Rough Bergomi revisited - exact and minimal-variance hedging for VIX and European options, and exact calibration to multiple smiles

Martin Forde

21st April 2024

Abstract

For a generalized rough Bergomi-type model, we formally show how to replicate a VIX option with dynamic trading in a VIX future, and a European option with dynamic trading in the underlying and a VIX future, using the Clark-Ocone formula from Malliavin calculus. As a by-product we also compute the minimal variance hedge for a European call when we can only dynamically hedge with the underlying which is relevant in practice since dynamic trading with a VIX future will incur a larger bid-offer spread, and these results are easily extended to mixed/two-factor rough Bergomi models which give better fits to VIX smiles in practice. This builds on the work of Keller-Ressel[KR22] who derives asymptotic approximations for the latter, and as a by-product we obtain a variant of the classical Bass martingale (in this case a path-dependent rough local-stochastic volatility model) with an exact fit to target laws $\mu_1, \mu_2, ...$ at multiple maturities $T_1 < T_2 < ..$ (with $\mu_1, \mu_2, ..$ in convex order) where the volatility process can be characterized explicitly using the Clark-Ocone formula. We also explain how to adapt the well known Renault-Touzi[RT96] conditioning trick to reduce the sample variance of Monte Carlo estimates for the European call hedge at each time instant.[1]

1 Introduction

The Rough Bergomi (rBergomi) model introduced in [BFG16] has been a popular, tractable and much cited rough volatility model. For the rBergomi model and the original Rough Fractional Stochastic Volatility (RFSV) model driven by fBM in [GJR18], the log of the instantaneous variance process V is Gaussian so the VIX index is approximately log-normally distributed and hence produces VIX smiles which are almost flat, not the concave upward-sloping smiles that we typically see in practice. A skewed Rough Bergomi model with a linear combination of exponential terms with two different ν-values as discussed in [Guy21] (see also [JMP21] for extensions/variations e.g. using two different H-values and [AL24]) can often fit a single-maturity short-maturity VIX smile very well but if we do this it typically struggles to achieve sufficient at-the-money skew for options on the SPX itself with the same maturity (see [Guy21] and we have also seen this phenomenon first hand in testing).

Asymmetric GARCH(1,1) (also known as QGARCH) models with i.i.d. symmetric or skewed t-distributed residuals typically fit daily historical returns data much better than standard rough volatility models across a wide range of assets when we apply goodness of fit tests (Kolmogorov-Smirnov, Shapiro-Wilks etc) to the residuals implied by daily returns using maximum likelihood estimates for the model parameters which are easily computed (see e.g. [F23II], [NPP14]). These techniques are well known in the econometrics literature, but aside from [F23] we have not seen any articles which examine maximum likelihood estimates and p-values for rough models. Unfortunately the distribution of the MLEs when we run synthetic simulations of the QGARCH(1,1) model with the fitted MLE parameters are typically much wider than we would ideally like. The solution to this issue in principle is just to use decades of data to reduce the sample variance of the MLEs (assuming the MLEs are consistent estimators) but obviously the further we go back in time in practice the more likely that the dynamics of the asset will have changed, and using intraday data is too time-inhomogenous due to markets opening/closing, lunch etc. For many assets an excellent fit is obtained just using the usual symmetric t-distribution for the residuals but for the SPX we typically need the non-symmetric t-distribution which has an additional asymmetry parameter.

Rough volatility models are typically much better than the aforementioned 1-day timestep GARCH(1,1) models at fitting observed option prices, specifically the steep short-maturity implied volatility skews we observe in practice at e.g. 1 month maturity and VIX option smiles, so there appears to be something of a disparity between option prices and historical behaviour of the assets they are written which may lead to statistical arbitrage opportunities.

The quadratic rough Heston model introduced in [GR20] is complete as it is driven by a single Brownian motion, and there is an explicit formula for sampling the VIX (cf. chapter 6.2 in [Rom22b]), which is obtained via the solution to a linear VIE in terms of the resolvent of the fractional kernel of the Z process. Using the Gamma kernel $K(t) = e^{-\lambda t} t^{\alpha-1}$, the model often has an uncanny ability to fit close-to-1-month and 2 month SPX and 1month VIX smiles simultaneously very well with only 5 parameters (α, a, c, λ and θ, setting $b = 0$ W.L.O.G.) and also fitting

[1] We thank Alan Lewis as always for many stimulating discussions.

Z_0, but calibrated $H=\alpha-\frac{1}{2}$ values can bounce around from as high as 0.14 to as low as 0.04 between Jan 2023 to Jan 2024 so out-of-sample fits typically do not work well. Exact hedging for the (non-quadratic) rough Heston model is formulated in [ER18], where the call option price satisfies an infinite-dimensional PDE in terms of a Frechet derivative with respect to the entire forward variance cuve $\xi_t(u)$ which evolves as $d\xi_t(u)=\frac{1}{\lambda}f^{\alpha,\lambda}(u-t)\nu\sqrt{V_t}dB_t$ where $f^{\alpha,\lambda}$ is the Mittag-Leffler function, but it is rather difficult to implement this in practice since the call payoff also has to be re-written as a Fourier integral involving complex exponential contracts, each of which can then in turn be replicated with dynamic trading in the underlying and a continuum of forward variance contracts.

Section 2.2 in Keller-Ressel[KR22] gives a concise background on the mean-variance hedge so we do not repeat this here, and derives asymptotic approximations for the mean variance hedge using the original SABR formula for the SABR model with $\beta=1$ and the recent rough SABR formula from [FG22] for the rough Bergomi case (see also section 4.1 in [Schw95] for the original formulation of the discrete-time variance optimal hedge). For the QGARCH(1,1) model discussed above, we can also use deep learning to approximate the mean-variance hedge by exploiting the Markov nature of the model, essentially just adding an extra dimension to existing code which uses deep hedging to approximate the classical Black-Scholes hedging strategy. One can also attempt to price options with transaction costs using deep hedging with exponential indifference pricing but for this we need to keep track of the agent's risky wealth as additional state variable (see many articles by Buehler et al. on this theme).

[CL21] shows how to calibrate a (one-dimensional) Bass[Bass83] martingale to given marginals at two different maturities; fitting a single maturity is elementary, but jointly fitting to two maturities requires an iterative fixed point scheme of the form $F^{n+1}=\mathcal{A}F^n$ for some non-linear integral operator $\mathcal{A}$ (see Theorem 2.1 in [CL21]) where $\mathcal{A}$ is a map from the space of distribution functions on $\mathbb{R}$ to itself. The aforementioned fixed point scheme just requires numerically computing two Gaussian convolution integrals, inverting a cdf and then iterating the procedure, for which [AMP23] establish existence and uniqueness (and linear convergence) results, and (in our experience) the scheme converges very quickly in practice.

The one-dimensional Bass martingale is also the solution to the martingale optimization problem $\inf_{X\in\mathcal{M}^c:X_t=X_0+\int_0^t\sigma_s dW_s:X_T\sim\mu}\mathbb{E}(\int_0^T(\sigma_t-1)^2dt)$, (where $\mathcal{M}^c$ is the space of continuous martingales) which is clearly also the solution to $\sup_{X\in\mathcal{M}^c:X_t=X_0+\int_0^t\sigma_s dW_s:X_T\sim\mu}\mathbb{E}(\int_0^T\sigma_t dt)$ (see e.g. introduction of [BST23] and section 1.3 in [BBHK20]); hence the Bass martingale is a *stretched Brownian motion*[2], which (formally at least) can also be dualized as $\sup_{f\in C_b(\mathbb{R})}(-\int_{\mathbb{R}}fd\mu+\inf_{\sigma\in\mathcal{A}}(\mathbb{E}(f(X_T)+\int_0^T(\sigma_t-1)^2dt))$ (for a suitable space of adapted processes $\mathcal{A}$) in the spirit of [GLOW22],[GLW22], which leads to a HJB equation for the inner inf. The [GLOW22] methodology can in principle be generalized to work with a simple rough reference model using a variational approach, but one ends up with seemingly intractable non-standard FBSDE.

2 Hedging VIX options

Let W denote a standard Brownian motion and $\mathcal{F}_t=\mathcal{F}_t^W$, and consider a generalized Rough Bergomi model for a log stock price process $X_t=\log S_t$ for which the squared spot volatility V_t process satisfies

$$V_t = \xi_0(t)e^{Z_t-\frac{1}{2}\mathrm{Var}(Z_t)} \tag{1}$$

under a risk-neutral measure $\mathbb{Q}$, where $Z_t=\int_0^t\kappa(t-s)dW_s$ for some $\kappa\in L^2([0,T])$, so $\mathrm{Var}(Z_t)=\int_0^t\kappa(t-s)^2ds=\int_0^t\kappa(s)^2ds$. We can easily extend the results in this paper to the case when $V_t=\xi_0(t)e^{Z_t-\frac{1}{2}c^2\mathrm{Var}(Z_t)}$. A popular choice is the Gamma kernel: $\kappa(t)=t^{H-\frac{1}{2}}e^{-\theta t}$ for $H\in(0,\frac{1}{2}]$ and $\theta\geq 0$, where the roughness and ergodicity of Z are controlled by H and θ respectively. Then we can easily verify that $\xi_t(u):=\mathbb{E}(V_u|\mathcal{F}_t)$ satisfies

$$\xi_t(u) = \xi_0(u)\,e^{\int_0^t\kappa(u-r)dW_r-\frac{1}{2}\int_0^t\kappa(u-r)^2dr}$$

and

$$d\xi_t(u) = \kappa(u-t)\xi_t(u)dW_t \tag{2}$$

so $\xi_t(u)$ is a driftless time-inhomogenous Geometric Brownian motion for each u and $\xi_t(u)$ is an $\mathcal{F}_t$-martingale.

The VIX index is a well known estimator of future volatility, which is quoted in the market. Theoretically the value of the VIX index time is $t\leq T$ is given by $\mathrm{VIX}_t=\sqrt{\frac{1}{\Delta}\int_t^{t+\Delta}\xi_t(u)du}$ for some $\Delta>0$. Then we can consider an option on the VIX which pays

$$F = \phi(\mathrm{VIX}_T)$$

at time T (and we assume interest rates are zero for simplicity). For the specific case of a VIX call option, $\phi(x)=(x-K)^+$ and $\phi'(x)=1_{x>k}$, even though ϕ' is not Lipshitz, we can compute Malliavin derivatives using a suitable approximation procedure (see e.g. end of page 333 in Nualart[Nua06]).

[2] See [BST23], [BBST23] and [BBHK20] for more on this, and extension to higher dimensions and randomized X_0

From the Clark-Ocone formula, we have

$$F = \mathbb{E}(F) + \int_0^T \mathbb{E}(D_t^W F|\mathcal{F}_t)dW_t \tag{3}$$

where $D_t^W F$ is the Malliavin derivative of F with respect to W.

Recall that we compute $D_t^W F$ by perturbing W by a function $H(t)$, such that $\int_0^T h(t)^2 dt < \infty$ where $h(t) = H'(t)$ and $h \in L^2([0,T])$. We denote the perturbed value of F by $F(W + \varepsilon H)$. Then $D_t^W F$ is the (in general) random function such that

$$\lim_{\varepsilon\to 0} \frac{1}{\varepsilon}(F(W+\varepsilon H) - F(W)) = \int_0^T D_t^W F \cdot h(t)dt$$

if such a function exists. For our $F = \phi(\mathrm{VIX}_T)$ payoff here, formally using the chain rule, we see that

$$\begin{aligned} D_t^W F &= \phi'(\mathrm{VIX}_T) D_t^W \mathrm{VIX}_T = \phi'(\mathrm{VIX}_T) D_t^W \sqrt{\mathrm{VIX}_T^2} = \phi'(\mathrm{VIX}_T)\frac{1}{2}(\mathrm{VIX}_T^2)^{-\frac{1}{2}} D_t^W(\mathrm{VIX}_T^2) \\ &= \frac{\phi'(\mathrm{VIX}_T)}{2\mathrm{VIX}_T} \cdot D_t^W \frac{1}{\Delta}\int_T^{T+\Delta} \xi_T(u)du \\ &= \frac{\phi'(\mathrm{VIX}_T)}{2\mathrm{VIX}_T} \cdot \frac{1}{\Delta}\int_T^{T+\Delta} D_t^W \xi_T(u)du\,. \end{aligned} \tag{4}$$

Using that $\xi_t(u) = \xi_0(u)\, e^{\int_0^t \kappa(u-r)dW_r - \frac{1}{2}\int_0^t \kappa(u-r)^2 dr}$ we see that

$$\lim_{\varepsilon\to 0} \frac{1}{\varepsilon}(\xi_T(u)(W+\varepsilon H) - \xi_T(u)(W)) = \xi_T(u) \cdot \int_0^T \kappa(u-r)h(r)dr\,.$$

Then we can just read off $D_r^W \xi_T(u)$ as whatever function is in front of $h(r)$; in this case

$$D_t^W \xi_T(u) = \xi_T(u)\kappa(u-t)$$

hence

$$D_t^W F = \frac{\phi'(\mathrm{VIX}_T)}{2\mathrm{VIX}_T} \cdot \frac{1}{\Delta}\int_T^{T+\Delta} \xi_T(u)\kappa(u-t)du \tag{5}$$

and recall that $\mathrm{VIX}_T^2 = \frac{1}{\Delta}\int_T^{T+\Delta}\xi_T(u)du$. There are two integrals in this expression, which can can be computed using Gauss-Legendre quadrature; specifically we need to jointly sample $\xi_T(u)$ at the n-point Gaussian-Legendre quadrature abcissae values $(u_i^n)_{i=1}^n$ values for the interval $[T, T+\Delta]$ and $\log \xi_T(u)$ are jointly Gaussian, so in principle we can use the Cholesky decomposition for this although in practice this often fails because the covariance matrix for this is close to singular, so we resort to short time steps instead.

If $\phi(x) = x$, then a VIX call is just a VIX future, so theoretically we can replicate a VIX option using a VIX future, by holding

$$\frac{\mathbb{E}(\frac{\phi'(\mathrm{VIX}_T)}{2\mathrm{VIX}_T} \cdot \frac{1}{\Delta}\int_T^{T+\Delta}\xi_T(u)\kappa(u-t)du|\mathcal{F}_t^W)}{\mathbb{E}(\frac{1}{2\mathrm{VIX}_T} \cdot \frac{1}{\Delta}\int_T^{T+\Delta}\xi_T(u)\kappa(u-t)du|\mathcal{F}_t^W)}$$

VIX futures at each time instant t. This may be desirable in practice since the bid-offer spread on VIX futures (in percentage terms) may be lower than for VIX options or a variance swap synthetically replicated with a finite number of Europeans. As of 22 Dec 2023, the bid-ask spread on VIX futures was \$0.05 with the VIX index itself at 13.50, and the spread on close-to-the-money VIX options was \$0.03 to \$0.04.[3]

3 Hedging European options

Now consider the standard Rough Bergomi model for a log stock price process X_t:

$$\begin{aligned} X_t &= -\frac{1}{2}\int_0^t V_s ds + \int_0^t \sqrt{V_s}(\rho dW_s + \bar{\rho} dB_s) \\ V_t &= V_0 e^{\int_0^t \kappa(t-s)dW_s - \frac{1}{2}\int_0^t \kappa(s)^2 ds} \end{aligned} \tag{6}$$

[3]Data obtained from CBOE data services and Charles Schwab.

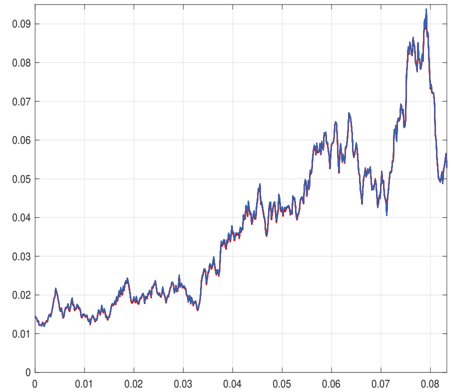

Figure 1: Here we have plotted the price of a VIX call option (blue) and the wealth process for the Clark-Ocone replication strategy (red) for a standard rough Bergomi model with $\kappa(t) = \nu t^{H-\frac{1}{2}}$ and $V_0 = .04$, $H = 0.1$, $\eta = 1$, $T = \frac{1}{12}$ and strike $K = .2$, and we see that the paths are more or less indistinguishable. We used 500 time steps for the single "outer" Monte Carlo path and 200 time steps with 250,000 paths and antithetic sampling for the nested Monte Carlo at each time point to compute $\mathbb{E}_t(D_t^W F)$ (see Eq (5)). We have used 20 Gauss-Legendre quadrature points to compute the VIX here, and we see that the blue and red lines are almost indistinguishable.

where $\bar{\rho} = \sqrt{1-\rho^2}$ and B is another Brownian motion independent of W, and we now define $\mathcal{F}_t := \mathcal{F}_t^{W,B}$, and we assume $\rho \in [-1, 0]$ which ensures that S is a true $\mathcal{F}_t$-martingale (see Gassiat[Gass19] for details). Now let $\Phi(W, B) := \phi(X_T)$, i.e. the payoff of a general European-type option. Then

$$\lim_{\varepsilon \to 0} \frac{1}{\varepsilon}(\Phi(W, B + \varepsilon H) - \Phi(W, B)) \quad = \quad \bar{\rho} \int_0^T \sqrt{V_t} h(t) dt$$

so we can read off that

$$D_t^B X_T \quad = \quad \bar{\rho}\sqrt{V_t}\,. \tag{7}$$

Similarly

$$\lim_{\varepsilon \to 0} \frac{1}{\varepsilon}(V_t(W + \varepsilon H) - V_t(W)) \quad = \quad V_t \int_0^t \kappa(t-r) h(r) dr$$

so $D_r^W V_t = \kappa(t-r) 1_{r \le t} V_t$ and $D_r^W \sqrt{V_t} = \frac{1}{2} V_t^{-\frac{1}{2}} \kappa(t-r) 1_{r \le t} V_t = \frac{1}{2}\kappa(t-r) 1_{r \le t} \sqrt{V_t}$. Thus

$$\begin{aligned} D_r^W X_T &= -\frac{1}{2} \int_0^T D_r^W V_s ds + \int_0^T D_r^W (\rho \sqrt{V_s} dW_s + \bar{\rho}\sqrt{V_s} dB_s) \\ &= -\frac{1}{2} \int_0^T \kappa(s-r) 1_{r \le s} V_s ds + \frac{1}{2}\bar{\rho} \int_0^T \kappa(s-r) 1_{r \le s} \sqrt{V_s} dB_s + \frac{1}{2}\rho \int_0^T \kappa(s-r) 1_{r \le s} \sqrt{V_s} dW_s + \rho\sqrt{V_r} \end{aligned}$$

so

$$D_t^W X_T \quad = \quad -\frac{1}{2} \int_t^T \kappa(s-t) V_s ds + \frac{1}{2}\bar{\rho} \int_t^T \kappa(s-t) \sqrt{V_s} dB_s + \frac{1}{2}\rho \int_t^T \kappa(s-t) \sqrt{V_s} dW_s + \rho\sqrt{V_t}\,. \tag{8}$$

Then from the two-dimensional Clark-Ocone formula, we have

$$F \quad = \quad \mathbb{E}(F) + \int_0^T \phi_t dB_t + \int_0^T \psi_t dW_t \tag{9}$$

where $F = \phi(X_T)$, $\phi_t = \mathbb{E}(D_t^B F | \mathcal{F}_t)$ and $\psi_t = \mathbb{E}(D_t^W F | \mathcal{F}_t)$, and hence

$$C_t \quad := \quad \mathbb{E}(F|\mathcal{F}_t) \quad = \quad \mathbb{E}(F) + \int_0^t \phi_s dB_s + \int_0^t \psi_s dW_s \tag{10}$$

and recall that $\mathcal{F}_t := \mathcal{F}_t^{W,B}$. From the chain rule, we know that

$$D_t^B F \quad = \quad \phi'(X_T) D_t^B X_T \quad , \quad D_t^W F \quad = \quad \phi'(X_T) D_t^W X_T \tag{11}$$

and we derived explicit expressions for $D_t^B X_T$ and $D_t^B X_T$ in (7) and (8) above.

Then using that

$$\begin{aligned} dS_t &= S_t\sqrt{V_t}(\rho dW_t + \bar\rho dB_t) \\ dC_t &= \psi_t dW_t \,+\, \phi_t dB_t \end{aligned}$$

we see that

$$d\langle C,S\rangle_t \;=\; S_t\sqrt{V_t}(\rho\psi_t \,+\, \bar\rho\phi_t)dt$$

so the minimal variance stock holding at time t is

$$\theta_t \;=\; \frac{d\langle C,S\rangle_t}{d\langle S_t\rangle} \;=\; \frac{\rho\psi_t \,+\, \bar\rho\phi_t}{S_t\sqrt{V_t}}$$

(see also section 10.4 in [CT04] for general background on mean variance hedging and application to exponential Lévy models).

3.1 Variance reduction for computing the hedge amount using Monte Carlo

A European call option corresponds to $\phi(x) = (e^x - e^k)^+$, and from the tower property we can reduce the sample variance of the numerical estimation of $\mathbb{E}(D_t^W F|\mathcal{F}_t)$ with Monte Carlo by conditioning on B (similar to the classic Renault-Touzi[RT96] conditioning trick) as

$$\mathbb{E}(D_t^W F|\mathcal{F}_t) \;=\; \mathbb{E}(\mathbb{E}(D_t^W F|\mathcal{F}_T^W)|\mathcal{F}_t)) \;=\; \mathbb{E}(\mathbb{E}(\phi'(X_T)D_t^W X_T)|\mathcal{F}_t)$$

and we can use that X_T and $D_t^W X_T$ are bivariate Normal conditioned on $\mathcal{F}_T^W$ to compute this expectation explicitly in terms of the Erf function in e.g. Mathematica (we omit the details here for the sake of brevity). As usual this trick is more effective when $|\rho|$ is smaller, and we gain no benefit when $|\rho| = 1$. We can also use antithetic sampling, even if $|\rho| = 1$. It does not appear trivial to adapt this to the two-maturity case as in [CL21] since the proof of the main result their relies heavily on the Markov structure of the problem.

4 Exact calibration to single or multiple smiles - a rough Bergomi Bass model

If $|\rho| < 1$ and ϕ is chosen so $F = \phi(X_T)$ has a given target law μ on $(0,\infty)$ with a strictly positive density [4] with $\int_0^\infty x\mu(x)dx = S_0$ then setting $S_t = \mathbb{E}(F|\mathcal{F}_t)$ in (10) yields a martingale price process $(S_t)_{t\in[0,T]}$ with $S_T \sim \mu$. In particular, since $D_t^B X_T = \bar\rho\sqrt{V_t}$ (Eq (7)) and $D_t^W X_T$ both include a $\sqrt{V_t}$ term (Eq (8)), and $D_t^B F = \phi'(X_T)D_t^B X_T$ (see Eq (11)), we see this model has a rough volatility component if $\kappa(t) \sim const. \times t^{\frac{1}{2}-H}$ (for $H \in (0,\frac{1}{2})$) as $t \to 0$, because in this case $\log V_t$ is a Gaussian process which is $H - \varepsilon$ Hölder continuous for $\varepsilon \in (0,H)$. We can view the $\mathbb{E}(\phi'(X_T)|\mathcal{F}_t)$ term in $\phi_t = \mathbb{E}(D_t^B F|\mathcal{F}_t)$ as a local volatility component since it can be re-written in terms of X_t. We can then also compute exact or mean-variance hedge quantities for options on X_T for this model using the same computations as Section 3.

If we wish to fit a rough Bergomi Bass model to two target densities μ and ν at maturities T and T_2 (both with mean S_0, with $0 < T < T_2$ and μ and ν in convex order), we first require a coupling $\pi \in \mathrm{Cpl}(\mu,\nu)$ in the martingale transport $\mathrm{MT}(\mu,\nu)$ of μ and ν, i.e. such that $\int y\pi_x(dy) = x$ where $\pi(dx,dy) = \pi_x(dy)\mu(dx)$, so $\mathbb{E}(S_{T_2}|S_T) = S_T$ (see e.g. section 2.2 of [BST23] for clarification on this notation) which gives us a a conditional distribution function $F_{S_{T_2}|S_T}$ for S_{T_2} given S_T. A viable/sensible choice for π could be to use the two-maturity Bass martingale discussed in [CL21] (see also [BBHK20]), or the Carr Local Variance Gamma model[Carr09].

By adapting the approach used in [BG24], a martingale model consistent with the two marginals here then takes the form

$$\begin{aligned} S_t &= \mathbb{E}(\phi(X_T)|\mathcal{F}_t^{W,B}) \qquad (t \in [0,T]) \\ S_t &= \mathbb{E}(F^{-1}_{S_{T_2}|S_T}(F_{\tilde X_{T_2-T}}(\tilde X_{T_2-T}), S_T)|\mathcal{F}_{t-T}^{\tilde W,\tilde B}) \qquad (t \in (T,T_2]) \end{aligned} \tag{12}$$

where $\tilde X$ is the log stock price for another rough Bergomi model of the form in (1) (independent of X, also driven by two independent Brownians $\tilde W$ and $\tilde B$). S is continuous on $[0,T_2]$, and in particular at $t = T$ since $\mathbb{E}(S_{T_2}|S_T) = S_T$ by construction. Note that the instantaneous variance process V for this model will not in general be continuous at T, but this is also the case for the standard two-maturity Bass martingale in [CL21], and we can also apply the Clark-Ocone formula to S for $t \in (T,T_2]$, and we can extend this construction to n maturities using the same conditional sampling trick.

[4] we can do this by setting $\phi(x) = F_\mu^{-1}(F_{X_T}(x))$, where F_μ is the distribution function of μ and F_{X_T} is the distribution function of X_T; then ϕ is strictly monotonically increasing because F_μ and F_{X_T} are strictly monotonically increasing, since μ has a strictly positive density by assumption and X_T has a strictly positive density when $|\rho| < 1$ because $X_T|V_{0\le t\le T}$ is conditionally Gaussian.

5 The SABR model

We now consider a classical SABR model with $\beta = 1$:

$$dS_t \;=\; S_t Y_t dW_t \quad , \quad dY_t \;=\; \nu Y_t dB_t$$

under a risk-neutral measure $\mathbb{Q}$, where W, B are two standard Brownians with $dW_t dB_t = \rho dt$. Then we know that the mean-variance hedge for a call option is given by

$$\theta_t \;=\; \frac{d\langle C, S\rangle_t}{dS_t^2} \;=\; \frac{S_t^2 Y_t^2 C_S(S_t, Y_t, t) + \rho S_t Y_t \nu Y_t C_y(S_t, Y_t, t)}{S_t^2 Y_t^2} \;=\; C_S(S_t, Y_t, t) + \frac{\rho\nu C_y(S_t, Y_t, t)}{S_t}$$

where $C(S, y, t) := \mathbb{E}^{\mathbb{Q}}((S_T - K)^+ | S_t = S, Y_t = y)$. We can avoid directly having to estimate $C_S(S_t, Y_t, t)$ with Monte Carlo and numerical finite differences (which will lead to a noisy estimate in practice) by instead appealing to the spatial homogeneity property of the model:

$$C(\lambda S, \lambda K) \;=\; \lambda C(S, K)$$

where here $C(S, K)$ denotes the price of a call option as a function of the initial stock price S and strike K with all other parameters fixed. If we differentiate this expression with respect to λ and set $\lambda = 0$, we get

$$SC_S + KC_K \;=\; C$$

and $C_K = -\mathbb{E}^{\mathbb{Q}}(1_{S_T > K})$ is minus the price of a digital option, so $C_S = (C - KC_K)/S$. We also note that

$$\frac{\partial S_t(\omega)}{\partial Y_0} \;=\; S_t \int_0^t e^{\nu B_s - \frac{1}{2}\nu^2 s}(\rho dB_s + \bar{\rho} dW_s)$$

so $\frac{\partial}{\partial Y_0}(S_t - K)^+ = 1_{S_t > K} \frac{\partial S_t}{\partial Y_0}$ (since S_t admits a density because S_t is conditionally log-normal if we condition on $(Y_s)_{0 \le s \le t}$) so $\mathbb{P}(S_t = K) = 0$, and we can (formally) use that

$$\frac{\partial}{\partial Y_0}\mathbb{E}((S_T - K)^+) \;=\; \mathbb{E}(\frac{\partial}{\partial Y_0}(S_T - K)^+) \;=\; \mathbb{E}(1_{S_t > K}\frac{\partial S_t}{\partial Y_0})$$

to compute the left hand side by computing the right hand side.

References

[AMP23] Acciaio, B., A.Marini, G.Pammer, "Calibration of the Bass Local Volatility model", preprint

[AL24] Abi-Jaber, E. and X.Li, "Volatility models in practice: Rough, Path-dependent or Markovian?", preprint, 2024.

[BBHK20] Backhoff-Veraguas, J., M.Beiglböck, M.Huesmann and S.Källblad, "Martingale Benamou–Brenier: A Probabilistic Perspective", *The Annals of Probability*, Vol. 48, No. 5, 2258–2289, 2020

[BBST23] Backhoff-Veraguas, J., M.Beiglböck, W.Schachermayer, and B.Tschiderer, "The structure of martingale Benamou–Brenier", preprint, 2023.

[BST23] Backhoff-Veraguas, J., W.Schachermayer, and B.Tschiderer, "The Bass functional of martingale transport", preprint, 2023

[Bass83] Bass, R.F., "Skorokhod imbedding via stochastic integrals", in J. Azéma and M. Yor, editors, Semin. Probab. XVII 1981/82 — Proceedings, volume 986 of Lecture Notes in Math., pages 221–224. Springer, Berlin, Heidelberg, 1983

[BFG16] Bayer, C., P.Friz and J.Gatheral, "Pricing Under Rough Volatility", *Quantitative Finance*, 16(6), 887-904, 2016

[BG24] Bourgey, F. and J.Guyon, "Fast Exact Joint S&P 500/VIX Smile Calibration in Discrete and Continuous Time", *Risk*, Feb 2024

[Carr09] Carr, P., "Local Variance Gamma Option Pricing Model", presentation Bloomberg/Courant Institute, April 28, 2009.

[CN17] Carr, P. and S.Nadtochiy, "Local Variance Gamma and Explicit Calibration to Option Prices", *Mathematical Finance*, 27(1):151–193, 2017

[CL21] Conze, A. and P.Henry-Labordère, "Bass Construction with Multi-Marginals: Lightspeed Computation in a New Local Volatility Model", preprint, 2021.

[CT04] Cont, R. and P. Tankov, "Financial modelling with Jump Processes", Chapman&Hall, 2004

[ER18] El Euch, O. and M.Rosenbaum, "Perfect hedging in rough Heston models", *The Annals of Applied Probability*, 28(6), 2018, pp. 3813-3856

[F23] Forde, M. (2023), "Statistical issues and calibration problems under rough and Markov volatility", presentation, Feb 2023, King's College London, `https://nms.kcl.ac.uk/martin.forde/Talk.pdf`

[F23II] Forde, M. (2023), "The QGARCH(1,1) model", `https://nms.kcl.ac.uk/martin.forde/QGARCH11-Notes.pdf`

[FG22] Fukasawa, M. and J.Gatheral, "A rough SABR formula", *Frontiers of Mathematical Finance*, 1(1), 81-97, 2022.

[Gass19] Gassiat, P., "On the martingale property in the rough Bergomi model", *Electron. Commun. Probab.*, 24: 1-9, 2019

[GR20] Gatheral, J. and M.Rosenbaum, "The quadratic rough Heston model and the joint S&P 500/VIX smile calibration problem", `risk.net`, 2020.

[GLOW22] Guo, I., G.Loeper, J.Obłój, S.Wang, "Optimal transport for model calibration", `risk.net`, Jan 2022

[GLW22] Guo, I., G.Loeper, S.Wang (2022), Calibration of Local-Stochastic Volatility Models by Optimal Transport, *Mathematical Finance*, 32(1), 46-77.

[GJR18] Gatheral, J., T.Jaisson and M.Rosenbaum, "Volatility is rough", *Quantitative Finance*, 18(6), 2018

[Guy21] Guyon, J., "The Joint S&P 500/VIX Smile Calibration Puzzle Solved", Seminar, `https://www.youtube.com/watch?v=pvq-rfajFRs&t=1265s`, 2021

[JMP21] Jacquier, A., A.Muguruza and A.Pannier, "Rough multifactor volatility for SPX and VIX options", preprint 2021

[JMP18] Jacquier, A., C.Martini and A.Muguruza, "On VIX futures in the rough Bergomi model", *Quantitative Finance*, 18(1), 2018, 45-61

[KR22] Keller-Ressel, M., "Bartlett's Delta revisited: Variance-optimal hedging in the lognormal SABR and in the rough Bergomi model", preprint, 2022

[NPP14] Nugroho, D.B., B.A.Pamungkas, and H.A.Parhusip, "Volatility Fitting Performance of QGARCH (1,1) Model with Student-t, GED, and SGED Distributions", *ComTech: Computer, Mathematics and Engineering Applications*, 11(2), December 2020, 97-104

[Nua06] Nualart, D., "The Malliavin Calculus and Related Topics", Springer, 2006

[RT96] E.Renault and N.Touzi (1996), Option hedging and implied volatilities in a stochastic volatility model, *Mathematical Finance*, 6(3), 279–302.

[Rom22b] S.Rømer (2022), Empirical analysis of rough and classical stochastic volatility models to the SPX and VIX markets, *Quantitative Finance*, 22(10), 1805-1838.

[Schw95] Schweizer, M., "Variance-optimal hedging in discrete time", *Mathematics of Operations Research* 20(1),1-32.

[VZ19] Viens, F. and J.Zhang, "A Martingale Approach for Fractional Brownian Motions and Related Path Dependent PDEs", *Ann. Appl. Probab.*, 29(6): 3489-3540, 2019