v & v^2 \end{pmatrix} \begin{pmatrix} \psi_1 \\ \psi_2 \end{pmatrix} = \psi_1^2 + 2v\psi_1\psi_2 + v^2\psi_2^2.$$

Then the Riccati-Volterra eq becomes:

$$\begin{aligned} \psi = \begin{pmatrix} \psi_1 & \psi_2 \end{pmatrix} &= u\tilde{K} + (f + \frac{1}{2}A(\psi)) * \tilde{K} \\ &= (u_1, u_2) \begin{pmatrix} 1 & 0 \\ 0 & K \end{pmatrix} + ((f_1, 0) + \frac{1}{2}(0, \psi_1^2 + 2v\psi_1\psi_2 + v^2\psi_2^2)) * \begin{pmatrix} 1 & 0 \\ 0 & K \end{pmatrix} \end{aligned}$$

which we can re-write as

$$\begin{aligned}\psi_1 &= u_1 + f_1 * 1 \\ \psi_2 &= u_2 K + \frac{1}{2}(\psi_1^2 + 2\nu\psi_1\psi_2 + \nu^2\psi_2^2) * K = u_2 K + I^\alpha F(\psi_1, \psi_2).\end{aligned}\quad (\text{B.6})$$

and

$$Y_0 = u_2 \xi_0(T) + (f * X)_0 + \frac{1}{2} \int_0^T \psi(T-s) a(X_0(s)) \psi(T-s)^T ds$$

and

$$\begin{aligned}\frac{1}{2} \psi(T-s) a(X^0(t)) \psi^T(T-s) &= \xi_0(t) (\psi_1(T-s), \psi_2(T-s)) \begin{pmatrix} 1 & \nu \\ \nu & \nu^2 \end{pmatrix} \begin{pmatrix} \psi_1(T-s) \\ \psi_2(T-s) \end{pmatrix} \\ &= \xi_0(t) (\psi_1^2 + 2\nu\psi_1\psi_2 + \nu^2\psi_2^2)\end{aligned}$$

so

$$\begin{aligned}Y_0 &= u_2 \xi_0(T) + (f * X)_0 + \frac{1}{2} \int_0^T \psi(s) a(X_0(T-s)) \psi(s)^T ds \\ &= u_2 \xi_0(T) + (f * X)_0 + \frac{1}{2} \int_0^T \xi_0(T-s) (\psi_1(s)^2 + 2\nu\psi_1(s)\psi_2(s) + \nu^2\psi_2(s)^2) ds \\ &= u_2 \xi_0(T) + (f * X)_0 + I^1(\xi(T - (\cdot))) D^\alpha(\psi_2 - u_2 K)(T) \\ &= (f * X)_0 + I^1(\xi(T - (\cdot))) D^\alpha \psi_2(T)\end{aligned}$$

(where we have used (B.6) for the second equality and $(I^{1-\alpha}K)(t) = 1$ for the final equality.) which is the exponent in (B.5). Moreover

$$Y_t = \xi_0(T) + \int_0^t \psi(T-s) \sigma(X_s) dW_s - \frac{1}{2} \int_0^t \psi(T-s) a(X_s) \psi(T-s)^T ds$$

so

$$\begin{aligned}dY_t &= \psi(T-t) \sigma(X_t) dW_t - \frac{1}{2} \psi(T-t) a(X_t) \psi(T-t)^T dt \\ &= \sqrt{V_t} (\psi_1(T-t) + \nu\psi_2(T-t)) dW_t^2 - \frac{1}{2} V_t (\psi_1(T-t)^2 + \nu\psi_1(T-t))^2 dt.\end{aligned}$$

Then from Lemma 7.3 in [ALP19], e^Y is a genuine \mathcal{F}_t^W -martingale on $[0, T]$ if $\psi_1 + \nu\psi_2 \in L^\infty([0, T])$ and since f is integrable, $\psi_2 \in L^\infty$ implies $\psi_1 + \nu\psi_2 \in L^\infty$, and $\psi_2 \in L^\infty$ if $T^*(u) > T$ (where $T^*(u)$ is the explosion time for ψ_2) since the solution to the VIE for ψ_2 is continuous up to the explosion time. ■

3.3.3 Appendix C: Uniform moment bound

Lemma 3.3.3 (see also Lemma 3.1 in [ALP19] and Lemma A.1 in [JP20]). For $m \geq 2$

$$\sup_{t \leq T} \mathbb{E}(V_t^m) \leq c_{m,T}$$

for some finite constant $c_{m,T}$ which depends on m and T and the model parameters.

Proof. Setting $K(t) = t^{\alpha-1}/\Gamma(\alpha)$ and using the Burkholder-Davis-Gundy inequality applied to the martingale $M_u := \int_0^u K(t-s)\sqrt{V_s}dW_s$ at $t = u$, we see that

$$\begin{aligned} \mathbb{E}(V_t^m) &= \mathbb{E}((\xi_0(t) + \int_0^t K(t-s)\sqrt{V_s}dW_s)^m) \\ &\leq 2^m \xi_0(t)^m + 2^m C_m \mathbb{E}((\int_0^t K(t-s)^2 V_s ds)^{\frac{1}{2}m}) \\ &= 2^m \xi_0(t)^m + 2^m C_m \mathbb{E}((\int_0^t K(t-s)^{2-\frac{4}{m}} K(t-s)^{\frac{4}{m}} V_s ds)^{\frac{1}{2}m}) \\ &\leq 2^m \xi_0(t)^m + 2^m C_m \|K\|_2^{m-2} \int_0^t K(t-s)^2 \mathbb{E}(V_s^{\frac{1}{2}m}) ds \\ &\leq 2^m \xi_0(t)^m + 2^m C_m \|K\|_2^{m-2} \int_0^t K(t-s)^2 \mathbb{E}(a(1+V_s)^m) ds \end{aligned}$$

where we have used Hölder's inequality with $p = \frac{1}{2}m$ and $q = (1 - 1/p)^{-1}$ in the final line as in Appendix A.2 in [JP20], so $f(t) := \mathbb{E}(V_t^m)$ satisfies

$$f(t) \leq c + c \int_0^t K(t-s)^2 f(s) ds = c + c \int_0^t (t-s)^{\alpha_2-1} f(s) ds$$

for some constant $c > 0$ and $t \in [0, T]$, where $\alpha_2 = 2H$. Using Lemma 3.3.1, we know that

$$\begin{aligned} f(t) &\leq c - (r * c)(t) \\ &= c + c \int_0^t (t-s)^{\alpha_2-1} E_{\alpha_2, \alpha_2}(\tilde{c}(t-s)^\alpha) c ds < \infty \end{aligned}$$

where r is the resolvent of ct^{α_2-1} given by $\hat{r}(t) = -\tilde{c}t^{\alpha_2-1}E_{\alpha_2, \alpha_2}(\tilde{c}t^{\alpha_2})$ where $\tilde{c} = c\Gamma(\alpha_2)$. $E_{\alpha_2, \alpha_2}(\tilde{c}(t-s)^{\alpha_2})$ is bounded on $[0, t]$, so $f(t) \leq \text{const.} \times \int_0^t (t-s)^{\alpha_2-1} \tilde{c}s^{-\alpha_2} ds < \infty$ for all $s \in [0, t]$. ■

3.3.4 Appendix D: Asymptotics for VIX call options

From Jensen's inequality, we know that for any $q \geq 1$ we have

$$(\text{VIX}_T^2)^q = \left(\frac{1}{\Delta} \int_T^{T+\Delta} \xi_T(u) du\right)^q \leq \frac{1}{\Delta} \int_T^{T+\Delta} \xi_T(u)^q du$$

and hence

$$\begin{aligned} \mathbb{E}(\text{VIX}_T^{2q}) &\leq \frac{1}{\Delta} \int_T^{T+\Delta} \mathbb{E}(\xi_T(u)^q) du = \frac{1}{\Delta} \int_T^{T+\Delta} \mathbb{E}(\mathbb{E}(V_u | \mathcal{F}_T)^q) du \\ &\leq \frac{1}{\Delta} \int_T^{T+\Delta} \mathbb{E}(V_u^q) du \end{aligned} \quad (\text{D.1})$$

which will be needed further down.

- **Lower bound.** We first note that for x fixed and any $\delta \in (0, x)$, $e^{xT^{\frac{1}{2}-H}} \leq 1 + (x + \delta)T^{\frac{1}{2}-H}$ for T sufficiently small. Recall that $\text{VIX}_0^2 = \frac{1}{\Delta} \int_T^{T+\Delta} \xi_T(u) du$ and we set $k_{x,\delta} := \text{VIX}_0(x + \delta)$. We first note that for $\delta > 0$ and $T = T(\delta)$ sufficiently small, $e^{xT^{\frac{1}{2}-H}} \leq 1 + (x + \delta)xT^{\frac{1}{2}-H}$. Thus for $T = T(\delta)$ sufficiently small

$$\begin{aligned} &\mathbb{E}((\text{VIX}_T - \text{VIX}_0 e^{xT^{\frac{1}{2}-H}})_+) \\ &\geq \mathbb{E}((\text{VIX}_T - \text{VIX}_0(1 + (x + \delta)T^{\frac{1}{2}-H}))_+) \\ &= T^{\frac{1}{2}-H} \mathbb{E}((\frac{\text{VIX}_T - \text{VIX}_0}{T^{\frac{1}{2}-H}} - k_{x,\delta})_+) \\ &\geq \delta T^{\frac{1}{2}-H} \mathbb{E}(1_{\frac{\text{VIX}_T - \text{VIX}_0}{T^{\frac{1}{2}-H}} > k_{x,\delta} + \delta}) \\ &= \delta T^{\frac{1}{2}-H} \mathbb{P}(\text{VIX}_T > \text{VIX}_0 + T^{\frac{1}{2}-H}(k_{x,\delta} + \delta)) \\ &= \delta T^{\frac{1}{2}-H} \mathbb{P}(\text{VIX}_T^2 > \text{VIX}_0^2 + 2\text{VIX}_0(k_{x,\delta} + \delta)T^{\frac{1}{2}-H} + (k_{x,\delta} + \delta)^2 T^{1-2H}). \end{aligned}$$

But for $T = T(\delta)$ sufficiently small, the right hand side here is greater than or equal to

$$\begin{aligned} &\delta T^{\frac{1}{2}-H} \mathbb{P}(\text{VIX}_T^2 - \text{VIX}_0^2 > 2\text{VIX}_0(k_{x,\delta} + 2\delta)T^{\frac{1}{2}-H}) \\ &= \delta T^{\frac{1}{2}-H} \mathbb{P}(\frac{1}{\Delta} \int_T^{T+\Delta} \xi_T(u) du - \frac{1}{\Delta} \int_T^{T+\Delta} \xi_0(u) du > 2\text{VIX}_0(k_{x,\delta} + 2\delta)T^{\frac{1}{2}-H}). \end{aligned}$$

Then using the LDP and the continuity of J we see that

$$\begin{aligned} \liminf_{T \rightarrow 0} T^{2H} \log \mathbb{E}((\text{VIX}_T - \text{VIX}_0 e^{xT^{\frac{1}{2}-H}})_+) &\geq -J(2\text{VIX}_0(k_{x,\delta} + 2\delta)) \\ &= -J(2\text{VIX}_0^2 + 2\delta\text{VIX}_0 + 4\delta\text{VIX}_0). \end{aligned}$$

We then let $\delta \rightarrow 0$ and again use the continuity of the rate function $J(x)$ to obtain the required lower bound.

- Upper bound. From Hölder's inequality, we note that for $q > 1$

$$\begin{aligned}
& \mathbb{E}((\text{VIX}_T - \text{VIX}_0 e^{xT^{\frac{1}{2}-H}})_+) \\
& \leq \mathbb{E}((\text{VIX}_T - \text{VIX}_0(1 + xT^{\frac{1}{2}-H}))_+) \\
& = \mathbb{E}((\text{VIX}_T - \text{VIX}_0(1 + xT^{\frac{1}{2}-H}))_+ \mathbf{1}_{\text{VIX}_T \geq \text{VIX}_0(1 + xT^{\frac{1}{2}-H})}) \\
& \leq \mathbb{E}[(\text{VIX}_T - \text{VIX}_0(1 + xT^{\frac{1}{2}-H}))_+^q]^{\frac{1}{q}} \mathbb{E}(\mathbf{1}_{\text{VIX}_T \geq \text{VIX}_0(1 + xT^{\frac{1}{2}-H})})^{1-\frac{1}{q}}.
\end{aligned}$$

Thus

$$\begin{aligned}
& T^{2H} \log \mathbb{E}((\text{VIX}_T - \text{VIX}_0(1 + xT^{\frac{1}{2}-H}))_+) \\
& \leq \frac{T^{2H}}{q} \log \mathbb{E}[(\text{VIX}_T - \text{VIX}_0(1 + xT^{\frac{1}{2}-H}))_+^q] \\
& + T^{2H} (1 - \frac{1}{q}) \log \mathbb{P}(\text{VIX}_T \geq \text{VIX}_0(1 + xT^{\frac{1}{2}-H})) \\
& \leq \frac{T^{2H}}{q} \log \mathbb{E}(\text{VIX}_T^q) + T^{2H} (1 - \frac{1}{q}) \log \mathbb{P}(\text{VIX}_T \geq \text{VIX}_0(1 + xT^{\frac{1}{2}-H})) \\
& \leq \frac{T^{2H}}{q} \log(\mathbb{E}(\text{VIX}_T^{2q})^{\frac{1}{2}}) + T^{2H} (1 - \frac{1}{q}) \log \mathbb{P}(\text{VIX}_T \geq \text{VIX}_0(1 + xT^{\frac{1}{2}-H})) \\
& \leq \frac{T^{2H}}{q} \frac{1}{2} \log(\frac{1}{\Delta} \int_T^{T+\Delta} \mathbb{E}(V_u^q) du) + T^{2H} (1 - \frac{1}{q}) \log \mathbb{P}(\text{VIX}_T^2 \geq \text{VIX}_0^2(1 + xT^{\frac{1}{2}-H})^2) \\
& \quad \text{(by (D.1))} \\
& \leq \frac{T^{2H}}{q} \frac{1}{2} \log(\frac{1}{\Delta} \int_T^{T+\Delta} (\mathbb{E}(V_u^q)^{\frac{1}{q}})^q du) + T^{2H} (1 - \frac{1}{q}) \log \mathbb{P}(\text{VIX}_T^2 \geq \text{VIX}_0^2(1 + 2xT^{\frac{1}{2}-H})) \\
& \quad \text{(using Minkowski applied to } \mathbb{E}((V_u)^q)) \\
& \leq \frac{T^{2H}}{q} \frac{1}{2} \log(c_{q,T}^{\frac{1}{q}})^q + T^{2H} (1 - \frac{1}{q}) \log \mathbb{P}(\text{VIX}_T^2 \geq \text{VIX}_0^2(1 + 2xT^{\frac{1}{2}-H}))
\end{aligned}$$

for some finite constant $c_{q,T}$ depending on q and T , where we have used Lemma 3.3.3 in the final line. Letting $T \rightarrow 0$ in the final line and using the LDP and the continuity of J , and then letting $q \rightarrow \infty$, we see that

$$\limsup_{T \rightarrow 0} T^{2H} \log \mathbb{E}((\text{VIX}_T - \text{VIX}_0(1 + xT^{\frac{1}{2}-H}))_+) \leq -J(2\text{VIX}_0^2 x).$$

Chapter 4

The Riemann-Liouville field and its GMC as $H \rightarrow 0$, and skew flattening for the rough Bergomi model

4.1 Introduction

Gaussian multiplicative chaos (GMC) is a random measure on a domain of \mathbb{R}^d that can be formally written as $M_\gamma(dx) = e^{\gamma X_x - \frac{1}{2}\gamma^2 \mathbb{E}(X_x^2)} dx$ where X is a Gaussian field with zero mean and covariance $K(x, y) := \mathbb{E}(X_x X_y) = \log^+ \frac{1}{|y-x|} + g(x, y)$ for some bounded continuous function g . X is not defined pointwise because there is a singularity in its covariance, rather X is a random tempered distribution, i.e. an element of the dual of the Schwartz space \mathcal{S} under the locally convex topology induced by the Schwartz space semi-norms. For this reason, making rigorous sense of M_γ requires a regularizing sequence X^ε of Gaussian processes (with the singularity removed), see e.g. [BBM13] and [BM03] and Section 4.2.2 here for such a regularization in 1d based on integrating a Gaussian white noise over truncated triangular regions or page 17 in [RV10]. In most of the literature on GMC, the choice of X^ε is a martingale in ε , from which we can then easily verify that $M_\gamma^\varepsilon(A) = \int_A e^{\gamma X_x^\varepsilon - \frac{1}{2}\gamma^2 \text{Var}(X_x^\varepsilon)} dx$ is a martingale, and then obtain a.s. convergence of $M_\gamma^\varepsilon(A)$ using the martingale convergence to a random variable $M_\gamma(A)$ with $\mathbb{E}(M_\gamma(A)) = \text{Leb}(A)$, and with a bit more work we can verify that $M_\gamma(\cdot)$ defines a random measure (see page 18 in [RV10]).

If $\gamma^2 < 2d$, $M_\gamma^\varepsilon(dx) = e^{\gamma X_x^\varepsilon - \frac{1}{2}\gamma^2 \mathbb{E}((X_x^\varepsilon)^2)} dx$ tends weakly to a multifractal random measure M_γ with full support a.s. which satisfies the local multifractality property

$$\lim_{\delta \rightarrow 0} \frac{\log \mathbb{E}(M_\gamma([x, x + \delta]^d)^q)}{\log \delta} = \zeta(q) \quad (4.1)$$

for $q \in (1, q^*)$ (see Proposition 3.7 in [RV10]), where $\zeta(q^*) = 1$ ¹ and

$$\zeta(q) = dq - \frac{1}{2}\gamma^2(q^2 - q)$$

so $q^* = \frac{2}{\gamma^2}$ for $d = 1$, and $\mathbb{E}(M_\gamma([0, t])^q) = \infty$ if $q > q^*$, see Theorem 2.13 in [RV14] and Lemma 3 in [BM03]). M_γ is the zero measure for $\gamma^2 = 2d$ and $\gamma^2 > 2d$; in these cases a different re-normalization is required to obtain a non-trivial limit.

In the sub-critical case, using a limiting argument it can be shown that M_γ satisfies

$$\mathbb{E}\left(\int_D F(X, z)M_\gamma(dz)\right) = \mathbb{E}\left(\int_D F(X + \gamma^2 K(z, \cdot), z)dz\right) \quad (4.2)$$

for any measurable function F and any interval D . This comes from the Cameron-Martin theorem for Gaussian measures, the notion of *rooted measures* and the disintegration theorem (see [FS20]). (4.2) can be taken as the definition of GMC, and it uniquely determines M_γ as a measurable function of X , and hence also uniquely fix its law. GMC also has natural applications in Liouville Quantum Field Theory.

Continuing in the same vein as [NR18] (see also [HN20]), we consider a re-scaled Riemann-Liouville process $Z_t^H = \int_0^t (t-s)^{H-\frac{1}{2}} dW_s$ in the $H \rightarrow 0$ limit. Using Lévy's continuity theorem for tempered distributions, we show that Z^H tends weakly to an almost log-correlated Gaussian field Z as $H \rightarrow 0$, which is a random tempered distribution, i.e. a random element of the dual of the Schwartz space \mathcal{S} . From Theorem A in [JSW19], we know this field differs from a standard Bacry-Muzy field by a Hölder continuous Gaussian process, and we show that $\xi_\gamma^H(dt) = e^{\gamma Z_t^H - \frac{1}{2}\gamma^2 \text{Var}(Z_t^H)} dt$ tends to a Gaussian multiplicative chaos (GMC) random measure ξ_γ for $\gamma \in (0, 1)$ as $H \searrow 0$. Unlike standard constructions of GMC, our approximating sequence Z_t^H is not a martingale so we cannot appeal to the martingale convergence theorem. We later address the more difficult “ L^1 -regime” where $\gamma \in [1, \sqrt{2})$ using standard tightness/weak convergence arguments and comparing ξ_γ^H to a sequence of GMCs ξ_φ^H constructed in using a Gaussian white noise integrated over curved regions in the upper half plane under the Haar measure. A stronger L^1 convergence is established by appealing to the framework of [Sha16] where the existence of a GMC (suitably defined) is shown to be equivalent to the existence of a *Randomised Shift*.

These results have a natural application to the popular Rough Bergomi stochastic volatility model, since ξ_γ^H is the quadratic variation of the log stock price for this model and values of H as low as .03 have been reported in empirical studies of this model (see e.g. [FTW19]). Using the Riemann-Liouville GMC and Jacod's stable convergence theorem, we then prove the surprising result that the martingale component X_t of the log stock price for the Rough Bergomi model tends weakly to $B_{\xi_\gamma([0, t])}$ as $H \rightarrow 0$ where B is a Brownian

¹see Lemma 3 in [BM03] to see why the critical q value is q^*

motion independent of everything else, which means the smile for the rBergomi model with $\rho \leq 0$ is symmetric in the $H \rightarrow 0$ limit for $\gamma \in (0, \sqrt{2})$. [Ger20] shows that $\mathbb{E}(X_t^3)$ decays exponentially fast or blows up exponentially fast depending on whether γ is less than or greater than a critical $\gamma \approx 1.61711$ which solves $\frac{1}{4} + \frac{1}{2} \log \gamma - \frac{3}{16} \gamma^2 = 0$. We also define a $H = 0$ model with non-zero skew for which X_t / \sqrt{t} tends weakly to a non-Gaussian random variable X_1 with non-zero skewness as $t \rightarrow 0$.

4.2 The Riemann-Liouville process and its GMC as $H \rightarrow 0$

We work on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with filtration $(\mathcal{F}_t)_{t \geq 0}$ throughout, which satisfies the usual conditions. In this section we consider a re-scaled Riemann-Liouville process in the limit as $H \rightarrow 0$; To this end, let $(W_t)_{t \geq 0}$ denote a standard Brownian motion and consider the following family of re-scaled Riemann-Liouville processes:

$$Z_t^H = \int_0^t (t-s)^{H-\frac{1}{2}} dW_s \quad (4.3)$$

for $H \in (0, \frac{1}{2})$, for which $R_H(s, t) := \mathbb{E}(Z_s^H Z_t^H) = \int_0^{s \wedge t} (s-u)^{H-\frac{1}{2}} (t-u)^{H-\frac{1}{2}} du$. The integrand here is dominated by

$$h(u, s, t) = ((s-u)^{-\frac{1}{2}} \vee 1) \cdot ((t-u)^{-\frac{1}{2}} \vee 1) \quad (4.4)$$

which is integrable for $s < t$, so using the dominated convergence theorem, we find that

$$R_H(s, t) \rightarrow R(s, t) := \int_0^{s \wedge t} (s-u)^{-\frac{1}{2}} (t-u)^{-\frac{1}{2}} du$$

for $s \neq t$ as $H \rightarrow 0$ and $R_H(s, t) \rightarrow \infty$ for $s = t > 0$. We note also that $R(0, 0) = \lim_{n \rightarrow \infty} \int_0^0 nds = 0$ (from the definition of Lebesgue integration) and we also note that $R_H(0, 0) = 0$ so $\lim_{H \rightarrow 0} R_H(0, 0) = R(0, 0) = 0$. We can evaluate this integral to obtain

$$\begin{aligned} R(s, t) &:= 2 \tanh^{-1} \left(\frac{\sqrt{s}}{\sqrt{t}} \right) = \log \frac{1 + \frac{\sqrt{s}}{\sqrt{t}}}{1 - \frac{\sqrt{s}}{\sqrt{t}}} = \log \frac{\sqrt{t} + \sqrt{s}}{\sqrt{t} - \sqrt{s}} = \log \frac{(\sqrt{t} + \sqrt{s})^2}{t - s} \\ &= \log \frac{1}{t - s} + g(s, t) \end{aligned} \quad (4.5)$$

for $0 < s < t$, where

$$g(s, t) = 2 \log(\sqrt{s} + \sqrt{t}) \quad (4.6)$$

and note that $R(s, t) \geq 0$ for all $s, t \geq 0$.

$$\int_{[0, T]^2} R_H(s, t) ds dt \leq 2 \int_{[0, T]^2} \int_0^t ((s-u)^{-\frac{1}{2}} \vee 1) \cdot ((t-u)^{-\frac{1}{2}} \vee 1) du ds dt < \infty$$

so from the dominated convergence theorem, we have

$$\lim_{H \rightarrow 0} \int_{[0, T]^2} \phi_1(s) \phi_2(t) R_H(s, t) ds dt = \int_{[0, T]^2} \phi_1(s) \phi_2(t) R(s, t) ds dt \quad (4.7)$$

for any $\phi_1, \phi_2 \in \mathcal{S}$, where \mathcal{S} denotes the Schwartz space. Similarly, for any sequence $\phi_k \in \mathcal{S}$ with $\|\phi_k\|_{m, j} \rightarrow 0$ for all $m, j \in \mathbb{N}_0^n$ for any $n \in \mathbb{N}$ (i.e. under the Schwartz space semi-norm defined in Eq 1 in e.g. [BDW17])

$$\lim_{k \rightarrow \infty} \int_{[0, T]^2} \phi_k(s) \phi_k(t) R(s, t) ds dt = 0 \quad (4.8)$$

since $\mu(A) = \int_A R(s, t) ds dt$ is a bounded non-negative measure (since $\int_0^T \int_0^t R(s, t) ds dt = \int_0^T 2t dt = T^2 < \infty$), and the convergence here implies in particular that ϕ_k tends to zero pointwise, so we can use the bounded convergence theorem. Thus if we define

$$\begin{aligned} \mathcal{L}_{Z^H}(f) &:= \mathbb{E}(e^{i(f, Z^H)}) = e^{-\frac{1}{2} \int_{[0, T]^2} f(s) f(t) R_H(s, t) ds dt} \\ \mathcal{L}(f) &:= e^{-\frac{1}{2} \int_{[0, T]^2} f(s) f(t) R(s, t) ds dt} \end{aligned}$$

for $f \in \mathcal{S}$, and note at the moment that we do not have a process or field as a subscript in $\mathcal{L}(f)$ since we have not yet shown that this is the characteristic functional of a random field. Then from (4.7) and (4.8) and Lévy's continuity theorem for generalized random fields in the space of tempered distributions (see Theorem 2.3 and Corollary 2.4 in [BDW17]), we see that $\mathcal{L}_{Z^H}(f)$ tends to $\mathcal{L}_Z(f)$ pointwise and $\mathcal{L}(\cdot)$ is continuous at zero, then there exists a generalized random field Z (i.e. a random *tempered distribution*) such that $\mathcal{L}_Z = \mathcal{L}$ and Z^H tends to Z in distribution with respect to the strong and weak topology (see page 2 in [BDW17] for definition). Based on the right hand side of (4.5), we can say that Z is an *almost log-correlated Gaussian field* (LGF).

Remark 4.2.1 Since $g(s, t)$ (the correction to the logarithmic term in the covariance) is smooth away from $(0, 0)$, from Theorem A in [JSW19], we know that Z differs from the standard Bacry-Muzy field on $(0, T]$ with covariance $\log \frac{1}{|t-s|}$ by some Gaussian process G_t which is a.s. Hölder continuous on $(0, T]$.

4.2.1 Constructing a Gaussian multiplicative chaos from Z^H as $H \rightarrow 0$

We now define the family of random measures : $\xi_\gamma^H(dt) := e^{\gamma Z_t^H - \frac{1}{2} \gamma^2 \text{Var}(Z_t^H)} dt$.

Theorem 4.2.1 *Let $H_n \searrow 0$. Then for any $A \in \mathcal{B}([0, T])$ and $\gamma \in (0, 1)$, $\xi_\gamma^{H_n}(A)$ tends to some non-negative random variable $\xi_{\gamma, A}$ in L^2 (and hence also converges in probability), $\xi_\gamma([0, T])$ is a non-trivial random variable (i.e. has finite non-zero variance), and there exists a random measure ξ_γ on $[0, T]$ such that $\xi_\gamma(A) = \xi_{\gamma, A}$ a.s. for all $A \in \mathcal{B}([0, T])$. ξ_γ is the GMC associated with the family of process Z^H as $H \rightarrow 0$.*

Proof. We wish to show that $\mathbb{E}((\xi_\gamma^{H_n}[0, T] - \xi_\gamma^{H_m}[0, T]))^2 \rightarrow 0$, i.e. that $\xi_\gamma^{H_n}[0, T]$ is a Cauchy sequence in L^2 . To this end, we first note that

$$\begin{aligned} \mathbb{E}(\xi_\gamma^{H_n}([0, T])\xi_\gamma^{H_m}([0, T])) &= \mathbb{E}\left(\int_{[0, T]^2} e^{\gamma^2(Z_t^{H_n} + Z_s^{H_m}) - \frac{1}{2}\gamma^2\mathbb{E}((Z_t^{H_n})^2) - \frac{1}{2}\gamma^2\mathbb{E}((Z_s^{H_m})^2)} ds dt\right) \\ &= \int_{[0, T]^2} \mathbb{E}\left(e^{\gamma^2(Z_t^{H_n} + Z_s^{H_m}) - \frac{1}{2}\gamma^2\mathbb{E}((Z_t^{H_n})^2) - \frac{1}{2}\gamma^2\mathbb{E}((Z_s^{H_m})^2)}\right) ds dt \\ &= \int_{[0, T]^2} e^{\frac{1}{2}\gamma^2 R_{H_n}(t, t) + \frac{1}{2}\gamma^2 R_{H_m}(s, s) + \gamma^2\mathbb{E}(Z_t^{H_n} Z_s^{H_m}) - \frac{1}{2}\gamma^2 R_{H_n}(t, t) - \frac{1}{2}\gamma^2 R_{H_m}(s, s)} ds dt \\ &= \int_{[0, T]^2} e^{\gamma^2\mathbb{E}(Z_t^{H_n} Z_s^{H_m})} ds dt. \end{aligned}$$

The integrand here is bounded by $e^{\gamma^2 \int_0^{s \wedge t} h(u, s, t) du}$ (where $h(u, s, t)$ is defined in (4.4)) and is integrable on $[0, T]^2$, and $\mathbb{E}(Z_t^{H_n} Z_s^{H_m}) = \int_0^s (t-u)^{H_n - \frac{1}{2}} (s-u)^{H_m - \frac{1}{2}} du \rightarrow R(s, t)$ Lebesgue a.e. on $[0, T]^2$ as $n, m \rightarrow \infty$, so from the dominated convergence theorem we see that

$$\begin{aligned} \mathbb{E}(\xi_\gamma^{H_n}([0, T])\xi_\gamma^{H_m}([0, T])) &\rightarrow \int_{[0, T]^2} e^{\gamma^2 R(s, t)} ds dt \quad (n, m \rightarrow \infty) \\ &= 2 \int_{[0, T]} \int_{[0, t]} e^{\gamma^2 R(s, t)} ds dt \\ &= 2 \int_{[0, T]} \int_{[0, t]} \left(\frac{\sqrt{t} + \sqrt{s}}{\sqrt{t} - \sqrt{s}}\right) \gamma^2 ds dt \\ &= 2 \int_{[0, T]} t \int_{[0, 1]} \left(\frac{\sqrt{t} + \sqrt{tu}}{\sqrt{t} - \sqrt{tu}}\right) \gamma^2 du dt \\ &= 2 \int_{[0, T]} t \int_{[0, 1]} \left(\frac{1 + \sqrt{u}}{1 - \sqrt{u}}\right) \gamma^2 du dt = 2 \int_0^T t a_\gamma dt \\ &= a_\gamma T^2 < \infty \end{aligned} \tag{4.9}$$

for $\gamma \in (0, 1)$, where

$$a_\gamma := \int_{[0, 1]} \left(\frac{1 + \sqrt{u}}{1 - \sqrt{u}}\right) \gamma^2 du = \frac{2 \cdot {}_2F_1(2, -\gamma^2, 3 - \gamma^2, -1)}{(1 - \gamma)(1 + \gamma)(2 - \gamma^2)} \tag{4.10}$$

where ${}_2F_1(z)$ is the hypergeometric function, and using that $1 - \sqrt{u} \sim \frac{1}{2}(1 - u)$ as $u \rightarrow 1$, we can easily verify that $a_\gamma \rightarrow \infty$ as $\gamma \uparrow 1$. Hence

$$\begin{aligned} & \mathbb{E}((\xi_\gamma^{H_n}([0, T]) - \xi_\gamma^{H_m}([0, T]))^2) \\ = & \mathbb{E}(\xi_\gamma^{H_n}([0, T])^2) - 2\mathbb{E}(\xi_\gamma^{H_n}([0, T])\xi_\gamma^{H_m}([0, T])) + \mathbb{E}(\xi_\gamma^{H_m}([0, T])^2) \rightarrow 0 \end{aligned}$$

so $\xi_\gamma^{H_n}([0, T])$ converges in $L^2(\Omega, \mathcal{F}, \mathbb{P})$ to some a.s. non-negative random variable $\xi_{\gamma, [0, T]}$, and hence also converges in probability. Similarly, for any $A \in \mathcal{B}([0, T])$, we can trivially modify the argument above to show that

$$\mathbb{E}(\xi_\gamma^{H_n}(A)\xi_\gamma^{H_m}(A)) \rightarrow \int_A \int_A e^{\gamma^2 R(s,t)} ds dt \leq a_\gamma T^2 < \infty$$

so $\xi_\gamma^H(A)$ tends to some random variable $\xi_{\gamma, A}$ in L^2 , and hence in probability.

We also know that $\mathbb{E}(\xi_\gamma^{H_n}([0, T])) = T$ for all n and we have already established L^2 -convergence for $\xi_\gamma^{H_n}(A)$ as $n \rightarrow \infty$ which implies L^1 convergence, so (by Scheffe's lemma) $\mathbb{E}(\xi_{\gamma, [0, T]}) = T$, which further implies that $\mathbb{P}(\xi_{\gamma, [0, T]} > 0) > 0$ and (from the reverse triangle inequality)

$$|\mathbb{E}(\xi_{\gamma, [0, T]}^2)^{\frac{1}{2}} - \mathbb{E}((\xi_\gamma^{H_n}([0, T]))^2)^{\frac{1}{2}}| \leq \mathbb{E}((\xi_\gamma^{H_n}([0, T]) - \xi_\gamma^H([0, T]))^2) \rightarrow 0$$

so

$$\mathbb{E}(\xi_{\gamma, [0, T]}^2) = \lim_{H \rightarrow 0} \mathbb{E}((\xi_\gamma^H([0, T]))^2) = a_\gamma T^2$$

so in particular ξ_γ is not multifractal at zero, since the power is 2 here and not $\zeta(2)$. The L^2 -convergence also means that $\xi_\gamma^H[0, T] \rightarrow \xi_{\gamma, [0, T]}$ in L^q as $H \rightarrow 0$ for all $q \in [1, 2]$ which (again from the reverse triangle inequality) implies that

$$\lim_{H \rightarrow 0} \mathbb{E}(\xi_\gamma^H([0, T])^q) = \mathbb{E}(\xi_{\gamma, [0, T]}^q). \quad (4.11)$$

Given that $\mathbb{E}(\xi_{\gamma, [0, T]}) = T$ and $\text{Var}(\xi_{\gamma, [0, T]}) = \int_{[0, T]^2} e^{\gamma^2 R(s,t)} ds dt - T^2 > 0$ since $a_\gamma > 1$ for $\gamma \in (0, 1)$, we see that $\xi_{\gamma, [0, T]}$ is a non-trivial random variable.

For $A, B \in \mathcal{B}([0, T])$ disjoint, $\xi_{\gamma, A \cup B}^H = \xi_{\gamma, A}^H + \xi_{\gamma, B}^H$ a.s. since ξ_γ^H is a measure, and we know that both sides tend to $\xi_{\gamma, A \cup B}$ and $\xi_{\gamma, A} + \xi_{\gamma, B}$ in probability. But by a standard result, if $X_n \xrightarrow{P} X$ and $X_n \xrightarrow{P} Y$, then $X = Y$ a.s., hence

$$\xi_{\gamma, A \cup B} = \xi_{\gamma, A} + \xi_{\gamma, B} \quad (4.12)$$

a.s.

Similarly for any sequence $A_n \downarrow \emptyset$ with $A_n \in \mathcal{B}([0, T])$, $\mathbb{E}(\xi_{\gamma, A_n}) = \text{Leb}(A_n)$, so by Markov's inequality $\mathbb{P}(\xi_{\gamma}(A_n) > \delta) \leq \frac{\text{Leb}(A_n)}{\delta}$, so $\xi_{\gamma}(A_n)$ tends to zero in probability, and from (4.12), we know that $\xi_{\gamma}(A_n)$ is decreasing, and hence also tends to some random variable Y a.s. (and hence also in probability). Thus by the same standard result discussed above, $Y = 0$ a.s. Thus by Theorem 9.1.XV in [DV07] (see also the end of Section 4 on page 18 in [RV10]), there exists a random measure ξ_{γ} on $[0, T]$ such that $\xi_{\gamma}(A) = \xi_{\gamma, A}$ a.s. for all $A \in \mathcal{B}([0, T])$. ■

Remark 4.2.2 If we replace the definition of Z^H with the usual Riemann-Liouville process $Z_t^H = \sqrt{2H} \int_0^t (t-s)^{H-\frac{1}{2}} dW_s$, then adapting the arguments above, we see that

$$\mathbb{E}\left(\left(\int_A e^{\gamma^2 Z_t^H - \frac{1}{2}\gamma^2 \text{Var}(Z_t^H)} dt\right)^2\right) \rightarrow \text{Leb}(A)^2$$

as $H \rightarrow 0$, for all $A \in \mathcal{B}([0, T])$. But we know that the first moment of $\int_A e^{\gamma^2 Z_t^H - \frac{1}{2}\gamma^2 \text{Var}(Z_t^H)} dt$ is $\text{Leb}(A)$ as well, hence $\int_A e^{\gamma^2 Z_t^H - \frac{1}{2}\gamma^2 \text{Var}(Z_t^H)} dt \rightarrow \text{Leb}(A)$ in L^2 .

Remark 4.2.3 For $c \in (0, 1]$, $(W_c, \xi_{\gamma}([0, c]) \sim (\sqrt{c}W_1, c\xi_{\gamma}[0, 1])$, so in particular, $\xi_{\gamma}([0, (\cdot)])$ is a self-similar process, and we can easily verify $\xi_{\gamma}([0, c])$ is monofractal at zero, i.e. $\mathbb{E}(\xi_{\gamma}([0, c])^q) = c^q \mathbb{E}(\xi_{\gamma}([0, 1])^q)$.

4.2.2 Construction and properties of the usual Bacry-Muzy multifractal random measure (MRM) via Gaussian white noise on triangles

In this subsection we briefly describe the family of (stationary) Gaussian process used in [BM03]; the Bacry-Muzy multifractal random measure (MRM) is then the GMC associated with this family of processes as the l parameter tends to zero. Define $\omega_l(t)$ as in Eq 7 in [BBM13] with $\lambda = 1$ and $T = 1$, and set $\bar{\omega}_l(t) := \omega_l(t) - \mathbb{E}(\omega_l(t))$, so $\bar{\omega}_l(t) = \int_{(u,s) \in \mathcal{A}_l(t)} dW(u,s)$ where (in this subsection alone) $dW(u,s)$ is a two-dimensional Gaussian white noise with variance $s^{-2}duds$, and $\mathcal{A}_l(t) = \{(u,s) : |u-t| \leq (\frac{1}{2}s) \wedge T, s \geq l\}$ is the cone-like region defined in Eq 11 in [BM03] (for the special case when $f(l) = f^{(e)}(t)$ in their notation, see Eqs 12 and 15 in [BM03]). Then

$$K_l^T(s,t) := \mathbb{E}(\bar{\omega}_l(t)\bar{\omega}_l(s)) = \begin{cases} \log \frac{T}{\tau} & l \leq \tau \leq T \\ \log \frac{T}{l} + 1 - \frac{\tau}{l} & \tau \leq l \\ 0 & \tau > T \end{cases} \quad (4.13)$$

4.3 ξ_γ for the full sub-critical range $\gamma \in (0, \sqrt{2})$

where $\tau = |t - s|$, and one can easily verify that $K_l^T(s, t) \leq \log \frac{T}{\tau}$ (see Eq 25 in [BM03]). From a picture, we also see that $\mathbb{E}(\bar{\omega}_l(t) \bar{\omega}_{l'}(s)) = K_l(s, t)$ for $l > l'$ (i.e. the answer does not depend on l'), and $K_l^T(s, t) \nearrow \log \frac{T}{|t-s|}$ as $l \rightarrow 0$. We now define the measure

$$M_\gamma^{T,l}(dt) = e^{\gamma \bar{\omega}_l(t) - \frac{1}{2} \gamma^2 \text{Var}(\bar{\omega}_l(t))} dt \quad (4.14)$$

and we use $M_\gamma^l(dt)$ as shorthand for $M_\gamma^{1,l}(dt)$. One can easily verify that $M_\gamma^l(A)$ is a martingale with respect to the filtration $\mathcal{F}_l := \sigma(W(A, B) : A \subset \mathbb{R}^+, B \subseteq [l, \infty])$ (see e.g. subsection 5.1 in [BM03] and page 17 in [RV10]) and $\sup_l \mathbb{E}(M_\gamma^l(A)^q) < \infty$ (Lemma 3 i) in [BM03]), so from the martingale convergence theorem, $M_\gamma^{T,l}(A)$ converges to $M_\gamma^T(A)$ in L^q for $q \in (1, q^*)$, and from the reverse triangle inequality this implies that

$$\lim_{l \rightarrow 0} \mathbb{E}((M_\gamma^{T,l}(A))^q) = \mathbb{E}((M_\gamma^T(A))^q) \quad (4.15)$$

and M^T is perfectly multifractal, i.e. $\mathbb{E}(|M_\gamma^T([0, t])|^q) = c_{q,T} t^{\zeta(q)}$ (see e.g. Lemma 4 in [BM03]) for some finite constant $c_{q,T} > 0$, depending only on q and T . For integer $q \geq 1$, we also note that

$$\begin{aligned} \mathbb{E}(M_\gamma^T(A)^q) &= \int_A \dots \int_A e^{\gamma^2 \sum_{1 \leq i < j \leq q} \log \frac{T}{|u_i - u_j|}} du_1 \dots du_q \\ &= \int_A \dots \int_A e^{\gamma^2 q(q-1) \log T + \sum_{1 \leq i < j \leq q} \log \frac{1}{|u_i - u_j|}} du_1 \dots du_q \\ &= T^{\gamma^2 q(q-1)} \mathbb{E}(M_\gamma(A)^q) \end{aligned} \quad (4.16)$$

so we see that

$$c_{q,T} = c_q T^{\gamma^2 q(q-1)} \quad (4.17)$$

where $c_q = c_{q,1}$, and this also holds for non-integer q (see e.g. Theorem 3.16 in [Koz06]).

4.3 ξ_γ for the full sub-critical range $\gamma \in (0, \sqrt{2})$

4.3.1 The Sandwich lemma

We now look to extend the definition of ξ_γ to $\gamma \in (0, \sqrt{2})$. We will use the following standard result:

Theorem 4.3.1 (Kahane's Inequality) (see e.g. Appendix of [RV10]). *Let I be a bounded subinterval of \mathbb{R} and $(X(u))_{u \in I}$, $(Y(u))_{u \in I}$ be two centred continuous Gaussian processes with $\mathbb{E}[X(u)X(u')] \leq \mathbb{E}[Y(u)Y(u')]$ for all u, u' . Then, for all convex functions $F : \mathbb{R} \rightarrow \mathbb{R}$,*

we have:

$$\mathbb{E}[F(\int_I e^{X(u) - \frac{1}{2}\mathbb{E}(X(u)^2)} du)] \leq \mathbb{E}[F(\int_I e^{Y(u) - \frac{1}{2}\mathbb{E}(Y(u)^2)} du)].$$

Lemma 4.3.2 (The Sandwich lemma). Fix any τ and δ such that $0 < \tau < \tau + \delta < 1$. Then for $\tau \leq s \leq t \leq t + \delta$ and $H > 0$ sufficiently small, we can sandwich $R_H(s, t)$ as follows:

$$K_{l_*(H, \tau)}^{4\tau}(k) \leq R_H(s, t) \leq K_{l^*(H)}^4(k) \quad (4.18)$$

for $k = |t - s| < \delta$ for $0 < s < t < 1$, where $l_*(H, \tau) = \frac{1}{F_H(k^*)} > 0$ and $l^*(H) := 4e^{-\frac{1}{2H}} > 0$ (which both tend to zero as $H \rightarrow 0$), and $F_H(k) := R_H(\tau, \tau + k)$. Note the upper bound trivially holds for $s = 0$ as well, since $R_H(0, k) = 0$ and $K_{l^*}^T(k) \geq 0$. We also remind the reader that if $0 = s < t$, $R(s, t) = 0$ not $\log \frac{1}{t-0} + g(0, t) = \infty$.

Remark 4.3.1 The lower bound of the Sandwich lemma will only be used to prove the local multifractality of ξ_γ , and is not needed for everything else in this chapter.

Proof. We define $G_H(k) := R_H(\tau + \delta - k, \tau + \delta)$, and at this point we refer the reader to Appendix A for some basic properties of $G_H(k)$. Then choosing $l^* = l^*(H)$ such that $G_H(0) = \frac{(\tau + \delta)^{2H}}{2H} \leq \frac{1}{2H} = \log(\frac{4}{l^*})$, we see that

$$l^*(H) = 4e^{-\frac{1}{2H}} \downarrow 0 \quad \text{as} \quad H \rightarrow 0.$$

(A.1) implies that $G_H(k) \leq \log \frac{4}{k}$, and for $k \in [l^*, 4]$, $K_{l^*}^4(k) = \log \frac{4}{k}$, so in this case $G_H(k) \leq K_{l^*}^4(k)$. For $k \in (0, l^*)$, $K_{l^*}^4(k) = \log(\frac{4}{l^*}) + 1 - \frac{k}{l^*} > \log \frac{4}{l^*} \geq G_H(0) > G_H(k)$. Hence for both cases, we have the following upper bound:

$$G_H(k) = R_H(\tau + \delta - k, \tau + \delta) \leq K_{l^*(H)}^4(k).$$

From Appendix A, we recall that

$$R_H(s, k+s) = \int_0^s (u(k+u))^{H-\frac{1}{2}} du$$

and if we restrict attention to $A_\delta := \{(s, t) : t - s = k \text{ and } (s, t) \in [\tau, \tau + \delta]^2\}$ for $0 < \tau < \tau + \delta < 1$ with $k \in [0, \delta]$, then from Appendix A we know that $R_H(s, t)$ is maximized at $s = \tau + \delta - k$ taking the value $G_H(k)$ and minimized at $s = \tau$ with value $F_H(k)$ (see Figure 4.2). Thus

$$R_H(s, t) \leq G_H(k) \leq K_{l^*(H)}^4(k) \quad (4.19)$$

for $(s, t) \in [\tau, \tau + \delta]^2$ where $k = |t - s|$.

From the second part of Appendix A, we know that $F_0(k) := \log \frac{1}{k} + 2\log(\sqrt{\tau} + \sqrt{\tau + k}) > \log \frac{4\tau}{k}$ but we also know that $F_H(k) \uparrow F_0(k)$ uniformly on compact intervals away from zero (Dini's theorem), and for $H > 0$ $F_H(0) < \infty$ and $\log(\frac{4\tau}{k}) \rightarrow \infty$ as $k \rightarrow 0$, so from the aforementioned uniform convergence, we see that for $H > 0$ sufficiently small there exists a $k^* = k^*(H, \tau) > 0$ such that

$$F_H(k^*) = \log \frac{4\tau}{k^*} \quad (4.20)$$

(see top-right plot in Figure 4.2) with

$$F_H(k) \geq \log \frac{4\tau}{k} \quad \text{for } k \in [k^*, 4\tau] \quad , \quad F_H(k) \leq \log \frac{4\tau}{k} \quad \text{for } k \leq k^* . \quad (4.21)$$

Now set $l_* = l_*(H, \tau)$ such that $|F'_H(k^*)| = \frac{1}{l_*} \cdot l_* \in [\tau, \tau + \delta]$ for H sufficiently small, and $l_* \geq k^*$ since

$$\frac{1}{k^*} = \left| \frac{d}{dk} \log \frac{4\tau}{k} \right|_{k=k^*} > |F'_H(k^*)| \quad (4.22)$$

(see Figure 4.2 top right-plot). We now note the following:

- In the region $[k^*, l_*]$, $F_H(k) > \log(4\tau/k)$ so $F_H(k) > \log(4\tau/l_*) + 1 - k/l_*$ (since the latter is just the tangent line to $\log(4\tau/k)$ at $k = l_*$), see Figure 4.2 top right plot.
- At $k = k^*$, F_H is greater than said tangent and by construction has the same gradient as the tangent, i.e. $\frac{1}{l_*}$. Then as k decreases to zero, the gradient of F_H increases in absolute value (due to the convexity of F_H) so F_H is greater than the tangent line.

Thus $K_{l_*}^{4\tau}(k) = \log \frac{4\tau}{l_*} + 1 - \frac{k}{l_*} < F_H(k)$ for $k \in (0, l_*)$. We also see that $l_* \downarrow 0$ as $H \downarrow 0$, since $k^* \rightarrow 0$ as $H \rightarrow 0$. Thus, to sum up, we have shown that

$$G_H(k) = R_H(\tau + \delta - k, \tau + \delta) \leq K_{l_*(H)}^4(k)$$

and

$$K_{l_*(H, \tau)}^{4\tau}(k) \leq F_H(k) = R_H(\tau, \tau + k)$$

for $k \in [0, 4\tau]$. From Appendix A, we recall that $R_H(s, k+s) = \int_0^s (u(k+u))^{H-\frac{1}{2}} du$ and that if we restrict attention to A_δ for $0 < \tau < \tau + \delta < 1$ with $k \in [0, \delta]$, then $R_H(s, t)$ is minimized at $s = \tau$ with value $F_H(k)$. Thus

$$K_{l_*(H, \tau)}^{4\tau}(k) \leq F_H(k) \leq R_H(s, t) \leq G_H(k) \leq K_{l_*(H)}^4(k) \quad (4.23)$$

for $(s, t) \in [\tau, \tau + \delta]^2$ where $k = |t - s|$. ■

4.3.2 Existence of a limiting law for ξ_γ for $\gamma \in (0, \sqrt{2})$

Let P be an independently scattered infinitely divisible random measure (see [BM03] for details) with

$$\mathbb{E}(e^{iqP(A)}) = e^{\varphi(q)\mu(A)}$$

for $q \in \mathbb{R}$ where $\mu(du, dw) = \frac{1}{w^2}dwdu$ denotes the *Haar measure*. Here we restrict attention to the special case where $\varphi(q) = \frac{1}{2}\gamma^2q^2$, in which case $P(du, dw)$ is just γ times a Gaussian white noise with variance $\frac{1}{w^2}dudw$ (similar to Section 4.2.2). Let $A_t^H := \{0 \leq u \leq t, w \geq g_H(u, t)\}$ for a family of functions which satisfy the following condition: $g_H(\cdot, t) \geq 0$ with $g_H(u, t)$ increasing in t and H . We now define the process $\omega_t^H = P(A_t^H)$ for $t \geq 0$ with filtration

$$\mathcal{F}_H := \sigma(P(A \times B) : B \subseteq [H, \infty], A, B \in \mathcal{B}(\mathbb{R})) \quad (4.24)$$

(compare to a similar filtration on page 17 in [RV10]), and ω_t^H is a Gaussian process since $\varphi(q)$ is the characteristic function of a Gaussian with covariance

$$\mathbb{E}(\omega_s^H \omega_t^H) = \int_0^s \int_{g_H(u, t)}^\infty \frac{1}{w^2}dwdu = \int_0^s \frac{1}{g_H(u, t)}du$$

for $0 \leq s \leq t$, and differentiating with respect to s , we see that if g satisfies $\frac{1}{g_H(s, t)} = R_s^H(s, t)$ then (for H fixed) the Gaussian process ω^H has the same covariance as our process Z^H , and the explicit formula for g_H is given as

$$g_H(s, t) = \frac{1}{\gamma \Gamma(\frac{1}{2} + H) (t(1 + 2H) {}_2F_1(1, \frac{1}{2} - H, \frac{3}{2} + H, \frac{s}{t}) + s(1 - 2H) {}_2F_1(2, \frac{3}{2} - H, \frac{5}{2} + H, \frac{s}{t}))} 2s^{\frac{1}{2}-H} t^{\frac{3}{2}-H}$$

where ${}_2F_1(a, b, c, z)$ is the regularized hypergeometric function² (and in Appendix B we verify that Condition 1 above is satisfied). For $H = 0$ we have $g_0(s, t) = \frac{\sqrt{s(t-s)}}{\sqrt{t}}$. For $H_2 < H_1$, $\omega_t^{H_2} - \omega_t^{H_1} = P(A_t^{H_2} \setminus A_t^{H_1})$ and $\omega_t^H = P(A_t^H)$ are independent for any $H \geq H_1$, so ω_t^H is an \mathcal{F}_H -martingale (see (4.24) for definition of \mathcal{F}_H , and we refer to this as a backward martingale since the martingale evolves as H goes smaller not larger and we start the martingale at some $H > 0$), and from this one can easily verify that $\xi_\varphi^H(I)$ is also an \mathcal{F}_H -backward martingale for any Borel set I .

²we are using Mathematica's definition here

4.3 ξ_γ for the full sub-critical range $\gamma \in (0, \sqrt{2})$

Theorem 4.3.3 *Let ξ_ϕ^H denote the GMC of $\gamma\omega^H$ on $[0, 1]$. Then for any $q \in (1, q^*)$ and any interval $I \subseteq [0, 1]$, $\xi_\phi^H(I)$ tends to some non-negative random variable $\xi_{\phi, I}$ as $H \rightarrow 0$ a.s. and in L^q , and $\mathbb{E}(\xi_\phi^H(I)^q) \rightarrow \mathbb{E}(\xi_{\phi, I}^q)$.*

Proof. From the upper bound in the Sandwich Lemma $R_H(s, t) \leq K_{l^*(H)}^\theta(s, t)$ for $0 < s < t < 1$, where $\theta = 4 \cdot \sup(I)$ and $K_l^T(s, t)$ is the covariance of the model in [BM03], and $l^*(H) \downarrow 0$ as $H \downarrow 0$. Then from Kahane's inequality we have that

$$\mathbb{E}(\xi_\phi^H(I)^q) \leq \mathbb{E}(M_{l^*(H)}^\theta(I)^q) \quad (4.25)$$

where M_l^T is defined as in Section 4.2.2. Moreover, from Lemma 3 in [BM03] we know that $\sup_{l>0} \mathbb{E}(M_l^\theta(I)^q) < \infty$ for $q \in [1, q^*)$, so we have the uniform bound $\sup_{H>0} \mathbb{E}(\xi_\phi^H(I)^q) < \infty$.

From above we know that $\xi_\phi^H(I)$ is a \mathcal{F}^H -backwards martingale. Then (by Doob's martingale convergence theorem for continuous martingales) $\xi_\phi^H(I)$ tends to some random variable (which we call $\xi_{\phi, I}$) as $H \rightarrow 0$ a.s. and in L^q for $q \in [1, q^*)$. Moreover, from the reverse triangle inequality, the aforementioned L^q -convergence implies that

$$\mathbb{E}((\xi_\phi^H(I))^q) \rightarrow \mathbb{E}(\xi_{\phi, I}^q) \quad (4.26)$$

as $H \rightarrow 0$, for $q \in [1, q^*)$. ■

Theorem 4.3.4 *The laws of $\xi_\gamma^H([0, \cdot])$ on $C_0([0, 1])$ converge weakly as $H \rightarrow 0$ to the law of a non-decreasing process on $C_0([0, 1])$ which induces a non-atomic measure ξ_γ on $[0, T]$ with $\mathbb{E}(\xi_\gamma(A)) = \text{Leb}(A)$.*

Remark 4.3.2 In next section, we give a stronger result involving L^1 -convergence using Theorem 25 in [Sha16]) via generalized randomized shifts

Proof. Note that although $\mathbb{E}(\omega_s^H \omega_t^H) = \mathbb{E}(Z_s^H Z_t^H)$ this does not imply that $\mathbb{E}(\omega_s^H \omega_t^{H_2}) = \mathbb{E}(Z_s^H Z_t^{H_2})$ for $H \neq H_2$. However (crucially) ξ_ϕ^H (defined in Theorem 4.3.3) has the same law as our original ξ_γ^H measure for all $H > 0$, and the non-decreasing process $\xi_\phi^H([0, (\cdot)])$ and $\xi_\gamma^H([0, (\cdot)])$ have the same finite-dimensional distributions, so it suffices to prove weak convergence in law of the sequence $\xi_\phi^H([0, (\cdot)])$. Thus from the a.s. convergence in Theorem 4.3.3 and the bounded convergence theorem, we see that for n distinct time values $t_1, \dots, t_n \in [0, 1]$ and $u_1, \dots, u_n \in \mathbb{R}$

$$\lim_{H \rightarrow 0} \mathbb{E}(e^{\sum_{k=1}^n i u_k \xi_\phi^H([0, t_k])}) = \mathbb{E}(e^{\sum_{k=1}^n \xi_\gamma([0, t_k])}).$$

4.3 ξ_γ for the full sub-critical range $\gamma \in (0, \sqrt{2})$

So we have convergence of the finite-dimensional distributions of the process $\xi_\gamma^H([0, \cdot])$. Moreover, from the upper bound for the Sandwich lemma, for $0 < s < t < 1$ we have

$$\mathbb{E}(\xi_\gamma^H([s, t])^q) \leq \mathbb{E}((M_\gamma^{A, I^*(H)}([s, t]))^q) \nearrow \mathbb{E}((M_\gamma^A([s, t]))^q) = c_{q,4}|t-s|^\zeta(q).$$

Moreover, $\zeta(q) = 1 + (1 - \frac{1}{2}\gamma^2)(q-1) + O((q-1)^2)$, and hence $\zeta(q) > 1$ for $q > 1$ sufficiently small for $\gamma \in (0, \sqrt{2})$. Hence by Problem 2.4.11 in [KS91] (or Theorem 1.8 in chapter XIII in [RY99]) with $X_t^m := \xi_\gamma^H([0, t])$ and $H = 1/m$, the probability measures $\mathbb{Q}^H = \mathbb{P} \circ (X^m)^{-1}$ induced by the sequence of processes $\xi_\gamma^H([0, \cdot])$ on $C_0([0, 1])$ are tight under the usual sup norm topology. Thus by Proposition 2.4.15 in [KS91] (see also Theorem B.1.3 in [FH05] and page 1 in [BM16]), the sequence \mathbb{Q}^H converges weakly to a probability measure \mathbb{Q} on $C_0([0, 1])$. Moreover, since

$$\xi_\varphi^H([0, s]) \leq \xi_\varphi^H([0, t])$$

for $0 < s < t$, and we have a.s. convergence of both sides, so $\xi_\varphi([0, s]) \leq \xi_\varphi([0, t])$ and hence \mathbb{Q} is the law of a non-decreasing continuous process, which induces a measure on $[0, 1]$ which we call ξ_γ , with no atoms. We know that $\mathbb{E}(\xi_{\gamma, A}) = \text{Leb}(A)$, so $\mathbb{E}(\xi_\gamma(A)) = \text{Leb}(A)$. ■

4.3.3 Existence of the GMC measure for $\gamma \in (0, \sqrt{2})$ using the Shamov approximation theorem

In the previous section we have discussed the limiting law of the GMC associated to the $H = 0$. Whilst providing us with results on convergence in law this doesn't tell us anything about how the limiting GMC depends on the underlying probability space. More precisely, can we strengthen the aforementioned convergence in law to convergence in probability or even in L^1 ? To establish this we adopt the more abstract (but more powerful) framework of Shamov [Sha16].

For $H > 0$ we have the RL-GMC:

$$\xi_\gamma^H(A) = \int_A e^{\gamma Z_t^H - \frac{1}{2}\gamma^2 \text{Var}(Z_t^H)} dt$$

and let $\mathcal{H} := H_0^1 = \{f : \int_0^1 f'(t)^2 dt\}$ denote the Cameron-Martin space of W . Consider the following element of \mathcal{H} indexed by $t \in [0, 1]$:

$$Y^H(t)(s) := \gamma \left(\frac{t^{H+\frac{1}{2}} - (t-s)^{H+\frac{1}{2}}}{H+\frac{1}{2}} 1_{s < t} + \frac{t^{H+\frac{1}{2}}}{H+\frac{1}{2}} 1_{s \geq t} \right) \quad (4.27)$$

Being an element of \mathcal{H} we can write the following:

$$Z_t^H = \langle Y^H(t), W \rangle \quad (4.28)$$

$$\gamma^2 \text{Var}(Z_t^H) = \langle Y_{H(t)}, Y_{H(t)} \rangle \quad (4.29)$$

where the inner product is understood to be the standard inner product in \mathcal{H} . The first equation above should be understood as a Paley-Wiener integral (since W is a.s. not an element of \mathcal{H}).

We have thus written the process Z_t^H as a pair (Y^H, W) . This decomposition is a special case of the Maurey-Nikishin factorisation which can be applied to any Gaussian (see [Sha16]). Shamov defines the subcritical GMC associated to the pair (Y^H, W) as a random measure ξ_γ^H satisfying:

1. $\mathbb{E}(\xi_\gamma^H(dt)) = dt$.
2. The measure ξ_γ^H is measurable with respect to W so we can write $\xi_\gamma^H(dt) = \xi_\gamma^H(W, dt)$
3. For all $f \in \mathcal{H}$

$$\xi_\gamma^H(W + f, dt) = e^{\langle Y^H(t), f \rangle} \xi_\gamma^H(W, dt) \quad \text{a.s.}$$

In Theorem 17 of [Sha16], the existence of such a random measure is shown to be equivalent to the statement that $Y^H(t)$ is a so-called *Randomised Shift* i.e. that if we sample t independently from W using the Lebesgue measure then the distribution of $W + Y^H(t)$ is absolutely continuous with respect to that of W :

$$\text{Law}_{\mathbb{P} \otimes \text{Leb}}[W + Y(t)] \ll \text{Law}_{\mathbb{P}}[W].$$

In the case $H > 0$ then $Y^H(t) \in \mathcal{H}$ and this absolute continuity property is simply a consequence of the Cameron-Martin theorem (the existence of the GMC can easily be seen directly without this theorem since for $H > 0$ the covariances are non-singular).

At $H = 0$, $Y^H(t)$ is no longer in \mathcal{H} and so the standard Cameron-Martin theorem cannot be straightforwardly applied. In this case Y_H remains a linear, bounded map from \mathcal{H} into $L^0[0, 1]$ (in fact $L^2[0, 1]$) and so in the language of [Sha16] it is a generalised \mathcal{H} -valued function.

The question remains whether $Y^0(t)$ is a random shift and if so what is the relationship between ξ_γ and its approximating measures. This is answered (in general) by the Shamov[Sha16] Approximation Theorem (Theorem 25) which states that if we have a series of randomized shifts Y_n with associated GMCs denoted M_{Y_n} and kernels $K_{Y_n Y_n}(t, s) := \langle Y_n(t), Y_n(s) \rangle$ satisfying:

- The family of random variables $\{M_{Y_n}\}$ is uniformly integrable.
- There exists a generalized \mathcal{H} -valued function that is the limit of Y_n in the sense that

$$\forall f \in \mathcal{H} \quad : \quad \langle Y_n, f \rangle \xrightarrow{L^0(\text{Leb})} \langle Y, f \rangle$$

then Y is a randomized shift. If, furthermore

- The kernels K_{Y_n, Y_n} converge to $K_{Y, Y}$ in $L^0(\mu \otimes \mu)$ (i.e. convergence in measure).

then the sub-critical GMC M_Y (associated to Y with expectation μ) is the limit of M_{Y_n} , in the sense that:

$$\forall f \in L^1(\mu) \quad : \quad \int f(t) M_{Y_n}(W, dt) \xrightarrow{L^1} \int f(t) M_Y(W, dt).$$

We address each of these points in turn:

- **Uniform Integrability.** As in the proof of multifractality (see next section), for each $H > 0$ we can bound the covariance of the RL process by the covariance of an approximate Bacry-Muzy multifractal random walk (see section 4.2.2). By Kahane's inequality we can thus bound the p -th moment of our measure from above by the p -th moment of the Bacry-Muzy MRM which is shown in [BM03] to be uniformly bounded. Thus $\{M_{H_n}\}$ are uniformly integrable.
- **Convergence of the shifts.** The operator Y^H is (up to an unimportant factor $\Gamma(H + \frac{1}{2})^{-1}$) the RL fractional integral I^α of order $\alpha = H + \frac{1}{2}$. As is proved in Samko et al.[SKM93] (Theorem 2.6), the RL integrals form a semigroup in $L^p(0, 1)$ for $p \geq 1$, which is continuous in the uniform topology for all $\alpha > 0$ and strongly for all $\alpha \geq 0$, which in our context implies that for all $f \in \mathcal{H}$:

$$\lim_{H \rightarrow 0} \|\langle Y^H, f \rangle - \langle Y^0, f \rangle\|_{L^2} = \lim_{H \rightarrow 0} \|I^\alpha(f') - I^{\frac{1}{2}}(f')\|_{L^2} = 0.$$

Note that $f \in \mathcal{H}$ implies $f' \in L^2$, and convergence in L^2 implies convergence in measure.

- **Convergence of kernels.** These kernels are the same (up to a factor of γ) as the covariances $R_H(s, t)$ and $R(s, t)$. As discussed previously, away from the diagonal $\{s = t\}$, $R_H(s, t) \rightarrow R(s, t)$ pointwise and hence in measure.

Thus we have shown that

$$\int f(t) \xi_\gamma^H(dt) \xrightarrow{L^1} \int f(t) \xi_\gamma(dt)$$

for all $f \in L^1$.

Remark 4.3.3 From section 6.8 in [Sha16], we know that ξ_γ has no atoms.

4.3.4 Local multifractality

Proposition 4.3.5 For $\gamma \in (0, \sqrt{2})$, ξ_γ has the following locally multifractal behaviour away from zero:

$$\lim_{\delta \rightarrow 0} \frac{\log \mathbb{E}(\xi_\gamma([t, t + \delta])^q)}{\log \delta} = \zeta(q) \quad (4.30)$$

for $t \in (0, 1)$ and $q \in (0, q^*)$.

Proof. Applying Kahane's inequality and Sandwich Lemma for $q \in (1, q^*)$ we have

$$\mathbb{E}[(M_\gamma^{4\tau, l_*(H, \tau)}([\tau, \tau + \delta]))^q] \leq \mathbb{E}[(\xi_\gamma^H([\tau, \tau + \delta]))^q] \leq \mathbb{E}[(M_\gamma^{4, l^*(H)}([\tau, \tau + \delta]))^q] \quad (4.31)$$

where $M_\gamma^{T, l}$ is defined as in Section 4.2.2. Using the L^q convergence of $M_\gamma^{T, l}(A)$ in (4.15) and (4.26), we see that

$$\mathbb{E}[(M_\gamma^{4\tau}([\tau, \tau + \delta]))^q] \leq \mathbb{E}[(\xi_\gamma([\tau, \tau + \delta]))^q] \leq \mathbb{E}[(M_\gamma^4([\tau, \tau + \delta]))^q].$$

Then using the multifractality property of M_γ^T we see that:

$$c_{q, 4\tau} \delta^{\zeta(q)} = c_{q, 1} (4\tau)^{\gamma^2 q(q-1)} \delta^{\zeta(q)} \leq \mathbb{E}[(\xi_\gamma([\tau, \tau + \delta]))^q] \leq c_{q, 4} \delta^{\zeta(q)} = c_{q, 1} 4^{\gamma^2 q(q-1)} \delta^{\zeta(q)}$$

where we have used (4.17) in the final line. Taking the logarithm of the above inequality, dividing by $\log \delta$ and taking limits yields the local multifractality property for ξ_γ (recall that we are assuming that $\tau > 0$ here). ■

4.4 Application to the Rough Bergomi model - skew flattening/blowup as $H \rightarrow 0$

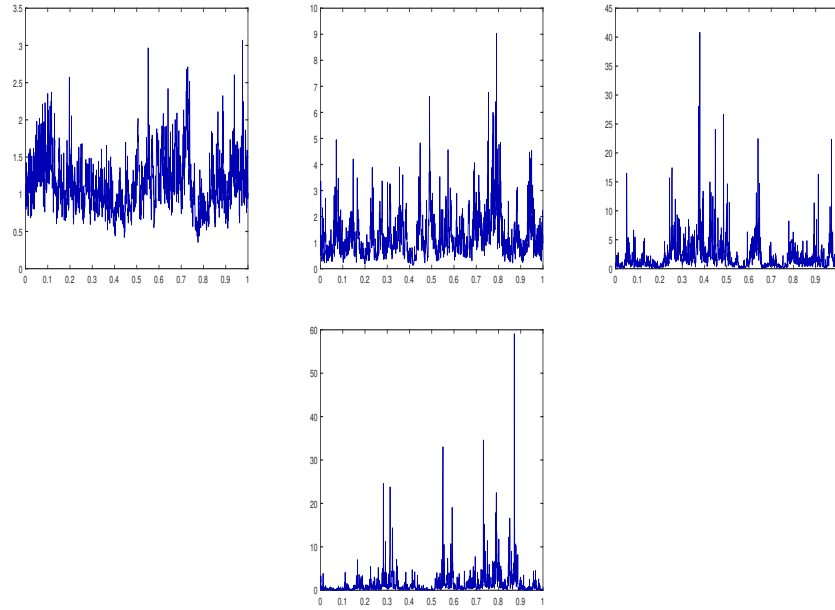


Fig. 4.1 Here we see simulations of ξ_γ using a spectral expansion for (from left to right) $\gamma = 0.125, 0.25, 0.375$ and 0.5 with $n = 1000$ eigenfunctions, 1000 time points, $H = 0$ and we have used Gauss-Legendre quadrature. For this range of γ -values, the first four raw sample moments are in very close agreement with the theoretical values for $H = 0$.

4.4 Application to the Rough Bergomi model - skew flattening/blowup as $H \rightarrow 0$

We consider the standard Rough Bergomi model for a stock price process X_t^H :

$$\begin{cases} dX_t^H = -\frac{1}{2}V_t^H dt + \sqrt{V_t^H} dW_t, \\ V_t^H = e^{\gamma Z_t^H - \frac{1}{2}\gamma^2 \text{Var}(Z_t^H)} \\ Z_t^H = \int_0^t (t-s)^{H-\frac{1}{2}} (\rho dW_s + \bar{\rho} dW_t^\perp) \end{cases} \quad (4.32)$$

where $\gamma \in (0, \sqrt{2})$, $|\rho| \leq 1$ and W, W^\perp are independent Brownian motions, and (without loss of generality) we set $\tilde{X}_0^H = 0$. We let $\tilde{X}_t^H = \int_0^t \sqrt{V_s^H} dW_s$ denote the martingale part of X^H .

Theorem 4.4.1 *For $\gamma \in (0, \sqrt{2})$, \tilde{X}^H tends to $B_{\xi_\gamma}^\perp$ stably (and hence weakly) in law on any finite interval $[0, T]$, where B^\perp is a Brownian motion independent of everything else.*

4.4 Application to the Rough Bergomi model - skew flattening/blowup as $H \rightarrow 0$

Corollary 4.4.2 *From the weak convergence of $\xi_\gamma^H([0, T])$ and the previous result we see that*

$$\begin{aligned} \lim_{H \rightarrow 0} \mathbb{E}(e^{ikX_t^H}) &= \mathbb{E}(e^{-\frac{1}{2}(ik+k^2)\xi_\gamma([0,t])}) \\ &= \mathbb{E}(e^{ik(-\frac{1}{2}\xi_\gamma([0,t]) + B_{\xi_\gamma([0,t])}^\perp)}) \end{aligned}$$

which (by a well known result in Renault&Touzi[RT96]) implies that implied volatility smile for the true Rough Bergomi model in (4.32) is symmetric in the log-moneyness $k = \log \frac{K}{S_0}$.

Remark 4.4.1 We call this the *skew flattening phenomenon*, so in particular \tilde{X}_t^H (for a single fixed t) tends weakly to some symmetric distribution μ .

Proof. From Theorem 4.2.1, we know that $\langle \tilde{X}^H \rangle_t$ tends to a random variable $\xi_\gamma([0, t])$ in L^2 (and hence in probability), and $\langle \tilde{X}^H, W \rangle_t = \rho \int_0^t \sqrt{V_u^H} du$. But

$$\begin{aligned} \mathbb{E}((V_t^H)^{\frac{1}{2}}) &= \mathbb{E}(e^{\frac{1}{2}(\gamma Z_t^H - \frac{1}{2}\gamma^2 \frac{1}{2H} t^{2H})}) \\ &= \mathbb{E}(e^{\frac{1}{2}\gamma Z_t^H - \frac{1}{2} \cdot \frac{1}{4}\gamma^2 \cdot \frac{1}{2H} + \frac{1}{2} \cdot \frac{1}{4}\gamma^2 \cdot \frac{1}{2H} - \frac{1}{2}\gamma^2 \frac{1}{4H} t^{2H}}) = e^{-\frac{1}{16H}\gamma^2 t^{2H}} \rightarrow 0 \end{aligned}$$

as $H \rightarrow 0$, so (by Markov's inequality) $\mathbb{P}(\sqrt{V_t^H} > \delta) \leq \frac{1}{\delta} \mathbb{E}(\sqrt{V_t^H}) \rightarrow 0$, so $\sqrt{V_t^H}$ tends to zero in probability, and hence

$$G_t := \langle \tilde{X}^H, W \rangle_t \xrightarrow{P} 0. \quad (4.33)$$

Moreover, for any bounded martingale N orthogonal to W

$$\langle \tilde{X}^H, N \rangle_t = 0. \quad (4.34)$$

Thus setting $Z_t = W_t$ and applying Theorem IX.7.3 in Jacod&Shiryaev[JS03] (see also Proposition II.7.5 and Definition II.7.8 in [JS03]), we can construct an extension $(\tilde{\Omega}, \tilde{\mathcal{F}}, (\tilde{\mathcal{F}}_t), \tilde{\mathbb{P}})$ of our original filtered probability space $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbb{P})$ and a continuous Z -biased $\tilde{\mathcal{F}}$ -progressive conditional PII martingale \tilde{X} on this extension (see Definition 7.4 in chapter II in [JS03] for definition), such that \tilde{X}^H converges stably (and hence weakly) to \tilde{X} (see Definition 5.28 in chapter XIII in [JS03] for definition of stable convergence) for which

$$\begin{aligned} \langle \tilde{X} \rangle_t &= \xi_\gamma([0, t]) \\ \langle \tilde{X}, M \rangle_t &= 0 \end{aligned}$$

4.4 Application to the Rough Bergomi model - skew flattening/blowup as $H \rightarrow 0$

for all continuous (bounded) martingales M with respect to the original filtration \mathcal{F}_t . From Proposition 7.5 and Definition 7.8 in Chapter 2 in [JS03], this means that $\tilde{X}_t = X'_t + \int_0^t u_s dW_s$ where X' is an $\tilde{\mathcal{F}}_t$ -local martingale and u is a predictable process on the original space $(\Omega, \mathcal{F}, \mathbb{P})$. One such M is $M_t = W_{t \wedge \tau_b \wedge \tau_{-b}}$, where $\tau_b = \inf\{t : W_t = b\}$, so we have a pair of continuous local martingales (M, X) with $\langle \tilde{X}, M \rangle_t = \langle \tilde{X}, W \rangle_t = \int_0^t u_s ds = 0$ for $t \leq \tau_b \wedge \tau_{-b}$, so in fact $u_t \equiv 0$. Then applying F.Knight's Theorem 3.4.13 in [KS91] with $M^{(1)} = X$ and $M^{(2)} = W$, if $T_t = \inf\{s \geq 0 : \langle X \rangle_s > t\}$, then X_{T_t} is a Brownian motion independent of W . Hence X has the same law as $B_{\xi_\gamma^\perp([0,t])}^\perp$ for any Brownian motion B^\perp independent of W . ■

4.4.1 $H \rightarrow 0$ behaviour for the usual rough Bergomi model

If we replace the definition of Z^H with the usual RL process $Z_t^H = \sqrt{2H} \int_0^t (t-s)^{H-\frac{1}{2}} ds$ (as is usually done), then from remark 4.2.2, we know that $\xi_\gamma^H(A)$ tends to $\text{Leb}(A)$ in L^2 for any Borel set $A \subseteq [0, 1]$, so adapting Theorem 4.4.1 for this case, we see that \tilde{X}^H tends weakly to a standard Brownian motion, which means the rough Bergomi model tends weakly to the Black-Scholes model in the $H \rightarrow 0$ limit.

4.4.2 A closed-form expression for $\mathbb{E}((\tilde{X}_t^H)^3)$

In this subsection we compute an explicit expression for the skewness of \tilde{X}_t^H (conditioned on its history), which (as a by-product) gives a more “hands-on” proof as to why the skew tends to zero as $H \rightarrow 0$, and also allows us to see how fast the skew decays.

We first note that (trivially) \tilde{X}^H has the same law as \tilde{X}^H defined by

$$\begin{cases} d\tilde{X}_t^H = \sqrt{V_t^H}(\rho dB_t + \bar{\rho} dW_t), \\ V_t^H = e^{\gamma Z_t^H - \frac{1}{2}\gamma^2 \text{Var}(Z_t^H)} \\ Z_t^H = \int_0^t (t-s)^{H-\frac{1}{2}} dB_s \end{cases} \quad (4.35)$$

where B is independent of W , and this is the version of the model we use in this subsection. We now replace the constant ρ with a time-dependent $\rho(t)$, and replace our original V_t^H process with

$$V_t^H = \xi_0(t) e^{\gamma Z_t^H - \frac{1}{2}\gamma^2 \text{Var}(Z_t^H)}$$

to incorporate a non-flat initial variance term structure.

Proposition 4.4.3

$$\begin{aligned} & \mathbb{E}_{t_0}((\tilde{X}_T^H - \tilde{X}_{t_0}^H)^3) \\ &= 3\gamma \int_{t_0}^T \int_0^t \rho(s) \xi_{t_0}^{\frac{1}{2}}(s) \xi_{t_0}(t) e^{\frac{1}{2}\gamma^2 \text{Cov}_{t_0}(Z_s^H Z_t^H) - \frac{1}{8}\gamma^2 \text{Var}_{t_0}(Z_s^H)} (t-s)^{H-\frac{1}{2}} ds dt \end{aligned} \quad (4.36)$$

where $\xi_{t_0}(t) = \xi_0(t) e^{\gamma \int_0^{t_0} (t-u)^{H-\frac{1}{2}} dB_u - \frac{\gamma^2}{4H} [t^{2H} - (t-t_0)^{2H}]}$. This simplifies to

$$\mathbb{E}((\tilde{X}_T^H)^3) = 3\rho\gamma\mathcal{W}_0^{\frac{3}{2}} \int_0^T \int_0^t e^{\frac{1}{2}\gamma^2 (R_H(s,t) - \frac{s^{2H}}{8H})} (t-s)^{H-\frac{1}{2}} ds dt < \infty \quad (4.37)$$

if $t_0 = 0$, ρ is constant and $\xi_0(t) = V_0$ for all t (i.e. flat initial variance term structure).

Proof. See Appendix C. ■

Remark 4.4.2 Using that $R_H(s,t) \rightarrow R^{\text{fBM}}(s,t)$ as $s,t \rightarrow 0$ (for $H > 0$ fixed), where $R^{\text{fBM}}(s,t) = \frac{1}{2H} \frac{1}{2} (t^{2H} + s^{2H} - |t-s|^{2H})$ is the covariance function of $\frac{1}{\sqrt{2H}} W^H$ where W^H is a standard (one or two-sided) fractional Brownian motion, we find that the exponent in (4.37) behaves like $\frac{1}{16H} (s^{2H} + 2t^{2H} - 2(t-s)^{2H})$ for $s < t$ as $s,t \rightarrow 0$, and thus can effectively be ignored, so (for ρ constant)

$$\begin{aligned} \mathbb{E}((\tilde{X}_T^H)^3) &\sim 3\rho\gamma\mathcal{W}_0^{\frac{3}{2}} \int_0^T \int_0^t e^{\frac{1}{2}\gamma^2 (R_H(s,t) - \frac{s^{2H}}{8H})} (t-s)^{H-\frac{1}{2}} ds dt \\ &= \frac{3\rho\gamma\mathcal{W}_0^{\frac{3}{2}}}{(H + \frac{1}{2})(H + \frac{3}{2})} T^{H+\frac{3}{2}} \quad (T \rightarrow 0). \end{aligned}$$

Remark 4.4.3 Note that \tilde{X}^H is driftless so (4.35) is only a toy model at the moment, but we easily adapt Proposition 4.4.3 and the two remarks above to incorporate the additional $-\frac{1}{2}\langle \tilde{X}^H \rangle_t$ drift term required to make $S_t = e^{\tilde{X}_t^H}$ a martingale. However, the relative contribution from this drift will disappear in the small-time limit, so we omit the tedious details, since rough stochastic volatility models are generally used (and considered more realistic) over small time horizons.

4.4.3 Convergence of the skew to zero

Corollary 4.4.4 For $\gamma \in (0, 1)$ and $0 \leq t \leq T \leq 1$, $\mathbb{E}_{t_0}((\tilde{X}_T^H - \tilde{X}_{t_0}^H)^3) \rightarrow 0$ a.s. as $H \rightarrow 0$.

4.4 Application to the Rough Bergomi model - skew flattening/blowup as $H \rightarrow 0$

Proof. For $T \leq 1$, using that $R_H(s, t) \uparrow R(s, t)$ and $(t-s)^{H-\frac{1}{2}} \uparrow (t-s)^{-\frac{1}{2}}$ we see that

$$\begin{aligned}
|\mathbb{E}_{t_0}((\tilde{X}_T^H - \tilde{X}_{t_0}^H)^3)| &\leq 3|\rho|\gamma \int_{t_0}^T \int_0^t \xi_{t_0}^{\frac{1}{2}}(s) \xi_{t_0}(t) e^{\frac{1}{2}\gamma^2(R_{t_0}(s,t) - \frac{s^{2H}}{8H})} (t-s)^{-\frac{1}{2}} ds dt \\
&\leq 3|\rho|\gamma \int_{t_0}^T \int_0^t \xi_{t_0}^{\frac{1}{2}}(s) \xi_{t_0}(t) e^{\frac{1}{2}\gamma^2(R(s,t) - \frac{s^{2H}}{8H}) - \frac{1}{2}\log(t-s)} ds dt \\
&\leq 3\bar{\xi}_{t_0}^{\frac{1}{2}}(s) \bar{\xi}_{t_0}(t) |\rho|\gamma \int_{t_0}^T \int_0^t e^{\frac{1}{2}(1+\gamma^2)\log\frac{1}{t-s} + \frac{1}{2}\gamma^2\bar{g}} ds dt \\
&\leq \text{const.} \times \mathbb{E}(M_{\sqrt{\frac{1}{2}(1+\gamma^2)}}([0, T])^2) < \infty
\end{aligned}$$

for $\gamma \in (0, 1)$ where $M_\gamma(dt)$ is the usual [BM03] GMC, and $R_0(s, t) = \mathbb{E}_{t_0}(Z_s Z_t) = \int_{t_0}^s (s-u)^{-\frac{1}{2}}(t-u)^{-\frac{1}{2}} du ds$, $\bar{g} = 2\log(2\sqrt{2})$, $\bar{\xi}_t = \sup_{0 \leq s \leq t} \xi_s$. The result follows from dominated convergence theorem. ■

4.4.4 Speed of convergence of the skew to zero

The proof in the previous subsection applies to $\gamma \in [0, 1]$ however the result remains true for γ up to a critical value $\gamma^* > \sqrt{2}$ see Theorem 2 in [Ger20]. This result also gives us the rate at which the skew converges to zero (or indeed blows up):

Proposition 4.4.5 (see [Ger20]). *Let $\rho(\cdot)$ be continuous and bounded away from zero with constant sign for t sufficiently small. Then*

$$-\lim_{H \rightarrow 0} H \log[\text{sgn}(\rho)\mathbb{E}((\tilde{X}_T^H)^3)] = \hat{r}(\gamma) = \begin{cases} \frac{1}{16}\gamma^2 & 0 \leq \gamma \leq 1, \\ \frac{1}{4} + \frac{1}{2}\log\gamma - \frac{3}{16}\gamma^2 & \gamma \geq 1 \end{cases} \quad (4.38)$$

$\hat{r}(\gamma)$ is negative for γ larger than the root of $\frac{1}{4} + \frac{1}{2}\log\gamma - \frac{3}{16}\gamma^2$ at ≈ 1.61711 , which makes the integral explode as $H \rightarrow 0$ for such values of γ .

Interestingly this critical value of $\gamma^* = 1.61711..$ is greater than $\sqrt{2}$.

4.4.5 A $H = 0$ model - pros and cons

We can circumvent the problem of vanishing skew, by considering a toy model of the form

$$X_t = \sigma(\rho W_t + \bar{\rho} B_{\bar{\xi}_\gamma([0, t])}^\perp) \quad (4.39)$$

4.4 Application to the Rough Bergomi model - skew flattening/blowup as $H \rightarrow 0$

where $\bar{\rho} = \sqrt{1 - \rho^2}$, W and $\xi_\gamma([0, t])$ are defined as in Section 4.3 with $\gamma \in (0, 1)$, and B^\perp is a Brownian motion independent of W . Then from the tower property we see that

$$\mathbb{E}(e^{ikX_t}) = \mathbb{E}(\mathbb{E}(e^{ik(\sigma\rho W_t + \sigma\bar{\rho}B_{\xi_\gamma([0,t])})} | W)) = \mathbb{E}(e^{ik\sigma\rho W_t - \frac{1}{2}k^2(\sigma\bar{\rho})^2\xi_\gamma([0,t])})$$

and (from Remark 4.2.3) we know that $\xi_\gamma([0, t]) \sim t\xi_\gamma([0, 1])$ (i.e. self-similarity), so

$$\mathbb{E}(e^{\frac{ik}{\sqrt{t}}X_t}) = \mathbb{E}(e^{ik\sigma\rho W_t/\sqrt{t} - \frac{1}{2}k^2(\sigma\bar{\rho})^2\xi_\gamma([0,t])/t}) = \mathbb{E}(e^{ik\sigma\rho W_1 - \frac{1}{2}k^2(\sigma\bar{\rho})^2\xi_\gamma([0,1])})$$

so X is self-similar: $X_t/\sqrt{t} \sim X_1$ for all $t > 0$, and X_1 (and hence X_t) has non-zero skewness for $\alpha \neq 0$; more specifically (see Appendix D for a derivation of the following):

$$\mathbb{E}\left(\left(\frac{X_t}{\sqrt{t}}\right)^3\right) = 3\sigma^3\rho(1 - \rho^2)\mathbb{E}[W_1\xi[0, 1]] \quad (4.40)$$

$$= 4\sigma^3\rho(1 - \rho^2)\gamma \quad (4.41)$$

and $\mathbb{E}(X_1^2) = \sigma^2$, and we can derive a similar (slightly more involved) expression for $\mathbb{E}(X_1^4)$. The ρ component achieves the goal of a $H = 0$ model with non-zero skewness, and one can establish the following small-time behaviour for European put options in the Edgeworth Central Limit Theorem regime:

$$\frac{1}{\sqrt{t}}\mathbb{E}((e^{x\sqrt{t}} - e^{X_t})^+) \sim e^{x\sqrt{t}}\mathbb{E}\left(\left(x - \frac{X_t}{\sqrt{t}}\right)^+\right) \sim \mathbb{E}\left(\left(x - \frac{X_t}{\sqrt{t}}\right)^+\right) \sim \mathbb{E}\left(\left(x - \bar{X}_1\right)^+\right)$$

and $\lim_{t \rightarrow 0} \hat{\sigma}_t(x\sqrt{t}, t) = C_B(x, \cdot)^{-1}(C(x))$ for $x > 0$, where $\hat{\sigma}_t(x, t)$ denotes the implied volatility of a European call option with strike $e^{x\sqrt{t}}$ maturity t and $S_0 = 1$ ($C_B(x, \sigma)$ is the Bachelier model call price formula). Hence we see the full smile effect in the small-time FX options Edgeworth regime unlike the $H > 0$ case where the leading order term is just Black-Scholes, followed by a next order skew term, followed by an even higher order convexity term.

We can go from a toy model to a real model adding back the usual $-\frac{1}{2}\langle X \rangle_t$ drift term for the log stock price X so $S_t = e^{X_t}$ is a martingale, and in this case we lose self-similarity for X but X_t/\sqrt{t} still tends weakly to a non-Gaussian random variable, and in particular $\lim_{t \rightarrow 0} \mathbb{E}\left(\left(\frac{X_t}{\sqrt{t}}\right)^3\right) = 4\sigma^3\rho\bar{\rho}^2\gamma$.³ This model overcomes two of the main drawbacks of the original Bacry et al. multifractal random walk, namely zero skewness and unrealistic small-time behaviour. However, the property in (4.41) does not appear to be time-consistent, since if we define $\eta_t^h := \mathbb{E}\left(\left(\frac{X_{t+h} - X_t}{\sqrt{h}}\right)^3 | \mathcal{F}_t\right)$ for $t > 0$, then $\mathbb{E}\left(\left(\eta_t^h\right)^2\right) = O(h^{-\gamma^2})$ (and not $O(1)$ as we would want), so we do not pursue this model further at the present time.

³We can also replace the ρW_t component of X with a second rBergomi component with a non-zero H -value, and derive similar results

4.5 Explicit spectral expansions for Z^H for $H \geq 0$

Following section 4.3 in [Gia15], we first briefly recall the classical Karhunen-Loève theorem. Let $(X_t)_{t \in [a,b]}$ be a centred continuous-parameter real-valued process defined on some probability space $(\Omega, \mathcal{F}, \mathbb{P})$, which is second order (i.e. $\mathbb{E}(X_t^2) < \infty$ for all $t \in [a, b]$) with continuous covariance function $K_X(s, t)$. Let

$$Z_k = \int_a^b X_t e_k(t)$$

where $\{e_k\}_{k=1}^\infty$ are the eigenfunctions of the Hilbert-Schmidt integral operator on $L^2([a; b])$ given by $(Af)(t) = \int_a^b K_X(s, t)f(s)ds$, which is an orthonormal basis for the space spanned by the eigenfunctions corresponding to the non-zero eigenvalues of A . Then $\mathbb{E}(Z_j Z_k) = \lambda_k \delta_{jk}$ for all j, k , $\mathbb{E}(Z_j) = 0$ for all j , and the series

$$\sum_{n=1}^{\infty} Z_n e_n(t)$$

converges to X_t in mean square, uniformly for $t \in [a, b]$. This expansion is often said to be *bi-orthogonal*, since the random coefficients Z_k are orthogonal in $L^2(\Omega, \mathcal{F}, \mathbb{P})$ and the eigenfunctions are orthogonal in $L^2([a, b])$. If X is Gaussian, then the Z_k 's are independent Gaussians.

The K-L expansion of standard Brownian motion on $[0, 1]$ is given by

$$W_t = \sum_{n=1}^{\infty} \sqrt{\lambda_n} \phi_n(t) Z_n$$

(see page 50 in [Gia15]), where Z_n is a sequence of i.i.d. standard Normals, and

$$\lambda_n = \frac{4}{(2n-1)^2 \pi^2} \quad , \quad \phi_n(t) = \sqrt{2} \sin\left(\left(n - \frac{1}{2}\right)\pi t\right). \quad (4.42)$$

λ_n and ϕ_n are the eigenvalues and eigenfunctions of the Hilbert-Schmidt covariance operator $R_{\frac{1}{2}} : L^2([0, 1]) \rightarrow L^2([0, 1])$ given by $R_{\frac{1}{2}}\theta(t) = \int_0^1 R_{\frac{1}{2}}(s, t)\theta(s)ds = \int_0^1 (s \wedge t)\theta(s)ds$, and ϕ_n forms an orthonormal basis of $L^2([0, 1])$ and $\sqrt{\lambda_n} \phi_n'$ forms an orthonormal basis of $L^2([0, 1])$.

We now recall the re-scaled Riemann-Liouville process from :

$$Z_t^H = \int_0^t (t-s)^{H-\frac{1}{2}} dW_s$$

with $H \in (0, 1)$ for which $R_H(s, t) := \mathbb{E}(Z_s^H Z_t^H) = \int_0^{s \wedge t} (s-u)^{H-\frac{1}{2}} (t-u)^{H-\frac{1}{2}} du$. We define the operator $K_H : L^2 \rightarrow C[0, 1]$ as $K_H f(t) = \int_0^t (t-s)^{H-\frac{1}{2}} f(s) ds$ for $H \in [0, \frac{1}{2})$, and set

$$X_t^n = \sum_{k=1}^n \sqrt{\lambda_k} K_H \phi_k'(t) Z_n. \quad (4.43)$$

Recall that $\sqrt{\lambda_n} \phi_n'$ forms an orthonormal basis of $L^2([0, 1])$, and we see that the covariance function of X^n is

$$R_{H,n}(s, t) = \sum_{k=1}^n \lambda_k K_H \phi_k'(s) K_H \phi_k'(t).$$

The aim of the next few subsections is to show that (4.43) converges to Z_t^H in appropriate sense for $H > 0$ and for $H = 0$. To do this, we first have to give some background on the Cameron-Martin and Reproducing Kernel Hilbert spaces associated with Gaussian fields.

4.5.1 The Cameron-Martin space of a log-correlated Gaussian field

Let

$$C(x, y) = \log \frac{1}{|y-x|} + g(x, y) \quad (4.44)$$

for some function g which is continuous and bounded away from $(0, 0)$. If the associated bilinear operator $C(\phi, \psi) := \int \int C(x, y) \phi_1(x) \phi_2(y) dx dy$ is positive definite (i.e. $C(\phi, \phi) \geq 0$ for all $\phi \in \mathcal{S}$) and continuous at zero (i.e. under the Schwartz space semi-norm defined in Eq 1 in e.g. [BDW17]) then the Minlos-Bochner theorem implies that C is the covariance of a centred Gaussian measure μ on the space \mathcal{S}' of *tempered distributions* which is the dual of the Schwartz space \mathcal{S} (see e.g. Theorem 2.1 in [BDW17] or page 8 in Janson[Jan97]). If $X \sim \mu$, then (due to the log term in (4.44)) we say that X is an *almost log-correlated Gaussian field*, which has the following covariance structure:

$$\mathbb{E}(X(\phi_1) X(\phi_2)) = \int \int C(x, y) \phi_1(x) \phi_2(y) dx dy. \quad (4.45)$$

\mathcal{S} is a Montel space and thus is reflexive, i.e. $(\mathcal{S}')'$ is isomorphic to \mathcal{S} using the canonical embedding of S into its bi-dual $(S')'$.

We can complete \mathcal{S} using the inner product $\langle \phi_1, \phi_2 \rangle_{R^\mu} := \int \int C(x, y) \phi_1(x) \phi_2(y) dx dy$ to form the Hilbert Space R^μ (see e.g. page 44 in Bogachev[Bog91]). If we define H_μ to

be the dual of R^μ i.e. the space of linear bounded functionals on R^μ with norm

$$\|h\|_{H_\mu} := \sup\{h(l) : l \in R^\mu, C(l,l) \leq 1\}$$

then H_μ is known as the *Cameron-Martin space* of X (again see page 44 in [Bog91]).

We now recall the standard Riesz Representation theorem:

Theorem 4.5.1 *Given a Hilbert space H , the map $\phi : H \rightarrow H^*$ defined by:*

$$\phi(h) = \langle h, \cdot \rangle$$

is an isometric isomorphism.

Any element of H_μ is a bounded linear operator on the Hilbert space R^μ , so (by Riesz) for any $h \in H_\mu$ there exists a unique $g \in R^\mu$ such that

$$h(f) = \langle f, g \rangle_{R^\mu} = \mathbb{E}(X(f)X(g)) = (Cg)(f) \quad \forall f \in R^\mu.$$

h and Cg are both bounded linear functions on R_μ that agree on every element of R_μ so $h = Cg$, so the Riesz map $\phi : H_\mu \rightarrow R^\mu$ here is given explicitly by $\phi(h) = C^{-1}h$.

4.5.2 Characterizing H_μ when $C = AA^*$

Suppose we have some Hilbert space E and we can factorise the covariance as $C = AA^*$

$$A : E \rightarrow H_\mu \quad , \quad A^* : R^\mu \rightarrow E$$

where A and A^* are bounded linear operators and A^* is the adjoint of each other in the sense that $(A\phi)(f) = \langle A^*f, \phi \rangle_E$ for all $\phi \in E$ and all $f \in R^\mu$ i.e. $A^*f = f(A(\cdot))$. Then we have the following:

Proposition 4.5.2 *H_μ is equal to the Hilbert space $H = AE$ with inner product $\langle Af, Ag \rangle_H := \langle f, g \rangle_E$ for all $f, g \in E$.*

Before proceeding with the proof, we recall a definition: given a set \mathcal{X} , a *Reproducing Kernel Hilbert Space* H (RKHS) is a Hilbert space of functions on \mathcal{X} admitting a reproducing kernel, i.e. there exists a positive definite function $K : \mathcal{X} \times \mathcal{X} \rightarrow \mathbb{R}$ such that:

- (i) $K(p, \cdot) \in H$ for all $p \in \mathcal{X}$.
- (ii) $f(p) = \langle K(p, \cdot), f \rangle_H$ for all $f \in H$ and $p \in \mathcal{X}$.

Proof. One proof can be found on pages 107-108 of [Bog91]. Alternatively, note that for all $f \in E$ and $g \in R^\mu$:

$$(Af)(g) := \langle A^*g, f \rangle_E = \langle AA^*g, Af \rangle_H = \langle Cg, Af \rangle_H. \quad (4.46)$$

But this is the statement that H is an RKHS on R^μ with kernel C . Such spaces are uniquely characterised by their kernel (this is the Aronszajn-Moore theorem) and so all that remains is to show that H_μ is also a RKHS with the same kernel. To this end, we note that any $h \in H_\mu$ can be written as $C\phi$ for some $\phi \in R^\mu$, so $\forall g \in R^\mu$ we have

$$h(g) = \langle C\phi, g \rangle = \langle C\phi, Cg \rangle_{H_\mu} = \langle Cg, h \rangle_{H_\mu}$$

which is the reproducing condition. ■

Our particular case of interest is of course for the Riemann-Liouville process (resp. field) where $(Af)(t) := \int_0^t (t-s)^{H-\frac{1}{2}} f(s) ds$ for $H \in [0, 1]$ which is a bounded linear map on $E = L^2([0, 1])$ (see Theorem 2.6 in Samko et al.[SKM93] for a proof of this) and $H_\mu = AL^2([0, 1])$. Moreover, using the Stone-Weierstrauss argument in Appendix A of [FZ17], we can also verify that H_μ is dense in L^2

As another simple example, we can trivially write $C = C\iota$ where ι is the identity mapping on R^μ . Let $E = R^\mu$, $A = C$ and $A^* = \iota$ we see the adjoint condition is simply $(C\phi)(f) = \langle \phi, f \rangle_{\mathbb{R}^\mu}$. By Proposition 4.5.2 we recover the result $H_\mu = CR^\mu$.

4.5.3 Karhunen-Loève type expansions

The Hilbert space structure allows us to expand our field in a basis. We have established a linear isomorphism $H_\mu \leftrightarrow R^\mu$ which is also an isometry. Thus if we have an O.N. basis of H_μ then there is a corresponding O.N. basis $\{A_k = C^{-1}e_k\}$ of R^μ . Then for all $\phi \in \mathcal{S}$ we have

$$\phi = \sum_{k=1}^{\infty} c_k^\phi A_k$$

where $c_k^\phi = \langle \phi, A_k \rangle_{R^\mu}$ and convergence is of course in the R^μ norm i.e. $(C(\phi - \phi_n), \phi - \phi_n) \rightarrow 0$, where $\phi_n = \sum_{k=1}^n \langle \phi, A_k \rangle A_k$ denotes the n th partial sum, or equivalently

$$X(\phi) = \sum_{k=1}^{\infty} c_k^\phi X(A_k)$$

in the sense that

$$\|X(\phi) - X^n(\phi)\|^2 = \langle C(\phi - \phi_n), \phi - \phi_n \rangle \rightarrow 0 \quad (4.47)$$

as $n \rightarrow \infty$, where $X^n = \sum_{k=1}^n c_k^\phi X(A_k)$. Using the isometry and the reproducing property in H we see that

$$c_k^\phi = \langle C\phi, A_k \rangle = \langle C\phi, CA_k \rangle_{H_\mu} = \langle C\phi, e_k \rangle_{H_\mu} = e_k(\phi)$$

by the reproducing property. In this sense we say that $X = \sum_{k=1}^\infty X(A_k)e_k$ and we can re-write the convergence in (4.47) as

$$\|X(\phi) - X^n(\phi)\|^2 = \mathbb{E}([(X, \phi) - (X^n, \phi)]^2) \rightarrow 0$$

where $X^n := \sum_{k=1}^n X(A_k)e_k$.

We also have that

$$\mathbb{E}(X(A_j)X(A_k)) = \langle CA_j, A_k \rangle = \langle A_j, A_k \rangle_{R^\mu} = \langle CA_j, CA_k \rangle_{H_\mu} = 1_{j=k}. \quad (4.48)$$

and we know that each $Z_k := X(A_k)$ is Gaussian so they must be i.i.d. standard Normals, so we can re-write our expansion as

$$X(\phi) = \sum_{k=1}^\infty e_k(\phi) Z_k.$$

Moreover, $X^n(\phi) = \sum_{k=1}^n e_k(\phi) Z_k$ is a discrete-time L^2 -martingale, so by the martingale convergence theorem $X_n(\phi)$ converges a.s. to $X(\phi)$, and hence $X_n \rightarrow X$ a.s. in the weak topology on \mathcal{S} (and the strong topology, see page 2 in [BDW17]).

4.5.4 Choice of basis and explicit computation of terms

The operator C is a Hilbert-Schmidt, compact, linear and self-adjoint operator on L^2 so by the spectral theorem we can form the O.N. basis (e_n) of eigenfunctions of C which is known as the *Karhunen-Loève basis*, which is frequently used in theoretical proofs but is not so useful in practice aside from a few special cases (e.g. Brownian motion and the Brownian bridge) since typically these eigenfunctions cannot be computed explicitly.

For this reason we appeal to Proposition 4.5.2 instead: let (\tilde{e}_k) be an O.N. basis of $L^2([0, 1])$. A is injective (due to the positive definiteness of C), so $e_k = A\tilde{e}_k$ is an O.N. basis

of $AL^2([0, 1])$. Then in this basis we have

$$X = \sum_{k=1}^{\infty} A\tilde{e}_k Z_k \quad (4.49)$$

and the $Z_k = X(A_k)$'s are i.i.d. Normals.

If we use the O.N. basis of $L^2([0, 1])$ given by $\sqrt{\lambda_n}\phi_k(t)$ from (4.42), then we can compute $K_H\phi_k'(t)$ explicitly as

$$K_H\phi_k'(t) = \frac{\sqrt{2}}{1+2H} (2n-1)\pi t^{\frac{1}{2}+H} {}_1F_1\left(\frac{3}{4} + \frac{1}{2}H, \frac{5}{4} + \frac{1}{2}H, -\frac{1}{16}(2n-1)^2\pi^2 t^2\right) \quad (H > 0)$$

and

$$\begin{aligned} K_H'\phi_k(t) &= \sqrt{2}\sqrt{2n-1}\pi \left[\cos\left(\frac{1}{2}(2n-1)\pi t\right) \text{FresnelC}(\sqrt{(2n-1)t}) \right. \\ &\quad \left. + \text{FresnelS}(\sqrt{(2n-1)t}) \sin\left[\frac{1}{2}(2n-1)\pi t\right] \right] \quad (H = 0) \end{aligned} \quad (4.50)$$

where ${}_pF_q$ is the generalized hypergeometric function⁴, $\text{FresnelC}(z) = \int_0^z \cos(\frac{1}{2}\pi t^2) dt$, $\text{FresnelS}(z) = \int_0^z \sin(\frac{1}{2}\pi t^2) dt$. No such explicit formulae are known for the standard K-L expansion for the RL process which requires knowledge of the eigenvalues of eigenfunctions of the covariance operator (see Gulisashvili et al.[GVZ19] for asymptotic results in this direction).

4.5.5 Using the spectral expansion to sample the GMC mass $\xi_\gamma^H([0, T])$ for $H \ll 1$ and $H = 0$

From the well known Selberg formula we have that

$$\mathbb{E}(\xi_\gamma^H([0, T])^q) = \int_{[0, T]} \dots \int_{[0, T]} e^{\gamma^2 \sum_{1 \leq i < j \leq q} R_H(u_i, u_j)} du_1 \dots du_q \quad (4.51)$$

for $q > 0$. In the following table we have tabulated the first four raw moments $\xi_\gamma([0, 1])$ for (using (4.51)) and their corresponding estimates using Monte Carlo simulation with the K-L expansion in (4.43) using (4.50) (which we denote by $\hat{\mu}_n$), with $n = 1000$ eigenfunctions, 1000 time steps and 1 million simulations for both cases, and Gaussian quadrature for the numerical integration, and we find that the exact and MC answers are in very close agreement (and similarly we get very close agreement for small positive H values, e.g.

⁴using Mathematica's definition

4.5 Explicit spectral expansions for Z^H for $H \geq 0$

$H = .0001$, for which the numbers are almost identical to those in the table). If we perform the same computations for e.g. $H = .0001$ using a traditional Cholesky decomposition with a simple Riemann sum, then we get nonsensical answers, which we have not tabulated here. Based on these results, our KL expansion clearly useful for pricing variance options under the rough Bergomi model with H small or zero (i.e. options on $\xi_\gamma^H([0, T])$); our expansion does not appear to work so well for the (driftless) rough Bergomi model with non-zero correlation defined in Eq 4.35 in estimating the third moment of the driftless log stock price: $\mathbb{E}((\tilde{X}_T^H)^3)$ (when compared to the analytical expression for this quantity given in 4.37).

γ	μ_1	μ_2	μ_3	μ_4	$\hat{\mu}_1$	$\hat{\mu}_2$	$\hat{\mu}_3$	$\hat{\mu}_4$
0.05	1	1.00502	1.01513	1.0305	1	1.005	1.0151	1.0305
0.1	1	1.02028	1.06216	1.12836	0.99999	1.0202	1.062	1.1281
0.15	1	1.04644	1.14633	1.31508	0.99999	1.0464	1.1462	1.3148
0.2	1	1.08466	1.27764	1.63646	0.99999	1.0845	1.2771	1.6351
0.25	1	1.13669	1.47323	2.1843	0.99999	1.1367	1.4734	2.1852
0.3	1	1.20512	1.76191	3.14796	0.99985	1.2042	1.7582	3.1336
0.35	1	1.29359	2.19289	4.94361	0.99975	1.2919	2.1849	4.9066
0.4	1	1.40729	2.85324	8.56902	1.0001	1.4078	2.8651	8.9761
0.45	1	1.5537	3.90481	16.6981	1	1.5532	3.8953	16.461
0.5	1	1.74375	5.66824	37.5977	1.0001	1.744	5.6369	35.0148

The following table performs the same computations as above but for $H = .03$ (empirical values as low as this are reported in Fukasawa et al.[FTW19]), and in the final column we compare against a traditional Cholesky scheme using Simpson's rule (also with 1000 time steps and 1 million simulations), and we see that our K-L expansion method outperforms the latter to an increasingly greater extent as γ increases. In the second column, we have computed the usual vol-of-variance parameter $\eta = \frac{\gamma}{\sqrt{2H}}$ corresponding to each choice of γ . Matlab was unable to compute a positive-definite 1000 point Cholesky decomposition when we tried using Gaussian quadrature instead of Simpson's rule (the former of course has non-equidistant time points), and also for $H = .05$, see final table below.

4.5 Explicit spectral expansions for Z^H for $H \geq 0$

γ	$\eta = \frac{\gamma}{\sqrt{2H}}$	μ_1	μ_2	μ_3	μ_4	$\hat{\mu}_1^H$	$\hat{\mu}_2^H$	$\hat{\mu}_3^H$	$\hat{\mu}_4^H$	$\hat{\rho}_1^{H,\text{chol}}$	$\hat{\rho}_2^{H,\text{chol}}$	$\hat{\rho}_3^{H,\text{chol}}$	$\hat{\rho}_4^{H,\text{chol}}$
0.05	0.204124	1	1.0043	1.01306	1.0263	1	1.0043	1.013	1.0262	1	1.0043	1.0131	1.0264
0.1	0.408248	1	1.01749	1.05345	1.10988	0.99999	1.0175	1.0534	1.1098	1	1.0176	1.0538	1.1107
0.15	0.612372	1	1.03993	1.125	1.26441	1	1.04	1.1253	1.267	1	1.0402	1.1259	1.2684
0.2	0.816497	1	1.0725	1.23489	1.52805	0.99997	1.0725	1.2349	1.5281	1.0001	1.0735	1.238	1.5353
0.25	1.02062	1	1.11644	1.39505	1.95613	1.0002	1.1174	1.3977	1.963	0.9999	1.1173	1.399	1.9678
0.30	1.22474	1	1.17353	1.62473	2.66892	0.99977	1.1727	1.6226	2.6646	0.99999	1.176	1.6352	2.7041
0.35	1.42887	1	1.24619	1.95509	3.90479	1.0002	1.2477	1.9635	3.951	0.99987	1.2505	1.976	3.9871
0.40	1.63299	1	1.3378	2.43775	6.17752	0.99988	1.3371	2.4352	6.1555	0.9998	1.3481	2.5016	6.5463
0.45	1.83712	1	1.45292	3.16099	10.6697	0.99971	1.4505	3.1351	10.2878	0.99937	1.4714	3.2987	11.6325
0.50	2.04124	1	1.59791	4.28211	20.4214	0.99966	1.5947	4.2561	20.1089	1.0003	1.6488	4.8197	28.5117

4.6 Appendix

4.6.1 Appendix A: Definition and properties of $F_H(k)$ and $G_H(k)$ for the Sandwich lemma

$R_H(s, t) = \int_0^{s \wedge t} (s-u)^{H-\frac{1}{2}}(t-u)^{H-\frac{1}{2}} du = \int_0^s u^{H-\frac{1}{2}}(t-s+u)^{H-\frac{1}{2}} du$ for $0 \leq s \leq t$, and note that the integrand is non-negative. Going forward we set $k = t - s$. We restrict $R_H(s, t)$ to $A_\delta := \{(s, t) : t - s = k, (s, t) \in [\tau, \tau + \delta]^2\}$ with $k \in (0, \delta)$ and $\delta \in (0, 1 - \tau)$, i.e. $R_H(s, k+s) = \int_0^s (u(k+u))^{H-\frac{1}{2}} du$. This expression is maximized at $s = \tau + \delta - k$ and minimized at $s = \tau$ for constant k (see Figure 4.2). Recall that $G_H(k) := R_H(\tau + \delta - k, \tau + \delta)$, we will now establish some basic properties of $G_H(k)$. From the analysis above: $G_H(k) = \int_0^{\tau+\delta-k} (u(k+u))^{H-\frac{1}{2}} du$. Taking the derivative with respect to k and using the Leibniz rule, we see that

$$G'_H(k) = -(\tau + \delta - k)^{H-\frac{1}{2}}(\tau + \delta)^{H-\frac{1}{2}} + (H - \frac{1}{2}) \int_0^{\tau+\delta-k} u^{H-\frac{1}{2}}(k+u)^{H-\frac{3}{2}} du$$

which is negative (since $H < \frac{1}{2}$), so $G_H(k)$ is decreasing in k . The integral term in the previous equation explodes as $k \downarrow 0$:

$$\int_0^{\tau+\delta-k} u^{H-\frac{1}{2}}(k+u)^{H-\frac{3}{2}} du \geq \int_0^{\tau+\delta-k} (k+u)^{2H-2} du = \frac{(\tau + \delta)^{2H-1}}{2H-1} - \frac{k^{2H-1}}{2H-1} \uparrow \infty.$$

Hence $G'_H(k) \rightarrow -\infty$ as $k \searrow 0$. Conversely, if we fix k and let $H \rightarrow 0$, we find that

$$\begin{aligned} G_H(k) &\uparrow G_0(k) = \log \frac{1}{k} + 2 \log(\sqrt{\tau + \delta - k} + \sqrt{\tau + \delta}) \quad (H \rightarrow 0) \\ &\leq g(k) := \log \frac{1}{k} + 2 \log(2\sqrt{\tau + \delta}) = \log \frac{1}{k} + \log(4(\tau + \delta)) \end{aligned}$$

with equality at $k = 0$ in the sense that both sides of the inequality are infinite. Thus

$$G_H(k) \leq G_0(k) \leq g(k) \leq \log \frac{4}{k} \quad (\text{A.1})$$

since $\tau + \delta < 1$ by assumption.

Similarly, we recall that $F_H(k) := R_H(\tau, \tau + k) = \int_0^\tau (\tau - u)^{H-\frac{1}{2}}(\tau + k - u)^{H-\frac{1}{2}} du$, so

$$\begin{aligned} F'_H(k) &= (H - \frac{1}{2}) \int_0^\tau (\tau - u)^{H-\frac{1}{2}}(\tau + k - u)^{H-\frac{3}{2}} du \geq (H - \frac{1}{2}) \int_0^\tau (\tau - u)^{2H-2} du \\ F''_H(k) &= (H - \frac{1}{2})(H - \frac{3}{2}) \int_0^\tau (\tau - u)^{H-\frac{1}{2}}(\tau + k - u)^{H-\frac{5}{2}} du \end{aligned}$$

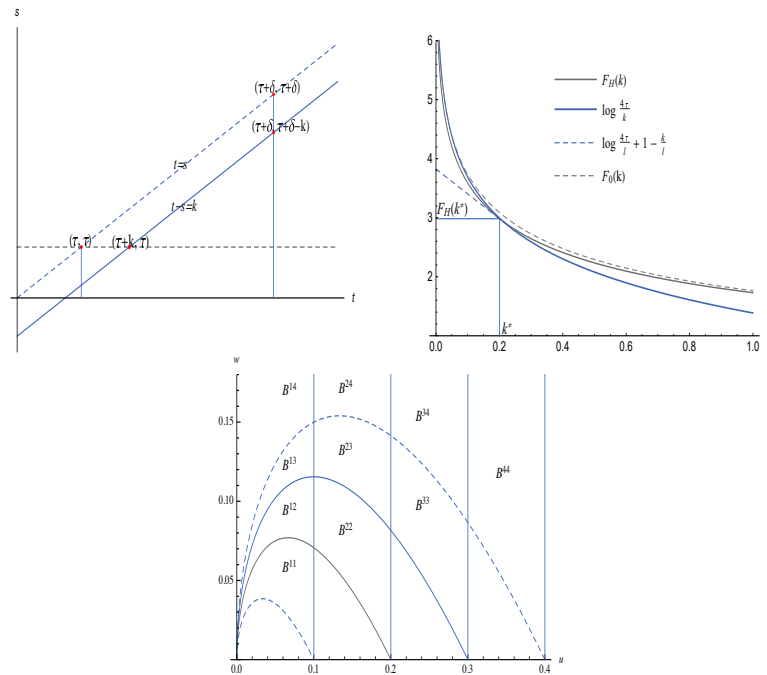


Fig. 4.2 Top left plot: $R(s,t)$ is maximized at $s = \tau + \delta - k$, and minimized at $s = \tau$. In the top right graphic, we have plotted the various quantities appearing in the lower bound part of the proof of the Sandwich Lemma with $H = .1$, $\tau = .95$ (of course in practice we care about much lower H -values but it is clearer to see what is going on here for a larger H -value so the curves are not so close to each other). Note the blue dashed line is tangential to the grey line at $k = k^*$, and the blue line has steeper slope than the grey line at this point. In the bottom graphic we we have plotted $g_H(s,t)$ for different t values for the RL process/field with $H = 0$ (left).

so $F_H(k)$ is decreasing and convex in k , and $F_H'(k) \searrow -\infty$ as $k \searrow 0$. $F_H(k)$ increases pointwise as $H \downarrow 0$ to $F_0(k) := \log \frac{1}{k} + 2 \log(\sqrt{\tau} + \sqrt{\tau+k})$. The second term is minimized at $k = 0$, so we define: $f(k) := \log \frac{4\tau}{k}$ and note that $f(k) < F_0(k)$.

4.6.2 Appendix B: Monotonicity properties of $g_H(s,t)$

The covariance of the RL process for $s < t < 1$ is $R(s,t) = \int_0^s (s-u)^{H-\frac{1}{2}} (t-u)^{H-\frac{1}{2}} du = \int_0^s u^{H-\frac{1}{2}} (t-s+u)^{H-\frac{1}{2}} du$. Differentiating this expression using the Leibniz rule we see that $R_s(s,t) = s^{H-\frac{1}{2}} t^{H-\frac{1}{2}} + (\frac{1}{2} - H) \int_0^s u^{H-\frac{1}{2}} (t-s+u)^{H-\frac{3}{2}} du$ and recall that $g_H(s,t) = \frac{1}{R_s(s,t)}$. Then we can infer monotonicity properties of g from R_s :

- By inspection R_s is a decreasing function of t , so g is increasing in t .
- For $0 < s < t$, $(t-s+u)^{H-\frac{1}{2}}$ is a smooth function of u on $[0,s]$ so the integral term in our expression for R_s is finite $\forall t > 0$. Thus $R_s(s,t)$ tends to $+\infty$ as $s \rightarrow 0$ so $g_H(0,t) = 0$ for $t > 0$.

- For $s = t > 0$ the first term in our expression for $R_s(s, t)$ is finite but the integral diverges, so we also have $g_H(t, t) = 0$.
- For $s, t \in (0, 1]^2$, $(st)^{H-\frac{1}{2}}$, $\frac{1}{2} - H$ and $u^{H-\frac{1}{2}}(t-s+u)^{H-\frac{3}{2}}$ are non-negative and decreasing in H , so $g_H(s, t)$ is increasing in H .
- By inspection, $g_H(s, t)$ is continuous for $s \in [0, t]$, and performing a Taylor series expansion of $\frac{\partial}{\partial s} g_H(s, t)(s, t)$ we can show that $\frac{\partial}{\partial s} g_H(s, t) \rightarrow -\infty$ as $s \searrow 0$ and $s \nearrow t$.

These properties can be seen in the bottom plot in Figure 4.2.

4.6.3 Appendix C: Proof of Proposition 4.4.3

We first recall that for any continuous martingale M , using Ito's lemma and integrating by parts we know that $\mathbb{E}(M_t^3) = 3\mathbb{E}(\int_0^t M_s d\langle M \rangle_s) = 3\mathbb{E}(M_t \langle M \rangle_t)$. Thus we see that

$$\begin{aligned}
 & \mathbb{E}_{t_0}((\tilde{X}_T^H - \tilde{X}_{t_0}^H)^3) \\
 &= 3\mathbb{E}_{t_0}((\tilde{X}_T^H - \tilde{X}_{t_0}^H)(\langle \tilde{X}_T^H \rangle - \langle \tilde{X}_{t_0}^H \rangle)) \\
 &= 3\mathbb{E}_{t_0}(\int_{t_0}^T \rho(s) \sqrt{V_s^H} dB_s \cdot \int_{t_0}^T V_t^H dt) \\
 &= 3\mathbb{E}_{t_0}(\int_{t_0}^T \rho(s) \xi_{t_0}^{\frac{1}{2}}(s) e^{\frac{1}{2}\gamma \int_{t_0}^s (s-u)^{H-\frac{1}{2}} dB_u - \frac{1}{2} \cdot \frac{1}{2} \gamma^2 \int_{t_0}^s (s-u)^{2H-1} du} dB_s \\
 & \quad \times \int_{t_0}^T \xi_{t_0}(t) e^{\gamma \int_{t_0}^t (t-u)^{H-\frac{1}{2}} dB_u - \frac{1}{2} \gamma^2 \int_{t_0}^t (t-u)^{2H-1} du} dt).
 \end{aligned}$$

So we (formally) need to compute

$$\begin{aligned}
 \delta I &= \mathbb{E}_{t_0}(e^{\frac{1}{2}\gamma \int_{t_0}^s (s-u)^{H-\frac{1}{2}} dB_u - \frac{1}{2} \cdot \frac{1}{2} \gamma^2 \int_{t_0}^s (s-u)^{2H-1} du} dB_s \cdot e^{\gamma \int_{t_0}^t (t-u)^{H-\frac{1}{2}} dB_u - \frac{1}{2} \gamma^2 \int_{t_0}^t (t-u)^{2H-1} du}) \\
 &= \mathbb{E}_{t_0}(e^{\gamma \int_{t_0}^t (t-u)^{H-\frac{1}{2}} dB_u + \frac{1}{2} \gamma \int_{t_0}^s (s-u)^{H-\frac{1}{2}} dB_u - (\dots)} dB_s)
 \end{aligned}$$

where (...) refers to the non-random terms. To this end, let $X = \gamma \int_{t_0}^t (t-u)^{H-\frac{1}{2}} dB_u + \frac{1}{2} \gamma \int_{t_0}^s (s-u)^{H-\frac{1}{2}} dB_u$ and $Y = dB_s$. Then $\mathbb{E}(XY) = \gamma(t-s)^{H-\frac{1}{2}} ds 1_{s < t}$ (since formally $\mathbb{E}(\frac{1}{2} \gamma \int_{t_0}^s (s-u)^{H-\frac{1}{2}} dB_u \cdot dB_s) = 0$, see end of proof for discussion on how to make this argument rigorous) and

$$\begin{aligned}
 \mathbb{E}(Ye^X) &= e^{\frac{1}{2}\mathbb{E}(X^2)} \mathbb{E}(XY) = e^{\frac{1}{2}V_H(s,t)} \gamma(t-s)^{H-\frac{1}{2}} ds 1_{s < t} \\
 \Rightarrow \delta I &= e^{-\frac{1}{2}\gamma^2 \int_{t_0}^t (t-u)^{2H-1} du - \frac{1}{2} \cdot \frac{1}{2} \gamma^2 \int_{t_0}^s (s-u)^{2H-1} du} e^{\frac{1}{2}V_H(s,t)} \gamma(t-s)^{H-\frac{1}{2}} ds 1_{s < t}
 \end{aligned}$$

where $V_H(s, t) = \gamma^2 \int_{t_0}^t [(t-u)^{H-\frac{1}{2}} + \frac{1}{2}(s-u)^{H-\frac{1}{2}} 1_{s < t}]^2 du$. Cancelling terms in the exponent, we see that δI simplifies to

$$\begin{aligned} \delta I &= e^{\frac{1}{2}\gamma^2 \int_{t_0}^s (s-u)^{H-\frac{1}{2}}(t-u)^{H-\frac{1}{2}} du - \frac{1}{8}\gamma^2 \int_{t_0}^s (s-u)^{2H-1} du} (t-s)^{H-\frac{1}{2}} ds 1_{s < t} \\ &= e^{\frac{1}{2}\gamma^2 \text{Cov}_{t_0}(Z_s^H Z_t^H) - \frac{1}{8}\text{Var}_{t_0}(Z_s^H)} (t-s)^{H-\frac{1}{2}} ds 1_{s < t} \end{aligned}$$

Then

$$\mathbb{E}_{t_0}((\tilde{X}_T^H - \tilde{X}_{t_0}^H)^3) = 3\mathbb{E}_{t_0} \int_{t_0}^T \int_{t_0}^T \rho(s) \xi_{t_0}^{\frac{1}{2}}(s) \xi_{t_0}(t) \delta I dt$$

and (4.36) and (4.37) follow. Finally we recall that a general stochastic integral $\int_0^t \phi_s dM_s$ with respect to a martingale M is defined as an L^2 -limit of $\int_0^t \phi_{\lfloor \frac{1}{n} \rfloor [ns]} dM_s$; using this construction we can rigourize the formal argument above with δI (we omit the tedious details for the sake of brevity).

4.6.4 Appendix D: Proof of skew formula

We require an expression for $\mathbb{E}[W_1 \xi_\gamma[0, 1]]$. Consider the following quantity:

$$\mathbb{E}[e^{\alpha W_1 - \frac{1}{2}\alpha^2} \xi_\gamma[0, 1](W)] = \tilde{\mathbb{E}}[\xi_\gamma[0, 1](W)] \tag{D.1}$$

$$= \mathbb{E}[\xi_\gamma[0, 1](W + \alpha h)] \tag{D.2}$$

$$= \mathbb{E}\left[\int_0^1 e^{\langle Y^0(t), \alpha h \rangle} \xi_\gamma[dt](W)\right] \tag{D.3}$$

$$= \int_0^1 e^{\langle Y^0(t), \alpha h \rangle} dt \tag{D.4}$$

where in the first line we have changed the measure, in the second we have applied the Cameron-Martin theorem and in the third and fourth we have used the Shamov definition of the GMC. Here h is simply the identity function.

Differentiating both sides with respect to α and then setting $\alpha = 0$ yields

$$\mathbb{E}[W_1 \xi_\gamma[0, 1]] = \int_0^1 \langle Y^0(t), h \rangle dt = \int_0^1 \gamma \frac{t^{1/2}}{1/2} dt = \frac{\gamma}{\frac{1}{2} \cdot \frac{3}{2}} \tag{D.5}$$

which gives the result.

References

- [A07] Andersen, Leif. (2007). Efficient Simulation of the Heston Stochastic Volatility Model. *J. Computat. Finance*. 11. 10.2139/ssrn.946405.
- [A19a] Abi Jaber, Eduardo. "Lifting the Heston model." *Quantitative Finance* 19.12 (2019): 1995-2013.
- [A19b] Abi Jaber, Eduardo, "Weak existence and uniqueness for affine stochastic Volterra equations with L^1 -kernels"
- [A20] Abi Jaber, Eduardo, "The characteristic function of Gaussian stochastic volatility models: an analytic expression." arXiv preprint arXiv:2009.10972 (2020).
- [AC01] Almgren, R. and N.Chriss, "Optimal execution of portfolio transactions", *J.Risk*, 3:5,39, 2001.
- [ACLP19] Abi Jaber, C.Cuchiero, E., M.Larsson, and S.Pulido, "A weak solution theory for stochastic Volterra equations of convolution type", 2019, to appear in *Annals of Applied Probability*.
- [AE19a] Abi Jaber, Eduardo, and Omar El Euch. "Multifactor approximation of rough volatility models." *SIAM Journal on Financial Mathematics* 10.2 (2019): 309-349.
- [AE19b] Abi Jaber, E. and O.El Euch, "Markovian structure of the Volterra Heston model", *Statistics&Probability Letters*, Volume 149, June 2019, Pages 63-72.
- [AGM18] Elisa Alòs, D.García-Lorite, A.Muguruza, "On smile properties of volatility derivatives and exotic products: understanding the VIX skew", preprint
- [ALP19] Abi Jaber, E., M.Larsson, and S.Pulido, "Affine Volterra processes", *Annals of Applied Probability*, Volume 29, Number 5 (2019), 3155-3200.
- [ALV07] Elisa Alòs, Jorge A. León, and Josep Vives. "On the short-time behavior of the implied volatility for jump-diffusion models with stochastic volatility" *Finance and Stochastics*, vol. 11, no. 4, 2007. doi:10.1007/s00780-007-0049-1

- [B05] Bergomi, Lorenzo, Smile Dynamics II (March 1, 2005). Available at SSRN: <https://ssrn.com/abstract=1493302> or <http://dx.doi.org/10.2139/ssrn.1493302>
- [BB21] Bayer, Christian, and Simon Breneis. "Makovian approximations of stochastic Volterra equations with the fractional kernel." arXiv preprint arXiv:2108.05048 (2021).
- [BBM13] Bacry, E., Rachel Baïle, and J.Muzy, "Random cascade model in the limit of infinite integral scale as the exponential of a nonstationary $1/f$ noise: Application to volatility fluctuations in stock markets", *PHYSICAL REVIEW E* 87, 042813, 2013.
- [BDFP22] Bourgey, Florian, et al. "Local volatility under rough volatility." arXiv preprint arXiv:2204.02376 (2022).
- [BDM01] Bacry, E., J.Delour, and J.Muzy, "Multifractal Random Walks", *Phys. Rev. E*, 64, 026103-026106, 2001.
- [BDM01b] Bacry, E., J.Delour, and J.Muzy, "Modelling Financial time series using multifractal random walks", *Physica A*, 299, 84-92, 2001.
- [BDW17] Bierme, H., O.Durieu, and Y.Wang, "Generalized Random Fields and Lévy's continuity Theorem on the space of Tempered Distributions", preprint, 2017.
- [Ben88] Gérard Ben Arous (1988) Methods de laplace et de la phase stationnaire sur l'espace de wiener, *Stochastics*, 25:3, 125-153, DOI: 10.1080/17442508808833536
- [Ber17] Berestycki, N., "Introduction to the Gaussian Free Field and Liouville Quantum Gravity", draft lecture notes, updated version, Dec 2017.
- [Ber17b] Berestycki, N., "An elementary approach to Gaussian multiplicative chaos", *Electr. Comm. Probab.*, vol. 22 (2017), no.27, 1-12.
- [BFG16] Christian Bayer, Peter Friz and Jim Gatheral (2016) Pricing under rough volatility, *Quantitative Finance*, 16:6, 887-904, DOI: 10.1080/14697688.2015.1099717
- [BFGHS18] Bayer, C., P.K.Friz, A.Gulisashvili, B.Horvath, B.Stemper, "Short-Time Near-The-Money Skew In Rough Fractional Volatility Models", to appear in *Quantitative Finance*.
- [BFN22] Bayer, Christian, Masaaki Fukasawa, and Shonosuke Nakahara. "On the weak convergence rate in the discretization of rough volatility models." arXiv preprint arXiv:2203.02943 (2022).

- [BHT20] Bayer, Christian, Eric Joseph Hall, and Raúl Tempone. "Weak error rates for option pricing under linear rough volatility." arXiv preprint arXiv:2009.01219 (2020).
- [Bil99] Billingsley, P., "Convergence of probability measures", Wiley, second edition, 1999.
- [BLP17] Bennedsen, M., Lunde, A. Pakkanen, M.S. Hybrid scheme for Brownian semistationary processes. *Finance Stoch* 21, 931–965 (2017). <https://doi.org/10.1007/s00780-017-0335-5>
- [BM16] Bogachev, V.I. and A.F. Miftakhov "On weak convergence of finite-dimensional and infinite-dimensional distributions of random processes", 2016.
- [BM03] Bacry, E., and J.Muzy, "Log-Infinately Divisible Multifractal Process", *Commun. Math. Phys.*, 236, 449-475, 2003.
- [BMO20] Belak, C., J.Muhle-Karbe and K.Ou, "Liquidation in Target Zone Models", *Market Microstructure and Liquidity*, Vol. 4 (2020), No. 03, pp. 1950010.
- [Bog91] Bogachev, V.,I. "Gaussian Measures", Moscow State University, Russia.
- [Bour20] Bourgey, F., "Stochastic approximations for financial risk computations", PhD thesis, L'Ecole Polytechnique, 2020.
- [Brun17] Brunner, H., "Volterra integral equations", Cambridge University Press, Cambridge, 2017.
- [BSV17] Bank, P., H.M.Soner and M.Voß, "Hedging with Temporary Price Impact", *Mathematics and Financial Economics*, 11(2), 215-239, 2017.
- [Bue19] Buehler, Hans, et al. "Deep hedging." *Quantitative Finance* 19.8 (2019): 1271-1291.
- [CK13] Cont, R. and T.Kokholm, "A Consistent Pricing Model for Index Options and Volatility Derivatives", *Mathematical Finance*, Volume23, Issue 2, April 2013, Pages 248-27
- [CMS76] Chambers, J., C.Mallows, and B.Stuck, "A method for simulating stable random variables", 71 (1976), pp. 340-344.
- [CPR17] Chavez, H.Prado and E.G.Reyes, "A Laplace transform approach to linear equations with infinitely many derivatives and zeta-nonlocal field equations", preprint, 2017

- [DeM18] De Marco, S., “Volatility derivatives in (rough) forward variance models”, presentation, May 2018.
- [Di58] Dinghas, A., “Zur Existenz von Fixpunkten bei Abbildungen vom Abel-Liouvilleschen Typus”, *Math. Z.*, 70, p. 174-189, 1958.
- [DJR19] Dandapani, A., P.Jusselin and M.Rosenbaum, *From quadratic Hawkes processes to super-Heston rough volatility models with Zumbach effect*, arxiv preprint.
- [DRSV17] Duplantier, D., R.Rhodes, S.Sheffield, and V.Vargas, “Log-correlated Gaussian Fields: An Overview”, *Geometry, Analysis and Probability*, August pp 191-216, 2017.
- [Dup93] Dupire, B. (1993) Pricing and Hedging with Smiles. Proceeding AFFI Conference, La Baule.
- [DV07] Daley, D.J., Vere-Jones, D., “An Introduction to the Theory of Point Processes” (second edition), Springer, 2007.
- [DZ98] Dembo, A. and O.Zeitouni, “Large deviations techniques and applications”, Jones and Bartlet publishers, Boston, 1998.
- [EFGR19] El Euch, O., M.Fukasawa, J.Gatheral and M.Rosenbaum, “Short-term at-the-money asymptotics under stochastic volatility models”, *SIAM Journal on Financial Mathematics*, *SIAM J. Finan. Math.*, 10(2), 491-511.
- [EFR18] El Euch, O., M.Fukasawa, and M.Rosenbaum, “The microstructural foundations of leverage effect and rough volatility”, *Finance and Stochastics*, 12 (6), p. 241-280, 2018.
- [EGR18] El Euch, O., Gatheral, J. and M.Rosenbaum, “Roughening Heston”, *Risk*, pp. 84-89, May 2019.
- [ER18] El Euch, O. and M.Rosenbaum, “Perfect hedging in Rough Heston models”, *Annals of Applied Probability*, 28 (6), 3813-3856, 2018.
- [ER19] El Euch, O. and M.Rosenbaum, “The characteristic function of Rough Heston models”, *Mathematical Finance*, 29(1), 3-38, 2019.
- [F11] Forde, M., Large-time asymptotics for an uncorrelated stochastic volatility model, “Statistics&Probability Letters”, 81(8), 1230-1232, 2011.
- [FFGS20] Forde, M., M.Fukasawa, S.Gerhold and B.Smith, “The Rough Bergomi model as $H \rightarrow 0$ - skew flattening/blow up and non-Gaussian rough volatility”, preprint, 2020.

- [FGP18a] Friz, P.K, P.Gassiat and P.Pigato, “Precise Asymptotics: Robust Stochastic Volatility Models”, preprint.
- [FGP18b] Friz, P., S.Gerhold and A.Pinter, “Option Pricing in the Moderate Deviations Regime”, *Mathematical Finance*, 28(3), 962-988, 2018.
- [FGS21] Forde, M, Gerhold, S, Smith, B. Small-time, large-time, and urn:x-wiley:09601627:media:mafi12290:mafi12290-math-0001 asymptotics for the Rough Heston model. *Mathematical Finance*. 2021; 31: 203– 241. <https://doi.org/10.1111/mafi.12290>
- [FH05] Fleming, T.R., D.P.Harrington “Counting Processes and Survival Analysis”, Wiley, 2005
- [FHT21] Fukasawa, Masaaki, Blanka Horvath, and Peter Tankov. "Hedging under rough volatility." arXiv preprint arXiv:2105.04073 (2021).
- [FJ11] Forde, M. and A.Jacquier, “The Large-maturity smile for the Heston model”, *Finance and Stochastics*, 15, 755-780, 2011.
- [FJ11b] Forde, M. and A.Jacquier, “Small-time asymptotics for an uncorrelated Local-Stochastic volatility model”, with A.Jacquier, *Appl. Math. Finance*, 18, 517-535, 2011.
- [FJL12] Forde, M., A.Jacquier and R.Lee, “The small-time smile and term structure of implied volatility under the Heston model”, *SIAM J. Finan. Math.*, 3, 690-708, 2012.
- [FJM11] Forde, M., A.Jacquier and A.Mijatovic, “A note on essential smoothness in the Heston model”, *Finance and Stochastics*, 15, 781-784, 2011.
- [FK16] Forde, M. and R.Kumar, “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, 2016.
- [FPSS11] Fouque, J.P., PAPANICOLAOU, G., Sircar, R. and Solna, K., “Multiscale Stochastic Volatility for Equity, Interest Rate and Credit Derivatives”, Cambridge press, 2011
- [FS20] Martin Forde, Benjamin Smith, The conditional law of the Bacry–Muzy and Riemann–Liouville log correlated Gaussian fields and their GMC, via Gaussian Hilbert and fractional Sobolev spaces, *Statistics Probability Letters*, Volume 161, 2020, 108732, ISSN 0167-7152, <https://doi.org/10.1016/j.spl.2020.108732>.

- [FS21] Forde, M., and B.Smith, “Rough Heston with jumps - joint calibration to SPX/VIX level and skew as $T \rightarrow 0$, and issues with the quadratic rough Heston model”, preprint
- [FSS22] Martin Forde, Leandro Sánchez-Betancourt Benjamin Smith (2022) Optimal trade execution for Gaussian signals with power-law resilience, *Quantitative Finance*, 22:3, 585-596, DOI: 10.1080/14697688.2021.1950919
- [FSV19] Forde, M., B.Smith and L.Viitasaari, “Rough volatility and CGMY jumps with a finite history and the Rough Heston model - small-time asymptotics in the $k\sqrt{t}$ regime”, preprint, 2019.
- [FSV21] Forde, M., B.Smith and L.Viitasaari, “Rough volatility and CGMY jumps with a finite history and the Rough Heston model - small-time asymptotics in the $k\sqrt{t}$ regime”, with B.Smith and L.Viitasaari, *Quantitative Finance*, 21(4), 541-563, 21(4), 2021.
- [FSS22] Martin Forde, Leandro Sánchez-Betancourt Benjamin Smith (2022) Optimal trade execution for Gaussian signals with power-law resilience, *Quantitative Finance*, 22:3, 585-596, DOI: 10.1080/14697688.2021.1950919
- [FTW19] Fukasawa, M., T.Takabatake, and R.Westphal, “Is Volatility Rough”, preprint, 2019.
- [Fuk17] Fukasawa, M., “Short-time at-the-money skew and rough fractional volatility”, *Quantitative Finance*, 17(2), 189-198, 2017.
- [Fuk21] Fukasawa, Masaaki. "Volatility has to be rough." *Quantitative Finance* 21.1 (2021): 1-8.
- [FU21] Fukasawa, Masaaki, and Takuto Ugai. "Limit distributions for the discretization error of stochastic Volterra equations." arXiv preprint arXiv:2112.06471 (2021).
- [FZ17] Forde, M. and H.Zhang, “Asymptotics for rough stochastic volatility models”, *SIAM J.Finan.Math.*, 8, 114-145, 2017.
- [G22] Gatheral, Jim, Efficient Simulation of Affine Forward Variance Models (June 29, 2021). *Risk*, February 2022, Available at SSRN: <https://ssrn.com/abstract=3876680> or <http://dx.doi.org/10.2139/ssrn.3876680>
- [G86] Gyöngy, I. (1986). Mimicking the one-dimensional marginal distributions of processes having an Itô differential. *Probab. Theory Relat. Fields*, 71(4):501–516
- [Gath21] Gatheral, J., “Diamond trees and the forest expansion”, Bloomberg Quant (BBQ) Seminar, January 2021 (joint work with E.Alòs, P.Friz and R.Radoičič).

- [Gas22] Gassiat, Paul. "Weak error rates of numerical schemes for rough volatility." arXiv preprint arXiv:2203.09298 (2022).
- [Ger20] Gerhold, S., "Asymptotic analysis of a double integral occurring in the rough Bergomi model", 2020, to appear in *Mathematical Communications*.
- [GGP19] Gerhold, S., C.Gerstenecker and A.Pinter, "Moment Explosions In The Rough Heston Model", published online in *Decisions in Economics and Finance*.
- [Gia15] Giambartolomei, G., "The Karhunen-Loève Theorem", PhD thesis, 2015.
- [GJ11] Gatheral, J. and A.Jacquier, "Convergence of Heston to SVT", *Quant.Finance*, 11(8), 1129-1132, 2011.
- [GJPSW19] Gerhold, S., Jacquier, A., Pakkanen, M., Stone, H., Wagenhofer, T. (2021). Pathwise large deviations for the rough Bergomi model: Corrigendum. *Journal of Applied Probability*, 58(3), 849-850. doi:10.1017/jpr.2020.109
- [GJR18] Jim Gatheral, Thibault Jaisson Mathieu Rosenbaum (2018) Volatility is rough, *Quantitative Finance*, 18:6, 933-949, DOI: 10.1080/14697688.2017.1393551
- [GJR20] Gatheral, J., P.Jusselin and M.Rosenbaum, "The quadratic rough Heston model and the joint S&P 500/VIX smile calibration problem", *RISK magazine*, 2020.
- [GK19] Gatheral, J. and M.Keller-Ressel, "Affine forward variance models", *Finance and Stochastics*, volume 23, pages 501-533, 2019.
- [GL14] Gao, K., and Lee, R., "Asymptotics of implied volatility to arbitrary order", *Finance Stoch.*, 18, 349-392, 2014.
- [GLS90] Gripenberg, G, S.O.Londen, and O.Staffans, "Volterra Integral and Functional Equations", Cambridge University Press, 1990.
- [GLW22] Guo, Ivan, Grégoire Loeper, and Shiyi Wang. "Calibration of local-stochastic volatility models by optimal transport." *Mathematical Finance* 32.1 (2022): 46-77.
- [GLWO20] Guo, Ivan, et al. "Joint modelling and calibration of SPX and VIX by optimal transport." arXiv preprint arXiv:2004.02198 (2020).
- [GMN17] Guyon, J., R.Menegaux and M.Nutz, "Bounds for VIX futures given S&P 500 smiles",
- [GSS12] Gatheral, J., A.Schied, and A.Slynko, "Transient linear price impact and Fredholm integral equations", *Math. Finance*, 22:445-474, 2012.

- [GR19] Gatheral, J. and R.Radoičić, “Rational Approximation of the Rough Heston Solution” *International Journal of Theoretical and Applied Finance*, Vol. 22, No. 3, 1950010, 2019
- [Guy21] Guyon, Julien, Dispersion-Constrained Martingale Schrödinger Problems and the Exact Joint SP 500/VIX Smile Calibration Puzzle (May 25, 2021). Available at SSRN: <https://ssrn.com/abstract=3853237> or <http://dx.doi.org/10.2139/ssrn.3853237>
- [Guy20] Guyon, J., “The Joint S&P 500/VIX Smile Calibration Puzzle Solved”, *RISK* magazine, 2020.
- [Guy20b] Guyon, J., “VIX-constrained Schrödinger bridges: joint S&P 500/VIX Smile Calibration calibration with continuous Stochastic Volatility Models”, presentation, Nov 2020.
- [Guy21] Guyon, J., “Dispersion-Constrained Martingale Schrödinger Problems and the exact joint S&P 500/Vix Smile Calibration puzzle”, preprint, 2021
- [Guy21b] Guyon, Julien, The smile of stochastic volatility: Revisiting the Bergomi-Guyon expansion (November 4, 2021). Available at SSRN: <https://ssrn.com/abstract=3956786> or <http://dx.doi.org/10.2139/ssrn.3956786> *Finance and Stochastics*, Volume 21, pages 593-630, 2017.
- [GVZ19] Gulisashvili, A., F.Viens, X.Zhang, “Extreme-Strike Asymptotics for General Gaussian Stochastic Volatility Models”, *Annals of Finance*, volume 15, pages 59-101, 2019.
- [H93] Heston, S.L. (1993) A Closed Solution for Options with Stochastic Volatility, with Application to Bond and Currency Options. *Review of Financial Studies*, 6, 327-343. <https://doi.org/10.1093/rfs/6.2.327>
- [HJM17] Horvath, B., Jacquier, A., Muguruza, A. (2017). Functional Central Limit Theorems for Rough Volatility. Big Data Innovative Financial Technologies Research Paper Series.
- [HMT19] Horvath, Blanka, Aitor Muguruza, and Mehdi Tomas. "Deep learning volatility." arXiv preprint arXiv:1901.09647 (2019).
- [HJT20] Horvath, B., A. Jacquier and P. Tankov, “Volatility options in rough volatility models”, *SIAM Journal on Financial Mathematics*, 11(2): 437-469, 2020.
- [HN20] Hager, P. and E. Neuman, “The multiplicative chaos of $H = 0$ Fractional Brownian Fields”, preprint.

- [HL19] Henry-Labordère, P., “From (Martingale) Schrodinger bridges to a new class of Stochastic Volatility Models’, preprint, 2019
- [HTZ21] Horvath, Blanka, Josef Teichmann, and Žan Žurič. "Deep hedging under rough volatility." *Risks* 9.7 (2021): 138.
- [Jan97] Janson, S., “Gaussian Hilbert spaces”, Cambridge University Press, 1997.
- [JMP21] Jacquier, A., Muguruza, A. and A.Pannier, “Rough Multifactor Volatility for SPX and VIX options’, preprint, 2021
- [JP20] Jacquier, A. and A.Pannier, “Large and moderate deviations for stochastic Volterra systems”, preprint, 2020.
- [JR16] Jaisson, T. and M.Rosenbaum, “Rough fractional diffusions as scaling limits of nearly unstable heavy tailed Hawkes processes”, *Annals of Applied Probability*, 26 (5), 2860-2882, 2016.
- [JR18] P.Jusselin and M.Rosenbaum, “No-arbitrage implies power-law market impact and rough volatility”, 2018, to appear in *Mathematical Finance*.
- [JS13] Jacod, J. and A.Shiryaev, “Limit theorems for stochastic processes”, volume 288, Springer Science&Business Media, 2013.
- [JPS18] Jacquier A, Pakkanen MS, Stone H, 2018, Pathwise large deviations for the rough Bergomi model, *Journal of Applied Probability*, Vol: 55, Pages: 1078-1092, ISSN: 0021-9002
- [JP20] Jacquier, A. and A.Pannier, “Large and Moderate Deviations for Stochastic Volterra Systems”, preprint
- [JS03] Jacod, J. and A.Shiryaev, “Limit theorems for Stochastic Processes”, Springer, second edition, 2003.
- [JS13] Jacod, J. and A.Shiryaev, “Limit theorems for stochastic processes”, volume 288, Springer Science&Business Media, 2013.
- [JSW19] Junnila, J., E.Saksman, C.Webb, “Decompositions of Log-Correlated Fields With application”, *Annals of Applied Probability*, Volume 29, Number 6 (2019), 3786-3820.
- [Kah85] Kahane, J.-P., “Sur le chaos multiplicatif”, *Ann. Sci. Math.*, Québec, 9 (2), 105-150, 1985.

- [Koz06] Kozhemyak, A., “Modélisation de séries financières á l’aide de processus invariants d’échelle. Application a la prediction du risque.”, thesis, Ecole Polytechnique, 2006,
- [KS91] Karatzas, I. and S.Shreve, “Brownian motion and Stochastic Calculus”, Springer-Verlag, 1991.
- [L04] Lee, Roger Avellaneda, Marco Carr, Peter Dembo, Amir Duffie, Darrell Hedegaard, Esben Papanicolaou, George Wu, Liuren. (2004). Option pricing by transform methods: extensions, unification and error control. *Journal of Computational Finance*. 7. 51-86.
- [LK07] Lord, R. and C.Kahl, “Optimal Fourier Inversion in Semi-Analytical Option Pricing”, *Tinbergen Institute Discussion Paper No. 2006-066/2*.
- [MF71] Miller, R.K. and A.Feldstein, “Smoothness of solutions Of Volterra Integral Equations with weakly singular kernels”, *Siam J. Math. Anal.*, Vol. 2, No. 2, 1971.
- [MW51] Mann, W.R. and F.Wolf, “Heat transfer between solids and gases under nonlinear boundary conditions”, *Quarterly of Applied Mathematics*, Vol. 9, No. 2, pp. 163-184, 1951.
- [NR18] Neuman, E. and M.Rosenbaum, “Fractional Brownian motion with zero Hurst parameter: a rough volatility viewpoint”, *Electronic Communications in Probability*, 23(61), 2018.
- [Olv74] Olver, F.W., “Asymptotics and Special Functions”, Academic Press, 1974.
- [Rom22] Rømer, Sigurd Emil, Hybrid multifactor scheme for stochastic Volterra equations with completely monotone kernels (May 3, 2022). Available at SSRN: <https://ssrn.com/abstract=3706253> or <http://dx.doi.org/10.2139/ssrn.3706253>
- [RO96] Roberts, C.A., and Olmstead, W.E., “Growth rates for blow-up solutions of nonlinear Volterra equations”, *Quart. Appl. Math.*, 54(1): 153-159, 1996.
- [RT96] Renault, E. and N.Touzi, “Option hedging and implied volatilities in a stochastic volatility model”, *Mathematical Finance*, 6(3):279-302, July 1996.
- [RV10] Robert, R. and V.Vargas, “Gaussian multiplicative chaos Revisited”, *Annals of Probability*, 38, 2, 605-631, 2010.
- [RV14] Rhodes, R. and V.Vargas, “Gaussian multiplicative chaos and applications: a review”, *Probab. Surveys*, 11, 315-392, 2014.

- [RY99] Revuz, D. and M.Yor, “Continuous martingales and Brownian motion”, Springer-Verlag, Berlin, 3rd edition, 1999.
- [Sha16] Shamov, A., “On Gaussian multiplicative chaos”, *Journal of Functional Analysis*, 270(9), 3224-3261, 2016.
- [SKM93] Samko.S., A.Kilbas and O.Marichev, “Fractional integrals and derivatives”, Gordon and Breach Science publishers, 1993.
- [Zha08] Zhang, X., “Euler schemes and large deviations for stochastic Volterra equations with singular kernels”, *Journal of Differential Equations*, 244:2226–2250, 2008