Weierstrass Institute for
Applied Analysis and Stochastics

Leibniz
Association

Rough volatility

Christian Bayer

23rd Winter school on Mathematical Finance, Soesterberg

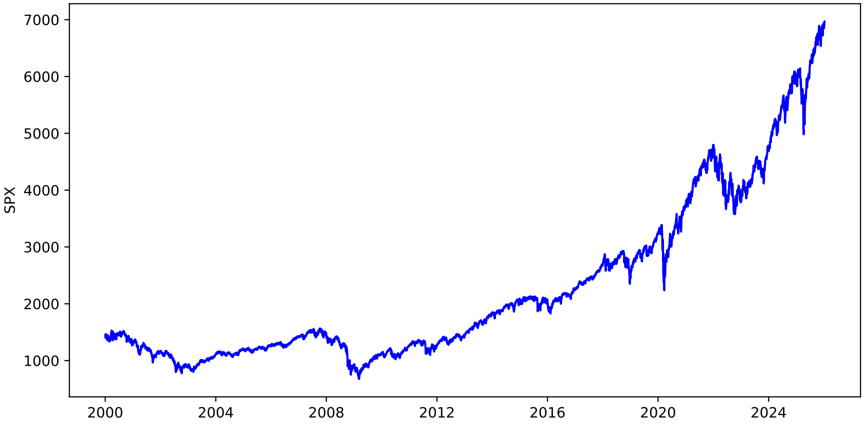


Figure: The S & P 500 index as an example of an equity market

Basic properties of a mathematical model for equity markets:

- Stochastic process $S = (S_t)_{t \geq 0}$ (asset price)
- S is a positive semimartingale (no arbitrage).

First guess: geometric Brownian motion (gBm).

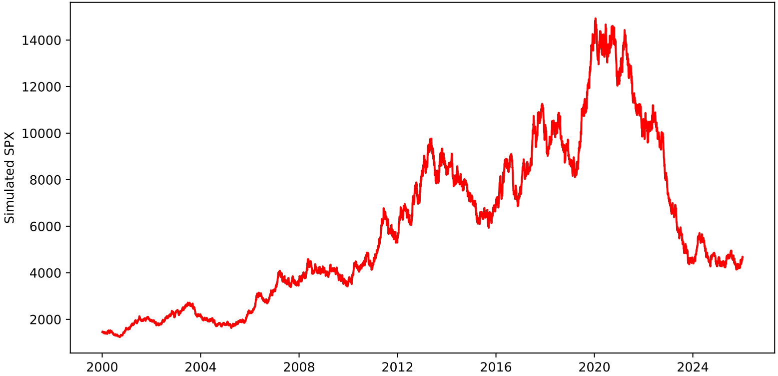


Figure: gBm with same initial value, drift and volatility (i.e., mean and standard devation of log-returns coincide with the sample versions of SPX over the plotted time period).

Basic properties of a mathematical model for equity markets:

- Stochastic process $S = (S_t)_{t \geq 0}$ (asset price)
- S is a positive semimartingale (no arbitrage).

First guess: geometric Brownian motion (gBm).

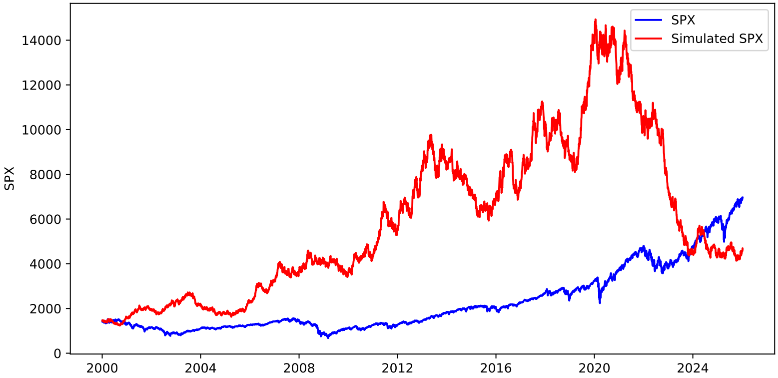


Figure: gBm with same initial value, drift and volatility (i.e., mean and standard devation of log-returns coincide with the sample versions of SPX over the plotted time period).

Is this a good model? Hard to tell from the price paths . . .

- Log-returns are not normally distributed (fat tails, skewness).
- Volatility is not constant (volatility clustering, leverage effect).

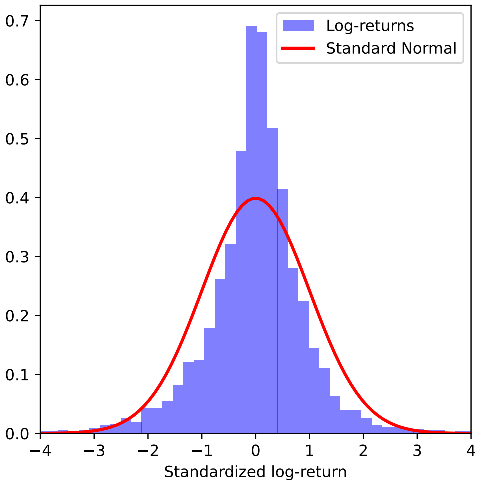


Figure: Histogram of standardized daily log-returns of SPX (2000–2026) compared to standard normal density.

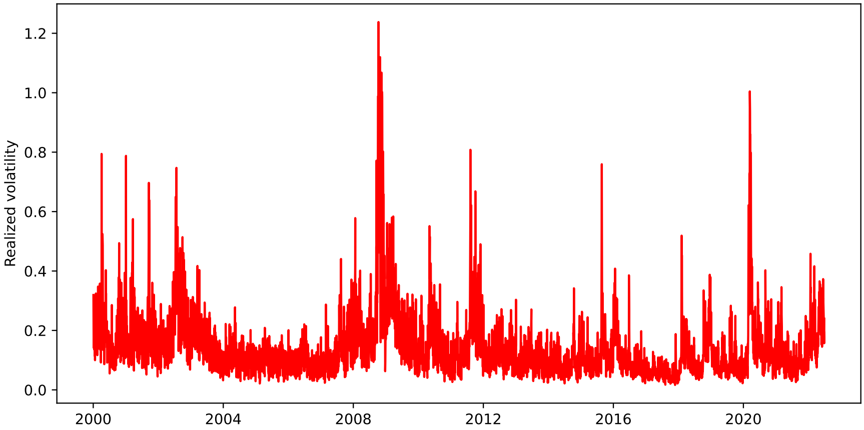


Figure: Annualized daily realized volatility of SPX (2000–2022, by Oxford Man institute).

Volatility is spiky

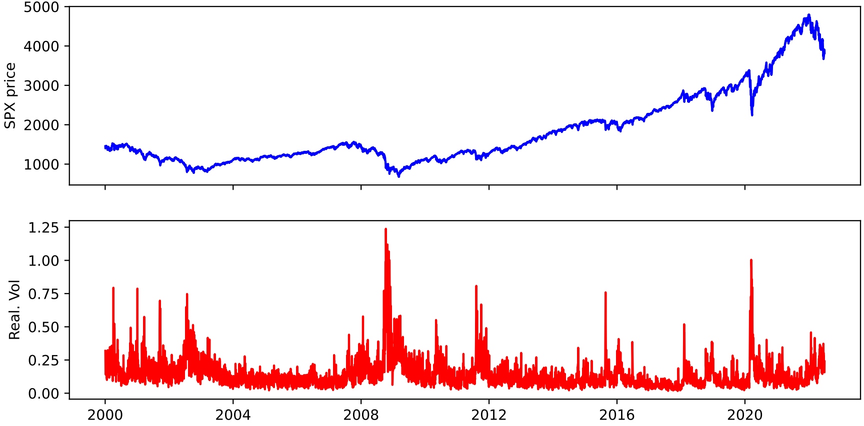


Figure: SPX and its realized volatility (2000–2022, by Oxford Man institute).

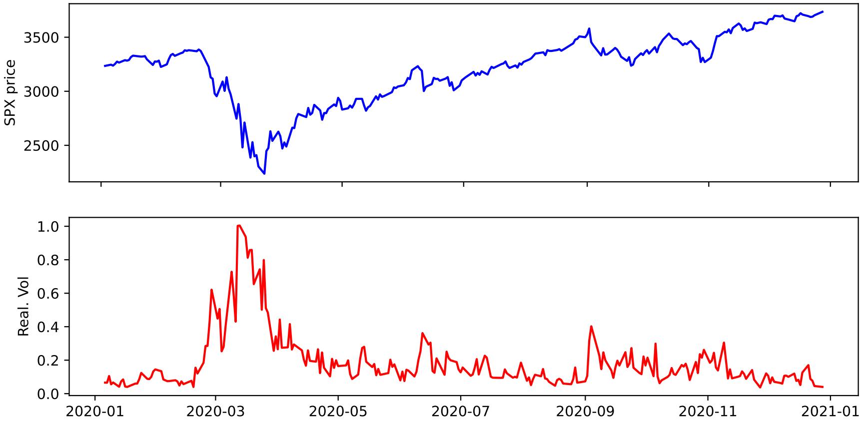


Figure: SPX and its realized volatility in 2020.

- Standard deviation of log-returns (volatility) is not constant.
- Volatility tends to increase when the price drops (leverage effect) – looks stochastic (negative correlation).
- gBm is not a good model for equity markets.

Caveats:

- Realized volatility is only an estimate of the "true" volatility, which cannot be observed.
- Market frictions (transaction costs, liquidity issues, etc.) are ignored.

Suggests stochastic volatility models:

$$dS_t = \mu S_t dt + \sqrt{v_t} S_t dZ_t, \quad dv_t = \ldots \text{ (stochastic)}.$$

Classical stochastic volatility models (Heston, SABR, ...):

$$\mathrm{d}v_t = a(v_t)\mathrm{d}t + b(v_t)\mathrm{d}W_t, \quad \mathrm{d}\langle Z, W\rangle_t = \rho \mathrm{d}t,\ \rho \in [-1, 0]$$

Rough volatility models: $v_t \sim W_t^H$, $H \in (0, 1/2)$, fractional Brownian motion (fBm).

Rough Bergomi model

$$v_t = \xi_0(t) \exp\left(\eta \sqrt{2H} \int_0^t (t-s)^{H-1/2} \mathrm{d}W_s - \frac{\eta^2}{2} t^{2H}\right)$$

Rough Heston model

$$v_t = v_0 + \frac{1}{\Gamma(H+1/2)} \int_0^t (t-s)^{H-1/2} \kappa(\theta - v_s)\mathrm{d}s + \frac{1}{\Gamma(H+1/2)} \int_0^t (t-s)^{H-1/2} \sigma \sqrt{v_s} \mathrm{d}W_s$$

Traditionally, a somewhat controversial topic!

- First suggested as model for asset prices by (Mandelbrot and Taylor '67)
- Based on statistical properties of asset returns – long-range dependence, self-similarity, further advocated in the late 90s/early 2000s (Comte and Renault '96).
- (Rogers '97) and (Cheridito '02) construct explicit arbitrage strategies – the fundamental theorem of asset pricing fails.

However,

- Arbitrage opportunities in fBm price models often rely on unrealistic trading strategies (e.g., continuous trading, perfect foresight), and disappear under (even asymptotic) transaction costs (Guasoni '06).
- Rough volatility models use fBm to model volatility, not prices directly – not traded!
- The price process remains a semimartingale, so no arbitrage opportunities arise from the use of fBm in this context.

- Empirical studies show that realized volatility behaves like a fractional Brownian motion with Hurst parameter $H \approx 0.1$ (Gatheral, Jaisson, Rosenbaum 2018).
- Rough volatility models are able to reproduce the steepness of the implied volatility surface for short maturities (Bayer, Friz, Gatheral 2016).
- Rough volatility models are consistent with microstructural properties of the market (Jaisson, Rosenbaum 2015).

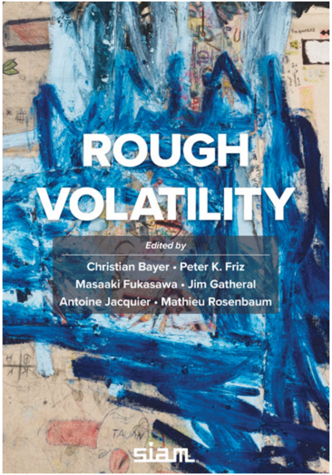
Challenges:

- Non-Markovian and non-semimartingale structure.
- Numerical methods for simulation and option pricing.

Rough volatility models are a fascinating intersection of advanced mathematics and practical finance.

- The process (S, v) is not a semi-martingale, so classical stochastic analysis does not apply (analysis of Volterra processes, Malliavin calculus, fractional calculus, rough paths, regularity structures, . . .).
- It is not a Markov process, either (path-dependent PDEs/BSPDEs, infinite-dimensional Markovian representations, path signature methods, deep learning, . . .).

1. Volatility is rough: empirical evidence
2. The rough Bergomi model
3. The rough Heston model
4. Multi-factor approximations
5. Diamond expansions



- Realized volatility time series exhibit scaling properties consistent with $H < 1/2$.
- Log-realized variance is approximately Gaussian.
- Smoothness of volatility paths is inconsistent with standard Brownian motion ($H = 1/2$).

We follow the excelent python code provided by Florian Bourgey:

https://github.com/fbourgey/RoughVolatilityWorkshop,

an adaptation of Jim Gatheral's original R code:

https://github.com/jgatheral/RoughVolatilityWorkshop.

We switch to a Jupyter notebook for the live demonstration.

All models under the risk-neutral measure $\mathbb{Q}$ with $r = 0$ (i.e., we consider forward prices):

$$\mathrm{d}S_t = S_t \sqrt{v_t}\mathrm{d}Z_t$$

Rough Bergomi model (B., Friz, Gatheral 2016):

$$v_t = \xi_0(t) \exp\left(\eta \widetilde{W}_t - \frac{\eta^2}{2} t^{2H}\right), \quad \widetilde{W}_t = \sqrt{2H} \int_0^t (t-s)^{H-1/2} \mathrm{d}W_s, \quad \mathrm{d}\langle Z, W\rangle_t = \rho \mathrm{d}t$$

Rough Heston model (Jaisson, Rosenbaum 2015):

$$v_t = v_0 + \int_0^t K(t-s)\kappa(\theta - v_s)\mathrm{d}s + \int_0^t K(t-s)\sigma\sqrt{v_s}\mathrm{d}W_s, \quad K(r) := \frac{r^{H-1/2}}{\Gamma(H+1/2)}, \quad \mathrm{d}\langle Z, W\rangle_t = \rho \mathrm{d}t$$

$$dS_t = S_t\sqrt{v_t}dZ_t, \quad d\langle Z, W\rangle_t = \rho dt$$
$$v_t = \xi_0(t)\exp\left(\eta\widetilde{W}_t - \frac{\eta^2}{2}t^{2H}\right), \quad \widetilde{W}_t = \sqrt{2H}\int_0^t (t-s)^{H-1/2}dW_s$$

- Initial forward variance curve $\xi_0(t) := \mathbb{E}_{\mathbb{Q}}[v_t \mid \mathcal{F}_0]$, $t > 0$, can be inferred from market data (variance swaps, log-strip formula).
- Parameters: volatility of volatility $\eta > 0$, Hurst parameter $H \in (0, 1/2)$, correlation $\rho \in [-1, 0]$.
- $\widetilde{W}_t$ is the Riemann-Liouville fractional Brownian motion.
- For $\rho \leq 0$, S is a $\mathbb{Q}$-martingale (Gassiat 2019), otherwise a strict local martingale.

Standard fBm Gaussian process W^H with $\mathbb{E}[W_t^H] \equiv 0$ and covariance

$$\mathbb{E}\left[W_s^H W_t^H\right] = \frac{1}{2}\left(|t|^{2H} + |s|^{2H} - |t-s|^{2H}\right), \quad 0 < H < 1$$

1. Trajectories are $H - \epsilon$ Hölder for any $\epsilon > 0$
2. Scaling property (self-similarity): $W_{at}^H = |a|^H \, W_t^H$
3. Increments are positively correlated for $H > 1/2$ and negatively correlated for $H < 1/2$
4. Increments are stationary.
5. Mandelbrot – Van Ness representation:

$$W_t^H = C_H \left(\int_{-\infty}^{0} \left((t-s)^{H-1/2} - (-s)^{H-1/2} \right) \mathrm{d}W_s + \int_0^t (t-s)^{H-1/2} \mathrm{d}W_s \right)$$

Exact simulation of Gaussian processes: for a grid $t_1 < \cdots < t_N$, let Σ be the corresponding covariance matrix, $\sigma_{i,j} = \mathbb{E}\left[W^H_{t_i} W^H_{t_j}\right]$. Then,

$$\left(W^H_{t_1}, \dots, W^H_{t_N}\right)^\top \overset{\text{Law}}{=} AZ, \quad AA^\top = \Sigma, \quad Z \sim \mathcal{N}(0, I_N)$$

- Compute A by Cholesky decomposition (nice for interpretability/coupling, but slow, $\mathcal{O}(N^3)$.)
- Computational cost of simulation: $\mathcal{O}(N^2)$ per sample.

If the grid is equidistant and the increments are stationary (e.g., for fractional Brownian motion), then one can use the circulant embedding method, with computational cost $\mathcal{O}(N \log N)$ per sample.

$X = (X_1, \dots, X_N)$ stationary Gaussian vector (e.g., increments of fBm on a uniform grid) with $\mathbb{E}[X_n] = 0$, $r_j := \mathbb{E}[X_n X_{n+j}]$.
Covariance matrix $\Sigma \in \mathbb{R}^{N\times N}$ can be embedded into a symmetric, circulant matrix $C \in \mathbb{R}^{2N\times 2N}$:

$$\Sigma = \begin{pmatrix} r_0 & r_1 & \cdots & r_{N-1} \\ r_1 & r_0 & \cdots & r_{N-2} \\ \vdots & \vdots & \ddots & \vdots \\ r_{N-1} & r_{N-2} & \cdots & r_0 \end{pmatrix}, \quad C = \begin{pmatrix} r_0 & r_1 & \cdots & r_{N-1} & 0 & r_{N-1} & r_{N-2} & \cdots & r_1 \\ r_1 & r_0 & \cdots & r_{N-2} & r_{N-1} & 0 & \cdots & r_2 & \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots & \\ r_1 & r_0 & \cdots & 0 & r_{N-1} & r_{N-2} & \cdots & r_0 & \end{pmatrix}$$

- $C = F_{2N}^* \Lambda F_{2N}$, where F_{2N} is the DFT matrix and $\Lambda = \operatorname{diag}(\sqrt{2N} F_{2N} c)$, $c := (r_0, r_1, \dots r_1)^\top$. Assume that C is positive-definite.
- Compute $\lambda = F_{2N} c$. If $\lambda \geq 0$, simulate $Z \sim \mathcal{N}(0, I_{2N})$, set $Y = F_{2N}^* \sqrt{\Lambda} F_{2N} Z$. Then $\Re(Y)_{1:N}$ has covariance Σ.

$$W_t^H = C_H \left(\int_{-\infty}^{0} \left((t-s)^{H-1/2} - (-s)^{H-1/2} \right) \mathrm{d}W_s + \int_0^t (t-s)^{H-1/2} \mathrm{d}W_s \right)$$

Annoying from a modelling perspective: need full history of W!

$$\widetilde{W}_t := C_H \int_0^t (t-s)^{H-1/2} \mathrm{d}W_s$$

Gaussian with mean $\mathbb{E}\left[\widetilde{W}_t\right] = 0$ and covariance

$$\mathbb{E}\left[\widetilde{W}_t \widetilde{W}_s\right] = 2Hs^{2H} \left[\frac{(t/s)^{-\gamma}}{1-\gamma} + \frac{\gamma (t/s)^{-(1+\gamma)}}{1-\gamma} + \frac{1}{2-\gamma} {}_2F_1\left(1, 1+\gamma, 3-\gamma, s/t\right) \right], \ 0 \le s \le t, \ \gamma := \frac{1}{2} - H$$

1. Trajectories are $H - \epsilon$ Hölder for any $\epsilon > 0$
2. Scaling property (self-similarity): $W_{at}^H = |a|^H \, W_t^H$
3. Increments are positively correlated for $H > 1/2$ and negatively correlated for $H < 1/2$

$$v_t = \xi_0(t) \exp\left(\eta \widetilde{W}_t - \frac{\eta^2}{2} t^{2H}\right), \quad \widetilde{W}_t = \sqrt{2H} \int_0^t (t-s)^{H-1/2} \mathrm{d}W_s, \quad \mathrm{d}\langle Z, W\rangle_t = \rho \mathrm{d}t$$

- The past $(W_s)_{-\infty<s\leq 0}$ is encoded in the forward variance curve $\xi_0(\cdot)$.
- Itô integral well-defined, assuming that $\mathbb{E}[v_t] = \xi_0(t)$ is integrable.
- Martingale iff $\rho \leq 0$, moment explosion in finite time.
- At-the money implied volatility skew behaves like

$$\psi(t) := \frac{\partial}{\partial k}\bigg|_{k=0} \sigma_{\text{imp}}(k,t) \sim \frac{\sqrt{2H}}{(H+1/2)(H+3/2)} \rho \eta t^{H-1/2} \text{ as } t \searrow 0$$

- Note that the skew flattens as $H \searrow 0$ in rough Bergomi (Forde, Fukasawa, Gerhold, Smith '22). Possible solution: log-modulated volatility (B., Harang, Pigato '21)

$$\widetilde{W}_t := \int_0^t K(t-s)\mathrm{d}W_s, \quad K(r) := C r^{H-1/2} \max\left(\zeta \log(1/r), 1\right)^{-p}, \quad p > 1, \quad H \geq 0.$$

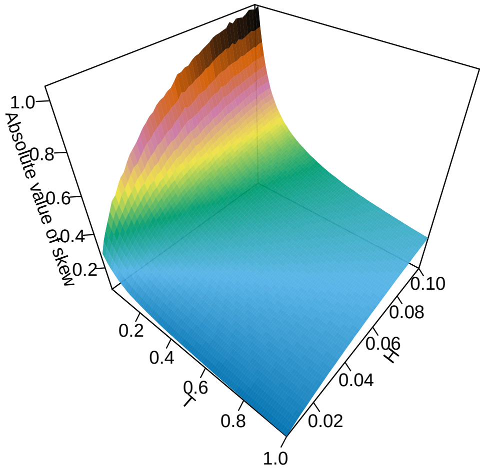


Figure: Rough Bergomi model.

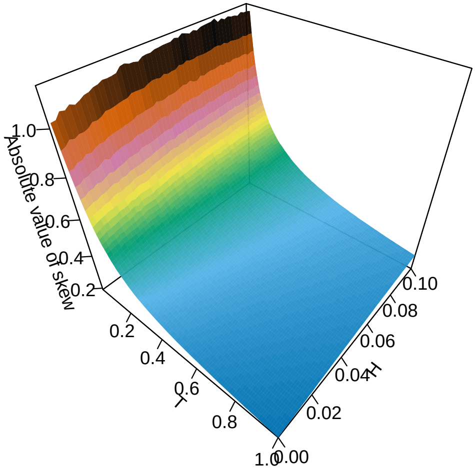


Figure: Log-modulated rough Bergomi.

Given realizations v_{t_i}, $Z_{t_i} = \rho W_{t_i} + \sqrt{1-\rho^2} W^{\perp}_{t_i}$, $i = 0, \dots, N$, on a grid $0 = t_0 < \cdots < t_N = T$, set $\overline{X}_0 := \log S_0$ and

$$\overline{X}_{t_{i+1}} := \overline{X}_{t_i} - \frac{1}{2} v_{t_i}(t_{i+1} - t_i) + \sqrt{v_{t_i}}(Z_{t_{i+1}} - Z_{t_i}), \quad i = 0, \dots, N-1,$$

$\overline{S}_{t_i} := \exp(\overline{X}_{t_i})$.

Convergence

As the mesh $\max_i(t_{i+1} - t_i) \searrow 0$, we have convergence of $\overline{S}$ to S with

- strong rate H (Neuenkirch and Shalaiko ’16);
- weak rate $\min(1/2 + 3H, 1)$ (conjectured), see (Bonesini, Jacquier, Pannier ’23+).

Exact simulation of $(v_{t_i}, Z_{t_i})_{i=0}^N$ or hybrid scheme of (Bennedsen, Lunde, Pakkanen ’17).

$$v_t = v_0 + \int_0^t K(t-s)\kappa(\theta - v_s)\mathrm{d}s + \int_0^t K(t-s)\sigma\sqrt{v_s}\mathrm{d}W_s, \quad K(r) := \frac{r^{H-1/2}}{\Gamma(H+1/2)}, \quad \mathrm{d}\langle Z, W\rangle_t = \rho \mathrm{d}t$$

- Example of a stochastic Volterra equation (SVE)
- Standard well-posedness theory of SVEs does not apply, due to $\sqrt{\cdot}$-term
- Uniqueness in law due to affine structure: there is a unique weak solution.
- The forward variance $\xi_t(T)$, $0 \le t \le T$, satisfies

$$\mathrm{d}\xi_t(T) = \frac{1}{\lambda} R_\kappa(T-t)\sigma\sqrt{v_t}\mathrm{d}W_t, \quad \xi_0(T) = v_0\left(1 - \int_0^T R_\kappa(s)\mathrm{d}s\right) + \theta\int_0^T R_\kappa(s)\mathrm{d}s,$$

where R_κ denotes the resolvent of the second kind of κK, i.e., it solves

$$\kappa K(t) - R_\kappa(t) = \int_0^t R_\kappa(t-s)\kappa K(s)\mathrm{d}s =: R_\kappa * \kappa K(t)$$

Let $X_t = \log S_t$. The characteristic function $\phi_t(u) = \mathbb{E}[e^{\mathrm{i}uX_T} \mid \mathcal{F}_t]$ is of the form

$$\phi_0(u) = \exp\left(\mathrm{i}uX_0 + \int_0^T \kappa\theta\psi(u,s)\mathrm{d}s + \psi(u,T)v_0\right).$$

Define

$$Q(u,z) := \frac{1}{2}\mathrm{i}u(\mathrm{i}u - 1) + \mathrm{i}\sigma\rho uz + \frac{1}{2}\sigma^2 z^2.$$

The function $\psi(u,\tau)$ satisfies the Riccati ODE ($\tau = T - t$):

$$\partial_\tau \psi(u,\tau) = Q(u,\psi(u,\tau)) - \kappa\psi(u,\tau), \quad \psi(u,0) = 0.$$

This system admits a closed-form solution.

For

$$Q(u,z) := \frac{1}{2}\mathrm{i}u(\mathrm{i}u-1) + \mathrm{i}\sigma\rho u z + \frac{\sigma^2}{2}z^2, \quad u \in \mathbb{R},$$

let $\psi(u,t)$, $u \in \mathbb{R}$, $t \in [0,T]$, denote the (unique!) solution of the Volterra – Riccati equation

$$\psi(u,\tau) = K * \left(Q(u,\psi(u,\cdot)) - \kappa\psi(u,\cdot)\right)(\tau), \quad \tau \in [0,T].$$

$$\phi_0(u) := \mathbb{E}\left[\exp\left(\mathrm{i}uX_T\right)\right] = \exp\left(\mathrm{i}uX_0 + \int_0^T \xi_t(s)Q(u,\psi(u,T-s))\mathrm{d}s\right)$$

Define the fractional integral and derivative operators by

$$I^{H+1/2}f(t) := K * f(t), \quad D^{H+1/2}f(t) := \frac{\mathrm{d}}{\mathrm{d}t} I^{H+1/2}f(t).$$

Note that $D^{H+1/2} \circ I^{H+1/2} = I^{H+1/2} \circ D^{H+1/2} = \mathrm{Id}$. The Riccati Volterra equation for ψ is equivalent to

$$D^{H+1/2}\psi(u,\cdot)(t) = Q(u,\psi(u,t)) - \kappa\psi(u,t),$$

and, additionally,

$$\phi_0(u) = \exp\left(\mathrm{i}uX_0 + \kappa\theta \int_0^T \psi(u,s)\mathrm{d}s + v_0 I^{1/2-H}\psi(u,\cdot)(T)\right).$$

$$\frac{\mathrm{d}}{\mathrm{d}t} y(t) = f(t, y(t)), \quad y(0) = y_0$$

- The Adams–Bashforth–Moulton method is a second order predictor–corrector method for finding approximations $(y_k)_{k=0}^N$ of the solution on a grid $(t_k)_{k=0}^N$, $t_0 = 0$, $t_N = T$.
- Based on discretization of $y(t_{k+1}) = y(t_k) + \int_{t_k}^{t_{k+1}} f(s, y(s))\,\mathrm{d}s$.

Prediction step: Approximate $s \mapsto f(s, y(s))$ as constant function (rectangle rule), corresponding to forward Euler approximation $y_{k+1}^P := y_k + \Delta t_k f(t_k, y_k)$.

Correction step: Approximate $s \mapsto f(s, y(s))$ as linear function (trapezoidal rule) $y(t_{k+1}) \approx y(t_k) + \frac{\Delta t_k}{2}(f(t_k, y(t_k)) + f(t_{k+1}, y(t_{k+1})))$. Plugging in the predictor gives

$$y_{k+1} := y_k + \frac{\Delta t_k}{2}\left(f(t_k, y_k) + f(t_{k+1}, y_{k+1}^P)\right).$$

$$y(t) = \frac{1}{\Gamma(\alpha)} \int_0^t (t-s)^{\alpha-1} f(s, y(s))\, \mathrm{d}s, \quad t \in [0,T],\ \frac{1}{2} < \alpha < 1,\ t_k := k\Delta t,\ \Delta t := T/N$$

Prediction step: Again, use rectangle rule, gives the Euler type approximation

$$\int_0^{t_{k+1}} (t_{k+1}-s)^{\alpha-1} g(s)\, \mathrm{d}s \approx \sum_{j=0}^{k} \int_{t_j}^{t_{j+1}} (t_{k+1}-s)^{\alpha-1} g(t_j)\, \mathrm{d}s = \sum_{j=0}^{k} \frac{(t_{k+1}-t_j)^\alpha - (t_{k+1}-t_{j+1})^\alpha}{\alpha} g(t_j).$$

$$y(t) = \frac{1}{\Gamma(\alpha)} \int_0^t (t-s)^{\alpha-1} f(s, y(s))\,\mathrm{d}s, \quad t \in [0,T],\ \frac{1}{2} < \alpha < 1,\ t_k := k\Delta t,\ \Delta t := T/N$$

Prediction step: Again, use rectangle rule, gives the Euler type approximation

$$y^P_{k+1} := \frac{1}{\Gamma(\alpha)} \sum_{j=0}^{k} b_{j,k+1} f(t_j, y_j),\ b_{j,k+1} := \frac{\Delta t^\alpha}{\alpha}\left((k+1-j)^\alpha - (k-j)^\alpha\right).$$

$$y(t) = \frac{1}{\Gamma(\alpha)} \int_0^t (t-s)^{\alpha-1} f(s, y(s))\, \mathrm{d}s, \quad t \in [0,T],\ \frac{1}{2} < \alpha < 1,\ t_k := k\Delta t,\ \Delta t := T/N$$

Prediction step: Again, use rectangle rule, gives the Euler type approximation

$$y^P_{k+1} := \frac{1}{\Gamma(\alpha)} \sum_{j=0}^{k} b_{j,k+1} f(t_j, y_j),\ b_{j,k+1} := \frac{\Delta t^\alpha}{\alpha}\left((k+1-j)^\alpha - (k-j)^\alpha\right).$$

Correction step: Use a trapezoidal rule, i.e., compute

$$\int_0^{t_{k+1}} (t_{k+1}-s)^{\alpha-1} g(s)\, \mathrm{d}s \approx \sum_{j=0}^{k} \int_0^{t_{k+1}} (t_{k+1}-s)^{\alpha-1} \left[g(t_j) + \frac{s-t_j}{\Delta t}\left(g(t_{j+1}) - g(t_j)\right)\right] \mathrm{d}s$$

$$y(t) = \frac{1}{\Gamma(\alpha)} \int_0^t (t-s)^{\alpha-1} f(s, y(s))\,\mathrm{d}s, \quad t \in [0,T],\ \frac{1}{2} < \alpha < 1,\ t_k := k\Delta t,\ \Delta t := T/N$$

Prediction step: Again, use rectangle rule, gives the Euler type approximation

$$y^P_{k+1} := \frac{1}{\Gamma(\alpha)} \sum_{j=0}^{k} b_{j,k+1} f(t_j, y_j),\ b_{j,k+1} := \frac{\Delta t^\alpha}{\alpha}\left((k+1-j)^\alpha - (k-j)^\alpha\right).$$

Correction step: Use a trapezoidal rule, i.e., compute (with $\widetilde{y}_j := y_j \mathbb{1}_{j\le k} + y^P_{k+1} \mathbb{1}_{j=k+1}$)

$$y_{k+1} := \frac{1}{\Gamma(\alpha)} \sum_{j=0}^{k+1} a_{j,k+1} f(t_j, \widetilde{y}_j),\ a_{j,k+1} := \frac{\Delta t^\alpha}{\alpha(\alpha+1)} \begin{cases} k^{\alpha+1} - (k-\alpha)(k+1)^\alpha, & j = 0, \\ (k-j+2)^{\alpha+1} + (k-j)^{\alpha+1} - 2(k-j+1)^{\alpha+1}, & \text{else} \\ 1, & j = k+1. \end{cases}$$

Theorem (Diethelm, Ford, Freed 2004)

Let $\frac{1}{2} \leq \alpha < 1$ *and assume that the solution* $y \in C^2([0,T])$. *Then we have*

$$\max_{0 \leq j \leq N} \left| y(t_j) - y_j \right| = \mathcal{O}(\Delta t), \quad |y(T) - y_N| = \mathcal{O}(\Delta t^{2-\alpha}).$$

- Rate > 1 except close to 0 – when the singularity hits hardest.
- Problem: smoothness of f does not imply smoothness of y. (E.g., fBm less smooth than driving Bm.) Indeed, simple examples for "f smooth $\Longleftrightarrow$ y non-smooth".
- Fractional Adams scheme converges for any $\alpha > 0$, and rate is increasing in α. Contrast: discretization of fractional derivative leads to scheme with rate decreasing (!) in α (no convergence for $\alpha \geq 2$).

- Option payoff f – in terms of log-price – with Fourier tranform $\widehat{f}$. E.g., for the call option:

$$f(x) = (\mathrm{e}^x - K)^+ , \quad \widehat{f}(z) = \frac{K^{1+\mathrm{i}z}}{\mathrm{i}z(1+\mathrm{i}z)}, \; \Im z > 1.$$

  (Note that for the put option we get the same formula, but for $\Im z < 0$.)

- Log-price has characteristic function $\phi_T(u) = \mathbb{E}[e^{\mathrm{i}uX_T}]$.

Pricing formula: For a dampening factor $R > 1$ chosen such that the integrand is integrable (on the contour $\mathbb{R} + \mathrm{i}R$),

$$\mathbb{E}[f(X_T)] = \frac{1}{2\pi} \int_{\mathbb{R}} S_0^{R-\mathrm{i}u} \phi_T(-\mathrm{i}R - u) \widehat{f}(u + \mathrm{i}R) \, \mathrm{d}u.$$

Compare the n-factor Bergomi model (2005)

$$v_t = \xi_0(t) \exp\left(\int_0^t \sum_{i=1}^n \omega_i e^{-\kappa_i(t-s)} \mathrm{d}W_s^i - \frac{1}{2} \int_0^t \sum_{i,j=1}^n \rho_{i,j} \omega_i \omega_j e^{-(\kappa_i+\kappa_j)(t-s)} \mathrm{d}s \right).$$

with the rough Bergomi model (2016):

$$v_t = \xi_0(t) \exp\left(\int_0^t \eta \sqrt{2H} (t-s)^{H-1/2} \mathrm{d}W_s - \frac{\eta^2}{2} t^{2H} \right).$$

If $W_1, \dots, W_n$ highly correlated and

$$\eta \sqrt{2H} r^{H-1/2} \approx \sum_{i=1}^n \omega_i e^{-\kappa_i r},$$

then n-factor Bergomi approximates rough Bergomi – or conversely.

Let $\widetilde{W}_t := \int_0^t K(t-s)\mathrm{d}W_s$ for a completely monotone kernel K with Laplace tranform μ, i.e.,

$$K(r) = \int_0^\infty \mathrm{e}^{-xr}\mu(\mathrm{d}x).$$

Example: fractional kernel

$$K(r) = \frac{r^{H-1/2}}{\Gamma(H+1/2)} \implies \mu(\mathrm{d}x) = w(x)\mathrm{d}x, \; w(x) = \frac{x^{-H-1/2}}{\Gamma(H+1/2)\Gamma(1/2-H)}$$

$$\widetilde{W}_t = \int_0^t K(t-s)\mathrm{d}W_s = \int_0^t \int_0^\infty \mathrm{e}^{-x(t-s)}\mu(\mathrm{d}x)\mathrm{d}W_s = \int_0^\infty \int_0^t \mathrm{e}^{-x(t-s)}\mathrm{d}W_s\,\mu(\mathrm{d}x) = \int_0^\infty Y_t^x \mu(\mathrm{d}x),$$

with $(Y^x)_{x\geq 0}$ denoting a family of Ornstein–Uhlenbeck processes indexed by their speed of mean reversion x – driven by a single Bm W.

Let $K(r)$ be a completely monotone kernel, $\widehat{K}(r) := \sum_{i=1}^{n} w_i e^{-x_i r}$. Consider

$$X_t = x_0 + \int_0^t K(t-s)b(X_s)\mathrm{d}s + \int_0^t K(t-s)\sigma(X_s)\mathrm{d}W_s,$$
$$\widehat{X}_t = x_0 + \int_0^t \widehat{K}(t-s)b(\widehat{X}_s)\mathrm{d}s + \int_0^t \widehat{K}(t-s)\sigma(\widehat{X}_s)\mathrm{d}W_s.$$

Under standard regularity conditions, Itô’s formula implies that (Alfonsi, Kebaier ’24):
$\widehat{X}_t = \sum_{i=1}^{n} w_i X_t^i$, where $X^{(n)} := (X^1, \ldots, X^n)$ solves

$$\mathrm{d}X_t^i = -x_i(X_t^i - x_0^i)\mathrm{d}t + b\left(\widehat{X}_t\right)\mathrm{d}t + \sigma\left(\widehat{X}_t\right)\mathrm{d}W_t, \quad X_0^i = x_0^i, \quad i = 1, \ldots, n,$$

provided that $x_0 = \sum_{i=1}^{n} w_i x_0^i$.

Theorem (Strong error; Alfonsi, Kebaier '24)

$$\mathbb{E}\left[\left|X_T - \widehat{X}_T\right|^2\right] \leq \text{const} \int_0^T \left|K(t) - \widehat{K}(t)\right|^2 \mathrm{d}t$$

In the specific case of affine Volterra processes (rough Heston), we can use the characteristic function to obtain estimates for the weak error.

Theorem (Weak error; B., Breneis '23; Abi Jaber, El Euch '19)

Under decay conditions of the Fourier transformation $\widehat{f}$ of the payoff f, we have

$$\left|\mathbb{E}\left[f(S_T)\right] - \mathbb{E}\left[f(\widehat{S}_T)\right]\right| \leq \text{const}\left(\int_0^T \left|K(s) - \widehat{K}(s)\right| \mathrm{d}s\right)^q.$$

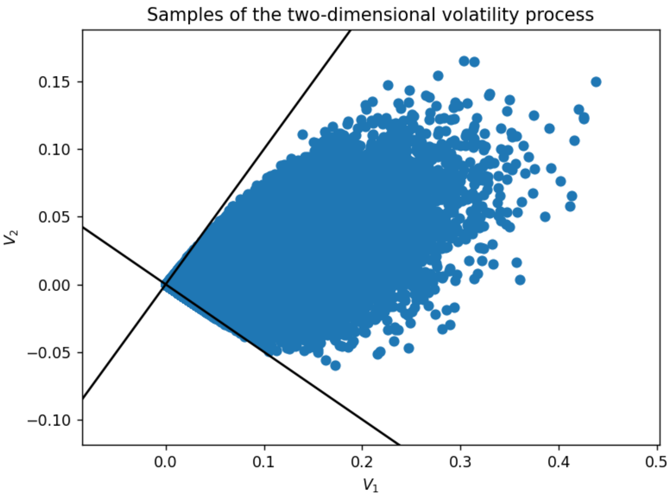
$$
\begin{aligned}
\mathrm{d}\widehat{S}_t &= \sqrt{\widehat{v}_t}\,\widehat{S}_t\,\mathrm{d}Z_t,\\
\mathrm{d}v_t^i &= -x_i\left(v_t^i - v_0^i\right)\mathrm{d}t + \left(\theta - \kappa v_t^i\right)\mathrm{d}t + \eta\sqrt{\widehat{v}_t}\mathrm{d}W_t,
\end{aligned}
$$

where

$$
\widehat{v}_t = v_0 + \int_0^t \widehat{K}(t-s)(\theta - \kappa\widehat{v}_s)\mathrm{d}s + \int_0^t \widehat{K}(t-s)\eta\sqrt{\widehat{v}_s}\mathrm{d}W_s = \sum_{i=1}^N w_i v_t^i,
$$

$$
\widehat{K}(t) := \sum_{i=1}^n w_i \mathrm{e}^{-x_i t}, \quad \sum_{i=1}^n w_i v_0^i = v_0
$$

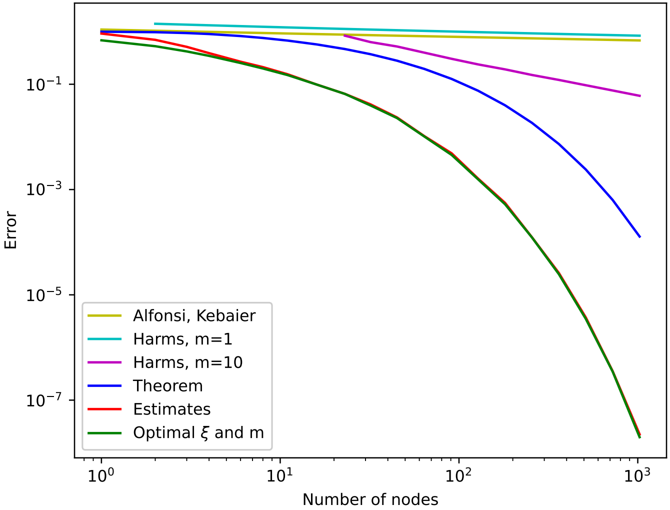
Figure: The domain of $v^{(n)}$ is a cone – explicit characterization of the state space due to (Abi Jaber, B., Breneis '24).



We consider the construction of Markovian approximations for the fractional kernel $K(r) = \frac{1}{\Gamma(H+1/2)} r^{H-1/2}$.

- Trapezoidal rule: Abi Jaber and El Euch 2019, Alfonsi and Kebaier 2021 choose (w_i), (x_i) inspired by the trapezoidal rule, giving an error asymptotic of (almost) n^{-H}.
- Gaussian quadrature: Harms 2021 chooses Gaussian quadrature of level m on k geometrically spaced intervals, giving an asymptotic error of (almost) $k^{-2Hm/3}$, where $n = km$.
  B. and Breneis optimize k, m for given n, obtaining an L^2 error (between kernels) of order $C_1 n \exp\left(-C_2 \left(\frac{1}{H} + \frac{1}{3/2-H}\right)^{-1/2} \sqrt{n}\right)$.
- Numerical minimization of the L^2 error between kernels.

Figure: Error bounds $\int_0^T \left|K(t) - \widehat{K}(t)\right|^2 \mathrm{d}t$ for the fractional kernel, $H = 0.1$, $T = 1$.



Theorem (B. and Breneis 2023+)

Using Gaussian quadrature on geometric grid or non-geometric grids, we obtain

$$\int_0^T \left|K(t) - K^n(t)\right| \mathrm{d}t \lesssim c_1 \exp\left(-c_2 \sqrt{(H+1/2)n}\right) \simeq c_3 \xi_n^{-H-1/2}.$$

- Geometric grid: $\xi_i = a\,(\xi_n/a)^{i/n}$, $\xi_0 = 0$.
- Non-geometric grid: $\xi_0 = 0$, $\xi_1 = a$, $\xi_{i+1} = \left(\frac{c+\xi_i^{\frac{1/2+H}{2m}}}{c-\xi_i^{\frac{1/2+H}{2m}}}\right)\xi_i$.
- Proofs based on careful estimation of the error built on classical error representations for Gaussian quadrature, linked with explicit optimization of the free parameters.
- The exponent remains bounded away from 0 even as $H \to 0$ – allowing extension to $H > -1/2$.

Choose K^n in order to minimize $\int_0^T \left|K(t) - \widehat{K}(t)\right| \mathrm{d}t$.

- Looks unpleasant, but for kernels K^n based on Gaussian quadrature:
  1. $K(t) \geq \widehat{K}(t)$;
  2. $\displaystyle\int_0^T \left|K(t) - \widehat{K}(t)\right| \mathrm{d}t = \frac{T^{H+1/2}}{\Gamma(H+3/2)} - \sum_{i=1}^{n} \frac{w_i}{x_i}\left(1 - \mathrm{e}^{-x_i T}\right)$.
- For general Markovian approximation, compute $\int_0^T \left|K(t) - \widehat{K}(t)\right| \mathrm{d}t$ by computing the locations t_k of sign change of $K(\cdot) - \widehat{K}(\cdot)$, i.e.,

$$\int_0^T \left|K(t) - \widehat{K}(t)\right| \mathrm{d}t = \sum_k \left|\int_{t_k}^{t_{k+1}} \left(K(t) - \widehat{K}(t)\right) \mathrm{d}t\right|.$$

$$\mathrm{d}v_t^i = -x_i \left(v_t^i - v_0^i \right) \mathrm{d}t + (\theta - \kappa \widehat{v}_t)\, \mathrm{d}t + \eta \sqrt{\widehat{v}_t} \mathrm{d}W_t$$

- Wide range of speeds of mean reversion, from ≈ 1 to $\approx 10^7$.
- All these scales are equally important for the solution due to the careful design of K^n.
- No clear separation of scales – works for rough Bergomi due to its simpler structure.
- Very often, $v \approx 0$, close to the singularity.
- Some v^i may become negative, provided that $\widehat{v} \geq 0$. Difficult to guarantee theoretically, though.

$$\mathrm{d}v_t^i = -x_i\left(v_t^i - v_0^i\right)\mathrm{d}t + (\theta - \kappa\widehat{v_t})\,\mathrm{d}t + \eta\sqrt{\widehat{v_t}}\mathrm{d}W_t$$

- Wide range of speeds of mean reversion, from ≈ 1 to $\approx 10^7$.
- All these scales are equally important for the solution due to the careful design of K^n.
- No clear separation of scales – works for rough Bergomi due to its simpler structure.
- Very often, $v \approx 0$, close to the singularity.
- Some v^i may become negative, provided that $\widehat{v} \geq 0$. Difficult to guarantee theoretically, though.

Failed attempts

- Drift-implicit Euler scheme
- Replacing fast components by conditional invariant distributions

$$\mathrm{d}v_t^i = -x_i\left(v_t^i - v_0^i\right)\mathrm{d}t + (\theta - \lambda\widehat{v_t})\,\mathrm{d}t + \eta\sqrt{\widehat{v_t}}\mathrm{d}W_t$$

- Standard diffusion process, well understood discretization methods, in principle.
- Stiffness largely due to very high speeds of mean reversion:

L^2-optimized ($n = 6$)

$x_n \approx 10^{25}$ $(H = 0.01)$, $x_n \approx 10^8$ $(H = 0.1)$

L^1 optimized ($n = 2$)

$x_n \approx 17.8$ $(H = 0.01)$, $x_n \approx 8.7$ $(H = 0.1)$

- Expect asymptotic behavior for time discretization schemes only when $\Delta t \ll 1/\xi_n$.
- Difficulties only arise in the drift. This suggests to use drift-implicit discretization or splitting.
- Splitting (with exact solution of the drift equation) works very well for rough Bergomi.

A weak, second order scheme for simulating

$$\mathrm{d}V_t = \alpha\sqrt{V_t}\mathrm{d}W_t.$$

Notation: $x := \overline{V}_k$, $z := \alpha^2\Delta t$, $A := \frac{3+\sqrt{3}}{4}$, update $\overline{V}_{k+1} = x_i$ with probability p_i, $i = 1, 2, 3$, where

$$p_1 := \frac{m_1 x_2 x_3 - m_2(x_2 + x_3) + m_3}{x_1(x_3 - x_1)(x_2 - x_1)}, \quad p_2 := \frac{m_1 x_1 x_3 - m_2(x_1 + x_3) + m_3}{x_2(x_3 - x_2)(x_1 - x_2)},$$

$$x_{1/3} := x + \left(A + \frac{3}{4}\right)z \mp \sqrt{\left(3x + \left(A + \frac{3}{4}\right)^2 z\right)z}, \quad x_2 := x + Az,$$

$$m_1 := x, \quad m_2 := x^2 + xz, \quad m_3 := x^3 + 3x^2 z + \frac{3}{2}xz^2.$$

$$\mathrm{d}v_t^i = -x_i\left(v_t^i - v_0^i\right)\mathrm{d}t + (\theta - \kappa\widehat{v_t})\,\mathrm{d}t + \eta\sqrt{\widehat{v_t}}\mathrm{d}W_t$$

Splitting

Drift: Let $D(\mathbf{v};\Delta t)$ denote the solution map at time Δt of

$$\mathrm{d}v_t^i = -x_i(v_t^i - v_0^i)\mathrm{d}t + (\theta - \kappa v_t)\mathrm{d}t, \quad v_0^i = v^i,\ \mathbf{v} = (v^i)_{i=1}^N,\ v_t = \sum_{i=1}^n w_i v_t^i,\ i = 1,\dots,n.$$

Note that $D(\mathbf{v};\Delta t)$ can be computed in closed form.

$$\mathrm{d}v_t^i = -x_i\left(v_t^i - v_0^i\right)\mathrm{d}t + (\theta - \kappa\widehat{v_t})\,\mathrm{d}t + \eta\sqrt{v_t}\mathrm{d}W_t$$

Splitting

Drift: Let $D(\mathbf{v};\Delta t)$ denote the solution map at time Δt of the drift.

Diffusion: Let $S(\mathbf{v};\Delta t)$ denote the solution map of

$$\mathrm{d}v_t^i = \eta\sqrt{v_t}\mathrm{d}W_t, \quad v_0^i = v^i, \quad \mathbf{v} = (v^i)_{i=1}^n,\ v_t = \sum_{i=1}^n w_i v_t^i,\ i = 1,\dots,n.$$

Note that v_t above solves $\mathrm{d}v_t = \alpha\sqrt{v_t}\mathrm{d}W_t$, $\alpha := \eta\sum_{i=1}^n w_i$. Let $\widetilde{S}(\mathbf{v};\Delta t)$ be the approximation to the solution due to Lileika–Mackevičius, and set

$$\widehat{S}(\mathbf{v};\Delta t) := \mathbf{v} + \frac{\widetilde{S}(\mathbf{v};\Delta t) - \sum_{i=1}^n w_i v^i}{\sum_{i=1}^N w_i}.$$

$$\mathrm{d}v_t^i = -x_i \left(v_t^i - v_0^i\right) \mathrm{d}t + (\theta - \kappa \widehat{v_t})\, \mathrm{d}t + \eta \sqrt{\widehat{v_t}}\mathrm{d}W_t$$

Splitting

Drift: Let $D(\mathbf{v}; \Delta t)$ denote the solution map at time Δt of the drift.

Diffusion: Let $\widehat{S}(\mathbf{v}; \Delta t)$ denote the approximate solution map of the diffusion part.

Splitting: The total one-step scheme is defined by Strang splitting, i.e.,

$$A(\mathbf{v}; \Delta t) := D\left(\widehat{S}\left(D\left(\mathbf{v}; \Delta t/2\right); \Delta t\right); \Delta t/2\right).$$

$$\mathrm{d}v_t^i = -x_i\left(v_t^i - v_0^i\right)\mathrm{d}t + (\theta - \kappa\widehat{v_t})\,\mathrm{d}t + \eta\sqrt{v_t}\mathrm{d}W_t$$

Splitting

Drift: Let $D(\mathbf{v};\Delta t)$ denote the solution map at time Δt of the drift.

Diffusion: Let $\widehat{S}(\mathbf{v};\Delta t)$ denote the approximate solution map of the diffusion part.

Splitting: The total one-step scheme is defined by Strang splitting, i.e.,

$$A(\mathbf{v};\Delta t) := D\left(\widehat{S}\left(D\left(\mathbf{v};\Delta t/2\right);\Delta t\right);\Delta t/2\right).$$

Stock price: Split into independent and correlated part, i.e.,

$$\mathrm{d}S_t = \rho S_t\sqrt{v_t}\mathrm{d}W_t + \sqrt{1-\rho^2}S_t\sqrt{v_t}\mathrm{d}W_t^{\perp} =: \mathrm{d}S_t^w + \mathrm{d}S_t^{\perp},$$

using that $\eta\int_s^t\sqrt{v_u}\mathrm{d}W_u = (v_t^1 - v_s^1) + \int_s^t[x_1(v_u^1 - v_0^1) - (\theta - \lambda v_u^N)]\mathrm{d}u.$

Trajectory of $(S, \sqrt{v})$

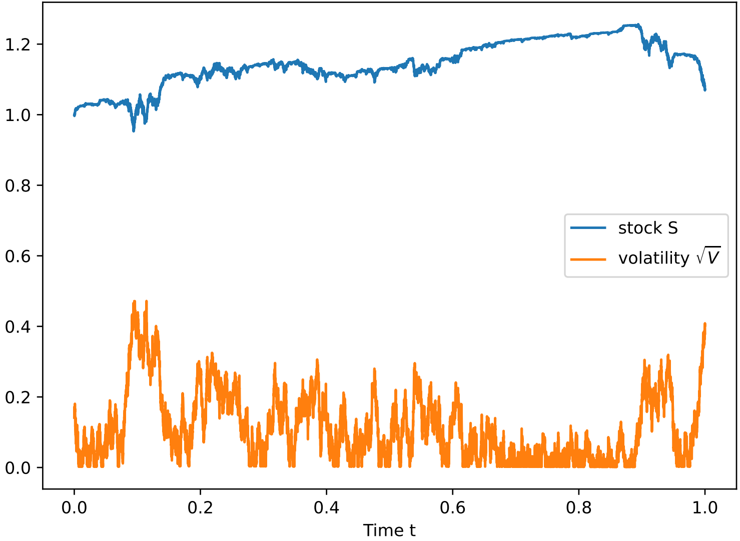


Figure: Sample trajectory of $(S, \sqrt{v})$ for $H = 0.1$, $N = 3$.

- Empirical results seem to indicate weak rate of convergence 2.
- Comparisons with Fourier pricing, drift-implicit Euler, and Gatheral's HQE scheme.
- HQE is using conditional moment matching inspired by Andersen's QE scheme for Heston.
- HQE empirically shows weak convergence with rate 1, but computational cost is quadratic in the number of time-steps.
- Further comparison with drift-implicit Euler scheme for the Markovian approximation. In our implementation, this is the fastest scheme. Convergence only for $\Delta t \ll \xi_n^{-1}$, seemingly with rate 1.

Smile of European call option $H = 0.1$, $N = 2$

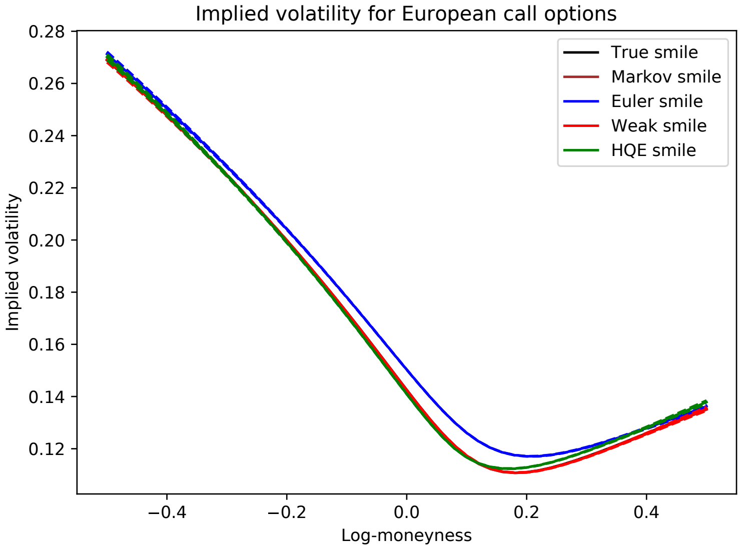


Figure: Implied volatility smile: $T = 1$, $H = 0.1$, $N = 2$, 32 timesteps.

Relative errors in the smile of European call options for $H = 0.1$, $N = 2$, $T = 1$

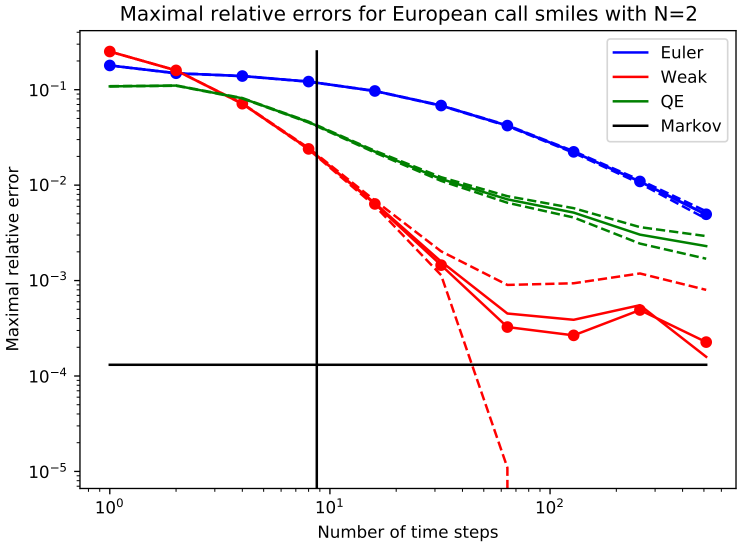


Figure: Maximal relative errors of implied volatility smiles for $T = 1$, $H = 0.1$, $N = 2$.

Relative errors in the smile of European call options for $H = 0.1$, $N = 3$, $T = 1$

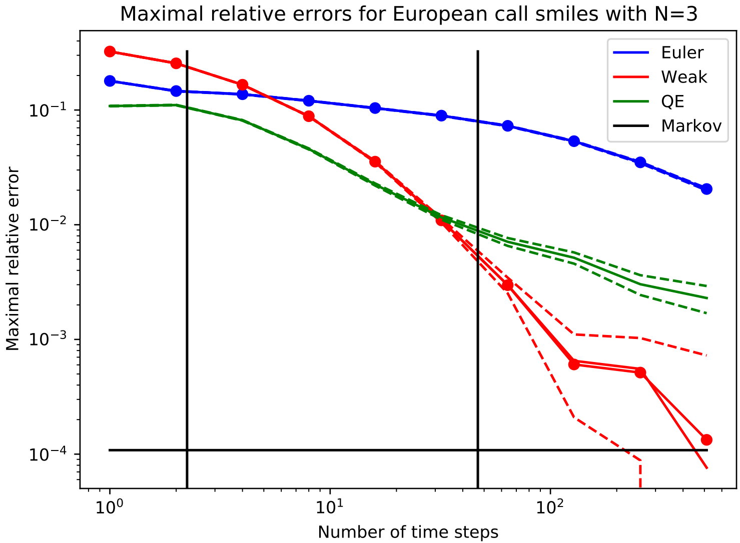


Figure: Maximal relative errors of implied volatility smiles for $T = 1$, $H = 0.1$, $N = 3$.

Relative errors in the smile of European call options for $H = -0.2$, $N = 2$, $T = 1$

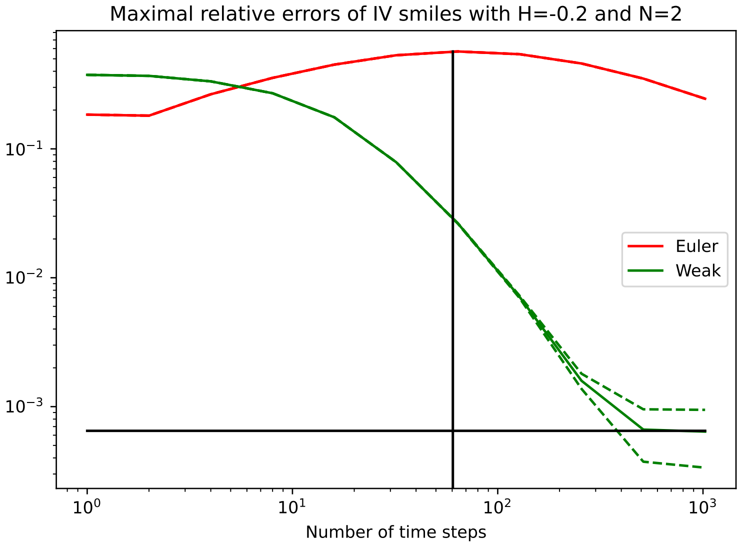


Figure: Maximal relative errors of implied volatility smiles for $T = 1$, $H = -0.2$, $N = 2$.

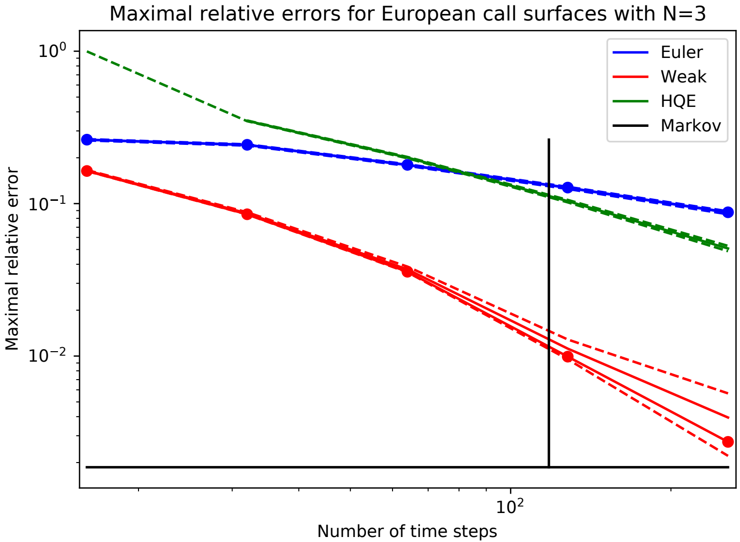


Figure: Maximal relative errors of implied volatility surface for $T = 1/16, 2/16, \ldots, 1$, $H = 0.1$, $N = 3$.

Markovian approximation discretized by the Lileika–Mackevičius scheme allows for efficient and accurate simulation of rough Heston.

- For simulation at fixed time T, $N = 2$ factors are sufficient. Simulation of a "normal surface" requires $N = 3$ factors.
- The maximal speed of mean reversion in these cases is at most ≈ 100.
- We recover the expected explosion of the ATM skew.
- Observed rates of convergence: 2 for the Lileika–Mackevičius scheme, 1 for the HQE scheme, possibly 1 for the drift-implicit Euler scheme.
- The computational cost of the Lileika–Mackevičius scheme for the Markovian approximation is considerably less than the cost of the HQE scheme.
- Open: Numerical analysis of the Lileika–Mackevičius scheme and the "full scheme", including the Markovian approximation.
- Open: American option pricing, but promising initial results.

In what follows, we shall always assume that processes are continuous and have enough integrability – i.e., exponential integrability.

Definition

Given two semi-martingales X, Y on $[0, T]$, then we define

$$(X \diamond Y)_t(T) \coloneqq \mathbb{E}\left[\langle X, Y\rangle_{t,T} \mid \mathcal{F}_t\right] = \mathbb{E}\left[\langle X, Y\rangle_T \mid \mathcal{F}_t\right] - \langle X, Y\rangle_t, \quad 0 \le t \le T.$$

- The diamond product is commutative: $X \diamond Y = Y \diamond X$.
- It is not associative: $X \diamond (Y \diamond Z) \neq (X \diamond Y) \diamond Z$.
- It only depends on the local martingale parts of X and Y.

Theorem (Friz, Gatheral, Radoiĉić '22)

Let Y be a martingale. Then, for $|a|, |b|$ small enough:

$$\log \mathbb{E}\left[e^{aY_T + b\langle Y\rangle_T} \mid \mathcal{F}_t \right] = aY_t + b\langle Y\rangle_t + \sum_{k=2}^{\infty} \mathbb{G}_t^k(T),$$

where

$$\mathbb{G}^2 := \left(\frac{1}{2}a^2 + b\right)(Y \diamond Y)_t(T), \quad \mathbb{G}^k := \frac{1}{2}\sum_{j=2}^{k-2} \mathbb{G}^{k-j} \diamond \mathbb{G}^j + (aY \diamond \mathbb{G}^{k-1}), \quad k > 2.$$

Example: If $b = -a^2/2$, then $e^{aY_t + b\langle Y\rangle_t}$ is a martingale and, hence, all the higher order terms have to vanish. And, indeed, they do.

$$\mathrm{d}S_t = S_t \sqrt{v_t}\mathrm{d}Z_t \Rightarrow X_t := \log S_t = \log S_0 - \frac{1}{2}\int_0^t v_s \mathrm{d}s + \int_0^t \sqrt{v_s}\mathrm{d}Z_s$$

with local martingale part $Y_t := \log S_0 + \int_0^t \sqrt{v_s}\mathrm{d}Z_s$. Hence, we have

$$aX_T + b\langle X\rangle_T = aY_T + \left(b - \frac{1}{2}a\right)\langle Y\rangle_T,$$

and the above theorem implies

$$\log \mathbb{E}\left[\mathrm{e}^{aX_T + b\int_0^T v_s \mathrm{d}s} \mid \mathcal{F}_t\right] = aX_t + b\langle X\rangle_t + \sum_{k=2}^{\infty} \mathbb{G}_t^k(T),$$

$$\mathbb{G}^2 := \left(\frac{1}{2}a(a-1) + b\right)(X \diamond X)_t(T), \quad \mathbb{G}^k := \frac{1}{2}\sum_{j=2}^{k-2} \mathbb{G}^{k-j} \diamond \mathbb{G}^j + (aX \diamond \mathbb{G}^{k-1}), \quad k > 2.$$

Implied volatility solves the equation

$$\int_0^\infty \frac{\mathrm{d}u}{u^2+1/4} \Re\left[\mathrm{e}^{-\mathrm{i}uk}\left(\mathrm{e}^{\phi(T;u-\mathrm{i}/2)} - \mathrm{e}^{-\frac{1}{2}(u^2+\frac{1}{4})\sigma_{\mathrm{impl}}(k,T)^2 T}\right)\right] = 0,$$

$$\phi_t(T;u) := \log\left(\mathbb{E}\left[\mathrm{e}^{\mathrm{i}uX_{t,T}} \mid \mathcal{F}_t\right]\right) = \sum_{k=2}^\infty \mathbb{G}_t^k(T;\mathrm{i}u).$$

Ansatz:

$$\phi_t(T;u) = \sum_{j=2}^\infty \mathbb{G}_t^k(T;\mathrm{i}u)\delta^{j-2}, \quad \sigma^2_{\mathrm{impl}}(k,T)T = \sum_{l=0}^\infty a_l(k,T)\delta^l$$

leads to

$$a_0(k,T) = M_0(T) := \int_0^T \xi_0(u)\mathrm{d}u, \quad a_1(k,T) = \left(\frac{k}{M_0(T)} + \frac{1}{2}\right)(X \diamond M)_0(T), \ldots,$$

which is the Bergomi–Guyon expansion (up to order 2)!