

**Weierstraß-Institut**  
**für Angewandte Analysis und Stochastik**  
**Leibniz-Institut im Forschungsverbund Berlin e. V.**

Preprint

ISSN 2198-5855

**Pricing under rough volatility**

Christian Bayer<sup>1</sup>, Peter K. Friz<sup>1,2</sup>, Jim Gatheral<sup>3</sup>

submitted: February 3, 2015

<sup>1</sup> Weierstrass Institute  
Mohrenstr. 39  
10117 Berlin  
Germany  
E-Mail: christian.bayer@wias-berlin.de  
peter.friz@wias-berlin.de

<sup>2</sup> Institute of Mathematics  
TU Berlin  
Strasse des 17. Juni 136  
10623 Berlin  
Germany  
E-Mail: friz@math.tu-berlin.de

<sup>3</sup> Baruch College  
City University of New York  
55 Lexington Avenue  
NY 10010  
USA  
E-Mail: jim.gatheral@baruch.cuny.edu

No. 2071  
Berlin 2015



---

2010 *Mathematics Subject Classification.* 91B25, 62P05.

*Key words and phrases.* Stochastic volatility, fractional Brownian motion, Bergomi model.

Edited by  
Weierstraß-Institut für Angewandte Analysis und Stochastik (WIAS)  
Leibniz-Institut im Forschungsverbund Berlin e. V.  
Mohrenstraße 39  
10117 Berlin  
Germany

Fax: +49 30 20372-303  
E-Mail: [preprint@wias-berlin.de](mailto:preprint@wias-berlin.de)  
World Wide Web: <http://www.wias-berlin.de/>

## Abstract

From an analysis of the time series of volatility using recent high frequency data, Gatheral, Jaisson and Rosenbaum [SSRN 2509457, 2014] previously showed that log-volatility behaves essentially as a fractional Brownian motion with Hurst exponent  $H$  of order 0.1, at any reasonable time scale. The resulting Rough Fractional Stochastic Volatility (RFSV) model is remarkably consistent with financial time series data. We now show how the RFSV model can be used to price claims on both the underlying and integrated volatility. We analyze in detail a simple case of this model, the rBergomi model. In particular, we find that the rBergomi model fits the SPX volatility markedly better than conventional Markovian stochastic volatility models, and with fewer parameters. Finally, we show that actual SPX variance swap curves seem to be consistent with model forecasts, with particular dramatic examples from the weekend of the collapse of Lehman Brothers and the Flash Crash.

## 1 Introduction

From an analysis of the time series of volatility using recent high frequency data, Gatheral, Jaisson and Rosenbaum [12] showed that log-volatility behaves essentially as a fractional Brownian motion with Hurst exponent  $H$  of order 0.1, at any reasonable time scale. The following stationary Rough Fractional Stochastic Volatility (RFSV) model was proposed:

$$\begin{aligned}\frac{dS_t}{S_t} &= \sigma_t dZ_t \\ \sigma_t &= \exp\{X_t\}, \quad t \in [0, T],\end{aligned}\tag{1.1}$$

where  $X_t$  is a fractional Ornstein-Uhlenbeck process (fOU process for short) satisfying

$$dX_t = \nu dW_t^H - \alpha(X_t - m)dt,$$

where  $m \in \mathbb{R}$  and  $\nu$  and  $\alpha$  are positive parameters, see [5]. Recall that sample paths of fractional Brownian motion  $W^H$  are  $(H - \varepsilon)$ -Hölder (and hence “rougher” than Brownian motion whenever  $H < 1/2$ ). The reversion time scale is understood to be very long so that  $\alpha T \ll 1$  for any reasonable time scale  $T$  of practical interest, in which case, the log-volatility behaves locally (at time scales smaller than  $T$ ) as a fractional Brownian motion (fBm). The RFSV model is remarkably consistent with financial time series data. Moreover, the RFSV model has a quantitative market microstructure-based foundation based on the modeling of order flow using Hawkes processes.

On the other hand, from the perspective of options pricing, it is well-known that conventional low-dimensional Markovian stochastic volatility models such as the Hull and White, Heston, and SABR models generate implied volatility surfaces whose shapes differ substantially from that of the empirically observed volatility surface. A typical such volatility surface generated from a “stochastic volatility inspired” (SVI) [11] fit to closing SPX option prices as of August 14, 2013<sup>1</sup> is shown in Figure 1.1. It is a stylized fact that, at least

---

<sup>1</sup>Closing prices of SPX options for all available strikes and expirations were sourced from OptionMetrics ([www.optionmetrics.com](http://www.optionmetrics.com)) via Wharton Research Data Services (WRDS).

in equity markets, although the level and orientation of the volatility surface do change over time, the general overall shape of the volatility surface does not change, at least to a first approximation. This suggests that it is desirable to model volatility as a time-homogenous process, *i.e.* a process whose parameters are independent of price and time.

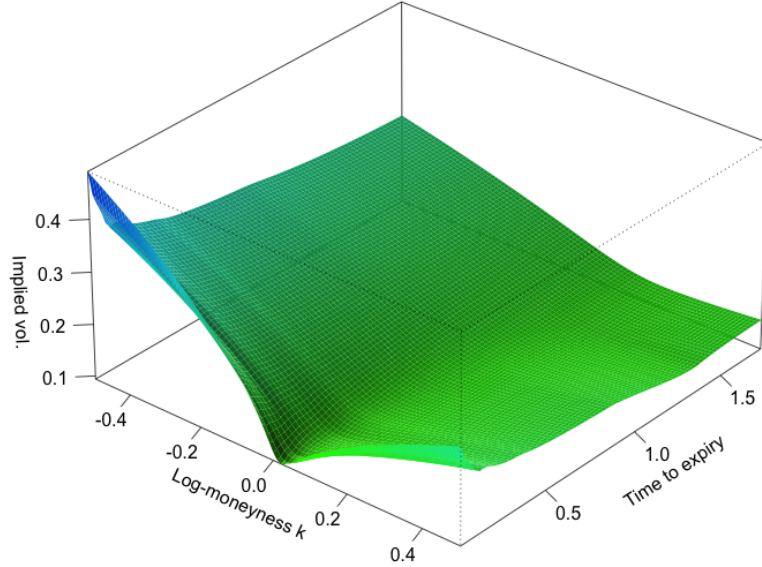


Figure 1.1: The SPX volatility surface as of August 14, 2013.

Given an implied volatility smile for a single expiration, little can be said about the process generating it; any process that generates uncertain realized volatility from inception to expiration and with some correlation between changes in volatility and returns of the underlying might suffice. To say more about the underlying process, the scaling of smiles with respect to time to expiration needs to be examined. In particular, one feature of the volatility surface that really does distinguish between models is the term structure of at-the-money (ATM) volatility skew defined as

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

where  $\tau = T - t$  denotes time to expiration. In conventional stochastic volatility models, the ATM volatility skew  $\psi(\tau)$  is constant for short dates and inversely proportional to  $\tau$  for long dates. Empirically, as shown in Figure 1.2, we observe that  $\psi(\tau)$  is proportional to  $1/\tau^\alpha$  for some  $0 < \alpha < 1/2$  over a very wide range of expirations.

Let  $v_u = \sigma_u^2$  denote instantaneous variance at time  $u > t$ . Then the forward variance curve is given by

$$\xi_t(u) = \mathbb{E}[v_u | \mathcal{F}_t], \quad u \geq t.$$

Bergomi and Guyon [3] derive a small noise expansion for the smile in a stochastic volatility model written in the following forward variance curve form:

$$\begin{aligned} \frac{dS_t}{S_t} &= \sqrt{\xi_t(t)} dZ_t \\ d\xi_t(u) &= \lambda(t, u, \xi_t) dW_t \end{aligned} \tag{1.2}$$

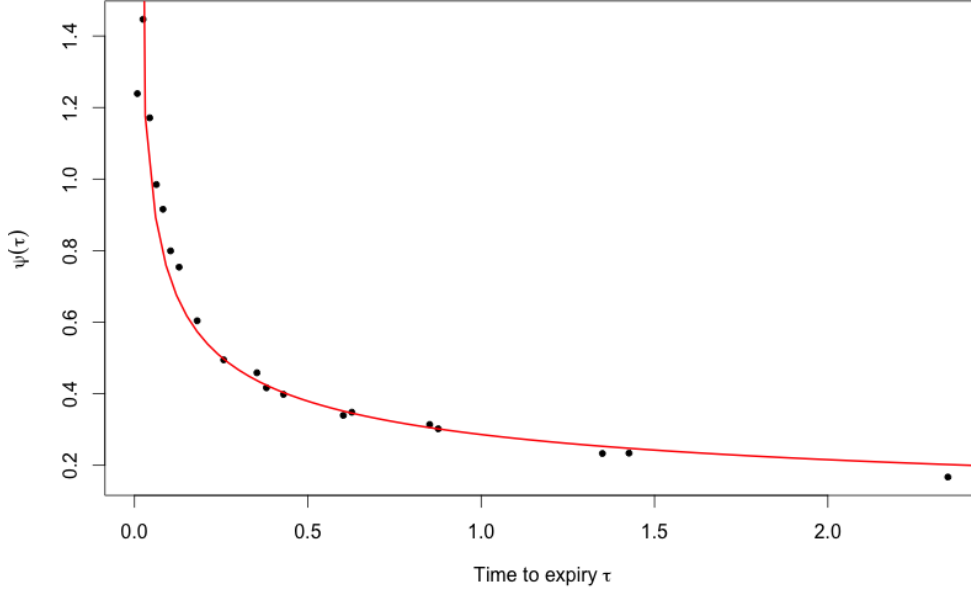


Figure 1.2: The black dots are non-parametric estimates of the S&P ATM volatility skews as of August 14, 2013; the red curve is the power-law fit  $\psi(\tau) = A\tau^{-0.407}$ .

where  $Z_t$  is a Brownian motion driving the asset price  $S_t$  and  $W_t$  is a (suitably correlated)  $d$ -dimensional Brownian motion driving the evolution of the forward variance curve. To first order in the volatility of volatility  $\lambda$ , The Bergomi-Guyon expansion takes the form

$$\sigma_{\text{BS}}(k, T) = \sigma_0(T) + \sqrt{\frac{w}{T}} \frac{1}{2w^2} C^{x\xi} k + O(\lambda^2) \quad (1.3)$$

where the log-strike  $k = \log K/S_0$ ,  $w = \int_0^T \xi_0(s) ds$  is total variance to expiration  $T$ , and

$$C^{x\xi} = \int_0^T dt \int_t^T du \frac{\mathbb{E}[dx_t d\xi_t(u)]}{dt}. \quad (1.4)$$

where  $x_t = \log S_t/S_0$ . Thus, given a stochastic model written in the forward variance curve form (1.2), we can easily (at least in principle) compute the term structure of ATM skew  $\psi(\tau)$  to first order in  $\lambda$ .

One well-known example of a stochastic volatility model expressed in forward variance curve form is the Bergomi model [2]. The  $n$ -factor Bergomi variance curve model may be written in the form

$$\xi_t(u) = \xi_0(u) \mathcal{E} \left( \sum_{i=1}^n \eta_i \int_0^t e^{-\kappa_i(u-s)} dW_s^{(i)} \right) \quad (1.5)$$

where  $\mathcal{E}(\cdot)$  denotes the stochastic exponential<sup>2</sup>.  $\xi_t(u)$  is thus a martingale in  $t$ , consistent with the fact that forward variances are tradable. As was pointed out by Bergomi, the entire forward variance curve

<sup>2</sup> For a continuous (semi)martingale  $Z$ , the stochastic exponential is classically defined as  $\mathcal{E}(Z)_t = \exp(Z_t - Z_0 - \frac{1}{2}[Z, Z]_{0,t})$ . If  $Z$  is a local martingale, then so is  $\mathcal{E}(Z)$ . On the other hand, for a zero-mean Gaussian random variable  $\Psi$ , one defines the "Wick exponential" as  $\mathcal{E}(\Psi) = \exp(\Psi - \frac{1}{2}\mathbb{E}[|\Psi|^2])$ . When  $\Psi$  is the increment of a Gaussian martingale - such as  $\int_0^t f(s)dW_s$  with deterministic integrand - the two notions coincide.

$\xi_t(\cdot) = \{\xi_t(u) : u > t\}$  is determined by  $n$ -factors, each of OU-type. Indeed, in the case  $n = 1$  (for notational simplicity only) one has

$$\xi_t(u) = \xi_0(u) \exp \left( \eta e^{-\kappa(u-t)} Y_t - \frac{1}{2} \eta^2 e^{-2\kappa(u-t)} \mathbb{E}[Y_t^2] \right)$$

in terms of a scalar OU process,

$$dY_t = -\kappa Y_t dt + dW_t.$$

To achieve a decent fit to the observed volatility surface, and to control the forward smile, we need at least two factors. In the two-factor case, there are 7 parameters:  $\eta_1, \eta_2, \kappa_1, \kappa_2$ , and the (constant) correlations  $\rho_{Z, W^{(1)}}, \rho_{Z, W^{(2)}}, \rho_{W^{(1)}, W^{(2)}}$ , in addition to the initial forward variance curve  $\xi_0(u)$ . When calibrating the two-factor Bergomi model to option prices, we find that it is already over-parameterized. Any combination of the parameters  $\eta_i, \kappa_i$ , and the various correlation parameters that gives a roughly  $1/\sqrt{T}$  term structure of ATM skew fits well enough. Moreover, the calibrated correlations between the Brownian increments  $dW_s^{(i)}$  tend to be high.

The Bergomi model generates a term structure of volatility skew  $\psi(\tau)$  that has the qualitative form

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

Indeed, it can be seen from the Bergomi-Guyon expansion that this functional form is related to the term structure of the autocorrelation functional  $C^{x\xi}$ , as defined in (1.4), which is in turn driven by the exponential kernels in the exponent in (1.5). To generate the empirically observed  $\psi(\tau) \sim \tau^{-\alpha}$  for some  $\alpha$ , it would be tempting to replace the exponential kernels in (1.5) with a power-law kernel. This would give a model of the form

$$\xi_t(u) = \xi_0(u) \mathcal{E} \left( \eta \int_0^t \frac{dW_s}{(u-s)^\gamma} \right) \quad (1.6)$$

with  $\xi_t(u)$  again a martingale in  $t$ . Assuming constant  $\xi_0(u) \equiv \sigma_0^2$ , and with the Wick interpretation of the stochastic exponential, we would have instantaneous stochastic volatility

$$\sigma_t = \sqrt{\xi_t(t)} = \sigma_0 \sqrt{\mathcal{E} \left( \eta \int_0^t \frac{dW_s}{(t-s)^\gamma} \right)} = \sigma_0 \exp \left\{ \frac{\eta}{2} V_t - \frac{\eta^2}{4} \mathbb{E}[V_t^2] \right\},$$

where  $V_t = \int_0^t \frac{dW_s}{(t-s)^\gamma}$  is known as ‘‘Volterra’’fractional Brownian motion with Hurst parameter  $H = 1/2 - \gamma$  and has, similar to classical fractional Brownian motion,  $(H - \varepsilon)$ -Hölder sample paths. We note a striking resemblance to the RFSV model (1.1). Moreover, by applying his Martingale expansion to a special case of a model originally proposed by Alós [1], Fukasawa [9] shows formally that the volatility skew generated by such models has the form

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

for small  $\tau$ .

In this paper, we show that the RFSV model does indeed lead naturally to a non-Markovian generalization of the Bergomi model, which we call the Rough Bergomi (rBergomi) model. This model fits the observed volatility surface markedly better than conventional Markovian stochastic volatility models, and with fewer parameters.

## 1.1 Main results and organization of the paper

Our paper is organized as follows. In Section 2, we show how the RFSV model leads naturally to an options pricing model. In Section 3, we analyze a special case of this model, the rBergomi model, where the change of measure from  $\mathbb{P}$  to  $\mathbb{Q}$  is deterministic. In Section 4, we show how to simulate the rBergomi model, and in Section 5 we show that volatility surfaces generated using the rBergomi model simulation are remarkably consistent with observed ones (at least on the two specific days presented). In Section 6, we examine consistency between the rBergomi model and the VIX options market, finding that in general, the rBergomi model is not consistent with the VIX options market. In Section 7, we compute coefficients of the Bergomi-Guyon expansion of the rBergomi model up to second order in volatility of volatility; sadly, we find that this asymptotic expansion does not converge with parameters of practical interest. In Section 8, we show that the evolution of market variance swap curves is consistent with forecasts obtained from the historical realized variance time series; we examine the cases of the collapse of Lehman Brothers and the Flash Crash in detail. Finally, in Section 9, we summarize and conclude. Some more detailed computations are relegated to the appendix.

## 2 Pricing under rough volatility

In [12], using RV estimates as proxies for daily spot volatilities, two startlingly simple regularities were uncovered. Firstly, consistent with many prior studies, distributions of increments of log volatility were found to be close to Gaussian. Second and more interestingly, for reasonable timescales of practical interest, the time series of volatility was found to be consistent with the simple model

$$\log \sigma_{t+\Delta} - \log \sigma_t = \nu (W_{t+\Delta}^H - W_t^H) \quad (2.1)$$

where  $W^H$  is fractional Brownian motion, which is simply the RFSV model (1.1) with  $\alpha = 0$ . This relationship was found to hold for all 21 equity indices in the Oxford-Man database, Bund futures, Crude Oil futures, and Gold futures. Perhaps this feature of the time series of volatility is universal?

Consider the Mandelbrot-Van Ness representation of fractional Brownian motion  $W^H$  in terms of Wiener integrals<sup>3</sup>:

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

where  $\gamma = 1/2 - H$  and the choice  $C_H = \sqrt{\frac{2H\Gamma(3/2-H)}{\Gamma(H+1/2)\Gamma(2-2H)}}$  ensures that

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

Substituting into (2.1) (and in terms of  $v_t = \sigma_t^2$ ), we obtain the following model for the evolution of  $v_u$  under the physical measure  $\mathbb{P}$ :

---

<sup>3</sup>Strictly speaking, this expression is only formal. The rigorous form, as used in the computations below, exploits the cancellation between the integrands as  $s \rightarrow -\infty$ .

$$\begin{aligned}
& \log v_u - \log v_t \\
&= 2\nu C_H \left\{ \int_{-\infty}^u \frac{dW_s^{\mathbb{P}}}{(u-s)^\gamma} - \int_{-\infty}^t \frac{dW_s^{\mathbb{P}}}{(t-s)^\gamma} \right\} \\
&= 2\nu C_H \left\{ \int_t^u \frac{1}{(u-s)^\gamma} dW_s^{\mathbb{P}} + \int_{-\infty}^t \left[ \frac{1}{(u-s)^\gamma} - \frac{1}{(t-s)^\gamma} \right] dW_s^{\mathbb{P}} \right\} \\
&=: 2\nu C_H \{M_t(u) + Z_t(u)\}. \tag{2.2}
\end{aligned}$$

Note that  $Z_t(u)$  is  $\mathcal{F}_t$ -measurable, whereas  $M_t(u)$  is independent of  $\mathcal{F}_t$ , and Gaussian with mean zero, and variance  $(u-t)^{2H}/(2H)$ . We introduce

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

which has the same properties as  $M_t(u)$ , only with variance  $(u-t)^{2H}$ . With  $\eta := 2\nu C_H/\sqrt{2H}$  we have  $2\nu C_H M_t(u) = \eta \tilde{W}_t^{\mathbb{P}}(u)$  and so

$$\mathbb{E}^{\mathbb{P}} [v_u | \mathcal{F}_t] = v_t \exp \left\{ 2\nu C_H Z_t(u) + \frac{1}{2} \eta^2 \mathbb{E} |\tilde{W}_t^{\mathbb{P}}(u)|^2 \right\}.$$

As a consequence, in terms of the (Wick) stochastic exponential<sup>4</sup>

$$\begin{aligned}
v_u &= v_t \exp \left\{ \eta \tilde{W}_t^{\mathbb{P}}(u) + 2\nu C_H Z_t(u) \right\} \\
&= \mathbb{E}^{\mathbb{P}} [v_u | \mathcal{F}_t] \mathcal{E} \left( \eta \tilde{W}_t^{\mathbb{P}}(u) \right). \tag{2.3}
\end{aligned}$$

This computation reveals that the conditional distribution of  $v_u$  depends on  $\mathcal{F}_t$  only through the variance forecasts  $\mathbb{E}^{\mathbb{P}} [v_u | \mathcal{F}_t]$ ,  $u > t$ <sup>5</sup>. In particular, to price options, one does not need to know  $\mathcal{F}_t$ , the entire history of the Brownian motion  $W_s^{\mathbb{P}}$  for  $s < t$ .

## 2.1 Pricing under $\mathbb{Q}$

We have a model (2.3) that accurately mimics the behavior of realized variance time series data, written under  $\mathbb{P}$ :

$$v_u = \mathbb{E}^{\mathbb{P}} [v_u | \mathcal{F}_t] \mathcal{E} \left( \eta \tilde{W}_t^{\mathbb{P}}(u) \right). \tag{2.4}$$

where in particular  $\mathbb{E}^{\mathbb{P}} [v_u | \mathcal{F}_t]$  is adapted to the filtration generated by  $W^{\mathbb{P}}$  which we assume is the same as the filtration generated by  $W^{\mathbb{Q}}$ . Consider some general change of measure

$$dW_s^{\mathbb{P}} = dW_s^{\mathbb{Q}} + \lambda_s ds, \tag{2.5}$$

<sup>4</sup> $\mathcal{E}(\Psi) = \exp(\Psi - \frac{1}{2}\mathbb{E}^{\mathbb{P}}[|\Psi|^2])$  where  $\Psi$  is zero-mean, Gaussian under  $\mathbb{P}$ . (To be fully consistent, we should write  $\mathcal{E} = \mathcal{E}^{\mathbb{P}}$ ).

<sup>5</sup>This is analogous to what happens in Comte, Coutin and Renault [6] in the context of their fractionally integrated square root model.



where  $\{\lambda_s : s > t\}$  has a natural interpretation as the price of volatility risk. We may then rewrite (2.4) as

$$\begin{aligned} v_u &= \mathbb{E}^{\mathbb{P}} [v_u | \mathcal{F}_t] \exp \left\{ \eta \sqrt{2H} \int_t^u \frac{1}{(u-s)^\gamma} dW_s^{\mathbb{P}} - \frac{\eta^2}{2} (u-t)^{2H} \right\} \\ &= \mathbb{E}^{\mathbb{P}} [v_u | \mathcal{F}_t] \mathcal{E} \left( \eta \tilde{W}_t^{\mathbb{Q}}(u) \right) \exp \left\{ \eta \sqrt{2H} \int_t^u \frac{\lambda_s}{(u-s)^\gamma} ds \right\}. \end{aligned} \tag{2.6}$$

The last term in the exponent obviously changes the marginal distribution of the  $v_u$ ; although the conditional distribution of  $v_u$  under  $\mathbb{P}$  is lognormal, it will not be lognormal in general under  $\mathbb{Q}$ .

### VIX smiles and the change of measure

In the case of SPX, it is obvious from the shape of VIX implied volatility smiles that the change of measure cannot be deterministic. If the change of measure were deterministic, it follows from (2.6) that  $v_u$  would be conditionally lognormal, VIX would also be approximately lognormal and so the VIX implied volatility smiles would be approximately flat, a well-known problem with the conventional Bergomi model. In contrast, we observe VIX smiles that are strongly upward sloping (see Figure 2.1 for example), reflecting the intuition that high volatility scenarios are priced more highly by the market than low volatility scenarios. Specifically, we conclude that the change of measure  $\lambda$  must be positively correlated with  $W^{\mathbb{Q}}$ .

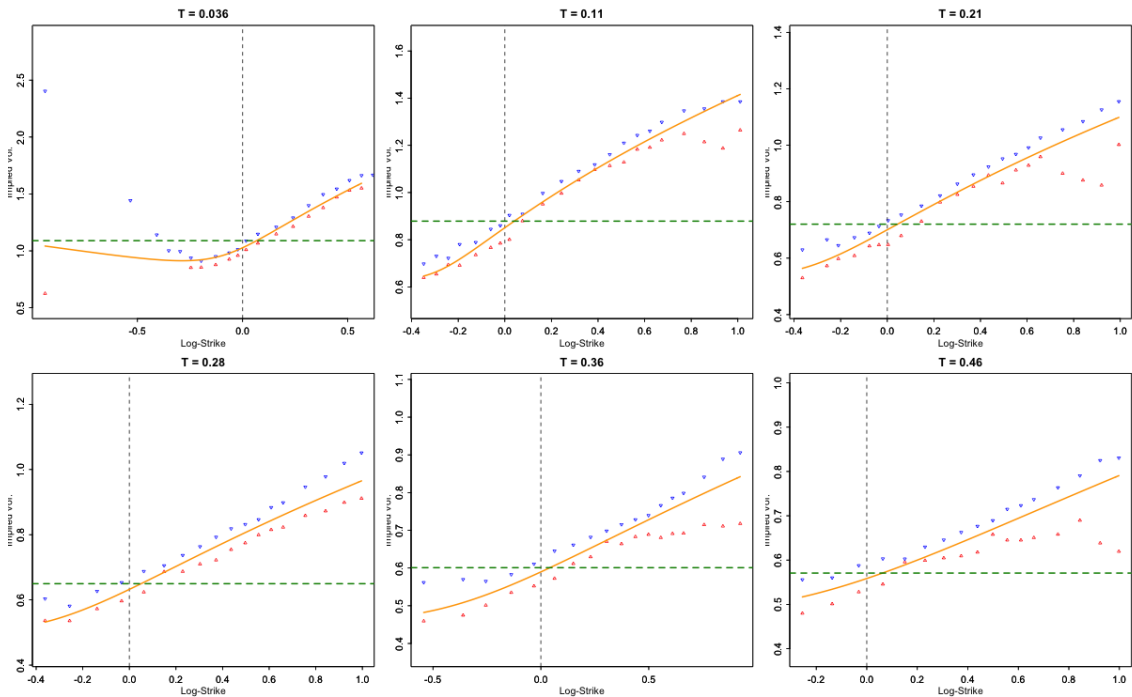


Figure 2.1: VIX implied volatility smiles as of February 4, 2010. Blue points are ask volatilities; red points are bid volatilities; orange lines are SVI fits; green dashed lines represent the VIX log-strike (VVIX).

### 3 The Rough Bergomi (rBergomi) model

Despite the inconsistency with VIX smiles pointed out above, 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 (2.6), we would have

$$\begin{aligned} v_u &= \mathbb{E}^{\mathbb{P}} [v_u | \mathcal{F}_t] \mathcal{E} \left( \eta \tilde{W}_t^{\mathbb{Q}}(u) \right) \exp \left\{ \eta \sqrt{2H} \int_t^u \frac{1}{(u-s)^\gamma} \lambda(s) ds \right\} \\ &= \xi_t(u) \mathcal{E} \left( \eta \tilde{W}_t^{\mathbb{Q}}(u) \right) \end{aligned} \quad (3.1)$$

where by definition,  $\xi_t(u) = \mathbb{E}^{\mathbb{Q}} [v_u | \mathcal{F}_t]$ . Moreover, the forward variance curve

$$\xi_t(u) = \mathbb{E}^{\mathbb{P}} [v_u | \mathcal{F}_t] \exp \left\{ \eta \sqrt{2H} \int_t^u \frac{1}{(u-s)^\gamma} \lambda(s) ds \right\}$$

is the product of two terms:  $\mathbb{E}^{\mathbb{P}} [v_u | \mathcal{F}_t]$  which depends on the history of the driving Brownian motion as explained earlier, and a term which depends on the price of risk  $\lambda(s)$ .

The model (3.1) is a non-Markovian generalization of the Bergomi model (1.5) which we might dub a *rough Bergomi* (or *rBergomi*) model. Specifically, this rBergomi model is non-Markovian in the instantaneous variance  $v_t$ :  $\mathbb{E}^{\mathbb{Q}} [v_u | \mathcal{F}_t] \neq \mathbb{E}^{\mathbb{Q}} [v_u | v_t]$  but is Markovian in the (infinite-dimensional) state vector  $\mathbb{E}^{\mathbb{Q}} [v_u | \mathcal{F}_t] = \xi_t(u)$ .

Note also that with (3.1), we have achieved the aim we set out in the introduction by replacing the exponential kernels in the Bergomi model (1.5) with a power-law kernel. We may therefore expect that the rBergomi model will generate a realistic term structure of ATM volatility skew.

The observed anticorrelation between price moves and volatility moves may be modeled naturally, just as in the conventional Bergomi model, 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  $\rho$  is the correlation between volatility moves and price moves.

#### 3.1 Re-interpretation of the conventional Bergomi model

According to [2], the conventional Bergomi model is a *market model*, by which it is meant that  $\xi_t(u)$  can be any given initial forward variance swap curve consistent with market prices. However, for the Bergomi model to properly describe the evolution of this curve,  $\xi_t(u) = \mathbb{E} [v_u | \mathcal{F}_t]$  should be consistent with the assumed dynamics; in this sense, a conventional  $n$ -factor Bergomi model is not self-consistent in general.

Viewed from the perspective of the rBergomi model however, the initial curve  $\xi_t(u)$  reflects the history  $\{W_s; s < t\}$  of the driving Brownian motion up to time  $t$ . The exponential kernels in the exponent of (1.5) approximate more realistic power-law kernels. The conventional two-factor Bergomi model is then justified in practice as a tractable Markovian engineering approximation to a more realistic rBergomi model.

## 4 Simulation of the rBergomi model

To simplify notation, we set the origin of the simulation to be  $t = 0$  and drop the explicit reference to the pricing measure  $\mathbb{Q}$ . From (3.1), the model to be simulated is

$$\begin{aligned} S_t &= S_0 \mathcal{E} \left( \int_0^t \sqrt{v_u} dZ_u \right) \\ v_u &= \xi_0(u) \mathcal{E} \left( \eta \sqrt{2H} \int_0^u \frac{1}{(u-s)^\gamma} dW_s \right) = \xi_0(u) \mathcal{E} \left( \eta \tilde{W}_u \right). \end{aligned}$$

where  $\tilde{W}$  is a Volterra process<sup>6</sup> with the scaling property  $\text{Var}[\tilde{W}_u] = u^{2H}$ . So far  $\tilde{W}$  behaves just like fBm. However, the dependence structure is different. Specifically, for  $v > u$ ,

$$\mathbb{E} [\tilde{W}_v \tilde{W}_u] = u^{2H} G \left( \frac{u}{v} \right)$$

where, for  $x \geq 1$ ,

$$\begin{aligned} G(x) &= 2H \int_0^1 \frac{ds}{(1-s)^\gamma (x-s)^\gamma} \\ &= \frac{1-2\gamma}{1-\gamma} x^\gamma {}_2F_1(1, \gamma, 2-\gamma, x) \end{aligned} \quad (4.1)$$

where  ${}_2F_1(\cdot)$  denotes the confluent hypergeometric function.

**Remark 4.1.** *The dependence structure of the Volterra process  $\tilde{W}$  is markedly different from that of fBm with the Molchan-Golosov kernel. In particular, for small  $H$ , correlations drop precipitously as the ratio  $u/v$  moves away from 1.*

We also need covariances of the Brownian motion  $Z$  with the Volterra process  $\tilde{W}$ . With  $v \geq u$ , these are given by

$$\mathbb{E} [\tilde{W}_v Z_u] = \rho D_H \left\{ v^{H+1/2} - (v-u)^{H+1/2} \right\}$$

and

$$\mathbb{E} [Z_v \tilde{W}_u] = \rho D_H u^{H+1/2}$$

where for future convenience, we have defined the constant,

$$D_H = \frac{\sqrt{2H}}{H+1/2}.$$

These two formulae may be conveniently combined as

$$\mathbb{E} [\tilde{W}_v Z_u] = \rho D_H \left\{ v^{H+1/2} - (v - \min(u, v))^{H+1/2} \right\}.$$

Lastly, of course, for  $v \geq u$ ,  $\mathbb{E} [Z_v Z_u] = u$ .

With  $m$  the number of time steps and  $n$  the number of simulations, our rBergomi model simulation algorithm may then be summarized as follows.

<sup>6</sup>This is identical up to a constant factor to the definition of [7].

- Construct the joint covariance matrix for the Volterra process  $\tilde{W}$  and the Brownian motion  $Z$  and compute its Cholesky decomposition.
- For each time, generate iid normal random vectors and multiply them by the lower-triangular matrix obtained by the Cholesky decomposition to get a  $m \times 2n$  matrix of paths of  $\tilde{W}$  and  $Z$  with the correct joint marginals.
- With these paths held in memory, we may evaluate the expectation under  $\mathbb{Q}$  of any payoff of interest.

The simulation procedure we have described is unsurprisingly very slow because of the high number of matrix-vector multiplications with a lower-triangular but otherwise dense matrix. We leave the search for faster simulation techniques based on the specific structure of the problem, including the specific choice of the correlation structure between  $Z$  and  $\tilde{W}$  for future research.

## 5 Consistency of the rBergomi model with the SPX volatility surface

As explained above, our simulation of the rBergomi model is very slow and this effectively rules out optimization in practice. However, the model parameters  $H$ ,  $\eta$  and  $\rho$  have very direct interpretations.  $H$  controls the decay of the term structure of volatility skew for very short expirations whereas the product  $\rho\eta$  sets the level of the ATM skew for longer expirations. Keeping the product  $\rho\eta$  roughly constant but decreasing  $\rho$  (so as to make it more negative) has the effect of pushing the minimum of each smile towards higher strikes. Thus, it is possible to guess parameters. Moreover, as we will show below,  $H$  and  $\eta$  may be estimated from historical data. We will now show that on two particular days in history, the rBergomi model was surprisingly consistent with the observed volatility surface. Fits for other days we tried are not always as impressive as these two but nevertheless visibly superior to fits of conventional Markovian stochastic volatility models.

### 5.1 Parameter estimation from the time series of realized variance

Both the roughness parameter (or Hurst parameter)  $H$  and the volatility of volatility  $\eta$  should be the same under  $\mathbb{P}$  and  $\mathbb{Q}$ .

In [12], we estimated the RFSV model (1.1) on the Oxford-Man realized variance dataset obtaining the historical effective parameter estimates  $H_{eff} \approx 0.14$  and volatility of volatility  $\nu_{eff} \approx 0.3$ . Recall however that the instantaneous volatility  $\sigma_t$  is not observed; rather we observe the realized variance  $\frac{1}{\delta} \int_0^\delta \sigma_t^2 dt$  where  $\delta$  corresponds to a trading day from the open to the close, roughly  $3/4$  of a whole day from close to close. Following the computation in Appendix C of [12], we may use these historical estimates to approximate the roughness and volatility of volatility corresponding to instantaneous volatility. This gives  $H \approx 0.05$  and  $\nu \approx 1.7$ . From Section 2, we have the relationship

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

which yields the estimate  $\eta \approx 2.5$ .

## 5.2 Estimation of the variance swap curve

Variance swaps are actively traded so in principle, computation of the forward variance swap curve should be straightforward. In practice however, it is not easy to obtain high quality variance swap quote data and in any case, the bid/ask spread is wide. We thus choose to proxy the value of a  $\tau$ -maturity variance swap by the value of a  $\tau$ -expiration log contract as explained for example in Chapter 11 of [10]. To price the log contract for a particular expiration  $\tau$  requires us to know the prices of  $\tau$ -expiration options for all strikes; of course prices are only quoted for a finite number of strikes. We therefore choose to interpolate and extrapolate observed implied volatilities using the arbitrage-free SVI parameterization of the volatility surface as explained in [11]. For any given day, we obtain the closing prices of SPX options for all available strikes and expirations from OptionMetrics ([www.optionmetrics.com](http://www.optionmetrics.com)) via Wharton Research Data Services (WRDS). Having estimated variance swaps to each expiration, we interpolate total variances using a monotonic spline to estimate variance swaps for intermediate dates. This allows us in turn to estimate the forward variance swap curve.

One subtlety is that by choosing SVI to interpolate and extrapolate, we may be assuming a smile that is inconsistent with the one generated by the rBergomi model, and therefore that the forward variance curve may not be accurate. The practical effect of this is that at-the-money implied volatilities are not matched in the first pass, with good agreement for very short expirations but rather less good agreement as time to expiry increases. A simple iteration on the forward variance curve soon reaches a fixed point that achieves consistency between model ATM volatilities and market ATM volatilities.

## 5.3 Fits to two specific days in history

### February 4, 2010

For our first comparison of the model to SPX options data, we choose February 4, 2010, a day when the ATM volatility term structure happened to be pretty flat. With guessed parameters  $H = 0.07$ ,  $\eta = 1.9$ ,  $\rho = -0.9$ , we obtain the impressive fit shown in Figure 5.1. Only three parameters to get a very good fit to the whole SPX volatility surface, including the shortest dated smile (Figure 5.2).

In Figures 5.3 and 5.4, we see just how well the rBergomi model can match empirical skews and vols. Recall also that the parameters we used are just guesses!

### August 14, 2013

For our second comparison, we choose a date just prior to an expiration date for SPX options. Rather than choosing the last Thursday of trading, we examine the volatility surface as of the close on the final Wednesday prior to expiration so that the shortest expiration smile is more meaningful; the latest such date available to us in our OptionMetrics data set is Wednesday August 14, 2013. With guessed parameters  $H = 0.05$ ,  $\eta = 2.3$ ,  $\rho = -0.9$ , we obtain the fit shown in Figure 5.5. Once again, only three parameters to get a very good fit to the whole SPX volatility surface, including the shortest dated smile (from options with only one day of trading left).

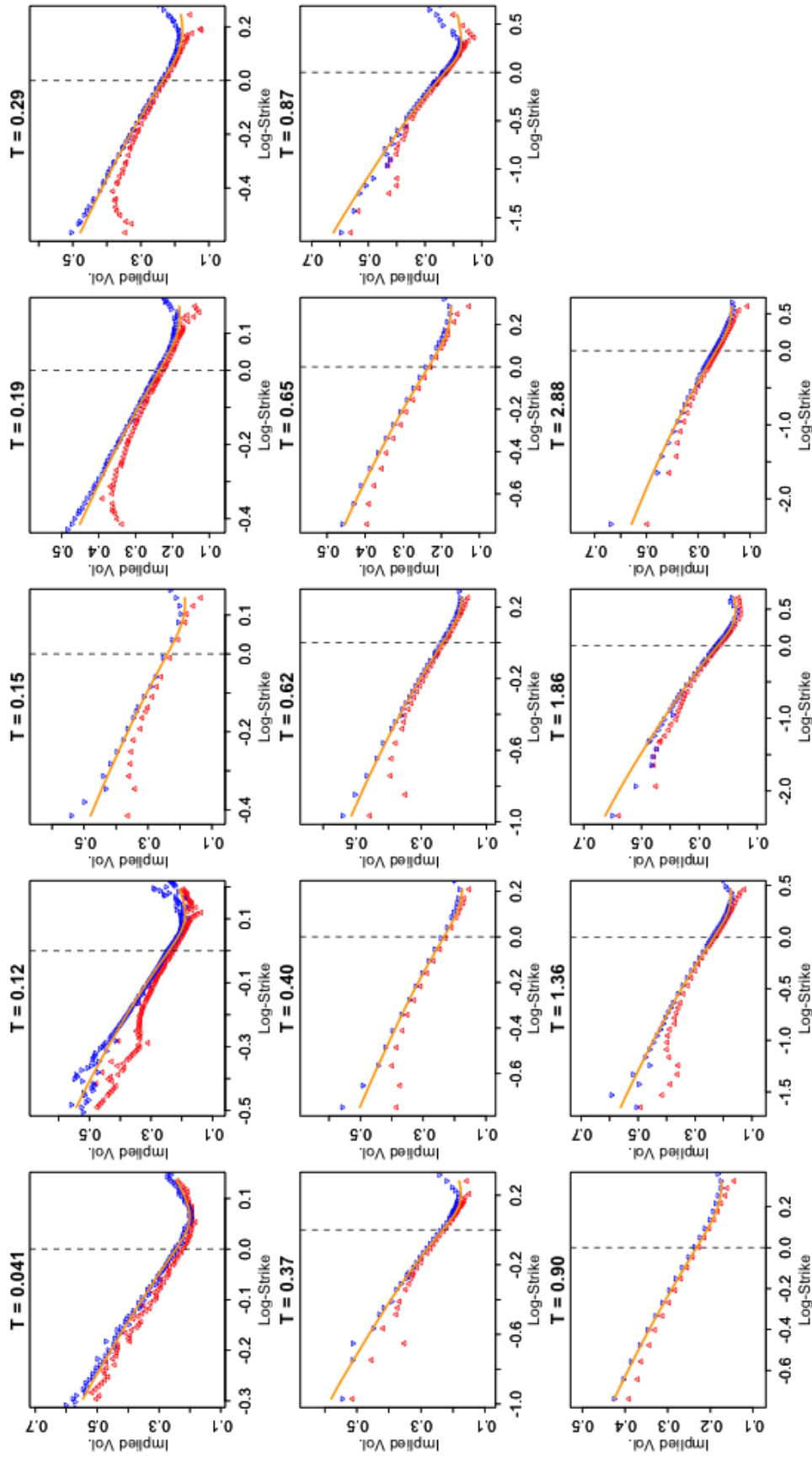


Figure 5.1: SPX smiles as of February 4, 2010: Red and blue points represent bid and offer SPX implied volatilities; orange smiles are from the rBergomi simulation.

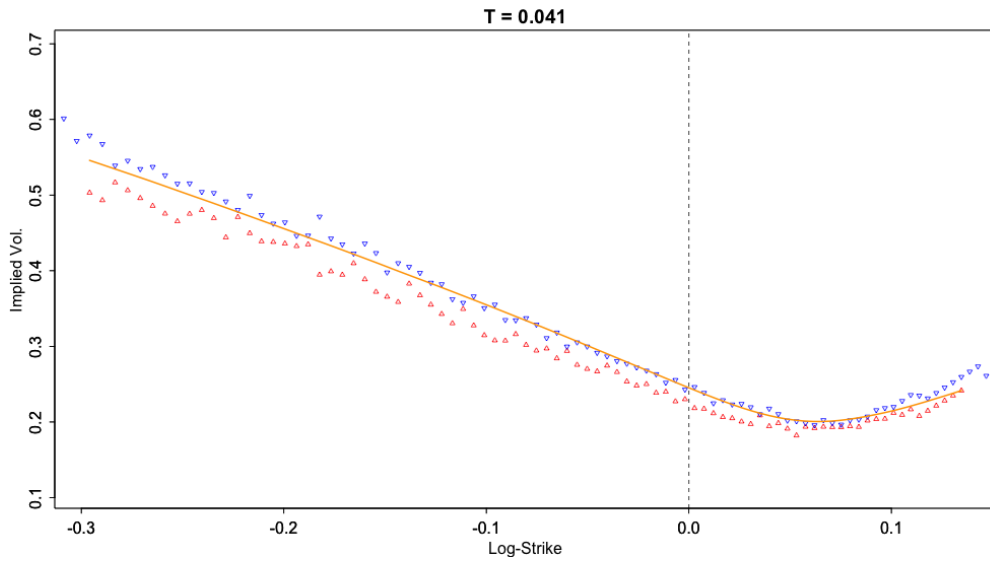


Figure 5.2: Shortest dated smile as of February 4, 2010: Red and blue points represent bid and offer SPX implied volatilities; orange smile is from the rBergomi simulation.

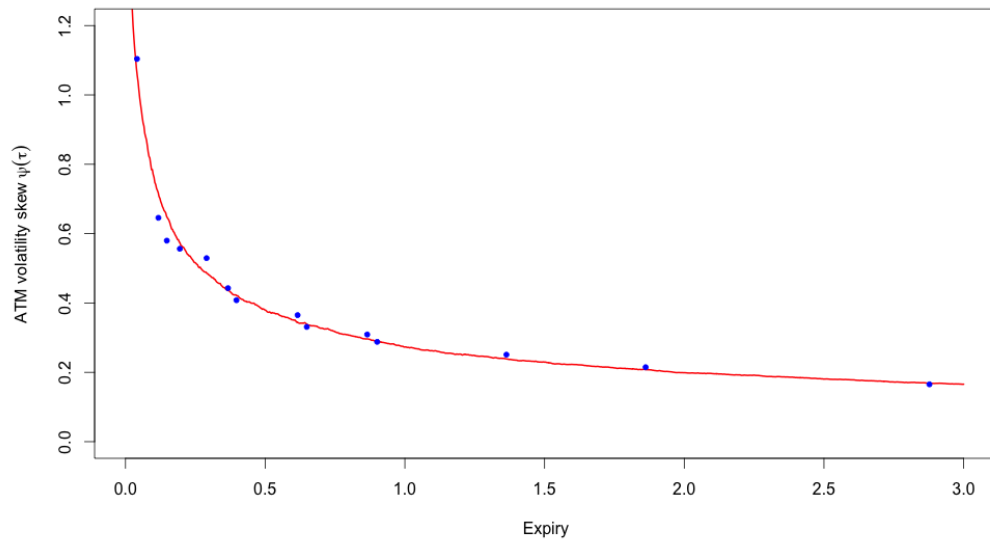


Figure 5.3: As of February 4, 2010: Blue points are empirical skews; the red line is from the rBergomi simulation.

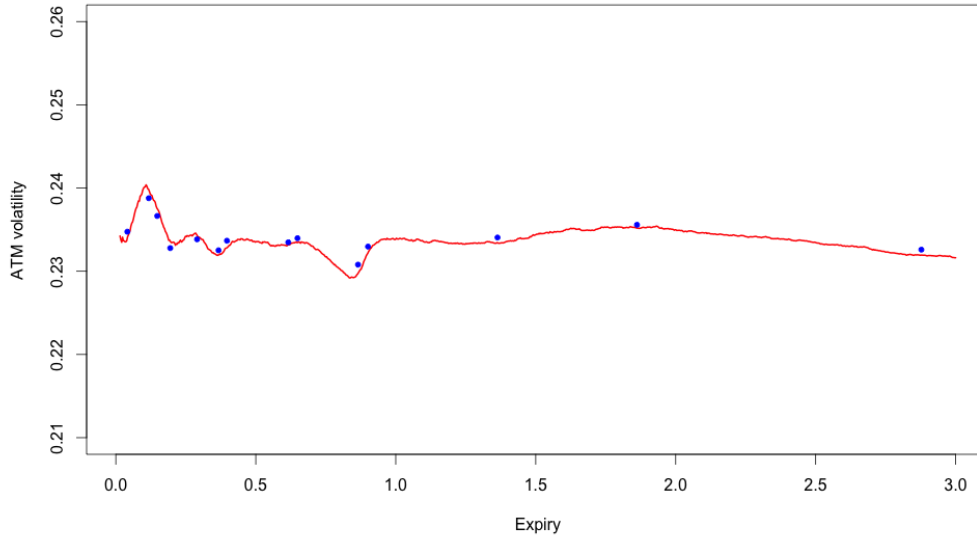


Figure 5.4: As of February 4, 2010: Blue points are empirical ATM volatilities; the red line is from the rBergomi simulation.

### 5.3.1 Jump-like behavior of the rBergomi price process

It has often been claimed that jumps are required to explain the observed extreme short-dated smile in SPX. In particular, in [4], Carr and Wu determine whether or not there are jumps in the asset process, and if so, whether such jumps are finite or infinite activity. They determine based on their analysis that jumps are required to generate the smiles observed in SPX. However, the class of processes that Carr and Wu consider is too restrictive, excluding models like rBergomi where the out-of-the-money smile explodes as time to expiration  $\tau \rightarrow 0$ . It is apparent from Figures 5.1 and 5.5 that the rBergomi model (where the price process is continuous) generates smiles consistent with those observed empirically *even for very short expirations*; there is no need for jumps.

## 6 The rBergomi model and VIX options

We pointed out earlier in Section 2.1 that observed VIX smiles are inconsistent with the rBergomi model. Nevertheless, even if the rBergomi model is misspecified, it may be possible to impute its parameters  $H$  and  $\eta$  by examining the term structure of VIX variance swaps<sup>7</sup>; if VIX corresponds to volatility, then VIX of VIX should correspond to “volatility of volatility”.

<sup>7</sup>The VIX log-strip forms the basis for the VVIX (VIX of VIX) index computation. Indeed, following CBOE ([www.cboe.com](http://www.cboe.com)), the VVIX term structure is computed every day ( $t$ ) for various maturities ( $T$ ) of VIX options using the usual log-strip formula that is used for the construction of VIX. More specifically, given  $T > t$  and assuming that VIX options with expiry  $T$  are traded, the  $VVIX_{t,T}$  is given by

$$VVIX_{t,T}^2(T-t) = -2\mathbb{E}_t \left[ \log \sqrt{\zeta(T)} - \log \sqrt{\zeta(t)} \right],$$

where  $\zeta(s)$  denotes the square of VIX at  $s$  and  $\mathbb{E}_t \log \sqrt{\zeta(T)}$  can be expressed in terms of put and call prices on VIX with expiry  $T$ . The usual VVIX index (at a given  $t$ ) then corresponds to  $VVIX_{t,t+\Delta}$  for  $\Delta$  equal to one month.



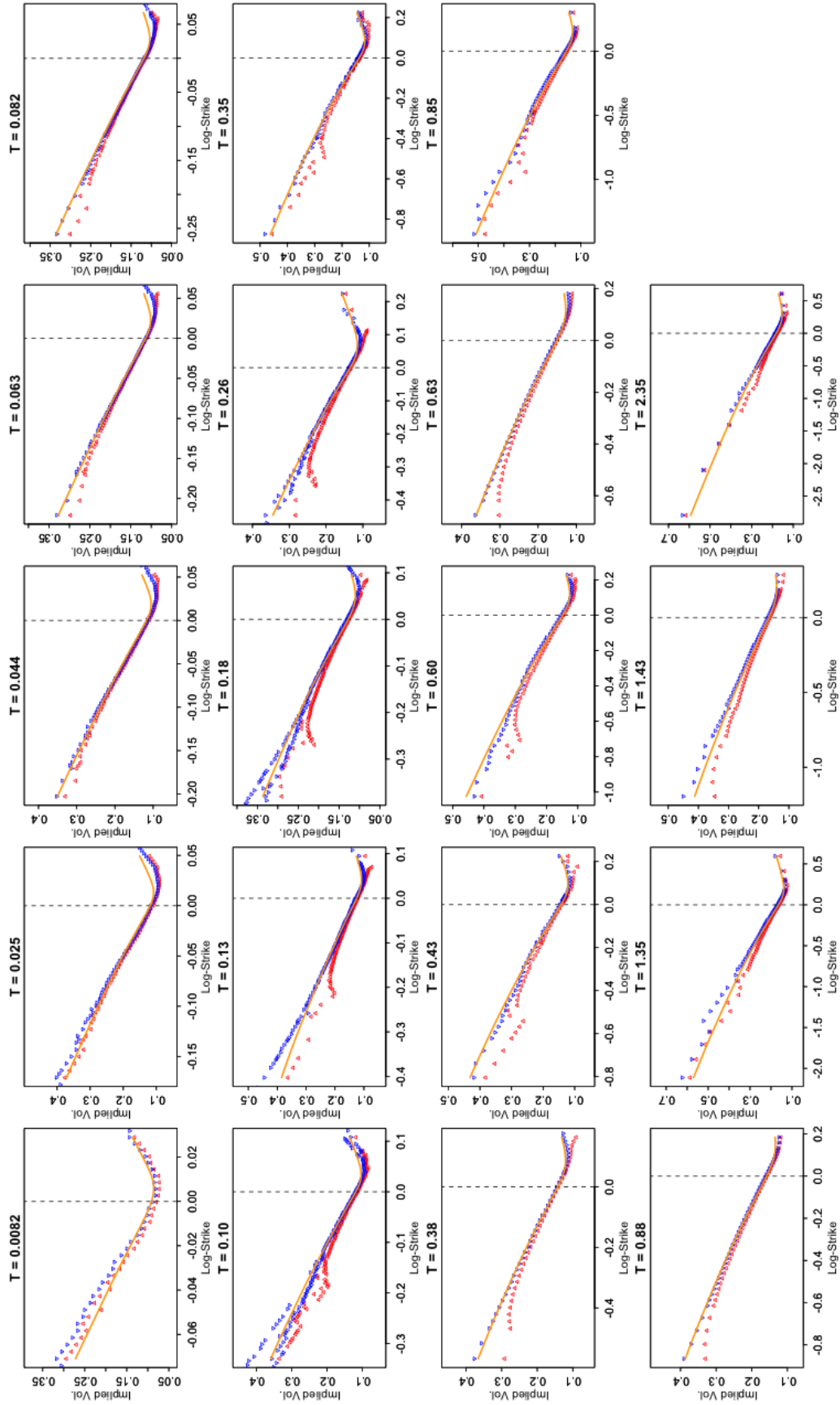


Figure 5.5: SPX smiles as of August 14, 2013: Red and blue points represent bid and offer SPX implied volatilities; orange smiles are from the rBergomi simulation.

Denote the terminal value of the VIX futures by  $\sqrt{\zeta(T)}$ . Then, by definition<sup>8</sup>,

$$\zeta(T) = \frac{1}{\Delta} \int_T^{T+\Delta} \mathbb{E}[v_u | \mathcal{F}_T] du.$$

where  $\Delta$  is one month. In the rBergomi model,

$$v_u = \xi_t(u) \mathcal{E} \left( \eta \sqrt{2H} \int_t^u \frac{dW_s}{(u-s)^\gamma} \right)$$

with  $\gamma = 1/2 - H$ . Instantaneous variances  $v_u$  are thus lognormally distributed. It should therefore be a good approximation (and so it turns out) to assume that the VIX payoff and its square  $\zeta(T)$  are also lognormally distributed. In that case, the terminal distribution of  $\zeta(T)$  is completely determined by  $\mathbb{E}[\zeta(T) | \mathcal{F}_t]$  and  $\text{Var}[\log \zeta(T) | \mathcal{F}_t]$ .

It is immediate that

$$\mathbb{E}[\zeta(T) | \mathcal{F}_t] = \frac{1}{\Delta} \int_T^{T+\Delta} \xi_t(u) du.$$

To estimate the conditional variance of  $\zeta(T)$ , we approximate the arithmetic mean by the geometric mean as follows:

$$\zeta(T) \approx \exp \left\{ \frac{1}{\Delta} \int_T^{T+\Delta} \mathbb{E}[\log v_u | \mathcal{F}_T] du \right\}.$$

After some computation detailed in Appendix B, we obtain

$$\text{Var}[\log \zeta(T) | \mathcal{F}_t] \approx \eta^2 (T-t)^{2H} f^H \left( \frac{\Delta}{T-t} \right)$$

where

$$f^H(\theta) = \frac{D_H^2}{\theta^2} \int_0^1 \left[ (1+\theta-x)^{1/2+H} - (1-x)^{1/2+H} \right]^2 dx. \quad (6.1)$$

It is straightforward to show that  $f^H(\theta) \rightarrow 1$  as  $\theta \rightarrow 0$  which is the limit in which  $\zeta(T) \rightarrow v_T$ . In Appendix B we show further how to express  $f^H(\theta)$  explicitly in terms of the hypergeometric function. However, the above form (6.1) is more convenient for computation.

The VIX variance swaps ( $VVIX^2$ ) are then given by

$$\begin{aligned} VVIX_{t,T}^2(T-t) &\approx \text{Var} \left[ \log \sqrt{\zeta(T)} \middle| \mathcal{F}_t \right] \\ &\approx \frac{1}{4} \eta^2 (T-t)^{2H} f^H \left( \frac{\Delta}{T-t} \right). \end{aligned} \quad (6.2)$$

VIX variance swaps may also be estimated directly from market prices of options on VIX using the log-strip in the usual way as

$$VVIX_{t,T}^2(T-t) = -2 \mathbb{E} \left[ \log \sqrt{\zeta(T)} - \log \sqrt{\zeta(t)} \middle| \mathcal{F}_t \right]. \quad (6.3)$$

By comparing the model VVIX term structure (6.2) with the market VVIX term structure (6.3), we can in principle fix the model parameters  $H$  and  $\eta$ .

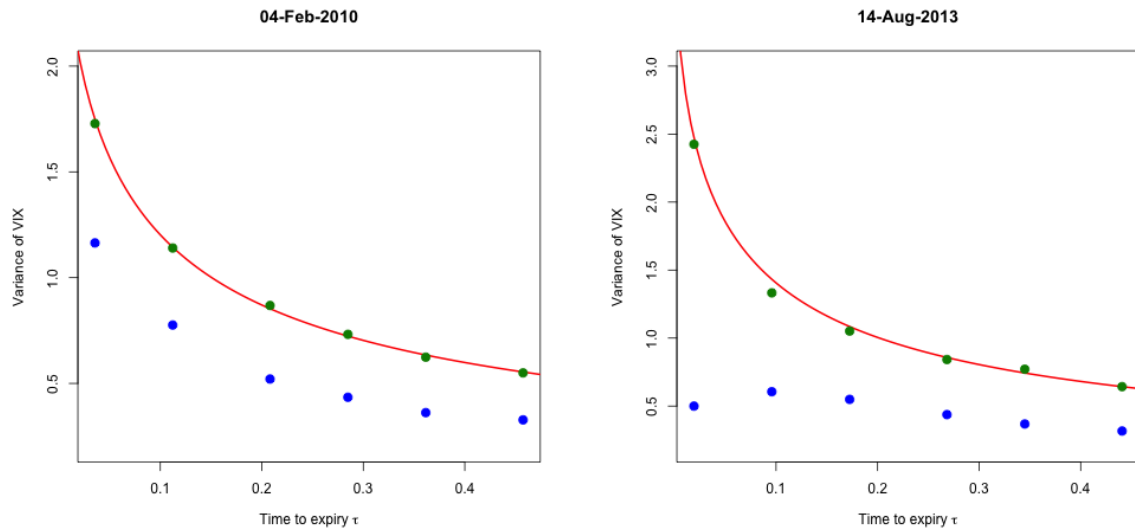


Figure 6.1: Empirical VVIX term structure data (blue points) and rBergomi (using (6.2), red line and from simulation, green points) estimates of  $\text{Var}(VIX)$  as of February 4, 2010 (left) and August 14, 2013 (right).

## 6.1 The VVIX term structure in practice

Recall that the parameters we used to obtain the reasonably impressive SPX fits of Section 5.3 were just guessed; specifically these parameters were as follows:

	$H$	$\eta$
February 4, 2010	0.07	1.9
August 14, 2013	0.05	2.3

In Figure 6.1, we show plots of equation (6.2) with the above parameters, superimposed on the empirically estimated variances of VIX. At least for February 4, 2010, given the qualitative agreement between the shape of the curve (6.2) and empirical estimates, it might be possible to argue some consistency of the rBergomi model with observation; for August 14, 2013 however, there is not even qualitative agreement. Whether this disagreement between model and market should be ascribed to a misspecified model, wrong market prices, or indeed both of these, is left for future research.

## 7 Volatility of volatility expansion

As simulation of the rBergomi model is so slow, one potential alternative is to estimate rBergomi parameters using an asymptotic expansion. Given a model expressed in variance curve form, Bergomi and Guyon [3] derive the following expression for the Black-Scholes implied volatility smile to second order in volatility of volatility:

$$\sigma_{BS}(k, t) = \hat{\sigma}_T + \mathcal{S}_T k + \mathcal{C}_T k^2 + O(\eta^3) \quad (7.1)$$

<sup>8</sup>See Chapter 11 of [10] for more details.

where

$$\begin{aligned}
\hat{\sigma}_T &= \sqrt{\frac{w}{T}} \left\{ 1 + \frac{1}{4w} C^{x\xi} + \frac{1}{32w^3} \left( 12(C^{x\xi})^2 - w(w+4)C^{\xi\xi} + 4w(w-4)C^\mu \right) \right\} \\
\mathcal{S}_T &= \sqrt{\frac{w}{T}} \left\{ \frac{1}{2w^2} C^{x\xi} + \frac{1}{8w^3} \left( 4wC^\mu - 3(C^{x\xi})^2 \right) \right\} \\
\mathcal{C}_T &= \sqrt{\frac{w}{T}} \frac{1}{8w^4} \left( 4wC^\mu + wC^{\xi\xi} - 6(C^{x\xi})^2 \right)
\end{aligned} \tag{7.2}$$

and

$$w = \int_0^T \xi_0(s) ds$$

is total variance to expiration  $T$ . The autocorrelation functionals  $C^{x\xi}$ ,  $C^{\xi\xi}$ , and  $C^\mu$  have the following explicit expressions:

$$\begin{aligned}
C^{x\xi} &= \int_0^T dt \int_t^T du \frac{\mathbb{E}[dx_t d\xi_t(u)]}{dt} \\
C^{\xi\xi} &= \int_0^T dt \int_t^T ds \int_t^T du \frac{\mathbb{E}[d\xi_t(s) d\xi_t(u)]}{ds} \\
C^\mu &= \int_0^T dt \int_t^T du \frac{\mathbb{E}[dx_t d\xi_t(u)]}{dt} \frac{\delta C_t^{x\xi}}{\delta \xi_t(u)}
\end{aligned} \tag{7.3}$$

where the notation  $\delta/\delta\xi_t(u)$  denotes a functional derivative.

In the case of the rBergomi model (3.1), we have

$$\begin{aligned}
\frac{dS_t}{S_t} &= \sqrt{\xi_t(t)} dZ_t \\
\frac{d\xi_t(u)}{\xi_t(u)} &= \eta \sqrt{2H} \frac{dW_t}{(u-t)^\gamma}
\end{aligned}$$

with  $\mathbb{E}[dZ_t dW_t] = \rho dt$  so that

$$\frac{\mathbb{E}[dx_t d\xi_t(u)]}{dt} = \rho \eta \sqrt{2H} \sqrt{\xi_t(t)} \frac{\xi_t(u)}{(u-t)^\gamma}.$$

The various autocorrelation functionals (7.3) may then be computed; explicit computations are presented in Appendix A.

## 7.1 Special case: Flat variance curve

The special case  $\xi_0(u) = \bar{\sigma}^2$ ,  $u \geq 0$  where the initial forward variance curve is flat, is particularly instructive. First, from (A.1), we have

$$\begin{aligned}
C^{x\xi} &= \rho \eta \sqrt{2H} \int_0^T \sqrt{\xi_t(s)} ds \int_s^T \xi_t(u) \frac{du}{(u-s)^\gamma} \\
&= \rho \eta \bar{\sigma}^3 E_H T^{H+3/2}.
\end{aligned} \tag{7.4}$$

where we have further defined

$$E_H = \frac{D_H}{H + 3/2}.$$

Also,  $w = \bar{\sigma}^2 T$ . Substituting back into (7.2) and then (7.1) gives, to first order in  $\eta$ ,

$$\sigma_{\text{BS}}(k, t) = \bar{\sigma} + \frac{\rho \eta}{2} E_H \frac{1}{T^{1/2-H}} \left( k + \frac{w}{2} \right) + O(\eta^2) \quad (7.5)$$

In particular, we see that to first order in  $\eta$ , the term structure of at-the-money volatility skew is given by

$$\psi(\tau) = \left. \frac{\partial \sigma_{\text{BS}}^2(k, \tau)}{\partial k} \right|_{k=0} = \frac{\rho \eta}{2} E_H \frac{1}{\tau^\gamma}$$

with  $\gamma = 1/2 - H$ . Similarly, substituting  $\xi_0(u) = \bar{\sigma}^2$  into (A.2) and (A.3) respectively gives the terms required for computation of the second order contribution:

$$\begin{aligned} C^{\xi\xi} &= \eta^2 \bar{\sigma}^4 D_H^2 \frac{T^{2+2H}}{2+2H} \\ C^\mu &= \frac{1}{2} \rho^2 \eta^2 \bar{\sigma}^4 D_H^2 \left\{ 1 + \frac{\Gamma(H+3/2)^2}{\Gamma(2H+3)} \right\} \frac{T^{2+2H}}{2+2H}. \end{aligned}$$

It follows that to second order in  $\eta$ , the term structure of at-the-money volatility skew is given by

$$\psi(\tau) = \frac{\rho \eta}{2} E_H \frac{1}{\tau^\gamma} + \frac{1}{4} \rho^2 \eta^2 \bar{\sigma} \tau^{2H} \left[ \frac{D_H^2}{1+H} \left\{ 1 + \frac{\Gamma(H+3/2)^2}{\Gamma(2H+3)} \right\} - \frac{3}{2} E_H^2 \right]. \quad (7.6)$$

## Numerical test

The dimensionless Bergomi-Guyon expansion parameter is  $\lambda = \eta T^H$ . When  $H$  is very small,  $\lambda \sim \eta$  for all reasonable expirations; with  $H < 0.1$  as in Section 5.3,  $\lambda \sim 1.9$  which is not small enough for the asymptotic expansion to converge, even at-the-money. With the much smaller value  $\eta = 0.4$ , we see in Figure 7.1 very good agreement between the Bergomi-Guyon asymptotic skew formula (7.6) and the simulation.

We thus conclude that both our Bergomi-Guyon computations and the simulation are likely to be correct. Sadly however, the Bergomi-Guyon expansion does not converge with values of  $\eta$  consistent with the SPX volatility surface, so the Bergomi-Guyon expansion is not useful in practice for calibration of the rBergomi model.

## 8 Forecasting the variance swap curve

Recall that in the RFSV model (1.1),  $\log v_t \approx 2\nu W_t^H + C$  for some constant  $C$ . In [14], it is shown, assuming  $H \in (0, 1/2)$ ,  $\Delta > 0$ , that  $W_{t+\Delta}^H$  is conditionally Gaussian with conditional expectation<sup>9</sup>

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

and conditional variance

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

<sup>9</sup>Trivially  $\mathbb{E}[W_{t+\Delta}^H | \mathcal{F}_t] = W_t^H$  when either  $\Delta = 0$  or  $H = 1/2$ . This corresponds to the singular behavior of the integrand, as either  $\Delta \rightarrow 0$  or  $H \rightarrow 1/2$ , with the necessary compensation given by  $\Delta^{H+1/2} \cos(H\pi) \sim 0$  in these regimes.

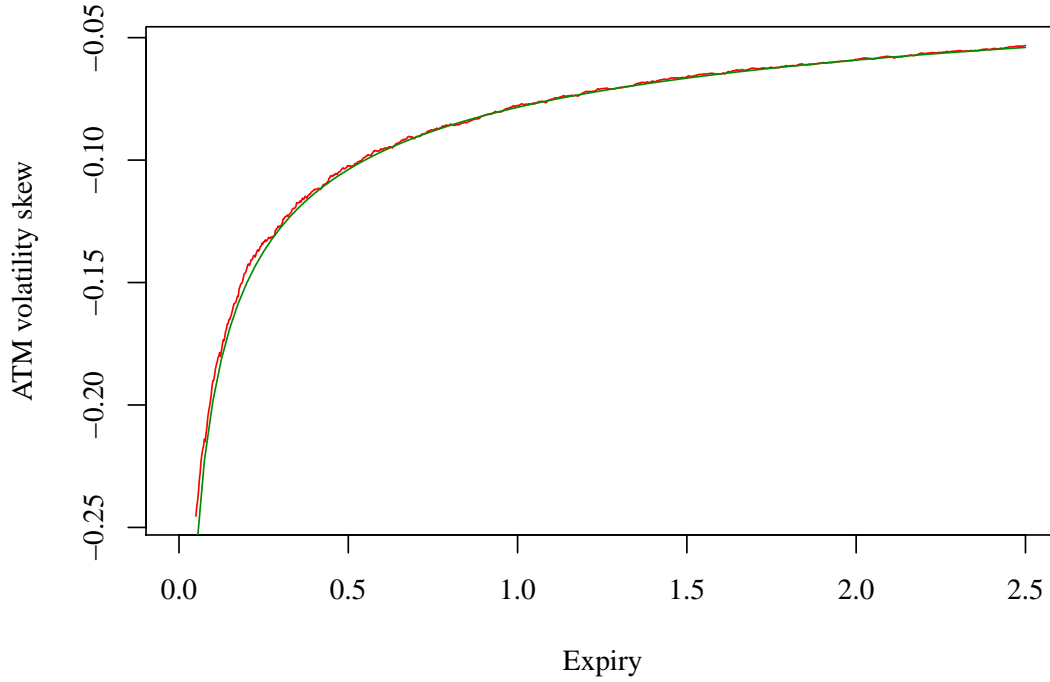


Figure 7.1: The Bergomi-Guyon second order ATM skew approximation is in green; ATM skews from Monte Carlo simulation are in red. Parameters used were  $H = 0.1, \eta = 0.4, \rho = -0.85, \bar{\sigma} = 0.235$ .

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

and conditional variance

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

where

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

Thus, we obtain the following natural form for the RFSV predictor of the variance:

$$\mathbb{E}^{\mathbb{P}} [v_{t+\Delta} | \mathcal{F}_t] = \exp \left\{ \mathbb{E}^{\mathbb{P}} [\log(v_{t+\Delta}) | \mathcal{F}_t] + 2c\nu^2 \Delta^{2H} \right\} \quad (8.1)$$

where

$$\mathbb{E}^{\mathbb{P}} [\log v_{t+\Delta} | \mathcal{F}_t] = \frac{\cos(H\pi)}{\pi} \Delta^{H+1/2} \int_{-\infty}^t \frac{\log v_s}{(t-s+\Delta)(t-s)^{H+1/2}} ds. \quad (8.2)$$

The fair value of a  $\tau$ -maturity variance swap is given (approximately) by

$$\mathcal{V}_t(\tau) = \frac{1}{\tau} \int_t^{t+\tau} \mathbb{E}^{\mathbb{Q}} [v_s | \mathcal{F}_t] ds$$

where  $\mathbb{Q}$  is the risk neutral measure. If it were possible to ignore the change of measure so that

$$\mathbb{E}^{\mathbb{Q}} [v_s | \mathcal{F}_t] = \mathbb{E}^{\mathbb{P}} [v_s | \mathcal{F}_t],$$

it would be possible to forecast variance swap curves using (8.1). In fact, we will see that from the data,  $\mathbb{Q}$  is close to  $\mathbb{P}$  in this sense. We now proceed to compare forecast and actual variance swaps curves.

SPX variance curve forecasts are formed using the predictor (8.1) from the time series of daily realized variance estimates from same Oxford-Man dataset that was used in [12].

As for market variance swap curves, although there is an active market, it is not easy to obtain high quality variance swap quote data and in any case, the bid/ask spread is wide. We thus choose to proxy the value of a  $\tau$ -maturity variance swap by the value of a  $\tau$ -expiration log contract as explained for example in Chapter 11 of [10]. To price the log contract for a particular expiration  $\tau$  requires us to know the prices of  $\tau$ -expiration options for all strikes; of course prices are only quoted for a finite number of strikes. We therefore choose to interpolate and extrapolate observed implied volatilities using the arbitrage-free SVI parameterization of the volatility surface as explained in [11]. With closing prices of SPX options for all available strikes and expirations sourced from OptionMetrics ([www.optionmetrics.com](http://www.optionmetrics.com)) via Wharton Research Data Services (WRDS), we follow the procedure just described to compute proxy variance swap curves each day from January 4, 1996 to August 30, 2013, a total of 4,443 days. We also need a sufficient history to be able to compute a forecast. We end up with 2,681 days of forecast and actual variance swap curves from Jan 3, 2003 to August 31, 2013.

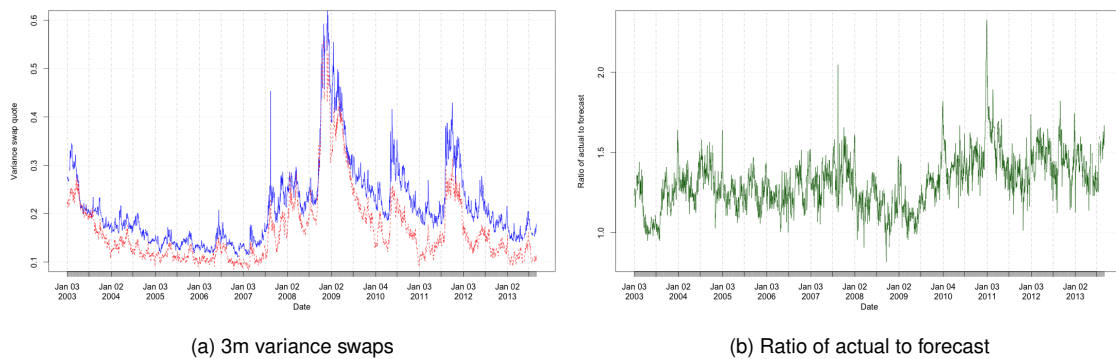
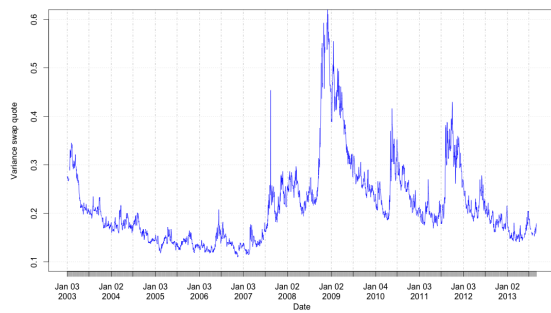


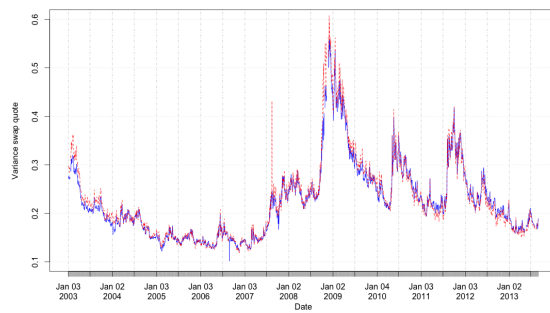
Figure 8.1: Plot (a) shows actual (proxy) 3-month variance swap quotes in blue vs forecast in red. Plot (b) shows the ratio between 3-month actual variance swap quotes and 3-month forecasts.

Plotting actual versus forecast 3-month variance swap curves in Figure 8.1, we immediately see that the actual variance swap curve is a factor (of roughly 1.4) higher than the forecast one, which we may attribute to overnight movements of the index. Recall that RV estimates are intraday from open to close. Realized variance forecasts must therefore be rescaled to obtain close-to-close realized variance forecasts as explained for example in [8] or alternatively using an econometric model such as the HEAVY model of [13].

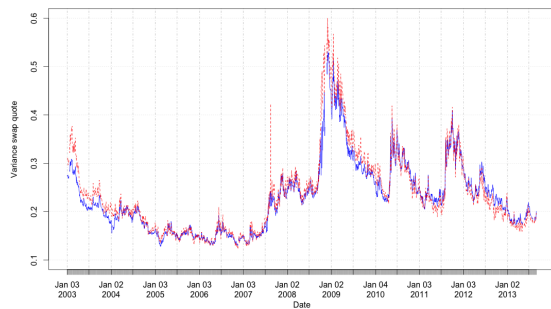
In Figure 8.2, we see that the 6-month, 9-month, and 12-month forecasts rescaled by the 3-month ratio of actual to forecast seem to be very consistent with actual variance swap quotes. This implies that although we can only forecast variance swap curves up to a factor, we can accurately forecast their shapes. We now demonstrate this further with two dramatic examples where the variance swap curve moved significantly from one day to the next. We will see that in both these cases, the evolution of the variance swap curve seems to be consistent with our model paradigm.



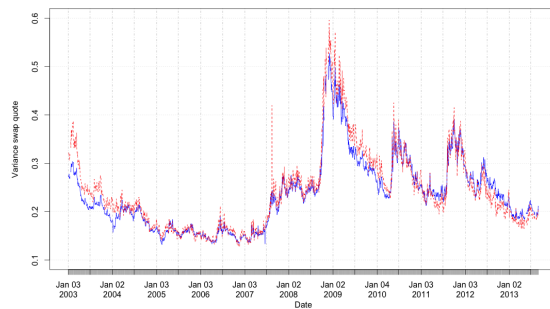
(a) Actual 3m variance swaps



(b) Actual vs rescaled forecast 6m variance swaps



(c) Actual vs rescaled forecast 9m variance swaps



(d) Actual vs rescaled forecast 12m variance swaps

Figure 8.2: Plot (a) shows actual (proxy) 3-month variance swap quotes. The other 3 figures show actual variance swap quotes for 6, 9, and 12 month respectively in blue with forecast variance swap quotes multiplied by the 3-month actual to forecast ratio in red.

## 8.1 The collapse of Lehman Brothers

As an interesting experiment<sup>10</sup>, consider the evolution of the S&P variance swap curve over the weekend of the collapse of Lehman Brothers. As of the market close on Friday September 12, 2008, it was generally expected that Lehman Brothers would be rescued over the weekend. As of the market close on Monday September 15 however, there had been no rescue and the market was in crisis. In Figure 8.3, we plot the actual variance swaps curves as of the Friday and Monday market closes together with forecast curves rescaled by the 3-month ratio as of the close on Friday September 12 (which was 1.29). Perhaps surprisingly, it appears that most of the evolution of the variance swap curve may be explained by a single extra data point – intraday realized variance from the open to the close of trading on Monday September 15, 2008.

## 8.2 The Flash Crash

In the so-called Flash Crash of Thursday May 6, 2010, major US equity indices suddenly dropped by about 10% intraday only to recover within 30 minutes or so. Consequently, intraday realized variance

<sup>10</sup>Suggested by Peter Leoni of KU Leuven.



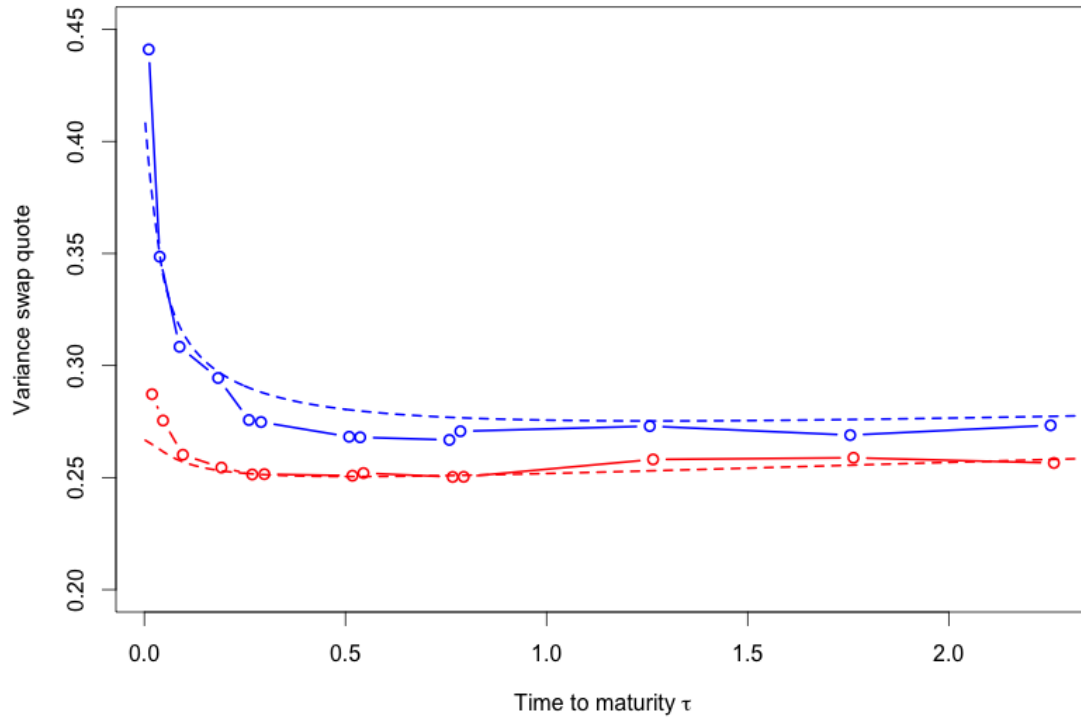


Figure 8.3: S&P variance swap curves as of September 12, 2008 (red) and September 15, 2008 (blue). The dashed curves are RFSV model forecasts rescaled by the 3-month ratio (1.29) as of the Friday close.

for May 6 was much higher than normal. In Figure 8.4, using the same methodology as in Section 8.1, we plot the actual variance swap curves as of the Wednesday and Friday market closes together with forecast curves rescaled by the 3-month ratio as of the close on Wednesday May 5 (which was 2.52). We see that the actual variance curve as of the close on Friday is consistent with a forecast from the time series of realized variance that *includes* the anomalous price action of Thursday May 6. In Figure 8.5 we see that actual variance swap curve as of the following Monday close is no longer consistent with the forecast. However, if we drop the May 6 datapoint, we get a forecast that is much closer to the actual variance swap curve. The obvious explanation is that volatility traders realized over the weekend that the anomalous intraday price action of the Flash Crash should not influence future realized variance projections, adjusting index option quotes accordingly.

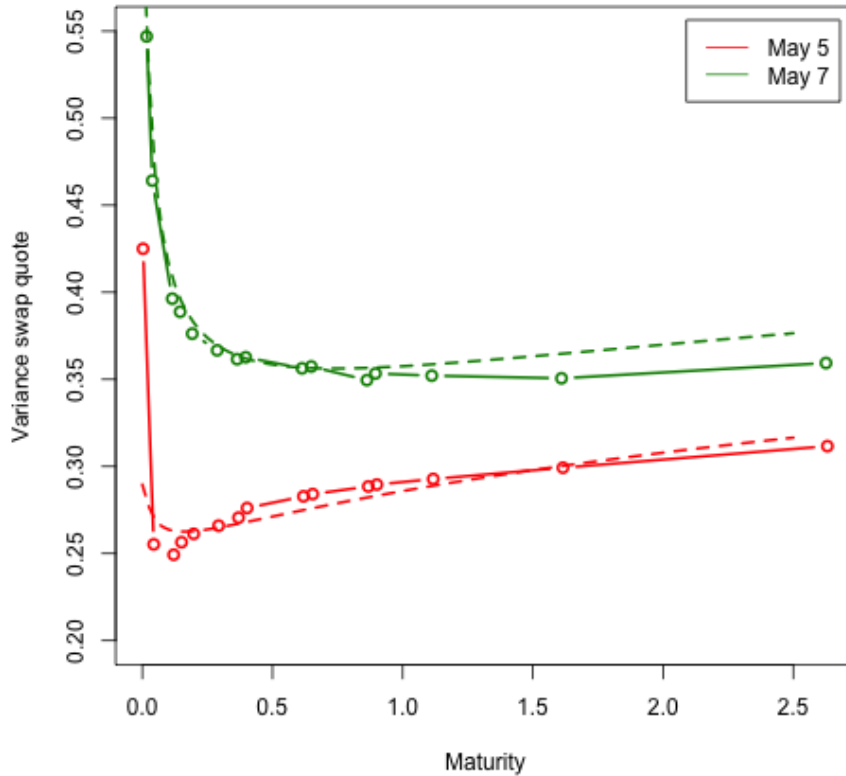


Figure 8.4: S&P variance swap curves as of May 5, 2010 (red) and May 7, 2010 (green). The dashed curves are RFSV model forecasts rescaled by the 3-month ratio (2.52) as of the close on Wednesday May 5.

## 9 Summary and conclusions

The Rough Fractional Stochastic Volatility (RFSV) model of [12] is remarkably consistent with the time series of realized volatility for a wide range of different underlying assets. In this paper, we have shown that this model written under the physical measure  $\mathbb{P}$  leads naturally to an options pricing model under  $\mathbb{Q}$  that is remarkably consistent with the observed shape of the implied volatility surface in the particular case of SPX. A special case of this model where we assume a deterministic change of measure between  $\mathbb{P}$  and  $\mathbb{Q}$  turns out to be a non-Markovian extension of the well-known Bergomi model, which we consequently dub the Rough Bergomi (or rBergomi) model. The rBergomi model is particularly tractable and seems to fit the SPX volatility surface very well, despite our lack at this stage of an efficient computational algorithm. We computed terms Bergomi-Guyon expansion up to second order in volatility of volatility but the expansion parameter  $\lambda = \eta \tau^H \approx 2$  required to fit SPX option prices is too big for this asymptotic expansion to be valid. However, we do not see agreement between the rBergomi model and the term structure of VIX volatility (VVIX).

Taken together, the present work and the econometric analysis of [12] offer a (perhaps even the first) promising paradigm for the understanding of asset price formation all the way from a basic microstruc-

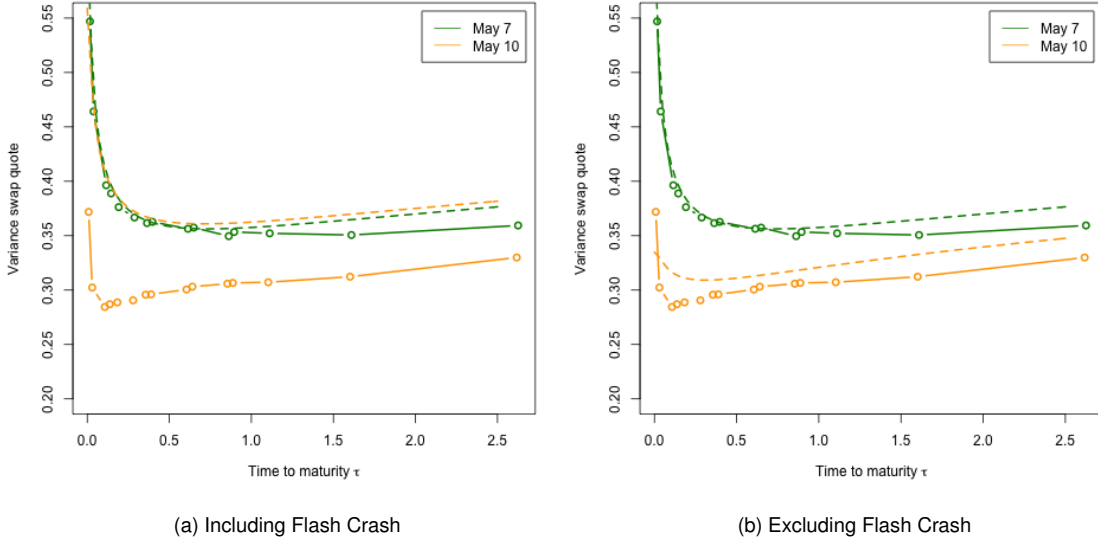


Figure 8.5: May 7 variance curves are in green; May 10 variance curves are in orange. Solid lines are actual curves and dashed lines are forecast curves. In plot (a), we see that the May 10 actual curve is completely inconsistent with a forecast based on a realized variance dataset that includes the Flash Crash. In contrast, in plot (b), we see that the Monday May 10 actual curve is consistent with a forecast using a dataset that excludes the Flash Crash.

ture description at the order book level to option pricing. Not least, our framework allows for accurate prediction of the volatility surface from high-frequency price data. More efficient computational methods and a more thorough investigation of the market implied change of measure  $d\mathbb{Q}/d\mathbb{P}$  are left for further research.

## A Computation of Bergomi-Guyon autocorrelation functionals in the rBergomi model

### Computation of $C^{x\xi}$

$$\begin{aligned}
C_t^{x\xi} &= \int_t^T ds \int_s^T du \frac{\mathbb{E}[dx_s d\xi_s(u)]}{ds} \\
&= \rho\eta\sqrt{2H} \int_t^T ds \int_s^T \mathbb{E}\left[\sqrt{\xi_s(s)} \xi_s(u) \middle| \mathcal{F}_t\right] \frac{du}{(u-s)^\gamma} \\
&= \rho\eta\sqrt{2H} \int_t^T \sqrt{\xi_t(s)} ds \int_s^T \xi_t(u) \frac{du}{(u-s)^\gamma} + \mathcal{O}(\eta^3). \tag{A.1}
\end{aligned}$$

### Computation of $C^{\xi\xi}$

By definition,

$$\begin{aligned}
C^{\xi\xi} &= \int_0^T dt \int_t^T du \int_t^T ds \frac{\mathbb{E}[d\xi_t(s) d\xi_t(u)]}{dt} \\
&= \int_0^T dt \int_t^T du \int_t^T ds \frac{\eta^2 2H}{(u-t)^\gamma (s-t)^\gamma} \xi_t(s) \xi_t(u) \\
&= \eta^2 2H \int_0^T dt \left( \int_t^T \frac{\xi_t(u)}{(u-t)^\gamma} du \right)^2 \\
&= \eta^2 2H \int_0^T dt \left( \int_t^T \frac{\xi_0(u)}{(u-t)^\gamma} du \right)^2 + \mathcal{O}(\eta^4). \tag{A.2}
\end{aligned}$$

### Computation of $C^\mu$

By definition,

$$C^\mu = \int_0^T dt \int_t^T du \frac{\mathbb{E}[dx_t d\xi_t(u)]}{dt} \frac{\delta C_t^{x\xi}}{\delta \xi_t(u)}$$

and from (A.1) above,

$$\begin{aligned}
\frac{\delta C_t^{x\xi}}{\delta \xi_t(v)} &= \rho \eta \sqrt{2H} \left\{ \int_t^T ds \sqrt{\xi_t(s)} \frac{1}{(v-s)^\gamma} \mathbf{1}_{v>s} + \frac{1}{2\sqrt{\xi_t(v)}} \int_v^T \xi_t(u) \frac{du}{(u-v)^\gamma} \right\} \\
&= \rho \eta \sqrt{2H} \left\{ \int_t^v ds \sqrt{\xi_t(s)} \frac{1}{(v-s)^\gamma} + \frac{1}{2\sqrt{\xi_t(v)}} \int_v^T \xi_t(u) \frac{du}{(u-v)^\gamma} \right\}.
\end{aligned}$$

Thus

$$\begin{aligned}
C^\mu &= \rho^2 \eta^2 2H \int_0^T \sqrt{\xi_t(t)} dt \int_t^T du \frac{\xi_t(u)}{(u-t)^\gamma} \\
&\quad \times \left\{ \int_t^u ds \sqrt{\xi_t(s)} \frac{1}{(u-s)^\gamma} + \frac{1}{2\sqrt{\xi_t(u)}} \int_u^T \xi_t(s) \frac{ds}{(s-u)^\gamma} \right\} \\
&= \rho^2 \eta^2 2H \int_0^T \sqrt{\xi_t(t)} dt \int_t^T \frac{du}{(u-t)^\gamma} \\
&\quad \times \left\{ \int_t^u \sqrt{\xi_t(s)} \frac{\xi_t(u)}{(u-s)^\gamma} ds + \frac{1}{2} \sqrt{\xi_t(u)} \int_u^T \frac{\xi_t(s)}{(s-u)^\gamma} ds \right\} \\
&= \rho^2 \eta^2 2H \int_0^T \sqrt{\xi_0(t)} dt \int_t^T \frac{du}{(u-t)^\gamma} \\
&\quad \times \left\{ \int_t^u \sqrt{\xi_0(s)} \frac{\xi_0(u)}{(u-s)^\gamma} ds + \frac{1}{2} \sqrt{\xi_0(u)} \int_u^T \frac{\xi_0(s)}{(s-u)^\gamma} ds \right\} + \mathcal{O}(\eta^4). \tag{A.3}
\end{aligned}$$

For any given initial forward variance curve  $\xi_0(u)$ , the above expressions for  $C^{x\xi}$ ,  $C^{\xi\xi}$  and  $C^\mu$  may be easily computed numerically.

## B Approximate variance of VIX

Let  $y_u = \log v_u$  and consider the following approximation of the arithmetic mean by the geometric mean:

$$\psi(T) = \frac{1}{\Delta} \int_T^{T+\Delta} \mathbb{E}[v_u | \mathcal{F}_T] du \approx \exp \left\{ \frac{1}{\Delta} \int_T^{T+\Delta} \mathbb{E}[y_u | \mathcal{F}_T] du \right\}.$$

Apart from  $\mathcal{F}_t$  measurable terms (abbreviated as “drift”), we have

$$\begin{aligned} \int_T^{T+\Delta} E[y_u | \mathcal{F}_T] du &= \eta \sqrt{2H} \int_t^T \frac{dW_s}{(u-s)^\gamma} du + \text{drift} \\ &= \eta \sqrt{2H} \int_t^T \int_T^{T+\Delta} \frac{du}{(u-s)^\gamma} dW_s + \text{drift} \\ &= \eta \frac{\sqrt{2H}}{1-\gamma} \int_t^T [(T+\Delta-s)^{1-\gamma} - (T-s)^{1-\gamma}] dW_s + \text{drift}. \end{aligned}$$

This gives

$$\begin{aligned} \text{Var}[\log \psi(T) | \mathcal{F}_t] &\approx \frac{\eta^2 D_H^2}{\Delta^2} \int_t^T \left[ (T+\Delta-s)^{1/2+H} - (T-s)^{1/2+H} \right]^2 ds \\ &= \eta^2 (T-t)^{2H} f^H \left( \frac{\Delta}{T-t} \right) \end{aligned}$$

where

$$D_H = \frac{\sqrt{2H}}{H+1/2}$$

and

$$f^H(\theta) = \frac{D_H^2}{\theta^2} \int_0^1 \left[ (1+\theta-x)^{1/2+H} - (1-x)^{1/2+H} \right]^2 dx.$$

To compute this integral explicitly, we use that, for  $\kappa = 1/2 + H$ ,

$$\begin{aligned} \int_t^T (T+\Delta-s)^\kappa (T-s)^\kappa ds &= \int_0^\tau (s+\Delta)^\kappa s^\kappa ds \\ &= \frac{\tau(\Delta\tau)^\kappa}{1+\kappa} {}_2F_1 \left( -(H+1/2), 3/2+H, 5/2+H, -\tau/\Delta \right). \end{aligned}$$

Thus, we get

$$\begin{aligned} \text{Var}[\log \psi(T) | \mathcal{F}_t] &\approx \frac{2H\eta^2}{\Delta^2} \left[ \frac{(\tau+\Delta)^{2(H+1)} - \Delta^{2(H+1)} + \tau^{2(H+1)}}{2(H+1)} \right. \\ &\quad \left. - 2 \frac{\tau(\tau\Delta)^{H+1/2}}{H+3/2} {}_2F_1 \left( -(H+\frac{1}{2}), H+\frac{3}{2}, H+\frac{5}{2}, -\frac{\tau}{\Delta} \right) \right] \end{aligned}$$

with  $\tau := T - t$ .

## References

- [1] E. Alòs, J. A. León, and J. Vives. On the short-time behavior of the implied volatility for jump-diffusion models with stochastic volatility. *Finance and Stochastics*, 11(4):571–589, Aug. 2007.
- [2] L. Bergomi. Smile dynamics II. *Risk October*, pages 67–73, 2005.
- [3] L. Bergomi and J. Guyon. Stochastic volatility’s orderly smiles. *Risk May*, pages 60–66, 2012.
- [4] P. Carr and L. Wu. What type of process underlies options? A simple robust test. *Journal of Finance*, 58(6):2581–2610, 2003.
- [5] P. Cheridito, H. Kawaguchi, and M. Maejima. Fractional Ornstein-Uhlenbeck processes. *Electron. J. Probab*, 8(3):14, 2003.
- [6] F. Comte, L. Coutin, and E. Renault. Affine fractional stochastic volatility models. *Annals of Finance*, 8(2-3):337–378, 2012.
- [7] F. Comte and E. Renault. Long memory continuous time models. *Journal of Econometrics*, 73(1):101–149, 1996.
- [8] F. Corsi, N. Fusari, and D. La Vecchia. Realizing smiles: Options pricing with realized volatility. *Journal of Financial Economics*, 107(2):284–304, 2013.
- [9] M. Fukasawa. Asymptotic analysis for stochastic volatility: Martingale expansion. *Finance and Stochastics*, 15(4):635–654, 2011.
- [10] J. Gatheral. *The volatility surface: A practitioner’s guide*. John Wiley & Sons, 2006.
- [11] J. Gatheral and A. Jacquier. Arbitrage-free SVI volatility surfaces. *Quantitative Finance*, 14(1):59–71, 2014.
- [12] J. Gatheral, T. Jaisson, and M. Rosenbaum. Volatility is rough. *Available at SSRN 2509457*, 2014.
- [13] D. Noureldin, N. Shephard, and K. Sheppard. Multivariate high-frequency-based volatility (heavy) models. *Journal of Applied Econometrics*, 27(6):907–933, 2012.
- [14] C. J. Nuzman and V. H. Poor. Linear estimation of self-similar processes via Lamperti’s transformation. *Journal of Applied Probability*, 37(2):429–452, 2000.