Statistical issues and calibration problems under rough and Markov volatility

(minor updates 29/03). Will briefly discuss

- **Statistical issues**: Problems with fitting rough models to high frequency historical time series
- **Option pricing**: calibrating Markov stochastic volatility models to European and/or VIX options using **HJB equations** and **duality** arguments

We consider a general stochastic volatility model of the form:

$$\begin{aligned} dS_t &= \mu S_t dt + S_t\sqrt{V_t}dB_t \\ V_t &= \xi e^{\nu Y_t} \end{aligned} \tag{1}$$

under the $\mathbb{P}$-measure, where Y is a **Gaussian process** of the form $Y_t = \int_{-\infty}^t k(t-s)dW_s$ (i.e. a general **stationary** Gaussian process (Y_t has same (centred normal) distribution for all t) or $Y_t = \int_0^t k(s,t)dW_s$ (e.g. fBM) and $dW_t dB_t = \rho dt$.

The covariance of Y is $R(s,t) = \mathbb{E}(Y_s Y_t) = \int_{...}^t k(u,s)k(u,t)du$, we can define the **covariance matrix** $\Sigma_{i,j} = R(i\Delta t, j\Delta t)$ of $Y_{t_1},...,Y_{t_N}$ with $t_i = \frac{iT}{N}$ for $1 \le i,j \le N$. We can **sample** Y at these time points as CZ, where $CC^\top = \Sigma$ and C is the (lower triangular) **Cholesky decomposition** of Σ, and Z is a column vector of N i.i.d standard Normals. Hence, from an **observed** V **path** at these time points and a given **parameter set** (H,ν,ξ) (or just (ν,ξ) for non-stationary case since in this case $Y_0 = 0$ so $\xi = V_0$), we can then extract the Z-vector implied by the Y path as

$$Z_i = C^{-1}\Big(\frac{1}{\nu}\log\frac{V_{t_i}}{\xi}\Big).$$

(Note for the non-stationary case, C^{-1} approximates the theoretical inversion formula which takes the form $W_t = \int_0^t \bar{k}(s,t)dY_s$). Note V is not observable in real life (see e.g. (2) below for further discussion on this). We can then perform standard statistical tests to see if this is indeed a sequence of i.i.d. standard Gaussians, e.g. the **Kolmogorov-Smirnov**(KS) which uses that $\sqrt{N}(\bar{Y}_N(F^{-1}(u))-u)$ has the same covariance $R^{br}(u,v) := \min(u,v) - uv$ as a **Brownian bridge** B^{br} on $[0,1]$ for all N, and tends weakly to B^{br} as $N \to \infty$, where $\bar{Y}_N(x) = \frac{1}{N}\sum_{i=1}^N 1_{Z_i \le x}$ is the **empirical cdf** of Z, and F is the cdf we are testing (standard Normal cdf in our case), and the cdf of the maximum of $|B^{br}|$ has an explicit (series) expansion (this is the **KS statistic**, which we call $\widehat{KS}$). In practice we make a finite sample adjustment to compute $\widehat{KS}$. The other test we commonly use is the **Shapiro-Wilks**(SW) test.

We can then find the parameter combination (H,ν,ξ) which maximizes the p**-value** $p := \mathbb{P}(KS \ge \widehat{KS}|H_0)$ (or to be more robust, we take the minimum of the p-values for the KS and SW tests), using a numerical **minimization scheme** (note I am not using the stock price path itself here or trying to estimate ρ or μ which is v difficult in practice, even for arithmetic BM). If the p-value is e.g. $\le \alpha$ (e.g. $\alpha = .01$ or $.05$), then we can reject the **null hypothesis** that the model is correct at the α**-significance level** in the usual way.

We have tested this approach with Y equal to fBM, and Y equal to the Bacry-Muzy-Wu [BMW22] **stationary fBM** with covariance function $R(s,t) = R(\tau) = \frac{1}{2H}(T^{2H} - \tau^{2H})^+$ (where $\tau = |t-s|$) (both of which are $H-\varepsilon$ **Hölder continuous** a.s., the latter can be obtained as a certain $\lambda \to 0$ limit of the OU process driven by fBM $\int_{-\infty}^t e^{-\lambda(t-s)}dW_s^H$ discussed in Pakkanen et al. [BCPV22]). For the latter, by expanding $R(\tau)$ around $H=0$ one can easily show that

$$\mathbb{E}\Big(\Big(\int_0^\tau V_s ds\Big)^2\Big) = const. \times \tau^{\zeta(2)} + O(H)$$

with $\zeta(2) < 2$. For the power $q=2$, this gives the correction to the Bacry-Muzy **multiscaling** behaviour. Y becomes **1/f noise** (a log correlated **Gaussian field**) in limit as $H \to 0$.

When applied to **SPX** (S&P 500 index) daily **realized variance** (estimated using **sums of squares** of **log returns** with **1min** time intervals during the 6.5 hour trading day, i.e. 390 intervals plus overnight period) this decade, the p-value is typically maximized with $H \in (0, .01]$, and good p-values can be achieved over time horizons larger than 2 years. The problem is that the **maximum likelihood estimate** (MLE) for H for the same data is typically between .15 to .30 (and so are estimates obtained using Generalized Method of Moments (**GMM**), e.g by looking at simple unbiased estimates for $\mathbb{E}(\sum_{i=1}^N |Y_{\Delta i}|^q)$ or $\mathbb{E}(\sum_{i=1}^N |Y_{i\Delta} - Y_{(i-1)\Delta}|^q))$ for $q > 0$, and drawing **log-log plots** i.e. log of this estimate vs log Δ to estimate H from the slope, see below (estimate for H will be **biased** in general), and if we generate **synthetic paths** from the fitted $H = .01$

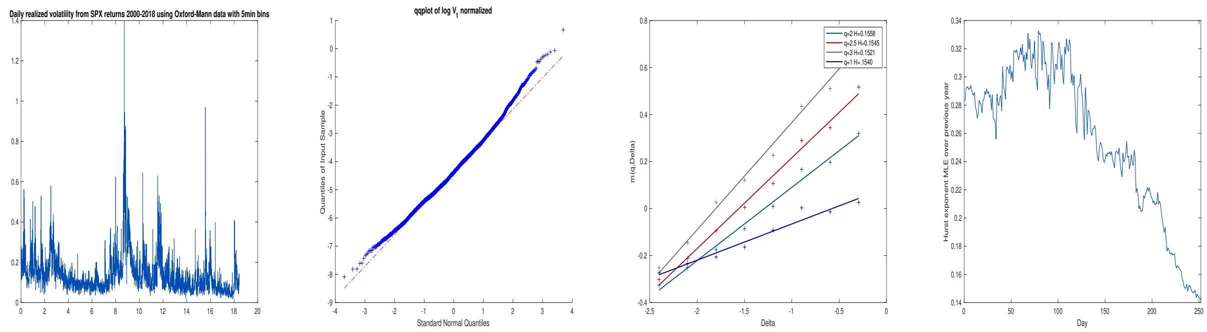

Figure 1: First three plots use Oxford Mann data from 2000-18. GMM plot in third plot, slopes of all four plots suggest H value of 0.15. Final plot is MLE of H with rolling 252 day window from Jan 2021-2023

model, we never see MLE or GMM estimates this large. Conversely, if we constrain H to be ≥ 0.15, the p-values are too low. So for these two models, the MLE is a sufficient statistic to be able to **reject** the null hypothesis that daily *realized variance* follows a model of the form above over the time horizons considered, even if the MLE itself may not be a good estimator for a moderate sample size. Realized and true integrated variance have the following relation

$$RV_t \quad \overset{\text{(d)}}{\approx} \quad \int_{t-1}^{t} V_s ds \,+\, (\frac{2}{m}\int_{t-1}^{t} V_s^2 ds)^{\frac{1}{2}} Z_t \tag{2}$$

as the **intraday sampling frequency** $m \to \infty$ (see e.g. Eq 25 in [BCPV22]) where the Z_t's are i.i.d Normals which are indept of W and B, which is a Central Limit Theorem result. We found that if we simulate the model above **intraday** with m time steps per day, and generate a synthetic sample path of daily realized variance, and then optimize the p-value for this, we typically still get H close to .01 but the estimated value of ν is way below the true value.

The p-values for a model like this where $\log V_t$ is Gaussian are typically a lot better than other well known rough models (**rough Heston** and **quadratic rough Heston**); these models are described by **Stochastic Volterra Equations** (SVEs) e.g. $V_t = V_0 + \int_0^t (t-s)^{H-\frac{1}{2}}\sqrt{V_s}dW_s$ which need to be approximated with **Euler-type schemes** to use this method, but this introduces additional non-negligible errors unless one is happy to just work with this discrete-time version of the process (otherwise would require a higher order **Milstein scheme** for SVEs, probably only feasible for a linear SVE of the form $V_t := V_0 + \int_0^t k(t-s)V_s dW_s$. If V is a standard 1d diffusion, **pointwise error** for Euler scheme is $O(\sqrt{N})$ but Milstein is $O(N)$, so the error in estimating the Gaussian deviates Z_i is $O(1)$ for an Euler scheme and $O(\frac{1}{\sqrt{N}})$ for Milstein; we can easily test this if e.g. V is GBM)). If $\rho = \pm 1$ (i.e. a **pure feedback** model like the quadratic rough Heston model or the Guyon **two-factor Markovian PDV model** we do not need to estimate V at all since it can be extracted from an observed stock price path since in this case $\sqrt{V_t}dW_t = dS_t/S_t$, but we are sceptical that this is the case since estimates of ρ tend to fluctuate wildly over different time windows (and sample skewness of log returns). The p-value method applied to the **diagonal QARCH model** or the Guyon-Lekeufack **discrete-time empirical PDV model** (both with fractional kernel $const. \times (t+\delta)^{-\alpha}$, both models discussed in [GL22]), using actual stock returns to compute V and then using S and V extract the Brownian increments implied by the stock history, leads to V paths which are far too large to be consistent with actual realized variance. The p-value method applied to the continuous-time **Markovian two-factor PDV model** in [GL22], leads to V paths which are way too smooth to the naked eye, because the volatility is overly dominated by the unusual drift term. One can also compute MLE estimates for the Euler discretization of a rough model by inverting the map from the discrete (estimated) V path (or observed S path in the $|\rho| = 1$ case) and expressing the joint density of V (resp. S) in terms of a the joint density of N i.i.d. Normals and the **Jacobian** of the map which can be computed using finite differences.

A **qq-plot** of (normalized) log-realized variance below of SPX suggests the **right tail** of $\log V$ is fatter than Gaussian (note these observations will not be i.i.d), so can consider e.g. letting $V_t = \xi e^{\nu Y_t + \beta Y_t^3}$ or replacing the driving Brownian motion W with a a Lévy process, but we still run into problems.

Calibrating Markov stochastic volatility models using HJB eqs

We now consider a Markov stochastic volatility model of the form

$$\begin{aligned} dS_t &= S_t\sqrt{V_t}dB_t \\ dV_t &= \kappa(\theta - V_t)dt + \beta(S_t, V_t, t)dW_t \end{aligned}$$

with $dW_t dB_t = \rho dt$. The **VIX index** at time t is (theoretically) defined as $\text{VIX}_t^2 := \frac{1}{\Delta}\mathbb{E}(\int_t^{t+\Delta} V_s ds|\mathcal{F}_t)$ which in practice is inferred from option prices, where $\Delta = 1/12$ i.e. 1 month. For a V process with a drift of this form, we can easily show that $\text{VIX}_t^2 = aV_t + b$, for two constants a and b that only depend on κ, θ and Δ, using the expression for $\mathbb{E}(V_t|V_s)$ for $s \le t$. This also means we can observe V directly if the model is correct and we know κ and θ.

We wish to find a $\beta(S, V, t)$ so this model calibrates to n **European call options** and m **VIX options** at a fixed maturity T with market prices c_i and c_j^v respectively. $u(S, V, t) = \sum_{i=1}^n w_i \mathbb{E}((S_T - K_i)^+|S_t = S) + \sum_{j=1}^m w_i^v \mathbb{E}((\text{VIX}_T - K_j^v)^+|S_t = S, V_t = v) + \alpha\mathbb{E}(\int_0^T (\beta(S_t, V_t, t) - \nu V_t^p)^2 dt|S_t = S, V_t = v)$ (for $p \in (0, 1]$) satisfies the **backward Kolmogorov eq**:

$$\mathcal{L}u \;:=\; u_t + \frac{1}{2}S^2 V u_{SS} + \kappa(\theta - v)u_v + \frac{1}{2}\beta^2 u_{vv} + \rho S\sqrt{v}\beta u_{Sv} + \alpha(\beta - \nu v^p)^2) \;=\; 0$$

with $u(x, T) = \sum_{i=1}^n w_i(x - K_i)^+ + \sum_{j=1}^m w_i(\sqrt{aV + b} - K_j^v)^+$. We can consider a more general model where

$$dV_t \;=\; \kappa(\theta - V_t)dt + \beta_t dW_t$$

where β_t is any $\mathcal{F}_t^W$-adapted process, we can ask how do we optimally choose β_t so as to minimize $\sum_{i=1}^n w_i \mathbb{E}((X_T - K_i)^+) + ... + \alpha\mathbb{E}(\int_0^T (\beta_t - \nu V_t^p)^2 dt)$.

This penalizes the "distance" from a **reference stoc vol model**, which is the **Hull-White model** when $\boldsymbol{p = 1}$ and **Heston** when $\boldsymbol{p = \frac{1}{2}}$. From standard **stochastic control theory**, formally at least we know the solution satisfies the **HJB eq**: $\min_\beta \mathcal{L}u = 0$. Then (assuming we can **interchange** the sup and inf)

$$\begin{aligned} & \sup_w \inf_{\beta \in \mathcal{A}} (-\sum_i \tilde{w}_i \tilde{c}_i + \sum_i \tilde{w}_i \mathbb{E}^\beta(f_i(X_T, V_T)) + \alpha\mathbb{E}^\beta(\int_0^T (\beta_t - \nu V_t^p)^2 dt) \\ = \; & \inf_{\beta \in \mathcal{A}} \sup_{w_i}(...) \\ = \; & \inf_{\beta \in \mathcal{A}\,:\,\mathbb{E}^\beta((X_T - K_i)^+ = c_i, \mathbb{E}^\beta((\sqrt{aV_T + b} - K_i)^+ = c_i^v} \alpha\mathbb{E}^\beta(\int_0^T (\beta_t - \nu V_t^p)^2 dt) \end{aligned} \tag{3}$$

(where we have aggregated the payoffs, weights and market prices of the European and VIX payoffs into single vectors f, $\tilde{w}$ and $\tilde{c}$ for ease of notation), where third line line follows because the **inner sup** in the middle line is $+\infty$ if for a particular $a \in \mathcal{A}$ the options are not correctly calibrated, since e.g. we can choose w_i to be arbitrarily large if $\sum_{i=1}^n w_i \mathbb{E}^\beta((X_T - K_i)^+) - \sum_{i=1}^n w_i c_i > 0$, and vice versa.

The inf of an empty set is infinity by convention, so if no admissible model exists, the final line in (3) is **infinity**, and from the general basic result that $\boldsymbol{\inf_Y \sup_X f(x, y) \ge \sup_X \inf_Y f(x, y)}$, even if duality doesn't hold, if the first line is infinity (which is what we actually compute by solving the HJB eq), then the final line is $+\infty$, so **no admissible model exists** (we cannot guarantee such a model exists unless e.g. we are only fitting European options and can use e.g. **Dupire model** or **Bass martingale**). Main ideas on this are due to **Guo,Loeper,Wang** et al. (see e.g. [GLOW22]), similar ideas for a specific case considered in separate articles by **Henry-Labordére** and Guyon using an **entropy** penalty function which only allows **Girsanov perturbations** from the reference model. All these articles do not appear to have noticed/used the simple relation $\text{VIX}_t^2 = aV_t + b$ to simplify this problem.

The final line in (3) is the **model-independent lower bound** for a contract which pays $\alpha\int_0^T (\beta_t - \nu V_t^p)^2 dt$ at time T, subject to matching the market prices of the n given European options and m VIX options. If duality holds i.e. $\inf\sup = \sup\inf$, this is also the **maximum subhedging** cost of this contract, using just cash, dynamic trading in X and a **static position** in the stock and VIX options. Note when we minimize here we are also including **rough models** since we do not assume a priori that β_t is Markov, but the optimal model is Markovian by the usual "rules" of how the HJB eq works, because the reference model is Markovian. One can generalize this methodology to use a **rough reference model**, but one ends up with an intractable non-standard **FBSDE**.

If we have an additional SPX options to fit at a **later maturity** $T_2 > T$, then we need a **nested PDE scheme** to solve this problem, i.e. we solve the PDE on $[T, T_2]$ with the boundary condition for the weights for the additional European options expiring at T_2. This then gives us a (non-trivial/non-explicit) boundary condition at time T and we then add-on the usual boundary condition for the SPX and VIX options expiring at the earlier maturity T.

References

[BCPV22] Bolko, A. E., K.Christensen, M.S.Pakkanen and B.Veliyev, "A GMM approach to estimate the roughness of stochastic volatility", *Journal of Econometrics*, to appear.

[GL22] Guyon, J. and J.Lekeufack, "Volatility is (mostly) path-dependent", preprint, 2022.

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

[BMW22] Wu, P., J-F.Muzy and E.Bacry, "From Rough to Multifractal volatility: the log S-fBM model", 2022, *Physica A: Statistical Mechanics and its Applications*, Volume 604, Oct 2022