Pricing of American Options under the Rough Heston model

Elizabeth Zúñiga

Supervisors:
Etienne Chevalier
Sergio Pulido

12th European Summer School in Financial Mathematics

LaMME and Fondation de Mathématiques Jacques Hadamard

September 5, 2019

Classical models in finance assume that the log-asset price X_t is given by:

$$dX_t = \mu_t dt + \sigma_t dW_t$$

where μ_t is the drift, W_t is a BM and σ_t is the volatility of the asset.

Classical models in finance assume that the log-asset price X_t is given by:

$$dX_t = \mu_t dt + \sigma_t dW_t$$

where μ_t is the drift, W_t is a BM and σ_t is the volatility of the asset.

- Stochastic volatility models (Heston model): These models reproduce correctly the term structure of ATM skew

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

for large maturities, but they fail to explain the convexity to very short maturities

Classical models in finance assume that the log-asset price X_t is given by:

$$dX_t = \mu_t dt + \sigma_t dW_t$$

where μ_t is the drift, W_t is a BM and σ_t is the volatility of the asset.

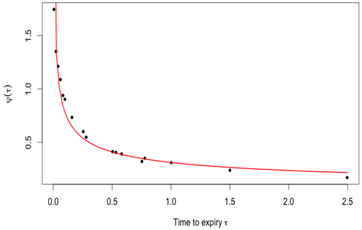
- Stochastic volatility models (Heston model): These models reproduce correctly the term structure of ATM skew

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

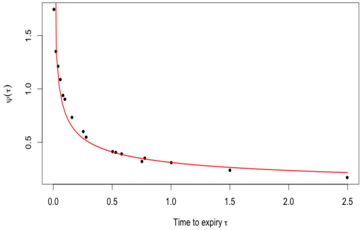
for large maturities, but they fail to explain the convexity to very short maturities

- Empirical studies indicate volatility is rougher than BM [Gatheral et al, 2014].

For small τ, in a model where the volatility is driven by a fractional Brownian motion with Hurst parameter H, $\psi(\tau) = \tau^{H-1/2}$ with $H = 0.1$. (see Figure of S&P ATM volatility skews)



For small τ, in a model where the volatility is driven by a fractional Brownian motion with Hurst parameter H, $\psi(\tau) = \tau^{H-1/2}$ with $H = 0.1$. (see Figure of S&P ATM volatility skews)



- New class of models: fractional stochastic volatility models (rough volatility models).
  - match roughness of time series data
  - fit implied volatility smiles remarkably well

Drawback: Loss of tractability, neither Markov nor semi-martingales.

The **Heston** model is a stochastic volatility model where:

$$dS_t = S_t \sqrt{V_t} dW_t$$

$$V_t = V_0 + \int_0^t \lambda(\theta - V_s) ds + \int_0^t \lambda \nu \sqrt{V_s} dB_s$$

λ, θ, V_0 and ν positive. W and B are two BM with correlation ρ.

The **Heston** model is a stochastic volatility model where:

$$dS_t = S_t \sqrt{V_t} dW_t$$

$$V_t = V_0 + \int_0^t \lambda(\theta - V_s)ds + \int_0^t \lambda\nu\sqrt{V_s} dB_s$$

λ, θ, V_0 and ν positive. W and B are two BM with correlation ρ.

Proposition (The characteristic function in the Heston model)

The characteristic function of the log-price $X_t = \log(S_t/S_0)$ satisfies:

$$\mathbb{E}[e^{iaX_t}] = \exp(g(a,t) + V_0 h(a,t)),$$

where h is the solution of the following Riccati equation

$$\partial_t h = \frac{1}{2}(-a^2 - ia) + \lambda(ia\rho\nu - 1)h(a,s) + \frac{(\lambda\nu)^2}{2} h^2(a,s), \quad h(a,0) = 0$$

and

$$g(a,t) = \theta\lambda \int_0^t h(a,s)ds$$

$$dS_t = S_t \sqrt{V_t} dW_t,$$

$$V_t = V_0 + \int_0^t \frac{(t-s)^{\alpha-1}}{\Gamma(\alpha)} (\lambda(\theta - V_s) ds + \lambda \nu \sqrt{V_s} dB_s)$$

where $\alpha \in (1/2, 1)$ governs the smoothness of the volatility.

$$dS_t = S_t \sqrt{V_t} dW_t,$$

$$V_t = V_0 + \int_0^t \frac{(t-s)^{\alpha-1}}{\Gamma(\alpha)} (\lambda(\theta - V_s) ds + \lambda\nu \sqrt{V_s} dB_s)$$

where $\alpha \in (1/2, 1)$ governs the smoothness of the volatility.

Proposition (Characteristic function of the rough Heston [El Euch and Rosenbaum (2016)])

The characteristic function of the log-price $X_t = \log(S_t/S_0)$ satisfies:

$$\mathbb{E}[e^{iaX_t}] = \exp(g_1(a,t) + V_0 g_2(a,t))$$

$$g_1(a,t) = \theta\lambda \int_0^t h(a,s) ds, \qquad g_2(a,t) = I^{1-\alpha} h(a,t)$$

where h is the solution of the following fractional Riccati equation:

$$D^\alpha h = \frac{1}{2}(-a^2 - ia) + \lambda(ia\rho\nu - 1)h(a,s) + \frac{(\lambda\nu)^2}{2} h^2(a,s), \quad I^{1-\alpha} h(a,0) = 0$$

$$dS_t = S_t\sqrt{V_t}dW_t,$$

$$V_t = V_0 + \int_0^t \frac{(t-s)^{\alpha-1}}{\Gamma(\alpha)}(\lambda(\theta - V_s)ds + \lambda\nu\sqrt{V_s}dB_s)$$

where $\alpha \in (1/2, 1)$ governs the smoothness of the volatility.

Proposition (Characteristic function of the rough Heston [El Euch and Rosenbaum (2016)])

The characteristic function of the log-price $X_t = \log(S_t/S_0)$ satisfies:

$$\mathbb{E}[e^{iaX_t}] = \exp(g_1(a,t) + V_0 g_2(a,t))$$

$$g_1(a,t) = \theta\lambda\int_0^t h(a,s)ds, \qquad g_2(a,t) = I^{1-\alpha}h(a,t)$$

where h is the solution of the following fractional Riccati equation:

$$D^{\alpha}h = \frac{1}{2}(-a^2 - ia) + \lambda(ia\rho\nu - 1)h(a,s) + \frac{(\lambda\nu)^2}{2}h^2(a,s), \quad I^{1-\alpha}h(a,0) = 0$$

Riemann-Liouville fractional integral and differential operators:

$$I^r f(t) = \frac{1}{\Gamma(r)}\int_0^t (t-s)^{r-1}f(s)ds, \quad D^r f(t) = \frac{1}{\Gamma(1-r)}\frac{d}{dt}\int_0^t (t-s)^{-r}f(s)ds$$

We call stochastic Volterra equations to the stochastic convolution equations of the form:

$$X_t = X_0 + \int_0^t K(t-s) b(X_s) ds + \int_0^t K(t-s) \sigma(X_s) dW_s, \tag{1}$$

We call stochastic Volterra equations to the stochastic convolution equations of the form:

$$X_t = X_0 + \int_0^t K(t-s)b(X_s)ds + \int_0^t K(t-s)\sigma(X_s)dW_s, \qquad (1)$$

- W is a multi-dimensional BM. $K \in L^2_{loc}(\mathbb{R}_+, \mathbb{R}^{d\times d})$ is called kernel.
- b and σ continuous and satisfy for some constant c_{LG} the linear growth conditions:

$$|b(x)| \vee |\sigma(x)| \leq c_{LG}(1+|x|), \quad x \in \mathbb{R}^d \qquad (2)$$

We call stochastic Volterra equations to the stochastic convolution equations of the form:

$$X_t = X_0 + \int_0^t K(t-s)b(X_s)ds + \int_0^t K(t-s)\sigma(X_s)dW_s, \qquad (1)$$

- W is a multi-dimensional BM. $K \in L^2_{loc}(\mathbb{R}_+, \mathbb{R}^{d\times d})$ is called kernel.
- b and σ continuous and satisfy for some constant c_{LG} the linear growth conditions:

$$|b(x)| \vee |\sigma(x)| \leq c_{LG}(1+|x|), \quad x \in \mathbb{R}^d \qquad (2)$$

We refer to the solutions of (1) as **affine Volterra processes** if $a(x) = \sigma(x)\sigma(x)^T$ and $b(x)$ are affine maps given by:

$$a(x) = A^0 + x_1 A^1 + \ldots x_d A^d \qquad b(x) = b^0 + x_1 b1 + \ldots x_d b^d$$

for some $d-$dimensional symmetric matrices A^i and vectors b^i.

Theorem (Bernstein–Widder theorem)

If K is completely monotone, there exists μ positive such that

$$K(t) = \int_0^{+\infty} e^{-t\gamma}\mu(d\gamma) \tag{3}$$

- In the fractional case with: $K(t) = \frac{t^{\alpha-1}}{\Gamma(\alpha)}$, $\mu(d\gamma) = \frac{\gamma^{-\alpha}}{\Gamma(\alpha)\Gamma(1-\alpha)}d\gamma$, $\alpha \in (1/2, 1)$.

Theorem (Bernstein–Widder theorem)

If K is completely monotone, there exists μ positive such that

$$K(t) = \int_0^{+\infty} e^{-t\gamma}\mu(d\gamma) \tag{3}$$

- In the fractional case with: $K(t) = \frac{t^{\alpha-1}}{\Gamma(\alpha)}$, $\mu(d\gamma) = \frac{\gamma^{-\alpha}}{\Gamma(\alpha)\Gamma(1-\alpha)}d\gamma$, $\alpha \in (1/2, 1)$.

We approximate $\mu(d\gamma)$ by $\sum_{i=1}^{n} c_i \delta_{\gamma_i}(d\gamma)$ [Carmona et al, 2000], then K by K^n

$$K^n(t) = \sum_{i=1}^{n} c_i e^{-\gamma_i t}.$$

For $n \in \mathbb{N}$, we define the following process:

$$X_t^n = X_0 + \int_0^t K^n(t-s)b(X_s^n)ds + \int_0^t K^n(t-s)\sigma(X_s^n)dW_s, \qquad t \in [0,T] \tag{4}$$

For $n \in \mathbb{N}$, we define the following process:

$$X_t^n = X_0 + \int_0^t K^n(t-s)b(X_s^n)ds + \int_0^t K^n(t-s)\sigma(X_s^n)dW_s, \qquad t \in [0,T] \tag{4}$$

For $n \in \mathbb{N}$ and $t \geq 0$, we have

$$X_t^n = X_0 + \sum_{i=1}^{n} c_i Y_t^{\gamma_i}$$

where for any $i \in \{1,...,n\}$, Y^{γ_i} is the affine process given by:

$$dY_t^{\gamma_i} = \left(-\gamma_i Y_t^{\gamma_i} + b_0 + b_1 X_0 + b_1 \sum_{j=1}^{n} c_j Y_t^{\gamma_j}\right) dt + \sigma \left(X_0 + \sum_{j=1}^{n} c_j Y_t^{\gamma_j}\right) dW_t \tag{5}$$

For $n \in \mathbb{N}$, we define the following process:

$$X_t^n = X_0 + \int_0^t K^n(t-s) b(X_s^n) ds + \int_0^t K^n(t-s)\sigma(X_s^n) dW_s, \qquad t \in [0,T] \tag{4}$$

For $n \in \mathbb{N}$ and $t \geq 0$, we have

$$X_t^n = X_0 + \sum_{i=1}^{n} c_i Y_t^{\gamma_i}$$

where for any $i \in \{1,...,n\}$, Y^{γ_i} is the affine process given by:

$$dY_t^{\gamma_i} = \left(-\gamma_i Y_t^{\gamma_i} + b_0 + b_1 X_0 + b_1 \sum_{j=1}^{n} c_j Y_t^{\gamma_j} \right) dt + \sigma \left(X_0 + \sum_{j=1}^{n} c_j Y_t^{\gamma_j} \right) dW_t \tag{5}$$

$Y = (Y^{(\gamma_i)})_{i=1}^n$ is affine, Markovian and can be simulated.

Proposition (Convergence in law of X^n to X: Abi Jaber and El Euch (2018))

Let $X^n = (X_t^n)_{t\, leT}$ continuous weak solutions to (4), with the kernel K^n such that there exist $\gamma > 0$ and $C > 0$:

$$\sup_{n\geq 1}\left(\int_0^{T-h}|K^n(u+h)-K^n(u)|^2 du + \int_0^h |K^n(u)|^2 du\right) \leq Ch^{2\gamma}, \tag{6}$$

and

$$\int_0^t |K(u)-K^n(u)|^2 du \to 0 \tag{7}$$

for any $t\in[0,T]$ as n goes to infinity. Then $(X^n)_{n\geq 1}$ is tight for the uniform topology and any point limit X is a solution of (1).

Proposition states the convergence in law of the family X^n to X whenever (1) admits a unique weak solution.

The adjusted forward process $(u_t)_{t\geq 0}$ and its factor-representation $(u_t^n)_{t\geq 0}$ are defined as follows:

$$u_t^n(x) = \mathbb{E}\left[X_{t+x}^n - \int_0^x K^n(x-s) b(X_{t+s}^n) ds \middle| \mathcal{F}_t \right]$$

The adjusted forward process $(u_t)_{t\geq 0}$ and its factor-representation $(u_t^n)_{t\geq 0}$ are defined as follows:

$$u_t^n(x) = \mathbb{E}\left[X_{t+x}^n - \int_0^x K^n(x-s)b(X_{t+s}^n)ds \middle| \mathcal{F}_t\right]$$

Lemma

Assume that K^n *satisfies* (6). *Let* $T > 0$ *and* $p > 2/\gamma$ *be such that* $sup_{n\geq 1, t\leq T}\mathbb{E}[|X_t^n|^p]$ *is finite. Then for every* $s, t \in [0,T]$ *and* $x, y \in [0,M]$

$$\mathbb{E}[|u_t^n(y) - u_s^n(x)|^p] \leq C(\max\,(|t-s|, |y-x|))^{p\gamma}$$

where $C = C(p, K, M, T)$. *As a consequence* u^n *admits a Hölder continuous modification of order* $\alpha \in [0, \gamma - p/2[$. *Denoting this version by* u^n, *we have*

$$\mathbb{E}\left[\left(\max_{(t,y)\neq(s,x)} \frac{|u_t^n(y) - u_s^n(x)|}{|(t-s, y-x)|^\alpha}\right)^p\right] < \infty$$

Proposition (Convergence of (u^n))

u^n *converges in law to* u *under the uniform topology when* n *goes to infinity.*

$$X_t = \ln(S_t) = \sqrt{V_t}dWt - \tfrac{1}{2}V_t dt,$$
$$u_t(x) = V_0 + \int_0^t K(t-s+x)\left(\lambda(\theta - Vs)ds + \eta\sqrt{V_s}dWs\right)$$

The conditional Fourier Laplace transform of the process $(\{u_t^n(.)\},(X_t^n)_{t\geq 0})$ is given by:

$$\mathbb{E}\left[exp\left(wX_T^n + \int_0^\infty h^n(x)u_T^n(x)dx\right)\Big|\mathcal{F}_t\right] = exp\left(wX_t^n + \phi^n(T-t,w) + \int_0^\infty \Psi^n(T-t,x,h^n,w)u_t^n(x)dx\right.$$

fractional Riccati equations:

$$\partial_t\phi^n(t,w) = \mathcal{R}_{\phi^n}\left(\psi^n(t,w)\right), \qquad \phi^n(0)=0$$

$$\Psi^n(t,x,h,w) = h^n(x-t)\mathbf{1}_{\{x\geq t\}} + \Im_{\Psi^n}\left(w,\psi^n(t-x,w)\right)\mathbf{1}_{\{x<t\}}, \quad \Psi^n(0,x)=h^n(x)$$

where

$$\mathcal{R}_{\phi^n}\left(\psi^n(t,w)\right) = \lambda\theta\psi^n(t,w), \quad \Im_{\Psi^n}(w,\psi^n(t)) = \frac{1}{2}(w^2-w) + \left(\rho\eta w - \lambda + \frac{\eta^2}{2}\psi^n(t,w)\right)\psi^n(t,w)$$

$$\psi^n(t) = \int_0^\infty h^n(x)K^n(t+x)dx + K * \Im_{\Psi^n}\left(\psi^n(.,w)\right)(t)$$

We denote the value of the American option at time i by U_i, $i = 0, \ldots .M$.

Theorem

We will suppose that $h^n \to h$ uniformly in n, $K^n \to K$ in $L^2_{loc}(\mathbb{R}+)$, (u^n, X^n) is tight and converges weakly to (u, X). Then,

$$U_0^n \to U_0$$

uniformly in n.

- Idea of the proof:
  - Approximate the American option with Bermuda (uniformly in n)
  - Show the convergence of the Bermuda: We prove the result by induction on the number of dates M of the Bermudan option. The convergence in law of the forward process is used alongside an argument of density (which comes from the Stone-Weierstrass theorem) that allows us to reduce the problem to payoffs of the form $exp(wX_T + \int h(x)u_T(x)dx)$.

Time grid: $t_k = k\Delta t$, $k = 1, \ldots, N$, $\Delta t = T/N$.

Time grid: $t_k = k\Delta t$, $k = 1, \ldots, N$, $\Delta t = T/N$.

Euler scheme for the log price $X_t = log(S_t)$

$$X_{t_{k+1}} = X_{t_k} + \left(r - \frac{V_{t_k}}{2}\right)\Delta t + \sqrt{(V_{t_k})^+ \Delta t}(\rho\mathcal{N}(0,1) + \sqrt{1-\rho^2}\tilde{\mathcal{N}}(0,1))$$

Time grid: $t_k = k\Delta t$, $k = 1, \ldots, N$, $\Delta t = T/N$.

Euler scheme for the log price $X_t = log(S_t)$

$$X_{t_{k+1}} = X_{t_k} + \left(r - \frac{V_{t_k}}{2}\right)\Delta t + \sqrt{(V_{t_k})^+ \Delta t}(\rho\mathcal{N}(0,1) + \sqrt{1-\rho^2}\tilde{\mathcal{N}}(0,1))$$

Explicit-implicit Euler scheme for the variance process V^n

$$V_{t_k} = u_0^n(t_k) + \sum_{i=1}^{n} c_i^n Y_{t_k}^i, \qquad Y_0^i = 0 \quad \text{for all} \quad i = 1, \ldots, n$$

$$Y_{t_{k+1}}^i = \frac{1}{1+\gamma_i \Delta t}\left(Y_{t_k}^i - \lambda V_{t_k}\Delta t + \eta\sqrt{(V_{t_k})^+ \Delta t}\right)\mathcal{N}(0,1), \quad i = 1, \ldots, n$$

with

$$u_0^n(t_k) = V_0 + \lambda\theta \sum_{i=1}^{n} c_i \left(\frac{1 - e^{-\gamma_i^n t_k}}{\gamma_i^n}\right)$$

Time grid: $t_k = k\Delta t$, $k = 1, \ldots, N$, $\Delta t = T/N$.

Euler scheme for the log price $X_t = log(S_t)$

$$X_{t_{k+1}} = X_{t_k} + \left(r - \frac{V_{t_k}}{2}\right)\Delta t + \sqrt{(V_{t_k})^+ \Delta t}(\rho \mathcal{N}(0,1) + \sqrt{1-\rho^2}\tilde{\mathcal{N}}(0,1))$$

Explicit-implicit Euler scheme for the variance process V^n

$$V_{t_k} = u_0^n(t_k) + \sum_{i=1}^{n} c_i^n Y_{t_k}^i, \qquad Y_0^i = 0 \quad \text{for all} \quad i = 1, \ldots, n$$

$$Y_{t_{k+1}}^i = \frac{1}{1+\gamma_i \Delta t}\left(Y_{t_k}^i - \lambda V_{t_k}\Delta t + \eta\sqrt{(V_{t_k})^+ \Delta t}\right)\mathcal{N}(0,1), \quad i = 1, \ldots, n$$

with

$$u_0^n(t_k) = V_0 + \lambda\theta \sum_{i=1}^{n} c_i \left(\frac{1 - e^{-\gamma_i^n t_k}}{\gamma_i^n}\right)$$

and parameters

$$c_i^n = \frac{(r_n^{1-\alpha} - 1) r_n^{(\alpha-1)(1+n/2)}}{\Gamma(\alpha)\Gamma(2-\alpha)} r_n^{(\alpha-1)i}, \qquad \gamma_i^n = \frac{1-\alpha}{2-\alpha}\frac{r_n^{2-\alpha} - 1}{r_n^{1-\alpha} - 1} r_n^{i-1-n/2}$$

Longstaff-Schwartz methodology.

$T = 1, \quad strike = 100, \quad \rho = -0.7, \quad \theta = 0.02, \quad \lambda = 0.3, \quad \nu = 0.3, \quad V_0 = 0.02 \quad \alpha = 0.6$

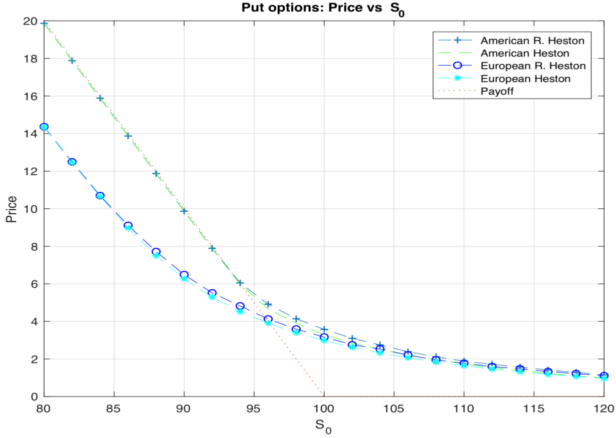
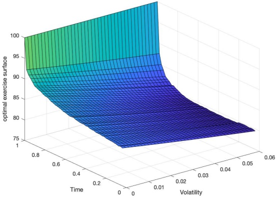
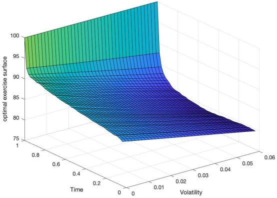


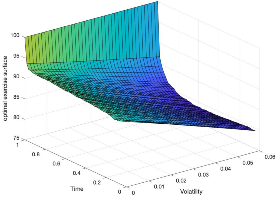
Figure: Put options with 10^5 simulations, 50 time step, 20 factors and $r_{20} = 2.5$



(a) $\alpha = 0.6$



(b) $\alpha = 0.8$



(c) $\alpha = 1$

FIGURE: Optimal exercise surface of American Put options in the Rough Heston

E.Abi Jaber and O. El Euch.
Markovian structure of the Volterra Heston model.
arXiv:1803.00477, 2018.

E. Abi Jaber, M. Larsson and S. Pulido
Affine Volterra processes.
The Annals of Applied Probability, 2019.

P. Carmona and L. Coutin and G. Montseny
Approximation of Some Gaussian Processes.
Statistical Inference for Stochastic Processes, volume: 3, number: 1, pages: 161-171, 2000.

P. Carr and Dilip B. Madan
Option Valuation Using the Fast Fourier Transform
Journal of Computational Finance, 1999. volume: 2, pages: 61-73.

G.Drimus
Options on realized variance by transform methods: A non-affine stochastic volatility model
September, 2009.

O. El Euch and M. Rosenbaum.
The characteristic function of rough Heston models.
Mathematical Finance, 2019.

D. Filipović and E. Mayerhofer E.
Affine Diffusion Processes: Theory and Applications.
arXiv:0901.4003v3. 2009.

Fukasawa and Masaaki
Asymptotic analysis for stochastic volatility: martingale expansion.
Finance and Stochastics, 2011. volume: 15, number: 4, pages: 635-654.

J. Gatheral, T. Jaisson and M. Rosenbaum
Volatility is rough.
Available at SSRN 2509457, 2014.

C. Li and C. Tao
On the fractional Adams method.
Computers and Mathematics with Applications. July, 2009.

D. Revuz and M.Yor.
Continuous martingales and Brownian motion
Springer-Verlag, 1991.