1 The Heston model

We consider the well known 1993 Heston stochastic volatility model for a stock price process S_t, defined by the following stochastic differential equations for a stock price process S_t:

$$\begin{cases} dS_t = rS_t dt \,+\, S_t\sqrt{V_t}dW_t^1 \,, \\ dV_t = \kappa(\theta - V_t)dt + \nu\sqrt{V_t}dW_t^2 \end{cases}$$

where W^1, W^2 are two correlated Brownian motions with $\mathbb{E}(W_t^1 W_t^2) = \rho t$, with $V_0 > 0$, $\kappa, \theta, \nu > 0, |\rho| \le 1$ and $2\kappa\theta > \nu^2$, which ensures that V cannot hit zero. The **no-arbitrage forward price** with maturity T is $F_t = S_t e^{r(T-t)}$ which (from Ito's lemma) satisfies $dF_t = F_t\sqrt{V_t}dW_t^2$, and $S_T = F_T$, so a no-arbitrage price for a call European option is $e^{-rT}\mathbb{E}((S_T - K)^+) = e^{-rT}\mathbb{E}((F_T - K)^+)$, so we can work with the driftless F_t instead of S_t without loss of generality.

- We cannot compute the density of $X_t = \log F_t$ exactly. However, there is a closed-form expression for the **characteristic function** $\phi(k) = \mathbb{E}(e^{ikX_t})$ of the form

$$\phi_t(k) \quad = \quad \mathbb{E}(e^{ikX_t}|X_0 = x, V_0 = v) \quad = \quad e^{ikx + vh(t) + g(t)} \tag{1}$$

  where g and h also depend on k, and we note that the exponent is **affine** in the v variable, i.e. linear plus a term which is independent of v.

- In general, if we know the characteristic function $\phi(k) := \mathbb{E}(e^{ikX})$ of a random variable X and $\phi \in L^1(\mathbb{R})$ i.e. $\int_{-\infty}^{\infty}|\phi(k)|dk < \infty$, we can compute the **density** of X using an **inverse Fourier transform**:

$$p(x) \quad = \quad \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-ikx}\phi(k)dk$$

  so we can use this general formula to compute the density of the log forward price X_t for any model for which its characteristic function $\phi_t(k) = \mathbb{E}(e^{ikX_t})$ is known (e.g. Heston and/or Levy models).

- A similar Fourier inversion formula can be used to price **European call options** at time zero:

$$C(K,T) \quad = \quad e^{-rT}\mathbb{E}((S_T - K)^+) \quad = \quad e^{-rT}(F_0 \;-\; \sqrt{F_0 K}\,\frac{1}{\pi}\int_0^{\infty}\frac{1}{k^2+\frac{1}{4}}\mathrm{Re}(e^{-ikx}\phi_T(k - \frac{1}{2}i))dk)$$

  where $x = \log\frac{K}{S_0}$, and Re(.) denotes the Real part. From this we can then compute the Black-Scholes **implied volatility** of a call option by solving $C(K,T) = C^{BS}(S_0, K, \sigma, T, r)$ for σ, using e.g. the **bisection method** or the **Newton-Raphson formula**, where $C^{BS}(S_0, K, \sigma, T, r)$ denotes the usual Black-Scholes call option formula

- To derive g and h in (1), we note that from the 2d version of the **Feynman-Kac formula**, $f(x,v,t) = \mathbb{E}(e^{ixX_T}|X_t = x, V_t = v)$ satisfies the PDE

$$f_t \;-\; \frac{1}{2}vf_x + \frac{1}{2}vf_{xx} + \kappa(\theta - v)f_v + \rho\nu v f_{xv} + \frac{1}{2}\nu^2 f_{vv} \quad = \quad 0$$

  with terminal condition $f(x,v,T) = e^{ikx}$. We can derive this by applying Ito's lemma to $f(X_t, V_t, t) = \mathbb{E}(e^{ixX_T}|X_t, V_t)$, and using that this process is a martingale, then letting $\tau = T - t$, so $u(x,v,\tau) = f(x,v,T-\tau)$ satisfies

$$u_\tau \quad = \quad -\frac{1}{2}vu_x + \frac{1}{2}vu_{xx} + \kappa(\theta - v)u_v + \rho\nu v u_{xv} + \frac{1}{2}\nu^2 u_{vv} \tag{2}$$

  with initial condition $u(x,v,0) = e^{ikx}$. With mild abuse of notation, we now change the definition of t to $t = \tau$ so $u_\tau = u_t$, so as to avoid using τ going forward.

- Now guessing the form of u as in (1), plugging into (2) (with $u_\tau = u_t$) and then equating coefficients in v and terms that do not contain v, we find that h and g must satisfy

$$h'(t) \quad = \quad \frac{1}{2}(-ik - k^2) \,+\, (\rho\nu ik - \lambda)h(t) \,+\, \frac{1}{2}\nu^2 h(t)^2 \quad , \quad g'(t) \quad = \quad \kappa\theta h(t) \tag{3}$$

  with $h(0) = 0$ and $g(0) = 0$ (Mathematica is very useful for doing these type of computations).

- We can extend the Heston model to the **Rough Heston** model for which

$$V_t = V_0 + \frac{1}{\Gamma(\alpha)}\int_0^t (t-s)^{\alpha-1}\kappa(\theta - V_s)ds + \frac{1}{\Gamma(\alpha)}\int_0^t (t-s)^{\alpha-1}\nu\sqrt{V_s}dW_s^2 \tag{4}$$

  for $\alpha = H + \frac{1}{2}$ with $H \in (0, \frac{1}{2}]$, where Γ is the Gamma function. The Rough Heston model reduces to the standard Heston model when $H = \frac{1}{2}$. H controls the **roughness** of the sample path of V, i.e. V_t is rougher than the standard Heston model when H is smaller. H is usually $\in (0, 0.1]$ in practice (see Figure 4 below for simulations).

- For the rough Heston model, we have to modify (1) to

$$\mathbb{E}(e^{ikX_t}) = e^{ikX_0 + V_0(I^{1-\alpha}h)(t) + (I^1 h)(t)} \tag{5}$$

  where $(I^r f)(t) = \frac{1}{\Gamma(r)}\int_0^t (t-s)^{r-1}f(s)ds$ denotes the rth **fractional integral** of a general function f, and now h satisfies the **Volterra Integral equation**:

$$\begin{aligned} h(t) &= I^{\alpha}(\frac{1}{2}(-ik - k^2) + (\rho\nu ik - \lambda)h + \frac{1}{2}\nu^2 h^2)(t) \\ &= \frac{1}{\Gamma(\alpha)}\int_0^t (t-s)^{\alpha-1}(\frac{1}{2}(-ik-k^2) + (\rho\nu ik - \lambda)h(s) + \frac{1}{2}\nu^2 h(s)^2)ds \end{aligned} \tag{6}$$

  which implies that $h(0) = 0$. Note that I^1 in (5) is just the usual integration operator, since $r - 1 = 0$ and $\Gamma(1) = 1$.

- We can simulate V for the rough Heston model with an **Euler**-type scheme as follows:

$$V_{(j+1)\Delta t} = V_0 + \frac{1}{\Gamma(\alpha)}\sum_{k=0}^{j-1}((j-k)\Delta t)^{H-\frac{1}{2}}(\kappa(\theta - V_{k\Delta})\Delta t + \sqrt{V_{k\Delta}}\Delta W_k) \tag{7}$$

  where ΔW_j is a sequence of i.i.d. $N(0, \Delta t)$ random variables, and we can use a similar double loop construction to solve (6) numerically (which does not require any random numbers). We can easily implement (7) in Matlab as:

```
for j=1:N
    t=j*dt;
    V(j)=V0;
    for k=1:j-1
        s=k*dt;
        V(j)=max(V(j)+c1*(t-s)^(H-.5)*(kappa*(theta-V(k))*dt+sqrt(V(k))*dW(k),0);
  end
end
```

  where $c_1 = \frac{1}{\Gamma(\alpha)}$, or similarly in **Python**, and the max is used to ensure that V stays non-negative. Note the code requires an additional **inner loop** which the usual Euler scheme for the $H = \frac{1}{2}$ case does not need, so the Monte Carlo for rough models like is slower, i.e. $O(N^2)$ not $O(N)$, which is an important issue in practice.

- If you write a function `f` in Matlab which takes a vector of model `parameters` as its argument which outputs the sum of squared errors between model and market option prices: $\sum_{i=1}^m \sum_{k=1}^n (\mathbb{E}((S_{T_i} - K_j)^+) - c_{ij})^2$ where c_{ij} is the market price of a European call option with strike K_j and maturity T_i using e.g. Monte Carlo, then one can use:

```
options = optimset('Display','iter','PlotFcns',@optimplotfval);
fminsearch(@f,parameters,options)
```

  to try and **calibrate** the vector of model parameters, where `parameters` is an initial guess for this vector. One can make this more interesting by also adding VIX options into the calibration.

  In **python**, one would use

```
res=scipy.optimize.minimize(error, x0=..., method='L-BFGS-B', tol=1E-5, options="maxiter":200)
```

  with an error function that looks something like

```
def error(x):
  yhat = model.predict(x)
  error = np.sum(np.power(yhat - ytarget, 2))
return error
```

```
IV = np.array([[0.3231,0.2987, ...],[0.2996,0.2935, ...]]
ytarget = np.array(list(IV[0])+list(IV[1])+list(IV[2])+list(IV[3])+list(IV[4])+list(IV[5])+list(IV[
```

to minimize the least squares error between the array of real implied vol values above, and those predicted by the NN

- We return now to the standard Heston model. The first eq in (3) is a **Riccati** ODE which is **non-linear** due to the h^2 term. If we now let $h(t) = -\frac{2}{\nu^2}\frac{f'(t)}{f(t)}$, we find that f satisfies the linear ODE:

$$f''(t) + (\kappa - ik\rho\nu)f'(t) + \frac{1}{4}(-k^2 - ik)\nu^2 f = 0 \tag{8}$$

with $f'(0) = 0$, and we can arbitrarily set $f(0) = 1$ since multiplying f by a constant will not affect $h(t)$. This linear ODE can be solved explicitly using the usual separation of variables method, with an ansatz of the form $f = e^{mt}$ for which we find that m satisfies the quadratic

$$m^2 + (\kappa - ik\rho\nu)m + \frac{1}{4}(-k^2 - ik)\nu^2 = 0$$

which has two roots $m = m_\pm$ in general. The general solution to (8) is then obtained as a linear combination of these solutions of the form $f = c_+e^{m_+t} + c_-e^{m_-t}$ and we choose the constants to satisfy the boundary conditions $f'(0) = 0, f(0) = 1$ from above as

$$c_+ + c_- = 1 \quad , \quad m_+c_+ + m_-c_- = 0\,.$$

- Now we have the solution for $f(t)$ we can get back to $h(t)$, and the second equation in (3) can then be integrate directly to get $g(t) = \kappa\theta\int_0^t h(s)ds$.
- ρ controls the **skewness** of the distribution of $X_t - \mathbb{E}(X_t)$, and the **asymmetry** of the **implied volatility smile** as a function of the **log-moneyness** $x = \log\frac{K}{S_0}$.
- If $\kappa > 0$, $V_\infty := \lim_{t\to\infty} V_t$ has a well defined density given by a **Gamma distribution**:

$$p_\infty(v) = \frac{\beta}{\Gamma(\alpha)}v^{\alpha-1}e^{-\beta v}$$

for $v > 0$, where here $\alpha = 2\kappa\theta/\nu^2$ and $\beta = 2\kappa/\nu^2$, and note this density is independent of V_0. We can use this to price VIX options in the limit as the maturity $T \to \infty$. p_∞ is the solution of

$$-\frac{\partial}{\partial y}(\kappa(\theta - y)p(y)) + \frac{\partial^2}{\partial y^2}(\frac{1}{2}\nu yp(y)) = 0\,.$$

- When the correlation $\rho = 0$, the implied volatility is symmetric in $x = \log\frac{K}{S_0}$.
- The density $p_t(v)$ of V_t at v for $0 < t < T$ is given by the **non-central chi square distribution**:

$$p_t(v) = ce^{-c(v_0e^{-\kappa t}+v)}(\frac{ve^{\kappa t}}{v_0})^{\frac{1}{2}q}\,\mathrm{BesselI}(q, 2c\sqrt{v_0e^{-\kappa t}v}) \tag{9}$$

where $q = 2\kappa\theta/\nu^2 - 1$ and $c = c(t) = 2\kappa/(\nu^2(1 - e^{-\kappa t}))$, and BesselI$(n, .)$ denotes the **modified Bessel function** of order n (for which there is an inbuilt function in Matlab), and $\mathbb{E}(f(V_T)) = \int_0^\infty f(v)p_T(v)dv$. This allows us to price **VIX** options with a single numerical integration without using Monte Carlo using an f function of the form $f(v) = \sqrt{av + b}$, see below. When $\kappa = 0$, the density simplifies to (12) below, but in this case the density **integrates to less than** 1 because there is non-zero probability of **absorption** at $V = 0$, and hence $\mathbb{P}(V_T = 0) > 0$ and $\mathbb{P}(V_T = 0) + \int_0^\infty p_T(v)dv = 1$. When V is close to zero, note that V essentially behaves like

$$dV_t = \kappa\theta dt$$

so if $\kappa\theta > 0$, this has the effect of pushing V back away from zero.

- We can simulate V in the usual way with an **Euler scheme**, or use a higher order **Milstein scheme**. More specifically, for a general one-dimensional SDE of the form

$$dX_t = \mu(X_t)dt + \sigma(X_t)dW_t$$

the higher order Milstein scheme for simulating X is as follows:

$$X_{t+\Delta t} = X_t + \mu(X_t)\Delta W_t + \sigma(X_t)\Delta W_t + \frac{1}{2}\sigma'(X_t)\sigma(X_t)((\Delta W_t)^2 - \Delta t)$$

where $\Delta W_t = W_{t+\Delta} - W_t$ (see end of document for proof). The first three terms on the right hand side are the usual **Euler scheme**, and the final term is the Milstein term.

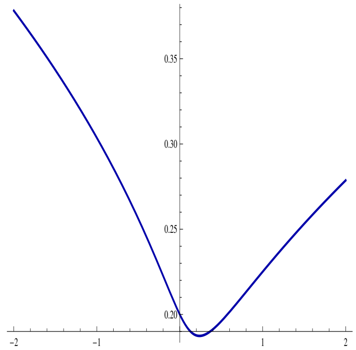

Figure 1: Here we have plotted the implied volatility $I(x)$ for a very small maturity T as function of the log moneyness $x = \log \frac{K}{S_0}$ for $V_0 = .04, \nu = .2, \rho = -.4$. Note that $I'(0) < 0$ if $\rho < 0$, and ρ is usually negative in practive

.

1.1 Pricing VIX options under the Heston model

The theoretical value VIX_t of the VIX index at time t is defined as

$$\mathrm{VIX}_t^2 \;=\; \frac{1}{\Delta}\mathbb{E}\Big(\int_t^{t+\Delta} V_u du|\mathcal{F}_t\Big) \;=\; \frac{1}{\Delta}\int_t^{t+\Delta}\mathbb{E}(V_u|\mathcal{F}_t)du \;=\; \frac{1}{\Delta}\int_t^{t+\Delta}\mathbb{E}(V_u|V_t)du \tag{10}$$

where the final equality follows because V is a **Markov process** (note this is not true for the **rough Heston model**). A VIX call option is a European call option on VIX_T for some maturity T, so the payoff of a VIX call option is $\max(\mathrm{VIX}_T - K, 0)$ at time T. Integrating the Heston SDE above we see that

$$\begin{aligned} V_t &= V_0 + \int_0^t \kappa(\theta - V_u)du + \int_0^t \nu\sqrt{V_u}dW_u^2 \\ \Rightarrow \quad \mathbb{E}(V_t) &= V_0 + \mathbb{E}\Big(\int_0^t \kappa(\theta - V_u)du\Big) \;=\; V_0 + \int_0^t \kappa(\theta - \mathbb{E}(V_u))du \end{aligned}$$

since the second term in the previous equation is a stochastic integral and thus has zero expectation. Differentiating we see that $f(t) = \mathbb{E}(V_t)$ satisfies the ordinary differential equation:

$$f'(t) \;=\; \kappa(\theta - f(t))$$

with initial condition $f(0) = V_0$, which has solution $f(t) = \theta + e^{-\kappa t}(V_0 - \theta)$. For (10), we need to be able to compute $\mathbb{E}(V_u|V_t)$. But since $\mathbb{E}(V_u|V_t = v) = \mathbb{E}(V_{u-t}|V_0 = v)$, we see that $\mathbb{E}(V_u|V_t) = \theta + e^{-\kappa(u-t)}(V_t - \theta)$ for $u \geq t$ i.e. we just replace t with $u - t$ and V_0 with V_t in $f(t)$, so setting $t = T$ in (10) we see that

$$\mathrm{VIX}_T^2 \;=\; \frac{1}{\Delta}\int_T^{T+\Delta}(\theta + e^{-\kappa(u-T)}(V_T - \theta))du\,.$$

We can compute the integral here explicitly since V_T does not depend on u, and we obtain

$$\mathrm{VIX}_T^2 \;=\; \frac{1 - e^{-\kappa\Delta}}{\kappa\Delta}V_T + \frac{\theta}{\kappa\Delta}(e^{-\kappa\Delta} + \kappa\Delta - 1) \;=\; F(V_T) \;=\; aV_T + b$$

for $\kappa > 0$.

Note this is just $F(V_T)$ for some affine function $F(v) = av + b$ for some a, b, so we can easily now estimate $\mathbb{E}(\max(\mathrm{VIX}_T - K, 0))$ (i.e. the VIX call price) using Monte Carlo as

$$\mathbb{E}(\max(\sqrt{aV_T + b} - K, 0)) \;=\; \int_0^\infty \sqrt{av + b}\; p(v_0, v, T)dv \tag{11}$$

where $p(.)$ is the density of V defined in (9), and one can estimate the integral on the right using e.g. **Gaussian quadrature** as a finite sum of the form $\sum_{i=1}^n w_i f(v_i)$ where $f(v) = \sqrt{av + b}\; p(v_0, v, T)$, and Matlab can compute the optimal weights w_i and v-values v_i. You do not need to simulate S itself to price this VIX option so we only need simulate the Brownian motion W^2 which drives V. One can also price the **VIX future** which pays VIX_T at time T, i.e. just set the strike $K = 0$.

Remark 1.1 Note that $b \geq 0$, so $\mathrm{VIX}_T^2 \geq b$ almost surely, which means that for $K \leq \sqrt{b}$, $\mathbb{E}(\max(\mathrm{VIX}_T - K, 0)) = \mathbb{E}(\mathrm{VIX}_T) - K$, i.e. the option is just worth its **intrinsic value** and hence has zero implied volatility. So for $b > 0$, the VIX implied volatility tends to zero as $K \to \sqrt{b}$ (see Figure 2).

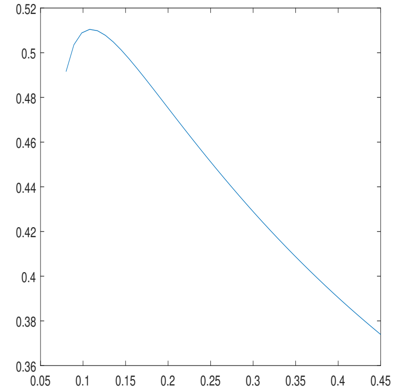

Figure 2: VIX smile for the parameters in the project. In this case $b = .0036$.

If $\kappa = 0$ and we impose that V is absorbed at zero, then the density of V_t is given explicitly by

$$p_t(v) \quad = \quad 2\sqrt{v_0}e^{-\frac{2(v+v_0)}{\nu^2 t}} \frac{\text{BesselI}(1, \frac{4\sqrt{vv_0}}{t\nu^2)})}{t\nu^2\sqrt{v}} \tag{12}$$

where $v_0 = V_0$ and $\text{BesselI}(n, .)$ denotes the **modified Bessel function** of order n (for which there is an inbuilt function in Matlab), and in this case $\text{VIX}_T^2 = V_T$ since V is a martingale so $\mathbb{E}(V_u|V_T) = V_T$. Note this density does not quite integrate to 1 since $\mathbb{P}(V_t = 0) > 0$ for $t > 0$, i.e. there is a **non-zero probability of absorption**. As the maturity $T \to 0$ for a VIX call option, the VIX implied volatility tends to

$$\hat{\sigma}(x) \quad = \quad \frac{\frac{1}{2}\nu x}{\sqrt{V_0}\,(e^x - 1)} \quad = \quad \frac{\nu}{\sqrt{V_0}}(\frac{1}{2} - \frac{1}{4}x + \frac{1}{24}x^2 + O(x^3)) \tag{13}$$

where $K = \sqrt{V_0}e^x$ is the strike of the VIX option, where the expression on the right is a **small log-moneyness expansion** (see plots below). Since the $O(x)$ term in this expansion is negative, we say that the smile has **negative skew**.

The F function above for the case $\kappa > 0$ remain unchanged if we generalize the model to the CEV-p model:

$$dV_t \quad = \quad \kappa(\theta - V_t)dt + \sigma V_t^p dW_u^2$$

for any $p \in [0, 1]$, except now we obviously have to simulate a different V process, and for the $\kappa = 0$ case we can extend the small-time analysis above.

One can then also compute the **VIX implied volatility** for VIX options by replacing S_0 by $\mathbb{E}(\text{VIX}_T)$ in the Black-Scholes formula, and then inverting to find the implied volatility in the usual way.

For the Rough Heston model

$$\begin{aligned} \text{VIX}_T^2 \quad &= \quad \frac{1}{\Delta}\int_T^{T+\Delta} \mathbb{E}(V_u|\mathcal{F}_T^W)du \\ &= \quad \frac{1}{\Delta}\int_T^{T+\Delta} (V_0 + \int_0^T \frac{\nu}{\Gamma(\alpha)}(u-s)^{H-\frac{1}{2}}\sqrt{V_s}dW_s^2)du \\ &= \quad V_0 + c_1 \int_0^T ((T+\Delta-s)^{\frac{1}{2}+H} - (T-s)^{\frac{1}{2}+H})\sqrt{V_s}dW_s^2 \end{aligned}$$

where $c_1 = \frac{\nu}{\Delta\Gamma(\alpha)(\frac{1}{2}+H)}$ and we have interchanged the order of integration, so we cane easily simulate VIX_T^2 if we have already simulated V using the Matlab code above.

2 Calibrating local volatility models to a finite number of European options

Consider a local volatility model for a stock price process X_t of the form

$$dX_t \quad = \quad \sigma(X_t, t)dW_t$$

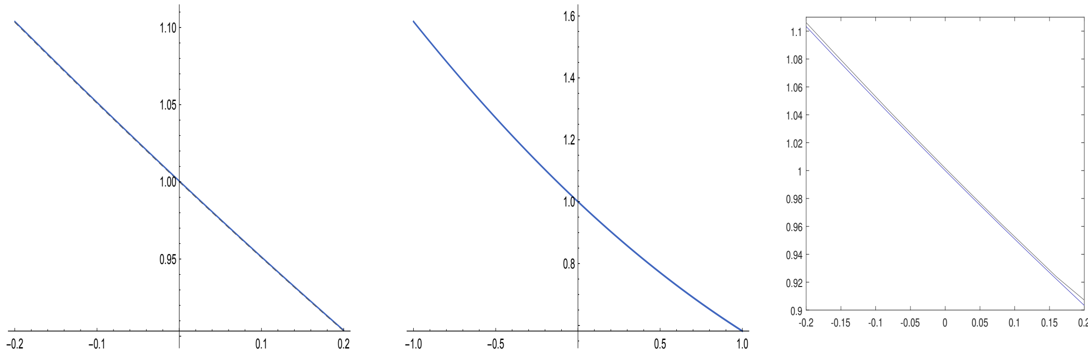

Figure 3: On the left we have plotted the asymptotic VIX implied volatility smile $\hat{\sigma}(x)$ in the $T \to 0$ limit as a function of x where $K = \sqrt{V_0}e^x$ (blue) verses the VIX implied volatility computed via numerical integration (grey dashed) over the density of V_T in (12) for $T = .001$, with $\kappa = 0$, $\nu = .4$ and $V_0 = .04$, and we see that both curves are almost indistinguishable over this range of x-values. In the second panel, we have re-plotted the $\hat{\sigma}(x)$ over a wider range of x-values. The final panel again plots $\hat{\sigma}(x)$ (blue) verses the values obtained from Monte Carlo (grey) in Matlab with $T = .004$, 5million simulations, 1000 time steps and using **antithetic** variables.

and assume interest rates are zero for simplicity. The price $u(x,t) = \mathbb{E}((X_T - K)^+ | X_t = x)$ of a European option under this model satisfies a similar equation to the Black-Scholes PDE:

$$u_t + \frac{1}{2}a(x,t)u_{xx} = 0$$

(also known as the **backwards Kolmogorov equation**) subject to $u(x,T) = (x-K)^+$, where $a(x,t) = \sigma(x,t)^2$. Similarly, if we let $v(x,t) = \sum_i w_i \mathbb{E}((X_T - K_i)^+) + \alpha \int_0^T (a(X_t,t) - \bar{\sigma})^2 dt$, then v satisfies

$$u_t + \frac{1}{2}a(x,t)u_{xx} + \alpha(a(x,t) - \bar{a})^2 = 0$$

subject to $u(x,T) = \sum_{i=1}^n w_i (x - K_i)^+$. For a more general model

$$dX_t = \sigma_t dW_t$$

where σ_t is any $\mathcal{F}_t^W$-adapted process for which $\mathbb{E}(\int_0^T \sigma_t^2 dt < \infty)$, we can ask how do we **optimally choose** σ_t so as to minimize $\sum_i w_i \mathbb{E}((X_T - K_i)^+) + \alpha \int_0^T (a_t - \bar{a})^2 dt$, where $a_t = \sigma_t^2$, $\bar{a} = \bar{\sigma}^2$ and $\bar{\sigma}$ is some reference fixed volatility level. Note this **penalizes** the model for being "too far" from a constant vol model with volatility $\bar{\sigma}$. The solution to this problem is then obtained via the solution to the **Hamilton-Jacobi-Bellman** PDE:

$$u_t + \min_{a \ge 0}(\frac{1}{2}a u_{xx} + \alpha(a - \bar{a})^2) = 0$$

again subject to $u(x,T) = \sum_{i=1}^n w_i (x - K_i)^+$ (note a is not allowed to be negative since $\sigma(X_t) = \sqrt{a(X_t,t)}$, hence the minimization over $a \ge 0$ only). When we compute the minimal a-value here by differentiating with respect to α and setting the answer to zero, we find that

$$a^* = \max(\bar{a} - \frac{u_{xx}(x,t)}{4\alpha}, 0) \tag{14}$$

and the PDE becomes

$$u_t + \frac{1}{2}\bar{a}u_{xx} - 1_{\bar{a} - \frac{u_{xx}}{4\alpha} > 0} \frac{u_x^2}{16\alpha} = 0 \tag{15}$$

which is **non-linear** due to the u_x^2 term, then the optimal $a(x,t)$ is given by (14). We can then try to solve this PDE **numerically** using an **explicit finite difference scheme** as in FM06, and since the scheme is explicit we must choose $\Delta t \le const. \times (\Delta x)^2$ for some constant $const.$. We can also solve this minimization problem using deep learning with neural networks in python (without using the HJB eq). If c_i denotes the market price of the K_i-strike call option, then if we then compute

$$\max_{w_i}(-\sum_{i=1}^n w_i c_i + u^w(X_0, 0))$$

where u satisfies (15), and note that (in general) we get a different solution to (15) for each choice of the vector $w = (w_1, .., w_n)$, so we have made the dependence of u on w explicit here. So we have an optimal w^*

and an optimal $a^*(x,t) = \bar{a} - \frac{u^{w^*}_{xx}(x,t)}{4\alpha}$. For this choice of a^*, the model is then consistent with the n option prices, i.e. $\mathbb{E}((X_T - K_i)^+) = c_i$ for $i = 1..n$, as long as such a model exists. To see why this is true, note that we can re-write as a max-min problem:

$$
\begin{aligned}
& \max_{w_i} \min_{a\in\mathcal{A}}(-\sum_{i=1}^n w_i c_i + \sum_{i=1}^n w_i \mathbb{E}((X_T - K_i)^+) + \alpha \int_0^T (a_t - \bar{a})^2 dt) \\
= \quad & \min_{a\in\mathcal{A}} \max_{w_i}(-\sum_{i=1}^n w_i c_i + \sum_{i=1}^n w_i \mathbb{E}((X_T - K_i)^+) + \alpha \mathbb{E}(\int_0^T (a_t - \bar{a})^2 dt) \\
= \quad & \min_{a\in\mathcal{A}\,:\,\mathbb{E}((X_T-K_i)^+=c_i} \alpha \mathbb{E}(\int_0^T (a_t - \bar{a})^2 dt)
\end{aligned}
$$

since the inner max in the middle equation is infinite if $\mathbb{E}((X_T - K_i)^+ \neq c_i$ for some i.

```
function f = FDSchemeExplicitGuoLoeperLocalVol(w)
T=1;
K=0.1; alpha=1; M=50000; N=100;
xmin=-3;xmax=3;abar=1;
u=zeros(M+1,N+1);ux=zeros(M+1,N+1);uxx=zeros(M+1,N+1);astar=zeros(M+1,N+1);X=linspace(xmin,xmax,N+1);
K=[-.1 0 .1];c=[.45 .4 .36];
u(1,:)=w(1)*max(X-K(1),0)+w(2)*max(X-K(2),0)+w(3)*max(X-K(3),0);
dt=T/M;
dx=(xmax-xmin)/N;
for i=1:M+1
   for j=2:N
      ux(i,j)=(u(i,j+1)-u(i,j-1))/(2*dx);
      uxx(i,j)=(u(i,j+1)-2*u(i,j)+u(i,j-1))/dx^2;
      astar(i,j)=abar-uxx(i,j)/(4*alpha);
   end
   for j=2:N
      u(i+1,j)=u(i,j)+.5*abar*uxx(i,j)*dt-heaviside(astar(i,j))*ux(i,j)^2/(16*alpha)*dt;
   end
   u(i,N+1)=u(i,N)+u(i,N)-u(i,N-1);
   u(i,1)=u(i,2)-(u(i,3)-u(i,2));
end
x0=0; f=-(u(end,fix((x0-xmin)/(xmax-xmin)*N)+1)-dot(c,w))
```

3 Lévy jump models

X is said to be a Lévy process if:

- X has independent increments;
- $X_t - X_s$ has the same distribution as X_{t-s} for any $0 < s < t$;
- X can only jump at random times.

Remark 3.1 Examples of Lévy processes: standard Brownian motion W_t (no jumps), the Poisson process N_t (pure jump process), the sum $X_t = W_t + N_t$, a jump diffusion process: $X_t = \mu t + \sigma W_t + \sum_{i=1}^{N_t} \xi_i$ where ξ_i is an i.i.d. sequence of random variables with some arbitrary distribution μ (this is a Poisson-type process if μ has a discrete distribution) and N_t is a Poisson process, with ξ_i, N and W all independent of each other, $(\tau_b)_{b\geq 0}$ (the first hitting time process for standard Brownian motion) where now b is the time variable.

A Lévy process has an associated **Lévy measure** $\nu(dx)$ which is such that for any n **disjoint** sets $A_1, A_1, ..., A_n$, the number of jumps of X whose size falls $A_1, ..., A_n$ over the interval $[0, t]$ is a vector of n independent Poisson random variables with parameters $\nu(A_1), ..., \nu(A_n)$, and recall that $\nu(A_i) = \int_{A_i} \nu(dz)$, which is $\int_{A_i} \nu(z)dz$ if $\nu(dz) = \nu(z)dz$.

When X is just a jump diffusion then $\nu(dx) = \lambda\mu(x)$, where $\mu(x)$ is the jump size distribution and λ is the **arrival rate** for the Poisson process. For **infinite-activity** Lévy processes, $\nu(A) = \int_A \nu(dx) = \infty$ if the set A includes zero, but finite otherwise (e.g. the CGMY process below), so the number of jumps of size $\leq x$ for any $x > 0$ is infinite, so we say the process has an infinite amount of "small" jumps.

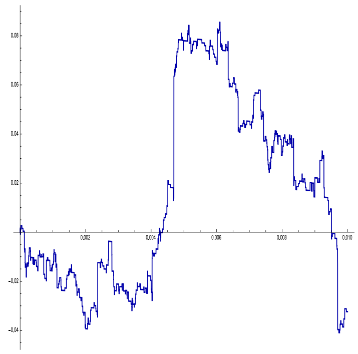

Figure 4: Monte Carlo simulation of the CGMY model with $C = 1$, $Y = 1.5$, $M = 3$, $G = 2$ and zero Brownian component.

Theorem 3.1 *Lévy-Khintchine representation). For any Lévy process X, its characteristic function can be written in the form*

$$\mathbb{E}(e^{iuX_t}) \quad = \quad \exp\left[t(i\gamma u - \frac{1}{2}\sigma^2 u^2 + \int_{-\infty}^{\infty}(e^{iux} - 1 - iux1_{|x|\leq 1})\nu(dx)\right]$$

for some (γ, σ, ν).

Notice that there is one additional term in the integral which is not there when X is a jump diffusion. We will not prove this here, but note that if $\nu(\mathbb{R}) = +\infty$, this term is needed to ensure that the integral here is finite.

7.3 Examples of Lévy processes

- The double exponential **Kou model** is a jump diffusion model where the jump arrive at Poisson rate λ and the jump size follows a two-sided exponential distribution, so

$$\nu(x) \quad = \quad \lambda[p\lambda_+ e^{-\lambda_+ x}1_{x>0} \, + \, (1-p)\lambda_- e^{-\lambda_- x}1_{x>0}] \, .$$

  This is a **finite activity** model, because $\nu(x)$ is just a multiple of a probability density, so $\nu(\mathbb{R}) < \infty$.

- The **CGMY** (Carr-Geman-Madan-Yor) model has a Lévy density of the form

$$\nu(x) \quad = \quad \frac{Ce^{-Mx}}{x^{1+Y}}\,\mathbf{1}_{\{x>0\}} \, + \, \frac{Ce^{Gx}}{|x|^{1+Y}}\,\mathbf{1}_{\{x<0\}}$$

  for $C, G, M > 0$ and $Y \in (0, 2)$, so we the see the jump rate tends to infinity as the jump size tends to zero. This is what we call an **infinite activity** model, because $\nu(x)$ is not a multiple of a pdf.

- The characteristic function (C.F.) of the CGMY model can be computed explicitly as

$$\phi_t(u) \quad = \quad \mathbb{E}(e^{ikX_t}) \quad = \quad e^{tC\Gamma(-Y)((M-ik)^Y+(G+ik)^Y-M^Y-G^Y)+ibkt-\frac{1}{2}\sigma^2k^2t}$$

  for $Y \neq 1$ and some constant b which controls the drift, and setting $ik = p$ for $p \in \mathbb{R}$, we can compute the **critical moments** p_+, p_- as $p_+ = M$ and $p_- = -G$. Note this C.F. is easier to compute than the Heston C.F. since it does not require solving ODEs. We can multiply e.g. the Heston+CGMY C.F.s to get the C.F. of the log stock price for a Heston model plus an additional independent CGMY component.

- To compute the density of $p_t(x)$ or the price of a European call option on $S_T = e^{X_T}$, we use the inverse Fourier transform formulae as before. For European call pricing, b has to be chosen so that $\mathbb{E}(S_t) = \mathbb{E}(e^{X_t}) = e^{rt}$, which ensures that $S_t e^{-rt}$ is a martingale (and hence no arbitrage) from the independent increments property (assuming $X_t = \log S_t$ here).

4 More background on the Rough Heston model

- For the general case $\alpha \neq 1$, the two integrands in (4) contain $(t-s)^{\alpha-1}$ terms (and thus depend on t), so this is not the integrated form of a standard SDE). The Rough Heston model is more realistic than Black-Scholes because volatility is now stochastic and **mean-reverting**, the model has fat tails i.e. $\mathbb{E}(e^{pX_t}) = \infty$ for some $p = p^*(t) < \infty$ sufficiently large (unlike the Black-Scholes model for which $\mathbb{E}(e^{pX_t}) = e^{\frac{1}{2}\sigma^2(p^2-p)t} < \infty$ for all $p \in \mathbb{R}$ when $r = 0$), and the Rough Heston model is more consistent with observed behaviour of traded European option prices, particularly at small maturities.

- Taking expectations of (4) and using that the expectation of the stochastic integral term is zero, we see that

$$\begin{aligned}\mathbb{E}(V_t) &= V_0 + \frac{1}{\Gamma(\alpha)}\int_0^t (t-s)^{\alpha-1}\lambda(\theta - \mathbb{E}(V_s))ds + \frac{1}{\Gamma(\alpha)}\mathbb{E}(\int_0^t (t-s)^{\alpha-1}\nu\sqrt{V_s}dB_s) \\ &= V_0 + \frac{\lambda}{\Gamma(\alpha)}\int_0^t (t-s)^{\alpha-1}(\theta - \mathbb{E}(V_s))ds. \end{aligned} \tag{16}$$

- Now let $\xi_t(u) = \mathbb{E}(V_u|\mathcal{F}_t)$. From the tower propery in FM01, we can easily verify that any process of the form $\mathbb{E}(X|\mathcal{F}_t)$ is a martingale, if X is random variable with $\mathbb{E}(|X|) < \infty$. Thus for our particular example here, $\xi_t(u)$ is a martingale in t with respect to $\mathcal{F}_t^B$, and

$$\begin{aligned}\xi_t(u) &= V_0 + \frac{1}{\Gamma(\alpha)}\int_0^u (u-s)^{\alpha-1}\lambda(\theta - \mathbb{E}(V_s|\mathcal{F}_t)ds + \frac{1}{\Gamma(\alpha)}\mathbb{E}((\int_0^t + \int_t^u)(u-s)^{\alpha-1}\nu\sqrt{V_s}dB_s) \\ &= V_0 + \frac{1}{\Gamma(\alpha)}\int_0^u (u-s)^{\alpha-1}\lambda(\theta - \mathbb{E}(V_s|\mathcal{F}_t)ds + \frac{1}{\Gamma(\alpha)}\int_0^t (u-s)^{\alpha-1}\nu\sqrt{V_s}dB_s\end{aligned}$$

  where we have used that $\int_0^t (u-s)^{\alpha-1}\nu\sqrt{V_s}dB_s$ is $\mathcal{F}_t$-measurable.

- If $\lambda = 0$, the 1st integral term on the right hand side is zero, and we can re-write the 2nd term in the differential form:

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

  where $\kappa(u-t) = \frac{\nu}{\Gamma(\alpha)}(u-t)^{\alpha-1}$.

- For general $\lambda \neq 0$, we can show that

$$d\xi_t(u) = \kappa(u-t)\sqrt{V_t}dB_t \tag{17}$$

  for some (more complicated) function κ for which we can compute a series expansion. Note that $\xi_t(t) = \mathbb{E}(V_t|\mathcal{F}_t) = \xi_t(t)$.

- $\xi_t(u)$ (considered as a function of u, for a fixed t) is known as the **forward variance curve** at time t, which moves up and down and tilts as time evolves, since (depending on κ) some parts of the curve are responsive than others to changes in W. $\xi_t(u)$ satisfies the **Markov property** in itself, since we only need to know $V_t = \xi_t(t)$ to be able to compute $d\xi_t(u)$. Note that V is not Markov in itself (see (20) below to see why).

- We can also consider other models which are not Rough Heston but for which (17) is still satisfied, and models of this form are known as **affine forward variance models**. We can integrate this relation to obtain

$$\xi_t(u) = \xi_0(u) + \int_0^t \kappa(u-s)\sqrt{V}_s dB_s \tag{18}$$

  so in particular

$$V_t = \xi_t(t) = \xi_0(t) + \int_0^t \kappa(t-s)\sqrt{V}_s dB_s \tag{19}$$

  which generalizes the Rough Heston model.

- Note we can either specify the dynamics for V (as we do for Rough Heston, in which case $\xi_0(u)$ is obtained by solving (16)) and $\kappa(u-t)$ can be computed as in Homework 2 q1, or much easier when $\lambda = 0$), in which case $\kappa(u-t) = \frac{1}{\Gamma(\alpha)}\nu(u-t)^{-\alpha}$), *or* we can specify the initial variance curve $\xi_0(t)$ and $\kappa(t-u)$ exogenously, and V_t is then given by (19).

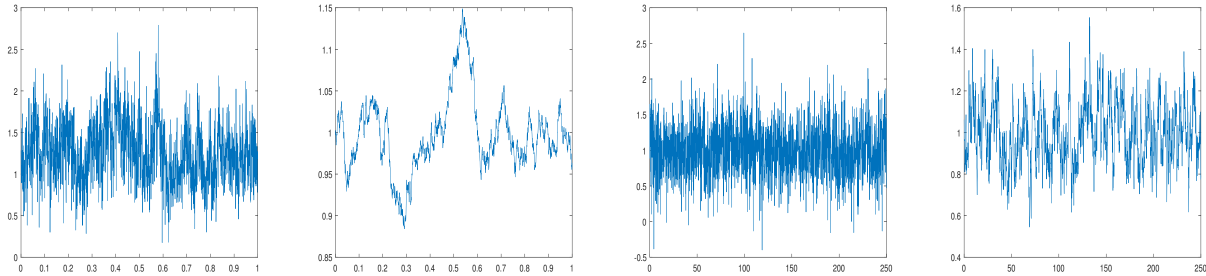

Figure 5: Here we have plotted a Monte Carlo simulation of V_t in Matlab for the rough Heston model with $\alpha = .55$ (first and third plots), i.e. α close to the lower limit of 0.5, and $\alpha = 1$ (second and final plot), with $\lambda = 1$, $\theta = V_0 = 1$ and $\nu = .4$. One can see that V becomes rougher as α becomes smaller

- Note that for general κ, (19) is no longer the Rough Heston model, but rather a more general *affine variance curve model*. Note that

$$V_t - V_s \quad = \quad \xi_0(t) \, - \, \xi_0(s) \, + \int_0^t \kappa(t-r)\sqrt{V}_r dB_r \, - \int_0^s \kappa(s-r)\sqrt{V}_r dB_r$$

  so V is not Markov in itself, since the right hand side depends on V going back to time zero, not just over $[s,t]$.

- Compare this to the **Rough Bergomi** model

$$d\xi_t(u) \quad = \quad \eta \xi_t(u) dB_t$$

  or the **standard Bergomi** model:

$$d\xi_t(u) \quad = \quad \sum_i \eta_i e^{-\lambda_i (T-u)} \xi_t(u) dB_t$$

  where we typically extract the initial variance curve $\xi_0(t)$ from the market prices of variance swaps which pay $\int_0^t V_s ds$, since $\xi_0(t) = \frac{d}{dt}\mathbb{E}(\int_0^t V_s ds)$.