Rough volatility: An overview

Jim Gatheral
(joint work with Christian Bayer, Peter Friz, Omar El Euch,
Masaaki Fukasawa, Thibault Jaisson, and Mathieu Rosenbaum)

Baruch College
The City University of New York

Advances in Financial Mathematics
Paris, January 10, 2017

Outline of this talk

- The term structure of the implied volatility skew
- A remarkable monofractal scaling property of historical volatility
- Fractional Brownian motion (fBm)
- The Rough Fractional Stochastic Volatility (RFSV) model
- Microstructural foundation
- The Rough Bergomi (rBergomi) model
- Fits to SPX
- Forecasting the variance swap curve
- Possible approaches to model calibration

The SPX volatility surface as of 15-Sep-2005

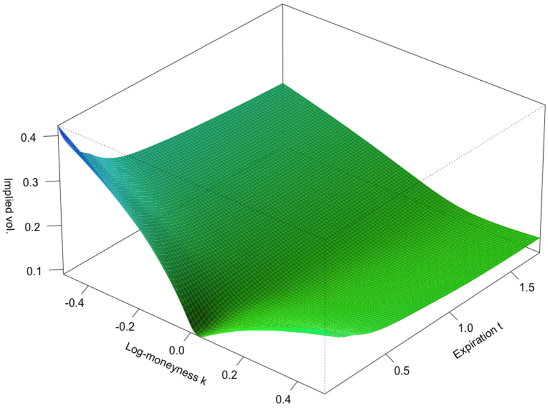

Figure 1: The SPX volatility surface as of 15-Sep-2005 (Figure 3.2 of The Volatility Surface).

Term structure of at-the-money skew

- Given one smile for a fixed expiration, little can be said about the process generating it.
- In contrast, the dependence of the smile on time to expiration is intimately related to the underlying dynamics.
    - In particular model estimates of the term structure of ATM volatility skew defined as

      $$\psi(\tau) := \left| \frac{\partial}{\partial k} \sigma_{\mathrm{BS}}(k, \tau) \right|_{k=0}.$$

      are very sensitive to the choice of volatility dynamics in a stochastic volatility model.

Term structure of SPX ATM skew as of 15-Sep-2005

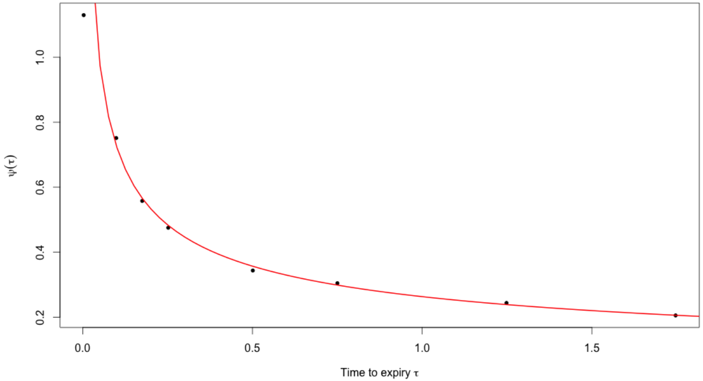

Figure 2: Term structure of ATM skew as of 15-Sep-2005, with power law fit $\tau^{-0.44}$ superimposed in red.

Stylized facts

- Although the levels and orientations of the volatility surfaces change over time, their rough shape stays very much the same.
  - It's then natural to look for a time-homogeneous model.
- The term structure of ATM volatility skew

$$\psi(\tau) \sim \frac{1}{\tau^{\alpha}}$$

with $\alpha \in (0.3, 0.5)$.

Motivation for Rough Volatility I: Better fitting stochastic volatility models

- Conventional stochastic volatility models generate volatility surfaces that are inconsistent with the observed volatility surface.
  - In stochastic volatility models, the ATM volatility skew is constant for short dates and inversely proportional to T for long dates.
  - Empirically, we find that the term structure of ATM skew is proportional to $1/T^{\alpha}$ for some $0 < \alpha < 1/2$ over a very wide range of expirations.
- The conventional solution is to introduce more volatility factors, as for example in the DMR and Bergomi models.
  - One could imagine the power-law decay of ATM skew to be the result of adding (or averaging) many sub-processes, each of which is characteristic of a trading style with a particular time horizon.

Bergomi Guyon

- Define the forward variance curve $\xi_t(u) = \mathbb{E}\left[v_u \middle| \mathcal{F}_t\right]$.
- According to [BG12], in the context of a variance curve model, implied volatility may be expanded as

$$\sigma_{\mathrm{BS}}(k,T) = \sigma_0(T) + \sqrt{\frac{w}{T}}\,\frac{1}{2\,w^2}\,C^{x\,\xi}\,k + O(\eta^2) \qquad (1)$$

where η is volatility of volatility, $w = \int_0^T \xi_0(s)\,ds$ is total variance to expiration T, and

$$C^{x\,\xi} = \int_0^T dt \int_t^T du\,\frac{\mathbb{E}\left[dx_t\,d\xi_t(u)\right]}{dt}. \qquad (2)$$

- Thus, given a stochastic model, defined in terms of an SDE, we can easily (at least in principle) compute this smile approximation.

The Bergomi model

- The n-factor Bergomi variance curve model reads:

$$\xi_t(u) = \xi_0(u)\, \exp\left\{ \sum_{i=1}^{n} \eta_i \int_0^t e^{-\kappa_i\,(t-s)}\, dW_s^{(i)} + \text{ drift } \right\}. \tag{3}$$

- The Bergomi model generates a term structure of volatility skew $\psi(\tau)$ that is something like

$$\psi(\tau) = \sum_i \frac{1}{\kappa_i\,\tau} \left\{ 1 - \frac{1 - e^{-\kappa_i\,\tau}}{\kappa_i\,\tau} \right\}.$$

  - This functional form is related to the term structure of the autocorrelation function.
  - Which is in turn driven by the exponential kernel in the exponent in (3).

Tinkering with the Bergomi model

- Empirically, $\psi(\tau) \sim \tau^{-\alpha}$ for some α.
- It's tempting to replace the exponential kernels in (3) with a power-law kernel.
- This would give a model of the form

$$\xi_t(u) = \xi_0(u) \exp\left\{\eta \int_0^t \frac{dW_s}{(t-s)^\gamma} + \text{ drift }\right\}$$

which looks similar to

$$\xi_t(u) = \xi_0(u) \exp\left\{\eta\, W_t^H + \text{ drift }\right\}$$

where W_t^H is fractional Brownian motion.

Motivation for Rough Volatility II: Power-law scaling of the historical volatility process

- The Oxford-Man Institute of Quantitative Finance makes historical realized variance (RV) estimates freely available at http://realized.oxford-man.ox.ac.uk. These estimates are updated daily.
- Using daily RV estimates as proxies for instantaneous variance, we may investigate the time series properties of v_t empirically.

SPX realized variance from 2000 to 2016

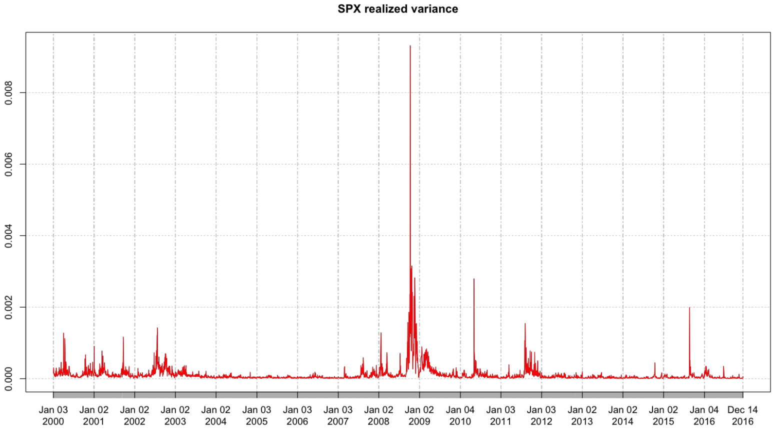

Figure 3: KRV estimates of SPX realized variance from 2000 to 2016.

The smoothness of the volatility process

- For $q \geq 0$, we define the qth sample moment of differences of log-volatility at a given lag Δ[1]:

$$m(q, \Delta) = \langle |\log \sigma_{t+\Delta} - \log \sigma_t|^q \rangle$$

- For example

$$m(2, \Delta) = \langle (\log \sigma_{t+\Delta} - \log \sigma_t)^2 \rangle$$

  is just the sample variance of differences in log-volatility at the lag Δ.

[1] $\langle \cdot \rangle$ denotes the sample average.

Scaling of $m(q, \Delta)$ with lag Δ

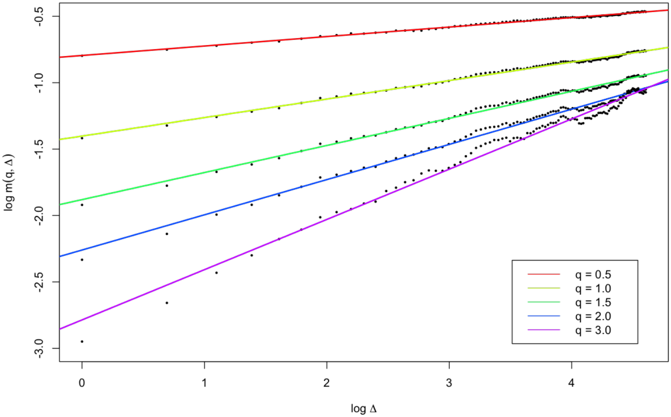

Figure 4: $\log m(q, \Delta)$ as a function of $\log \Delta$, SPX.

Scaling of ζ_q with q

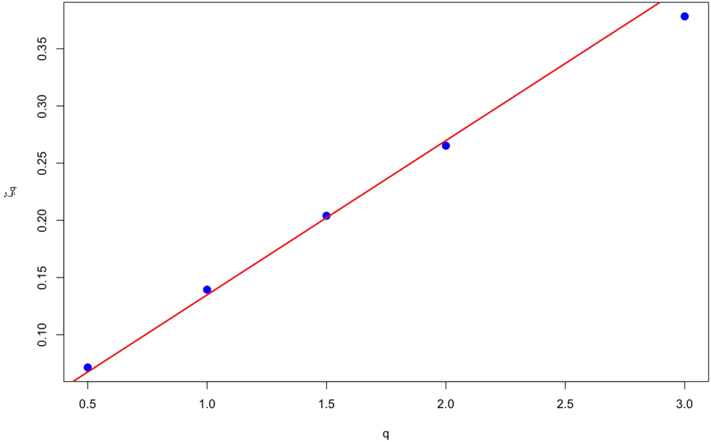

Figure 5: Scaling of ζ_q with q.

Monofractal scaling result

- From the log-log plot Figure 4, we see that for each q, $m(q, \Delta) \propto \Delta^{\zeta_q}$.
- And from Figure 5 the monofractal scaling relationship

$$\zeta_q = q\, H$$

with $H \approx 0.13$.

  - Note also that our estimate of H is biased high because we proxied instantaneous variance v_t with its average over each day $\frac{1}{T} \int_0^T v_t\, dt$, where T is one day.
  - On the other hand, the time series of realized variance is noisy and this causes our estimate of H to be biased low.

- A time series of H for SPX following the methodology of [BLP16] is shown in the next figure.

The time series of $\hat{\alpha} = H - \frac{1}{2}$ for SPX

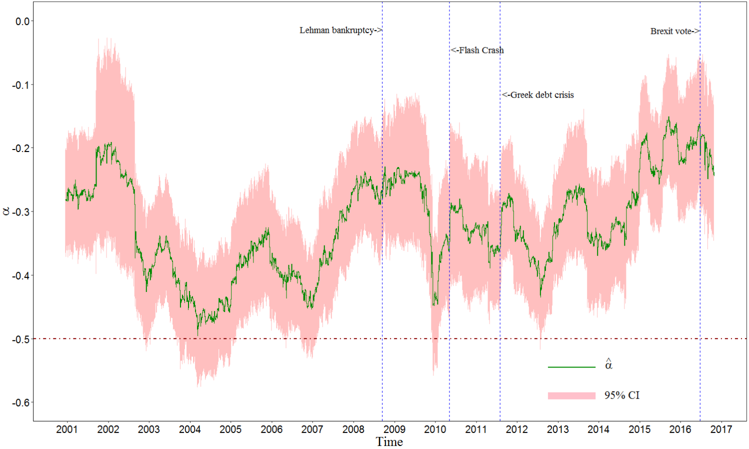

Figure 10: Half year rolling-window estimates of α on the realized variance measures of the daily volatility by variogram OLS regression (3.10) with $m = 3$. The pink area is the 95% confidence interval by bootstrap method with $B = 999$. The four vertical dashed blue lines indicate four periods of market turmoil: Lehman Brothers filing for bankruptcy, the Flash Crash, the first bailout during Greek debt crisis and the Brexit referendum.

Distributions of $(\log \sigma_{t+\Delta} - \log \sigma_t)$ for various lags Δ

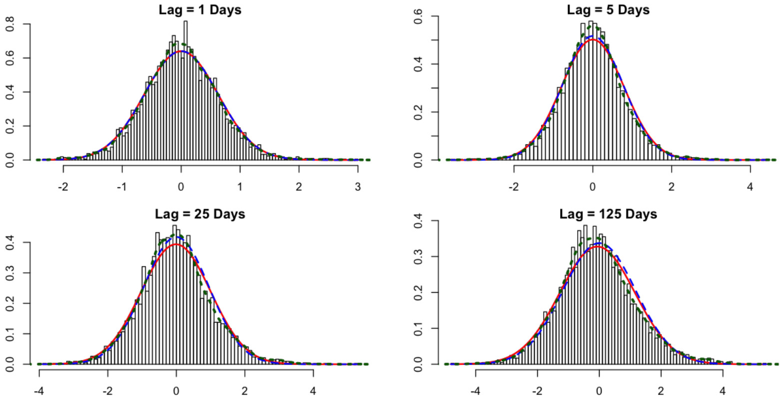

Figure 6: Histograms of $(\log \sigma_{t+\Delta} - \log \sigma_t)$ for various lags Δ; normal fit in red; $\Delta = 1$ normal fit scaled by $\Delta^{0.14}$ in blue.

Estimated H for all indices

Estimating the relationship

$$\langle (\log \sigma_{t+\Delta} - \log \sigma_t)^2 \rangle = \nu^2\, \Delta^{2\,H}$$

for all 21 indices in the Oxford-Man dataset yields:

Index	H	ν
SPX2.rk	0.13	0.32
FTSE2.rk	0.14	0.27
N2252.rk	0.11	0.33
GDAXI2.rk	0.15	0.28
RUT2.rk	0.12	0.33
AORD2.rk	0.08	0.36
DJI2.rk	0.13	0.32
IXIC2.rk	0.13	0.30
FCHI2.rk	0.13	0.29
HSI2.rk	0.10	0.28
KS11.rk	0.12	0.28
AEX.rk	0.14	0.29
SSMI.rk	0.18	0.22
IBEX2.rk	0.13	0.28
NSEI.rk	0.11	0.32
MXX.rk	0.09	0.33
BVSP.rk	0.11	0.31
GSPTSE.rk	0.12	0.31
STOXX50E.rk	0.12	0.34
FTSTI.rk	0.13	0.23
FTSEMIB.rk	0.13	0.30

Universality?

- In [GJR14], we compute daily realized variance estimates over one hour windows for DAX and Bund futures contracts, finding similar scaling relationships.
- We have also checked that Gold and Crude Oil futures scale similarly.
  - Although the increments $(\log \sigma_{t+\Delta} - \log \sigma_t)$ seem to be fatter tailed than Gaussian.
- In [BLP16] Bennedsen et al., estimate volatility time series for more than five thousand individual US equities, finding rough volatility in every case.

A natural model of realized volatility

- Distributions of differences in the log of realized volatility are close to Gaussian.
  - This motivates us to model σ_t as a lognormal random variable.
- Moreover, the scaling property of variance of RV differences suggests the model:

$$\log \sigma_{t+\Delta} - \log \sigma_t = \nu \left(W^H_{t+\Delta} - W^H_t \right) \tag{4}$$

where W^H is fractional Brownian motion.

- In [GJR14], we refer to a stationary version of (4) as the RFSV (for Rough Fractional Stochastic Volatility) model.

Fractional Brownian motion (fBm)

- *Fractional Brownian motion* (fBm) $\{W_t^H; t \in \mathbb{R}\}$ is the unique Gaussian process with mean zero and autocovariance function

$$\mathbb{E}\left[W_t^H W_s^H\right] = \frac{1}{2}\left\{|t|^{2H} + |s|^{2H} - |t-s|^{2H}\right\}$$

  where $H \in (0,1)$ is called the *Hurst index* or parameter.
- In particular, when $H = 1/2$, fBm is just Brownian motion.
  - If $H > 1/2$, increments are positively correlated.
  - If $H < 1/2$, increments are negatively correlated.

Apparent fractality of the volatility time series

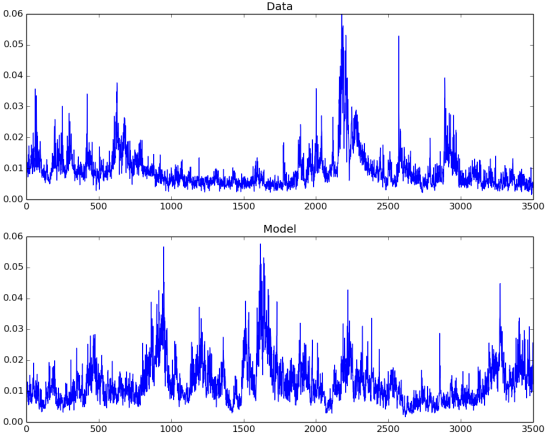

Figure 7: Volatility of SPX (above) and in the RFSV model (below).

Remarks on the comparison

- The qualitative features of simulated and actual graphs look very similar.
  - Persistent periods of high volatility alternate with low volatility periods.
- $H \sim 0.1$ generates very rough looking sample paths (compared with $H = 1/2$ for Brownian motion).
  - Hence *rough volatility*.
- On closer inspection, we observe fractal-type behavior.
  - The graph of volatility over a small time period looks like the same graph over a much longer time period.
- This feature of volatility has been investigated both empirically and theoretically in, for example, [BM03].
  - In particular, their Multifractal Random Walk (MRW) is related to a limiting case of the RSFV model as $H \to 0$.

A Hawkes model of price formation

In [EFR16], El Euch, Fukasawa and Rosenbaum consider a generalization of a simple model of price dynamics in terms of Hawkes processes due to Bacry et al. ([BM14]) with the following properties:

- Reflecting the high degree of endogeneity of the market, the L^1 norm of the kernel matrix is close to one (nearly unstable).
- No drift in the price process imposes a relationship between buy and sell kernels.
- Liquidity asymmetry: The average impact of a sell order is greater than the impact of a buy order.
- Splitting of metaorders motivates power-law decay of the Hawkes kernels $\varphi(\tau) \sim \tau^{-(1+\alpha)}$ (empirically $\alpha \approx 0.6$).

The scaling limit of the price model

They construct a sequence of such Hawkes processes suitably rescaled in time and space that converges in law to a Rough Heston process of the form

$$\begin{aligned} \frac{dS_t}{S_t} &= \sqrt{v_t}\, dZ_t \\ v_t &= v_0 + \frac{\lambda}{\Gamma(\alpha)} \int_0^t \frac{\theta - v_s}{(t-s)^{1-\alpha}}\, ds + \frac{\lambda\, \nu}{\Gamma(\alpha)} \int_0^t \frac{\sqrt{v_s}\, dW_s}{(t-s)^{1-\alpha}} \end{aligned}$$

with

$$d\langle Z, W\rangle_t = \rho\, dt.$$

- The correlation ρ is related to a liquidity asymmetry parameter.
- Rough volatility can thus be understood as relating to the persistence of order flow and the high degree of endogeneity of liquid markets.

The characteristic function

Define the fractional integral and differential operators:

$$I^{1-\alpha} f(t) = \frac{1}{\Gamma(1-\alpha)} \int_0^t \frac{f(s)}{(t-s)^\alpha}\, ds; \quad D^\alpha f(t) = \frac{d}{dt}\, I^{1-\alpha} f(t).$$

Remarkably, in [ER16], El Euch and Rosenbaum compute the following expression for the characteristic function of the Rough Heston model:

$$\phi_t(u) = \exp\left\{\theta\,\lambda \int_0^t h(u,s)\, ds + v_0\, I^{1-\alpha} h(u,t)\right\}$$

where $h(u,)$ solves the fractional Riccati equation

$$D^\alpha h(u,s) = -\frac{1}{2}\, u\,(u+i) + \lambda\,(i\,\rho\,\nu\,u - 1)\, h(u,s) + \frac{(\lambda\,\nu)^2}{2}\, h^2(u,s).$$

Representations of fBm

There are infinitely many possible representations of fBm in terms of Brownian motion. For example, with $\gamma = \frac{1}{2} - H$,

Mandelbrot-Van Ness

$$W_t^H = C_H \left\{ \int_{-\infty}^{t} \frac{dW_s}{(t-s)^\gamma} - \int_{-\infty}^{0} \frac{dW_s}{(-s)^\gamma} \right\}.$$

where the choice

$$C_H = \sqrt{\frac{2\,H\,\Gamma(3/2-H)}{\Gamma(H+1/2)\,\Gamma(2-2\,H)}}$$

ensures that

$$\mathbb{E}\left[W_t^H\,W_s^H\right] = \frac{1}{2}\left\{t^{2H} + s^{2H} - |t-s|^{2H}\right\}.$$

Pricing under rough volatility

Once again, the data suggests the following model for volatility under the real (or historical or physical) measure $\mathbb{P}$:

$$\log \sigma_t = \nu \, W_t^H .$$

Let $\gamma = \frac{1}{2} - H$. We choose the Mandelbrot-Van Ness representation of fractional Brownian motion W^H as follows:

$$W_t^H = C_H \left\{ \int_{-\infty}^{t} \frac{dW_s^{\mathbb{P}}}{(t-s)^\gamma} - \int_{-\infty}^{0} \frac{dW_s^{\mathbb{P}}}{(-s)^\gamma} \right\}$$

where the choice

$$C_H = \sqrt{\frac{2\,H\,\Gamma(3/2-H)}{\Gamma(H+1/2)\,\Gamma(2-2\,H)}}$$

ensures that

$$\mathbb{E}\left[W_t^H \, W_s^H\right] = \frac{1}{2}\left\{t^{2H} + s^{2H} - |t-s|^{2H}\right\}.$$

Pricing under rough volatility

Then

$$
\begin{aligned}
& \log v_u - \log v_t \\
= \; & \nu \, C_H \left\{ \int_t^u \frac{1}{(u-s)^\gamma} \, dW_s^{\mathbb{P}} + \int_{-\infty}^t \left[\frac{1}{(u-s)^\gamma} - \frac{1}{(t-s)^\gamma} \right] dW_s^{\mathbb{P}} \right\} \\
=: \; & 2 \, \nu \, C_H \left[M_t(u) + Z_t(u) \right]. \qquad (5)
\end{aligned}
$$

- Note that $\mathbb{E}^{\mathbb{P}}\left[M_t(u) \middle| \mathcal{F}_t \right] = 0$ and $Z_t(u)$ is $\mathcal{F}_t$-measurable.
- To price options, it would seem that we would need to know $\mathcal{F}_t$, the entire history of the Brownian motion W_s for $s < t$!

Pricing under $\mathbb{P}$

Let

$$\tilde{W}_t^{\mathbb{P}}(u) := \sqrt{2\,H} \int_t^u \frac{dW_s^{\mathbb{P}}}{(u-s)^{\gamma}}$$

With $\eta := 2\,\nu\,C_H/\sqrt{2\,H}$ we have $2\,\nu\,C_H\,M_t(u) = \eta\,\tilde{W}_t^{\mathbb{P}}(u)$ so denoting the stochastic exponential by $\mathcal{E}(\cdot)$, we may write

$$\begin{aligned} v_u &= v_t \exp\left\{\eta \tilde{W}_t^{\mathbb{P}}(u) + 2\,\nu\,C_H\,Z_t(u)\right\} \\ &= \mathbb{E}^{\mathbb{P}}\left[v_u \middle| \mathcal{F}_t\right] \mathcal{E}\left(\eta\,\tilde{W}_t^{\mathbb{P}}(u)\right). \end{aligned} \tag{6}$$

- The conditional distribution of v_u depends on $\mathcal{F}_t$ only through the variance forecasts $\mathbb{E}^{\mathbb{P}}\left[v_u \middle| \mathcal{F}_t\right]$,
- To price options, one does not need to know $\mathcal{F}_t$, the entire history of the Brownian motion $W_s^{\mathbb{P}}$ for $s < t$.

Pricing under $\mathbb{Q}$

Our model under $\mathbb{P}$ reads:

$$v_u = \mathbb{E}^{\mathbb{P}}\left[v_u \middle| \mathcal{F}_t \right] \mathcal{E}\left(\eta\, \tilde{W}_t^{\mathbb{P}}(u) \right). \tag{7}$$

Consider some general change of measure

$$dW_s^{\mathbb{P}} = dW_s^{\mathbb{Q}} + \lambda_s\, ds,$$

where $\{\lambda_s : s > t\}$ has a natural interpretation as the price of volatility risk. We may then rewrite (7) as

$$v_u = \mathbb{E}^{\mathbb{P}}\left[v_u \middle| \mathcal{F}_t \right] \mathcal{E}\left(\eta\, \tilde{W}_t^{\mathbb{Q}}(u) \right) \exp\left\{ \eta \sqrt{2\,H} \int_t^u \frac{\lambda_s}{(u-s)^{\gamma}}\, ds \right\}.$$

- Although the conditional distribution of v_u under $\mathbb{P}$ is lognormal, it will not be lognormal in general under $\mathbb{Q}$.
  - The upward sloping smile in VIX options means λ_s cannot be deterministic in this picture.

The rough Bergomi (rBergomi) model

Let's nevertheless consider the simplest change of measure

$$dW_s^{\mathbb{P}} = dW_s^{\mathbb{Q}} + \lambda(s)\, ds,$$

where $\lambda(s)$ is a deterministic function of s. Then from (32), we would have

$$\begin{aligned} v_u &= \mathbb{E}^{\mathbb{P}}\left[v_u \middle| \mathcal{F}_t \right] \mathcal{E}\left(\eta\, \tilde{W}_t^{\mathbb{Q}}(u)\right) \exp\left\{\eta \sqrt{2H} \int_t^u \frac{1}{(u-s)^{\gamma}} \lambda(s)\, ds\right\} \\ &= \xi_t(u)\, \mathcal{E}\left(\eta\, \tilde{W}_t^{\mathbb{Q}}(u)\right) \qquad (8) \end{aligned}$$

where the forward variances $\xi_t(u) = \mathbb{E}^{\mathbb{Q}}\left[v_u \middle| \mathcal{F}_t \right]$ are (at least in principle) tradable and observed in the market.

- $\xi_t(u)$ is the product of two terms:
  - $\mathbb{E}^{\mathbb{P}}\left[v_u \middle| \mathcal{F}_t \right]$ which depends on the historical path $\{W_s, s < t\}$ of the Brownian motion
  - a term which depends on the price of risk $\lambda(s)$.

Features of the rough Bergomi model

- The rBergomi model is a non-Markovian generalization of the Bergomi model:

$$\mathbb{E}\left[v_u \middle| \mathcal{F}_t\right] \neq \mathbb{E}[v_u | v_t].$$

- The rBergomi model is Markovian in the (infinite-dimensional) state vector $\mathbb{E}^{\mathbb{Q}}\left[v_u \middle| \mathcal{F}_t\right] = \xi_t(u)$.
- We have achieved our aim of replacing the exponential kernels in the Bergomi model (3) with a power-law kernel.
  - We may therefore expect that the rBergomi model will generate a realistic term structure of ATM volatility skew.

The stock price process

- The observed anticorrelation between price moves and volatility moves may be modeled naturally by anticorrelating the Brownian motion W that drives the volatility process with the Brownian motion driving the price process.
- Thus

$$\frac{dS_t}{S_t} = \sqrt{v_t}\, dZ_t$$

with

$$dZ_t = \rho\, dW_t + \sqrt{1-\rho^2}\, dW_t^{\perp}$$

where ρ is the correlation between volatility moves and price moves.

Hybrid simulation of BSS processes

- In [BFG16], we simulate the rBergomi model by generating paths of $\tilde{W}$ and Z with the correct joint marginals using Cholesky decomposition.
  - This is very slow!
- The rBergomi variance process is a special case of a Brownian Semistationary (BSS) process.
- In [BLP15], Bennedsen et al. show how to simulate such processes more efficiently.
  - Their hybrid BSS scheme is much more efficient than the exact simulation described above.
  - However, it is still not fast enough to enable efficient calibration of the Rough Bergomi model to the volatility surface.

Guessing rBergomi model parameters

- The rBergomi model has only three parameters: H, η and ρ.
- These parameters have very direct interpretations:
  - H controls the decay of ATM skew $\psi(\tau)$ for very short expirations
  - The product $\rho\,\eta$ sets the level of the ATM skew for longer expirations.
  - Keeping $\rho\,\eta$ constant but decreasing ρ (so as to make it more negative) pushes the minimum of each smile towards higher strikes.
- So we can guess parameters in practice.

SPX smiles in the rBergomi model

- In Figure 9, we show how well a rBergomi model simulation with guessed parameters fits the SPX option market as of August 14, 2013, one trading day before the third Friday expiration.
  - Options set at the open of August 16, 2013 so only one trading day left.
- rBergomi parameters were: $H = 0.05$, $\eta = 2.3$, $\rho = -0.9$.
  - Only three parameters to get a very good fit to the whole SPX volatility surface!
- Note in particular that the extreme short-dated smile is well reproduced by the rBergomi model.
  - There is no need to add jumps!

SPX smiles as of August 14, 2013

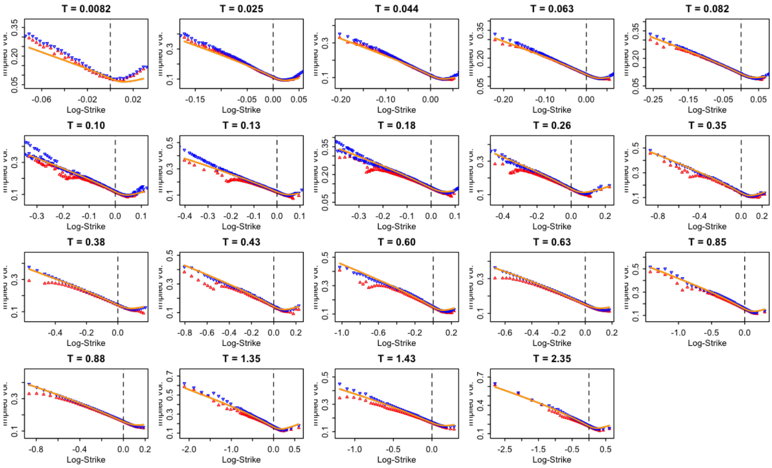

Figure 8: Red and blue points represent bid and offer SPX implied volatilities; orange smiles are from the rBergomi simulation.

The one-month SPX smile as of August 14, 2013

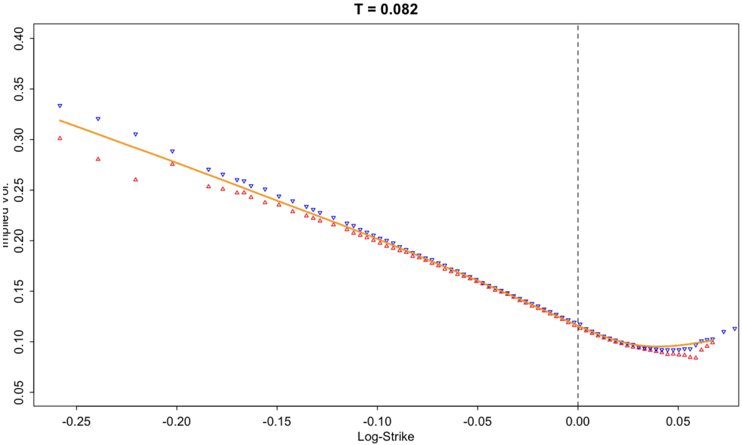

Figure 9: Red and blue points represent bid and offer SPX implied volatilities; the orange smiles is from the rBergomi simulation.

ATM volatilities and skews

In Figures 10 and 11, we see just how well the rBergomi model can match empirical ATM vols and skews. Recall also that the parameters we used are just guesses!

Term structure of ATM vol as of August 14, 2013

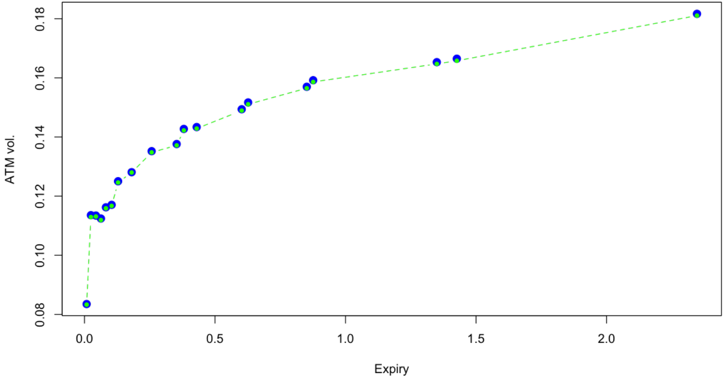

Figure 10: Blue points are empirical ATM volatilities; green points are from the rBergomi simulation. The two match very closely, as they should.

Term structure of ATM skew as of August 14, 2013

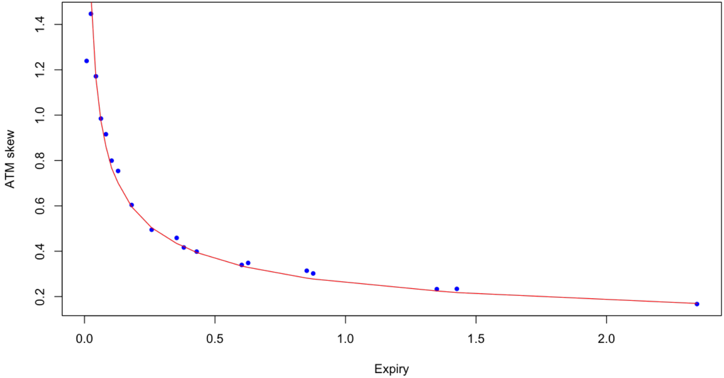

Figure 11: Blue points are empirical skews; the red line is from the rBergomi simulation.

The forecast formula

- In the RFSV model (4), $\log v_t \approx 2\,\nu\, W^H_t + C$ for some constant C.
- [NP00] show that $W^H_{t+\Delta}$ is conditionally Gaussian with conditional expectation

$$\mathbb{E}[W^H_{t+\Delta}|\mathcal{F}_t] = \frac{\cos(H\pi)}{\pi}\Delta^{H+1/2}\int_{-\infty}^{t}\frac{W^H_s}{(t-s+\Delta)(t-s)^{H+1/2}}ds$$

and conditional variance

$$\text{Var}[W^H_{t+\Delta}|\mathcal{F}_t] = c\,\Delta^{2H}.$$

where

$$c = \frac{\Gamma(3/2-H)}{\Gamma(H+1/2)\,\Gamma(2-2H)}.$$

The forecast formula

- Thus, we obtain

Variance forecast formula

$$\mathbb{E}^{\mathbb{P}}\left[v_{t+\Delta}\middle|\,\mathcal{F}_t\right] = \exp\left\{\mathbb{E}^{\mathbb{P}}\left[\log(v_{t+\Delta})\middle|\,\mathcal{F}_t\right] + 2\,c\,\nu^2\Delta^{2\,H}\right\} \tag{9}$$

where

$$\mathbb{E}^{\mathbb{P}}\left[\log v_{t+\Delta}\middle|\,\mathcal{F}_t\right] = \frac{\cos(H\pi)}{\pi}\Delta^{H+1/2}\int_{-\infty}^{t}\frac{\log v_s}{(t-s+\Delta)(t-s)^{H+1/2}}ds.$$

- [BLP16] confirm that this forecast outperforms the best performing existing alternatives such as HAR, at least at daily or higher timescales.

Calibration

- As mentioned earlier, calibration of the rBergomi model is not easy.
- We have investigated a number of approaches to calibration
  - Asymptotic expansions
  - Chebyshev interpolation
  - Moment matching
- So far, we cannot claim to have had real success with any of these approaches.

Calibration using Chebyshev interpolation

Christian Bayer and I tried calibrating the Rough Bergomi model to the volatility surface as follows:

- For a given set of 3 parameters, compute option prices using the hybrid BSS scheme [BLP15]. Compute a suitably chosen objective function.
- Following a suggestion of Kathrin Glau,
  - Repeat this 125 times on a 5x5x5 grid of Chebyshev knots.
  - Use Chebyshev interpolation to fill in the gaps.
  - Find the minimum of the objective.

- Despite that the hybrid scheme is very much faster than the Cholesky exact simulation scheme used in [BFG16], this procedure still took 2 hours running in parallel on 32 CPUs.
  - The problem is that over one million paths are needed to get Monte Carlo error down to a level that allows resolution of the minimum of the objective function.

- Another idea is to find some quantity, such as the variance swap, that is exactly computable in the model, and may be accurately estimated from market prices.

The Alòs decomposition formula

Following Elisa Alòs in [Alò12], let $X_t = \log S_t/K$ and consider the price process

$$dX_t = \sigma_t\,dZ_t - \frac{1}{2}\,\sigma_t^2\,dt. \tag{10}$$

Now let $F(X_t, w_t(T))$ (F_t for short) be some function that solves the Black-Scholes equation.

- Specifically,

$$-\partial_w F_t + \frac{1}{2}\left(\partial_{x,x} - \partial_x\right) F_t = 0. \tag{11}$$

- $w_t(T)$ is any approximation to the implied total variance $\mathcal{V}_t(T) = \mathbb{E}\left[\left.\int_t^T \sigma_s^2\,ds\right| \mathcal{F}_t\right]$ obtained by any method.

We now specify $w_t(T)$ (Bergomi-Guyon style) as:

$$w_t(T) = \int_t^T \mathbb{E}\left[\sigma_u^2 \middle| \mathcal{F}_t\right] du = \int_t^T \xi_t(u)\, du.$$

where the $\xi_t(u)$ are forward variances.

- $w_t(T)$ then represents the value of the static hedge portfolio (the log-strip) for a variance swap and is thus a *tradable asset* in the terminology of Fukasawa [Fuk14].
- For each u, $\xi_t(u)$ is a martingale in t so we may write

$$dw_t(T) = -\sigma_t^2\, dt + \int_t^T d\xi_t(u)\, du =: -\sigma_t^2\, dt + dM_t \quad (12)$$

where M is a martingale.

Applying Itô's Lemma to F, taking conditional expectations, simplifying using the Black-Scholes equation and integrating, we obtain

Theorem (The Itô Decomposition Formula of Alòs)

$$\begin{aligned}\mathbb{E}\left[F_T \middle| \mathcal{F}_t\right] &= F_t + \mathbb{E}\left[\int_t^T \partial_{x,w} F_s \, d\langle X, M\rangle_s \middle| \mathcal{F}_t\right] \\ &\quad + \frac{1}{2}\mathbb{E}\left[\int_t^T \partial_{w,w} F_s \, d\langle M, M\rangle_s \middle| \mathcal{F}_t\right]. \end{aligned} \tag{13}$$

- Note in particular that (13) is an *exact* decomposition.

Notation

We adopt the following notation for the Bergomi-Guyon autocorrelation functionals:

$$\begin{aligned} C_t^{XM}(T) &= \mathbb{E}\left[\int_t^T d\langle X, M\rangle_s \middle| \mathcal{F}_t\right] \\ C_t^{MM}(T) &= \mathbb{E}\left[\int_t^T d\langle M, M\rangle_s \middle| \mathcal{F}_t\right]. \end{aligned} \tag{14}$$

- In the notation of [BG12], $C_t^{XM}(T) = C^{x\xi}$ and $C_t^{MM}(T) = C^{\xi\xi}$.

Conditional variance of X_T

Consider

$$F_t = X_t^2 + w_t(T)\,(1 - X_t) + \frac{1}{4}\,w_t(T)^2.$$

- F_t satisfies the Black-Scholes equation and $F_T = X_T^2$.
  - $\partial_{x,w} F_t = -1$ and $\partial_{w,w} F_t = \frac{1}{2}$.
- Plugging into the Decomposition Formula (13) gives

$$\begin{aligned}
\mathbb{E}\left[X_T^2 \middle| \mathcal{F}_t\right] &= w_t(T) + \frac{1}{4}\,w_t(T)^2 - \mathbb{E}\left[\int_t^T d\langle Y, M\rangle_s \middle| \mathcal{F}_t\right] \\
&\quad + \frac{1}{4}\,\mathbb{E}\left[\int_t^T d\langle M, M\rangle_s \middle| \mathcal{F}_t\right] \\
&= w_t(T) + \frac{1}{4}\,w_t(T)^2 - C_t^{XM}(T) + \frac{1}{4}\,C_t^{MM}(T).
\end{aligned}$$

Volatility stochasticity

We can rewrite this as

Lemma

$$\zeta_t(T) := \mathrm{var}[X_T|\mathcal{F}_t] - w_t(T) = -C_t^{XM}(T) + \frac{1}{4}\,C_t^{MM}(T). \quad (15)$$

- Recall that in a stochastic volatility model, the variance of the terminal distribution of the log-underlying is not in general equal to the expected quadratic variation.
  - In the Black-Scholes model of course $\zeta_t(T) = 0$.
- We term the difference $\zeta_t(T)$ *volatility stochasticity* or just *stochasticity*.

Model calibration

Once again, equation (15) reads

$$\zeta_t(T) = -C_t^{XM}(T) + \frac{1}{4}\, C_t^{MM}(T).$$

- The LHS may be estimated from the volatility surface using the spanning formula.
  - $\zeta_t(T)$ is a tradable asset for each T.
  - We get a matching condition for each expiry T_i, $i \in \{1, ..n\}$.
- The RHS may typically be computed in a given model as a function of model parameters.
  - If so, we would be able calibrate such a model directly to tradable assets with no need for any expansion.

$\zeta_t(T)$ from the smile

Let

$$d_\pm(k) = \frac{-k}{\sigma_{\rm BS}(k,T)\sqrt{T}} \pm \frac{\sigma_{\rm BS}(k,T)\sqrt{T}}{2}$$

and following Fukasawa, denote the inverse functions by $g_\pm(z) = d_\pm^{-1}(z)$. Further define

$$\sigma(z) = \sigma_{\rm BS}(g_-(z),T)\,\sqrt{T}.$$

In terms of the implied volatility smile, it is a well-known corollary of Matytsin's characteristic function representation in [Mat00], that

$$w_t(T) = \int dz\, N'(z)\, \sigma^2(z) =: \bar{\sigma}^2.$$

Similarly, we can show that

$$\zeta_t(T) = \frac{1}{4} \int N'(z)\, \left[\sigma^2(z) - \bar{\sigma}^2\right]^2\, dz + \frac{2}{3} \int N'(z)\, z\, \sigma^3(z)\, dz. \tag{16}$$

Example: The Heston model

Consider the Heston model

$$dv_t = -\lambda\,(v_t - \theta)\,dt + \eta\,\sqrt{v}\,dW_t$$

with $d\langle W, Z\rangle_t = \rho\,dt$.

- As is typical in the Heston model, everything may be computed explicitly.
- With $\tau = T - t$,

$$w_t(T) = (v_t - \theta)\,\frac{1 - e^{-\lambda\,\tau}}{\lambda} + \theta\,\tau.$$

- Likewise we may compute both the LHS and RHS of

$$\zeta_t(T) = -C_t^{XM}(T) + \frac{1}{4}\,C_t^{MM}(T).$$

We find

$$C_t^{XM}(T) = \frac{\rho\,\eta}{\lambda^2}\left\{(v_t-\theta)\left[1-e^{-\lambda\tau}(1+\lambda\tau)\right] + \theta\left(e^{-\lambda\tau}-1+\lambda\tau\right)\right\}$$

$$C_t^{MM}(T) = \frac{\eta^2}{\lambda^3}\left\{\left(1-2\lambda\tau\,e^{-\lambda\tau}-e^{-2\lambda\tau}\right)(v_t-\theta) + \frac{1}{2}\theta\left[2\left(e^{-\lambda\tau}-1+\lambda\tau\right)-\left(1-e^{-\lambda\tau}\right)^2\right]\right\}.$$

- Compare with the small η Bergomi-Guyon expansion which gives only approximate expressions for ATM level, skew and curvature.

The Rough Bergomi model

The rBergomi model reads

$$
\begin{aligned}
S_t &= S_0\, \mathcal{E}\left(\int_0^t \sqrt{v_u}\, dZ_u\right) \\
v_u &= \xi_0(u)\, \mathcal{E}\left(\tilde{\eta} \int_0^u \frac{dW_s}{(u-s)^\gamma}\right).
\end{aligned}
$$

with $\gamma = \frac{1}{2} - H$ and $\tilde{\eta} = \eta \sqrt{2\,H}$. Then

$$
\begin{aligned}
\frac{dS_t}{S_t} &= \sqrt{\xi_t(t)}\, dZ_t, \\
\frac{d\xi_t(u)}{\xi_t(u)} &= \tilde{\eta}\, \frac{dW_t}{(u-t)^\gamma}
\end{aligned}
$$

with $\mathbb{E}[dZ_t\, dW_t] = \rho\, dt$.

$C_t^{XM}(T)$ and $C_t^{MM}(T)$ are then computed as:

$$C_t^{XM}(T) = \rho\,\tilde{\eta} \int_t^T ds\,\sqrt{\xi_t(s)} \int_s^T \xi_t(u)\,\exp\left\{\frac{\tilde{\eta}^2}{2}\,(s-t)^{2\,H}\left[G_\gamma\left(\frac{u-t}{s-t}\right)-\frac{1}{8H}\right]\right\}$$

and

$$C_t^{MM}(T) = 2\,\tilde{\eta}^2 \int_t^T \xi_t(v)\,dv \int_t^v \xi_t(u)\,du\left[\exp\left\{\tilde{\eta}^2\,(u-t)^{2H}\,G_\gamma\left(\frac{v-t}{u-t}\right)\right\}-1\right].$$

where for $y \geq 1$,

$$\begin{aligned} G_\gamma(y) &= \int_0^1 \frac{dr}{(y-r)^\gamma\,(1-r)^\gamma} \\ &= \frac{1}{(1-\gamma)\,(y-1)}\,y^{1-\gamma}\,{}_2F_1\left(1, 2-2\gamma; 2-\gamma; \frac{y}{y-1}\right). \end{aligned}$$

A numerical experiment

- We start with SPX options as of February 4, 2010, noting all strikes and expirations with nonzero bid prices.
- Starting from model with parameters chosen to more or less fit the observed smiles, for these strikes and expirations, we replace market option prices with model option prices and compute implied volatilities.
- We then check to see how consistent robust estimates of stochasticity from these (fake) market smiles are with known values.

Heston stochasticity: robust estimates vs exact

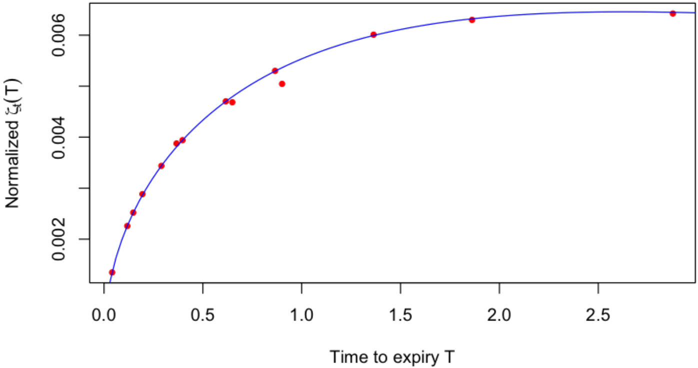

Figure 12: Plot of $\frac{\zeta_t(T)}{T^{3/2}}$ vs time to expiry. The blue line is the exact Heston formula, the red dots are robust estimates from the Heston implied volatility smiles using (16).

Remarks on the experiment

- In Figure 12, we note that some of the red points are off.
  - For these expirations, there are insufficient strikes to accurately estimate the integrals in

$$\zeta_t(T) = \frac{1}{4}\int N'(z)\,\left[\sigma^2(z) - \bar{\sigma}^2\right]^2\,dz + \frac{2}{3}\int N'(z)\,z\,\sigma^3(z)\,dz.$$

- Despite this, Heston parameters may be accurately recovered from the fake smiles.
- To generate Figure 12, we used flat extrapolation of the smile beyond available strikes, as in [Fuk12].
- What happens if we extrapolate using SVI?

Heston stochasticity: robust estimates vs exact

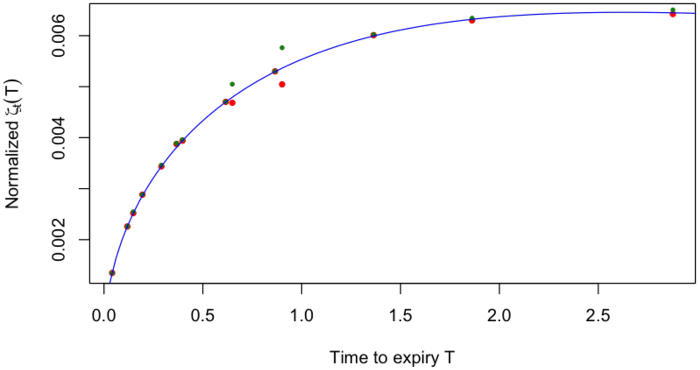

Figure 13: Plot of $\frac{\zeta_t(T)}{T^{3/2}}$ vs time to expiry. The blue line is the exact Heston formula, the red and green dots are robust estimates using flat and SVI extrapolation respectively. We note significant sensitivity to the extrapolation method.

rBergomi stochasticity: robust estimates vs exact

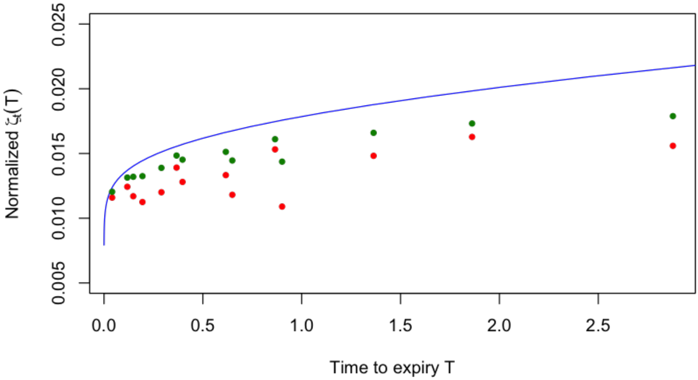

Figure 14: Plot of $\frac{\zeta_t(T)}{T^{3/2}}$ vs time to expiry. The blue line is the exact computation, the red and green dots are robust estimates using flat and SVI extrapolation respectively. We note even greater sensitivity to the extrapolation method.

One particular rBergomi volatility smile

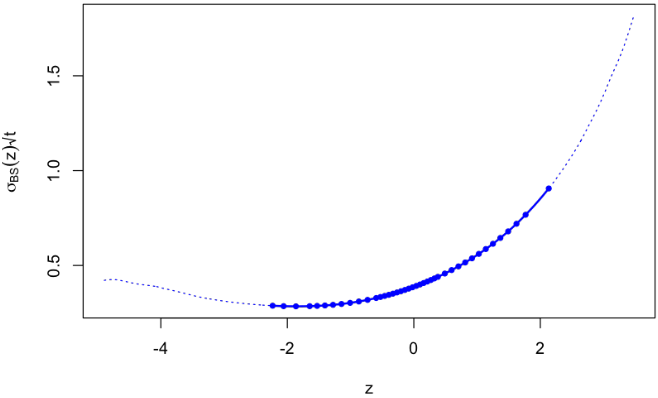

Figure 15: The fake rBergomi 22-Dec-2012 expiration smile (2.88 years) as of 04-Feb-2010. The blue points are market strikes; the dotted line is the model generated smile.

rBergomi stochasticity: robust estimates vs exact again

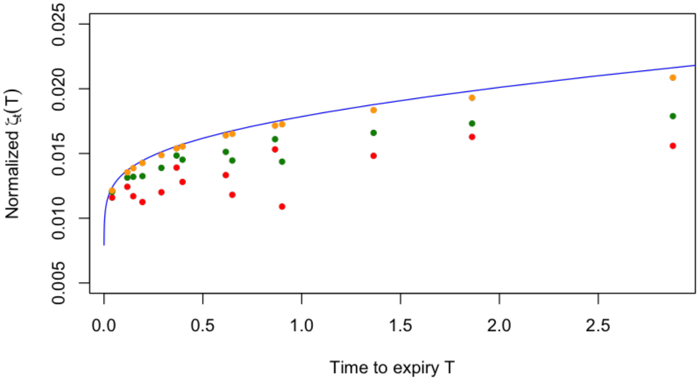

Figure 16: Plot of $\frac{\zeta_t(T)}{T^{3/2}}$ vs time to expiry. The blue line is the exact computation, the red and green dots are robust estimates using flat and SVI extrapolation respectively. The orange points use the whole smile in Figure 15.

Interim conclusion

- rBergomi stochasticity is very sensitive to the extrapolation method in practice.
    - There are insufficiently many strikes available in the market for robust estimation of rBergomi stochasticity.
    - Calibration of model parameters by matching model and market stochasticity would then need a very (unrealistically?) good smile extrapolation method.

- Though matching model and market stochasticity is a nice idea in theory, we have not yet found a smile extrapolation method to make it work in practice.

Plot integrands

- We now plot the various integrands for the fake rBergomi 22-Dec-2012 expiration smile to visualize sensitivity to the extrapolation method.
- Recall that the variance swap is given by

$$\bar{\sigma}^2 = \int dz\, N'(z)\, \sigma^2(z)$$

and stochasticity by

$$\begin{aligned} \zeta_t(T) &= \frac{1}{4} \int N'(z) \left[\sigma^2(z) - \bar{\sigma}^2\right]^2 dz + \frac{2}{3} \int N'(z)\, z\, \sigma^3(z)\, dz \\ &=: \frac{1}{4} I_4 + \frac{2}{3} I_3. \end{aligned}$$

The variance swap integrand

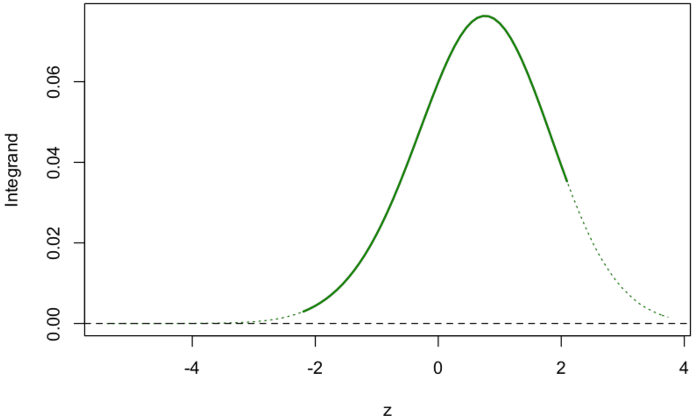

Figure 17: Plot of $N'(z)\,\sigma^2(z)$. The solid line corresponds to strikes available in the market. 10% of the integral is sensitive to extrapolation.

The stochasticity integrand I_4

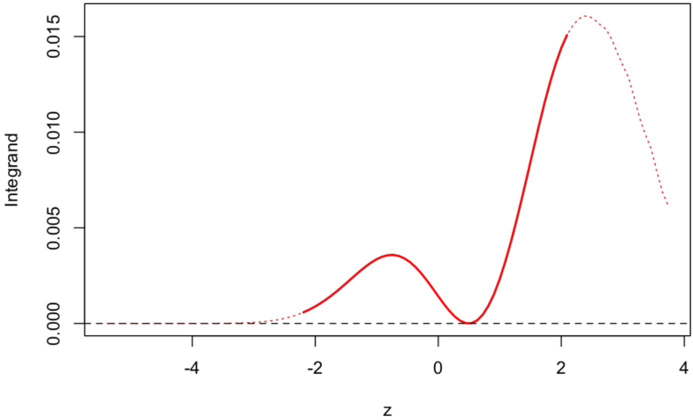

Figure 18: Plot of $N'(z)\left[\sigma^2(z) - \bar{\sigma}^2\right]^2$. The solid line corresponds to strikes available in the market. 28% of the integral is sensitive to extrapolation.

The stochasticity integrand I_3

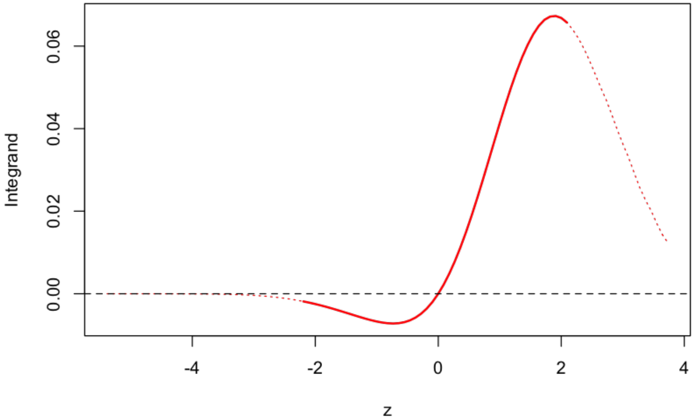

Figure 19: Plot of $N'(z)\, z\, \sigma^3(z)$. The solid line corresponds to strikes available in the market. 29% of the integral is sensitive to extrapolation.

Summary

- We uncovered a remarkable monofractal scaling relationship in historical volatility which now appears to be universal.
- This leads to a natural non-Markovian stochastic volatility model under $\mathbb{P}$.
- The resulting volatility forecast beats existing alternatives.
- The simplest specification of $\frac{d\mathbb{Q}}{d\mathbb{P}}$ gives a non-Markovian generalization of the Bergomi model.
    - The history of the Brownian motion $\{W_s, s < t\}$ required for pricing is encoded in the forward variance curve, which is observed in the market.
- This model fits the observed volatility surface surprisingly well with very few parameters.
- Efficient calibration of the model to the volatility surface remains an open problem.
    - Matching model and market stochasticity is still work in progress.

References

Elisa Alòs.
A decomposition formula for option prices in the Heston model and applications to option pricing approximation.
Finance and Stochastics, 16(3):403–422, 2012.

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

Lorenzo Bergomi and Julien Guyon.
Stochastic volatility's orderly smiles.
Risk May, pages 60–66, 2012.

Mikkel Bennedsen, Asger Lunde, and Mikko S Pakkanen.
Hybrid scheme for brownian semistationary processes.
arXiv preprint arXiv:1507.03004, 2015.

Mikkel Bennedsen, Asger Lunde, and Mikko S Pakkanen.
Decoupling the short-and long-term behavior of stochastic volatility.
Available at SSRN 2846756, 2016.

Emmanuel Bacry and Jean François Muzy.
Log-infinitely divisible multifractal processes.
Communications in Mathematical Physics, 236(3):449–475, 2003.

Emmanuel Bacry and Jean-François Muzy.
Hawkes model for price and trades high-frequency dynamics.
Quantitative Finance, 14(7):1147–1166, 2014.

Omar El Euch, Masaaki Fukasawa, and Mathieu Rosenbaum.
The microstructural foundations of leverage effect and rough volatility.
arXiv preprint arXiv:1609.05177, 2016.

Omar El Euch and Mathieu Rosenbaum.
The characteristic function of rough Heston models.
arXiv preprint arXiv:1609.02108, 2016.

Masaaki Fukasawa.
The normalizing transformation of the implied volatility smile.
Mathematical Finance, 22(4):753–762, 2012.

Masaaki Fukasawa.
Volatility derivatives and model-free implied leverage.
International Journal of Theoretical and Applied Finance, 17(01):1450002, 2014.

Jim Gatheral.
The volatility surface: A practitioner's guide.
John Wiley & Sons, 2006.

Jim Gatheral, Thibault Jaisson, and Mathieu Rosenbaum.
Volatility is rough.
Available at SSRN 2509457, 2014.

Andrew Matytsin.
Perturbative analysis of volatility smiles.
Columbia Practitioners Conference on the Mathematics of Finance, 2000.

Carl J Nuzman and Vincent H Poor.
Linear estimation of self-similar processes via Lamperti's transformation.
Journal of Applied Probability, 37(2):429–452, 2000.