Pricing in affine forward variance models

Jim Gatheral

Baruch College
Weissman School of Arts and Sciences
Master of Financial Engineering

Quantminds International,
November 9, 2022

Outline of this talk

- Affine forward variance (AFV) models
  - The rough Heston model
- The CGF and its associated Volterra integral equation
  - Numerical solution using the Adams scheme
  - Rational approximation of the rough Heston solution
- Simulation of AFV models
  - The HQE scheme
  - Numerical tests of convergence

Forward variance models

- Let S be a strictly positive continuous martingale.
- Then $X := \log S$ is a semimartingale with quadratic variation process $\langle X \rangle$.
- Following [BG12], it is natural to specify a model in forward variance form.

$$\begin{aligned} V_t\, dt &:= d\langle X \rangle_t \\ \xi_t(u) &= \mathbb{E}_t\left[V_u\right], \quad u > t. \end{aligned}$$

  - Forward variances are tradable assets (unlike spot variance).
  - We get a family of martingales indexed by their individual time horizons T.
- As noted in [BG12], all conventional finite-dimensional Markovian stochastic volatility models may be cast as forward variance models.

Affine CGF

Let $X_t = \log S_t$. According to Definition 2.2 of [GKR19], a forward variance model has an *affine cumulant generating function* determined by $g(t;u)$, if its conditional cumulant generating function is of the form

$$\log \mathbb{E}\left[e^{u(X_T - X_t)} \middle| \mathcal{F}_t \right] = \int_t^T \xi_t(s)\, g(T-s;u)\, ds. \tag{1}$$

When is a forward variance model affine?

Theorem 2.4 of [GKR19] states that a forward variance model has an affine CGF if and only if it takes the form

$$
\begin{aligned}
\frac{dS_t}{S_t} &= \sqrt{V_t}\, dZ_t \\
d\xi_t(u) &= \kappa(u-t)\, \sqrt{V_t}\, dW_t
\end{aligned}
$$

for some deterministic, non-negative decreasing kernel κ, which satisfies $\int_0^T \kappa(r)dr < \infty$ for all $T > 0$.

- Essentially, the only affine forward variance model is the Heston model, up to a choice of kernel.

Example: Rough Heston

With $\alpha = H + 1/2 \in (1/2, 1)$, the rough Heston model of [ER19] reads

$$V_u = \theta_t(u) - \frac{1}{\Gamma(\alpha)} \int_t^u (u-s)^{\alpha-1} \lambda \, V_s \, ds + \frac{1}{\Gamma(\alpha)} \int_t^u (u-s)^{\alpha-1} \nu \sqrt{V_s} \, dW_s.$$

In the special case $\lambda = 0$, this model takes the forward variance form (by inspection):

Rough Heston with $\lambda = 0$

$$d\xi_t(u) = \frac{\nu}{\Gamma(\alpha)} (u-t)^{\alpha-1} \sqrt{V_t} \, dW_t.$$

Rough Heston with $\lambda > 0$

In the more general case $\lambda > 0$, the rough Heston model takes the forward variance form (see [GKR19]):

Rough Heston with $\lambda > 0$

$$\xi_t(u) = \nu\,(u-t)^{\alpha-1}\,E_{\alpha,\alpha}(-\lambda\,(u-t)^{\alpha})\,\sqrt{V_t}\,dW_t, \tag{2}$$

where $E_{\alpha,\beta}(\cdot)$ is the generalized Mittag-Leffler function and $\alpha = H + \frac{1}{2}$.

- Putting $\alpha = 1$ in (2) gives the exponential kernel - the classical Heston model.

Solving for $g(\cdot)$

$g(\cdot; u)$ in the definition (1) of the CGF is the unique global continuous solution of the convolution Riccati equation

$$g(\tau; u) = R_V\left(u, \int_0^\tau \kappa(\tau - s) g(s; u) ds\right) = R_V\Big(u, (\kappa \star g)(\tau; u)\Big) \tag{3}$$

where

$$R_V(u, w) = \frac{1}{2}(u^2 - u) + \rho\, u\, w + \frac{1}{2}\, w^2.$$

Convolution Riccati equation as a fractional ODE

- When the kernel is of the form $\kappa(\tau) \sim \tau^{\alpha-1}$, the convolution Riccati equation may be rewritten as a fractional ODE.
- For example, in the case of the rough Heston model (with $\lambda = 0$), with $\alpha = H + \frac{1}{2}$,

$$\begin{aligned} \nu\, h(\tau; u) &:= (\kappa \star g)(\tau; u) \\ &= \frac{\nu}{\Gamma(\alpha)} \int_0^\tau (\tau - s)^{\alpha-1}\, g(s; u)\, ds \\ &= \nu\, I^\alpha g(\tau; u). \end{aligned}$$

- Inverting this gives $g(\tau; u) = D^\alpha h(\tau; u)$.

Convolution Riccati equation as a fractional ODE

- The convolution Riccati equation then reads

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

- In the general case $\lambda > 0$, using that $g = D^{\alpha} h + \lambda\, h$, the convolution Riccati equation becomes

$$D^{\alpha} h(\tau; u) = \frac{1}{2}\, u\,(u-1) + (\rho\,\nu\, u - \lambda)\, h(\tau; u) + \frac{1}{2}\,\nu^2\, h(\tau; u)^2. \tag{4}$$

The Lewis formula

- Given the solution $g(\cdot)$ of the convolution Riccati equation, it is straightforward to price European options, for example using the Lewis formula [Lew00]:

$$C(S,K,t,T) = S - \sqrt{SK}\frac{1}{\pi}\int_0^\infty \frac{da}{a^2+\frac{1}{4}} \operatorname{Re}\left[e^{-iak}\varphi_t^T(a-i/2)\right],$$

where

$$\varphi_t^T(a) := \mathbb{E}_t\left[e^{\mathrm{i}aX_{t,T}}\right] = \exp\left\{\int_t^T \xi_t(s)\, g(T-s;\mathrm{i}a)\, ds\right\}.$$

Pricing of more exotic claims

- Let

$$\zeta_t(T) = \int_T^{T+\Delta} \xi_t(u)\, du = \mathbb{E}_t \int_T^{T+\Delta} v_u\, du = \mathbb{E}_t \langle X \rangle_{T,T+\Delta}.$$

  - In the case of SPX, $\zeta_T(T)$ is essentially the payoff of VIX_T^2.

- In Theorem 4.5 of [FGR22], we show that in AFV models, the joint MGF of X, $\langle X \rangle$ and $\zeta(T)$ is given by

$$\begin{aligned} &\log \mathbb{E}_t \left[e^{a\, X_T + b\, \langle X \rangle_{t,T} + c\, \zeta_T(T)} \right] \\ = \quad & a\, X_t + c\, \zeta_t(T) + (\xi \star g)(T - t; a, b, c, \Delta). \end{aligned}$$

- $g(\tau; a, b, c, \Delta)$ satisfies the convolution Riccati equation

$$\begin{aligned} & g(\tau; a, b, c, \Delta) \\ = \; & b - \tfrac{1}{2} a + \tfrac{1}{2}\left(1-\rho^2\right) a^2 \\ & \quad + \tfrac{1}{2}\left[\rho a + c\,\bar{\kappa}(\tau) + (\kappa \star g)(\tau; a, b, c, \Delta)\right]^2, \quad (5) \end{aligned}$$

with the boundary condition

$$g(0; a, b, c, \Delta) = b + \tfrac{1}{2}\, a(a-1) + \rho a c\, \bar{\kappa}(0) + \tfrac{1}{2}\, c^2\, \bar{\kappa}(0)^2,$$

where $\bar{\kappa}(\tau) = \int_\tau^{\tau+\Delta} \kappa(u)\, du$.

Exotic payoff examples

- Volatility swaps: $\sqrt{\langle X \rangle_{t,T}}$
- Options on variance: $(\langle X \rangle_{t,T} - K)^+$
- Corridor variance swaps (CVS): $\int_t^T \mathbb{1}_{\{L \le S_u \le H\}}\, d\langle X \rangle_u$
- Target volatility options (TVO): $\frac{(S_T - K)^+}{\sqrt{\langle X \rangle_{t,T}}}$
- Double digital calls (DDC): $\mathbb{1}_{\{S_T \ge K, \langle X \rangle_{t,T} \ge K_M\}}$

- Given the solution $g(\cdot)$ of the joint convolution Riccati equation (5), such claims may be valued using appropriate versions of the Lewis formula.
  - So how do we solve for $g(\cdot)$?

The fractional Adams scheme

The fractional Adams scheme of [DFF04] is for the numerical approximation of the solution of equations of the form

$$h(\tau) = \frac{1}{\Gamma(\alpha)} \int_0^\tau (\tau - s)^{\alpha-1} F(h(s))\, ds. \tag{6}$$

- In the case of the rough Heston model, the convolution Riccati equations (4) and (5) are of this type.

Rational approximation of the Heston solution

- The Adams scheme for solving the rough Heston fractional differential equation is slow!
- In [GR19], we showed how to approximate the solution of the rough Heston fractional Riccati equation by a rational function.
  - This approximation solution is just as fast as the classical Heston solution and appears to be more accurate than the Adams scheme for any reasonable number of time steps!
  - The rational approximation can be extended to the rough Heston model with $\lambda > 0$.

Rational approximation to the rough Heston solution

Wlog, set $\nu = 1$ and $x = t$. When $\lambda = 0$, the rough Heston fractional Riccati ODE (4) reads

$$\begin{aligned} D^\alpha h(x;a) &= -\frac{1}{2}\, a(a+\mathrm{i}) + \mathrm{i}\,\rho\, a\, h(x;a) + \frac{1}{2}\, h(x;a)^2 \\ &= \frac{1}{2}\, \left(h(x;a) - r_-\right) \left(h(x;a) - r_+\right) \end{aligned}$$

with

$$A = \sqrt{a\,(a+i) - \rho^2\, a^2}; \quad r_\pm = \{-\mathrm{i}\,\rho\, a \pm A\}\,.$$

The idea is to paste together short- and long-time expansions of the solution using a rational (Padé) approximation.

Short-time expansion

From (for example) the exponentiation theorem of [AGR2020], $h(x;a)$ can be written as

$$h(x;a) = \sum_{j=0}^{\infty} \frac{\Gamma(1+j\,\alpha)}{\Gamma(1+(j+1)\,\alpha)} \beta_j(a)\, x^{(j+1)\,\alpha}$$

with

$$\begin{aligned} \beta_0(a) &= -\frac{1}{2}\, a(a+\mathrm{i}) \\ \beta_k(a) &= \frac{1}{2} \sum_{i,j=0}^{k-2} \mathbb{1}_{\{i+j=k-2\}}\, \beta_i(a)\, \beta_j(a) \frac{\Gamma(1+i\,\alpha)}{\Gamma(1+(i+1)\,\alpha)} \frac{\Gamma(1+j\,\alpha)}{\Gamma(1+(j+1)\,\alpha)} \\ &\quad + \mathrm{i}\,\rho\, a\, \frac{\Gamma(1+(k-1)\,\alpha)}{\Gamma(1+k\,\alpha)} \beta_{k-1}(a). \end{aligned}$$

Solving the rough Heston Riccati equation for long times

- In analogy with the classical Heston solution, we expect that for a suitable range of a,

$$\lim_{x \to \infty} h(x;a) = r_- .$$

- In that case, for large x, we could linearize the fractional Riccati equation as follows.

$$\begin{aligned} D^\alpha h(x;a) &= \frac{1}{2}\left(h(x;a) - r_-\right)\left(h(x;a) - r_+\right) \\ &\approx -\frac{1}{2}\left(r_+ - r_-\right)\left(h(x;a) - r_-\right) \\ &= -A\left(h(x;a) - r_-\right). \end{aligned}$$

continued...

- The above linear fractional differential equation has the exact solution
$$h_\infty(a,x) = r_- \left[1 - E_\alpha(-A\,x^\alpha)\right],$$
where $E_\alpha(\cdot)$ is the Mittag-Leffler function.
- As $x \to \infty$,
$$E_\alpha(-A\,x^\alpha) = -\frac{1}{A}\frac{x^{-\alpha}}{\Gamma(1-\alpha)} + \mathcal{O}\left(|A\,x^\alpha|^{-2}\right).$$
- Thus, as $x \to \infty$,
$$h_\infty(a,x) - r_- = \frac{r_-}{A}\frac{x^{-\alpha}}{\Gamma(1-\alpha)} + \mathcal{O}\left(|A\,x^\alpha|^{-2}\right).$$

Large x expansion

- The form of the asymptotic solution motivates the following expansion of h for large x:

$$h(x;a) = r_- \sum_{k=0}^{\infty} \gamma_k \frac{x^{-k\alpha}}{A^k\,\Gamma(1-k\alpha)}.$$

- The coefficients γ_k satisfy the recursion

$$\begin{aligned}\gamma_1 &= -\gamma_0 = -1 \\ \gamma_k &= -\gamma_{k-1} + \frac{r_-}{2A} \sum_{i,j=1}^{\infty} \mathbb{1}_{\{i+j=k\}}\, \gamma_i\, \gamma_j \frac{\Gamma(1-k\alpha)}{\Gamma(1-i\alpha)\,\Gamma(1-j\alpha)}.\end{aligned}$$

Rational approximation

- Now we have small- and large-x expansions we can compute global rational approximations to $h(x;a)$ of the form

$$h^{(m,n)}(x;a) = \frac{\sum_{i=1}^{m} p_m y^m}{\sum_{j=0}^{n} q_n y^n}$$

with $y = x^\alpha$ that match these expansions up to order m and n respectively.

- Only the diagonal approximants $h^{(n,n)}$ are admissible approximations of h.

Adams and Padé smiles compared

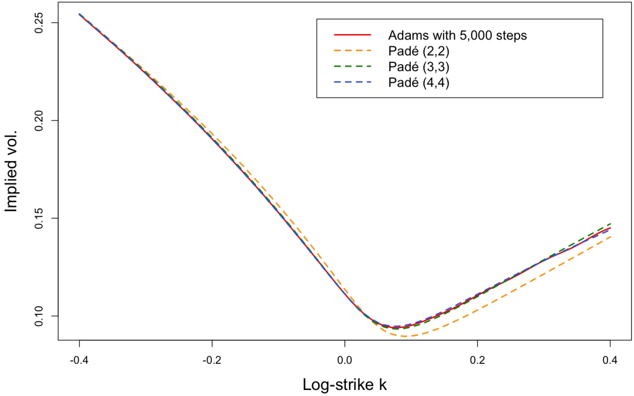

Figure 1: The red curve is the Adams smile with 5,000 steps. The Padé approximations seem to improve with increasing order.

Padé vs Adams smile errors

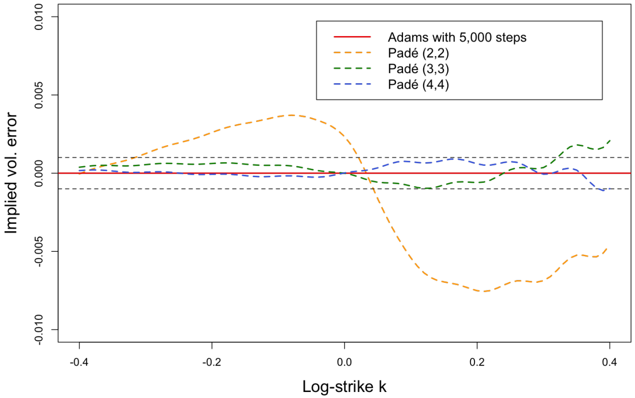

Figure 2: The red curve is the Adams smile with 5,000 steps. The Padé approximations seem to improve with increasing order.

Simulation

- The Adams scheme and the rational approximations apply only to models where the convolution Riccati equation can be recast as a fractional ODE.
- The simulation scheme prescribed in [Gat22] is applicable to any AFV model.
  - The scheme is inspired by [BLP17] and [And08].
- In the rough Heston case, we thus have two alternative ways to compute the volatility smile, so we can easily check convergence.

Discretization of the spot and variance processes

From the AFV dynamics,

$$d\xi_t(u) = \kappa(u-t)\,\sqrt{V_t}\,dW_t,$$

it follows that

$$\begin{aligned} V_T = \xi_T(T) &= \xi_0(T) + \int_0^T d\xi_s(T) \\ &= \xi_0(T) + \int_0^T \kappa(T-s)\,\sqrt{V_s}\,dW_s. \end{aligned} \tag{7}$$

- Wlog, let $t = 0$ and $\xi(u) = \xi_0(u)$. Let the time step $\Delta = T/N$ where N is the number of steps.
- As in [BLP17], we have the following exact decomposition of (7):

$$V_{n\Delta} = \xi(n\Delta) + \sum_{k=1}^{n} \int_{(k-1)\Delta}^{k\Delta} \kappa(n\Delta - s)\, \sqrt{V_s}\, dW_s.$$

Discretization of the v-process

- With simpler notation,

$$V_n = \xi_n + \sum_{k=1}^{n} \int_{(k-1)\Delta}^{k\Delta} \kappa(n\Delta - s) \sqrt{V_s}\, dW_s =: \hat{\xi}_n + u_n, \quad (8)$$

where the $\mathcal{F}_{n-1}$-adapted variable $\hat{\xi}_n$ is given by

$$\hat{\xi}_n = \mathbb{E}\left[V_n \middle| \mathcal{F}_{n-1} \right] = \xi_n + \sum_{k=1}^{n-1} \int_{(k-1)\Delta}^{k\Delta} \kappa(n\Delta - s) \sqrt{V_s}\, dW_s, \quad (9)$$

and the martingale increment u_n by

$$u_n = \int_{(n-1)\Delta}^{n\Delta} \kappa(n\Delta - s) \sqrt{V_s}\, dW_s. \quad (10)$$

The X-process

- We also need to simulate the nth increment of the component of the log-stock price process $X = \log S$ parallel to the volatility process[1],

$$\chi_n = \int_{(n-1)\Delta}^{n\Delta} \sqrt{V_s}\, dW_s. \tag{11}$$

- We then have the following discretization of the X process:

$$X_n = X_{n-1} - \frac{1}{4}\,(V_n + V_{n-1})\,\Delta + \sqrt{1-\rho^2}\,\sqrt{\bar{V}_n\,\Delta}\,Z_n^{\perp} + \rho\,\chi_n,$$

where $Z_n^{\perp}$ is standard normal, independent of χ_n and U_n.

[1] We write the increments as χ_n to emphasize that they should be approximately χ^2 distributed random variables.

Simulation step

- At each step, we need to generate (at least) three random variables: u_n, χ_n, and $\hat{\xi}_{n+1}$.

$$
\begin{aligned}
u_n &= \int_{(n-1)\Delta}^{n\Delta} \kappa(n\Delta - s)\sqrt{V_s}\, dW_s \\
\chi_n &= \int_{(n-1)\Delta}^{n\Delta} \sqrt{V_s}\, dW_s \\
\hat{\xi}_{n+1} &= \xi_{n+1} + \sum_{k=1}^{n} \int_{(k-1)\Delta}^{k\Delta} \kappa((n+1)\Delta - s)\sqrt{V_s}\, dW_s.
\end{aligned}
$$

The correlation matrix

- Because variances and covariances in an AFV model are linear in ξ, the correlation matrix takes the simple form.

$$R = \begin{pmatrix} 1 & \rho_{u\chi} & \rho_{u\xi} \\ \rho_{u\chi} & 1 & \rho_{\xi\chi} \\ \rho_{u\xi} & \rho_{\xi\chi} & 1 \end{pmatrix}. \tag{12}$$

  where the correlations $\rho_{u\chi}$, $\rho_{u\xi}$ and $\rho_{\xi\chi}$ are independent of n.
- In the case of the power-law kernel $\kappa(\tau) = \tilde{\eta}\, \tau^{\alpha-1}$, these correlations are functions of H only.

- In Figure 3, we plot these correlations as a function of H.

Plot of the correlation matrix in the power-law kernel case

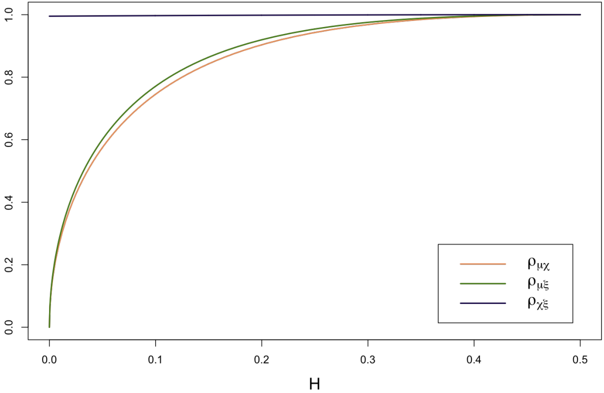

Figure 3: The correlations $\rho_{u\chi}$, $\rho_{u\xi}$, and $\rho_{\xi\chi}$ vs. H in the power-law kernel case.

A further approximation

- By assumption, the kernel behaves as a power-law kernel for Δ sufficiently small.
- Figure 3 suggests the approximations

$$\begin{aligned} \rho_{\xi\chi} &\approx 1 \\ \rho_{u\xi} &\approx \rho_{u\chi}. \end{aligned}$$

  - Moreover, when $H = 1/2$, $\rho_{u\chi} \approx 1$.
- Thus, when the model is Markovian ($H = 1/2$), we need only generate u_n at the nth time step; χ_n and $\hat{\xi}_{n+1}$ are perfectly correlated with u_n.
  - In practice, in the non-Markovian case ($H < 1/2$), we need only generate one other random variable.

Average values of the kernel

- Echoing the notation of [BLP17], let

$$b_j^{\star 2} = \frac{1}{\Delta} \int_0^{\Delta} \kappa(s + (j-1)\Delta)^2 \, ds. \tag{13}$$

- $b_j^{\star 2}$ thus gives the RMS average of the kernel at the jth lag.

The evolution of the forward variance curve

- The approximation

$$\int_{(k-1)\Delta}^{k\Delta} \kappa((n+1)\Delta - s)\,\sqrt{V_s}\,dW_s \quad\approx\quad b^{\star}_{n+1-k}\,\chi_k$$

gives

$$\hat{\xi}_{n+1} \quad\approx\quad \xi_{n+1} + \sum_{k=1}^{n} b^{\star}_{n+1-k}\,\chi_k.$$

- Similarly (though not needed for the algorithm), for $m > n$,

$$\mathbb{E}\left[\,V_m\middle|\,\mathcal{F}_n\right] \quad\approx\quad \xi_m + \sum_{k=1}^{n} b^{\star}_{m-k}\;\chi_k.$$

  - We see that the entire forward variance curve evolves according to the weighted historical path of the $X = \log S$ process.

The Andersen Quadratic Exponential (QE) scheme

- Naïve simulation of the V process leads to negative values
- Andersen's Quadratic Exponential (QE) scheme [And08] guarantees V positive in simulation of the classical Heston model.
  - Conditional means and variances are matched at each step.
- In [Gat22], we give a bivariate version of this scheme.

The hybrid QE (HQE) scheme

The HQE scheme

1. Given χ_k, for $k < n$, with ϵ very small, compute $\hat{\xi}_n = \max\left[\epsilon, \xi_n + \sum_{k=1}^{n-1} b^{\star}_{n-k+1}\,\chi_k\right]$.
2. Simulate χ_n and ε_n using the bivariate QE scheme
3. $V_n = \hat{\xi}_n + \frac{1}{\Delta}\mathcal{K}_0(\Delta)\,\chi_n + \varepsilon_n$.
4. Finally, $X_n = X_{n-1} - \frac{1}{4}\,(V_n + V_{n-1})\,\Delta + \sqrt{1-\rho^2}\,\sqrt{\bar{V}_n\,\Delta}\,Z_n^{\perp} + \rho\,\chi_n$.

Rough Heston parameters

- Consider the power-law kernel $\kappa(\tau) = \sqrt{2H}\,\eta\,\tau^{\alpha-1}$ with parameters roughly consistent with those found from calibration to SPX options on May 19, 2017 in [EGR19]:

$$\xi(u) = 0.025;\ H = 0.05;\ \eta = 0.8;\ \rho = -0.65. \tag{14}$$

- Note that the rough Heston kernel in [EGR19] takes the form

$$\kappa(\tau) = \nu\,\frac{\tau^{\alpha-1}}{\Gamma(\alpha)},$$

so ν in in [EGR19] and η are related as $\nu = \eta\,\sqrt{2H}\,\Gamma(\alpha)$.
  - $\eta = 0.8$ corresponds to $\nu \approx 0.4089$.

Richardson extrapolation

- It seems that the order of weak convergence of the fractional Adams scheme is one.
    - It therefore makes sense to use Richardson extrapolation to increase the order of convergence.

Definition 1 (Richardson extrapolation)

Let $\mathcal{S}_n$ denote an n-step approximation of the volatility smile according to some numerical scheme. Then the *n-step Richardson extrapolation* is given by

$$\mathcal{S}_n^R = \mathcal{S}_{2n} - \mathcal{S}_n.$$

- Adopting the $2,500$-step Richardson extrapolated Adams smile $\mathcal{S}_{2500}^R$ as our reference smile, we plot errors in the 200 step Adams and Padé approximated smiles in Figure 4.

Plots of smile and errors

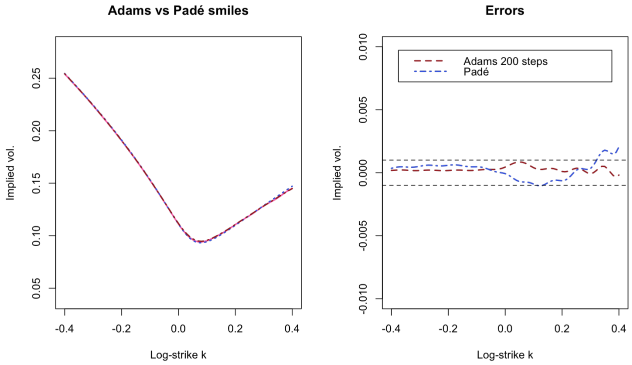

Figure 4: The 1-year rough Heston smile with parameters (14). The pink curve is the reference Adams smile S^R_{2500}. The blue and brown curves are from the Adams scheme with 200 steps and the Padé approximation respectively. The dashed horizontal lines indicate our target error band of $\pm 0.10\%$.

Convergence of the RSQE and HQE schemes

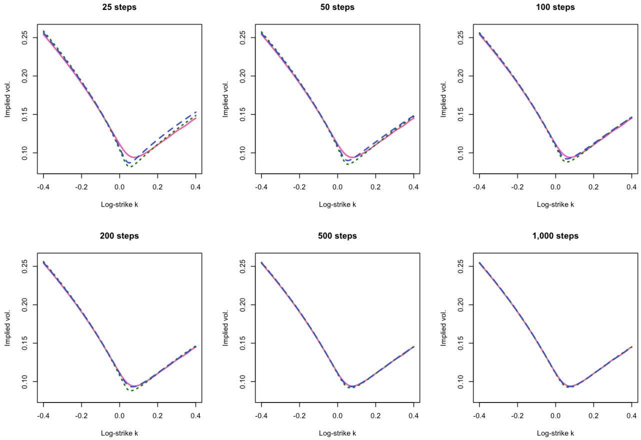

Figure 5: The 1-year rough Heston smile with parameters (14). The pink reference curve is the Adams reference smile $\mathcal{S}^R_{2500}$. The green-dotted and blue-dashed curves are from RSQE and HQE simulations with 10^6 paths.

Convergence of the HQE scheme

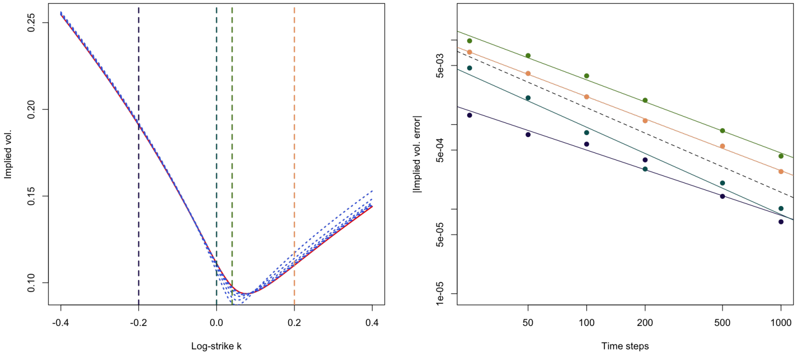

Figure 6: In the LH plot, the pink curve is the HQE smile $\mathcal{S}^R_{500}$. The blue dotted lines are HQE smiles S_n computed with $n \in \{25, 50, 100, 200, 500, 1000\}$. In the RH plot, we plot absolute implied volatility errors. All simulations are with 10^6 paths.

Convergence of Richardson extrapolated HQE smiles

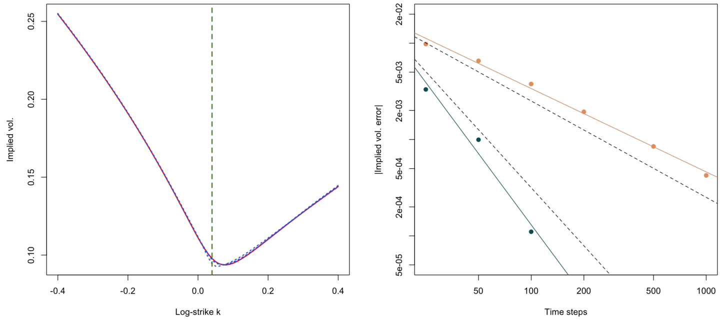

Figure 7: In the LH plot, the pink curve is the HQE smile $\mathcal{S}^R_{500}$. The blue dotted lines are the Richardson-extrapolated smiles S^R_n computed with $n \in \{25, 50, 100\}$. In the RH plot, we plot absolute implied volatility errors vs time steps for log-strike $k = 0.04$. We see evidence of order 2 weak convergence of Richardson-extrapolated smiles.

Summary

- In the case of the Mittag-Leffler (rough Heston) kernel, the convolution Riccati equation may be solved numerically using
  - the Adams scheme or
  - the rational approximation.
- Both of these methods can be used to compute the joint MGF of rough Heston.
  - The rational approximations are all much faster than the Adams scheme.
- The HQE scheme may be used to efficiently simulate AFV models for *any* choice of kernel.
- For path-dependent options such as barriers or lookbacks, simulation is the only choice.

References

Eduardo Abi Jaber, Martin Larsson, and Sergio Pulido
Affine Volterra processes
The Annals of Applied Probability, 29(5):3155–3200, 2019.

Elisa Alòs, Jim Gatheral, and Radoš Radoičić.
Exponentiation of conditional expectations under stochastic volatility.
Quantitative Finance, 20(1):13–27, 2020.

Leif Andersen.
Simple and efficient simulation of the Heston stochastic volatility model.
The Journal of Computational Finance, 11(3):1–42, 2008.

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

Mikkel Bennedsen, Asger Lunde, and Mikko S Pakkanen.
Hybrid scheme for Brownian semistationary processes.
Finance and Stochastics, 21(4): 931-965, 2017.

Kai Diethelm, Neville J. Ford, and Allan D. Freed
Detailed error analysis for a fractional Adams method.
Numerical Algorithms, 36(1) 31–52, 2004.

Omar El Euch and Mathieu Rosenbaum.
The characteristic function of rough Heston models.
Mathematical Finance, 29(1) 3–38, 2019.

Omar El Euch, Jim Gatheral, and Mathieu Rosenbaum.
Roughening Heston
Risk.net, May 2019.

Peter K Friz, Jim Gatheral and Radoš Radoičić.
Forests, cumulants, martingales.
The Annals of Probability, 50(4):1418–1445, 2022.

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

Jim Gatheral and Martin Keller-Ressel.
Affine forward variance models.
Finance and Stochastics, 23(3):501–533, 2019.

Jim Gatheral and Radoš Radoičić.
Rational approximation of the rough Heston solution.
International Journal of Theoretical and Applied Finance, 22(3):1950010–19, 2019.

Jim Gatheral
Efficient simulation of affine forward variance models.
risk.net, February 2022, Longer version available at https://www.ssrn.com/abstract=3876680.

Alan Lewis
Option Valuation under Stochastic Volatility with Mathematica Code.
Finance Press: Newport Beach, CA, 2000.