Heston Simulation

Daniele Marazzina
Finanza Computazionale @ Polimi

The Heston model is defined by the coupled two-dimensional SDE

$$dX(t)/X(t) = \sqrt{V(t)}dW_X(t)$$
$$dV(t) = k(\theta - V(t))dt + \epsilon\sqrt{V(t)}dW_V(t)$$

where

- k, θ and ϵ are strictly positive constants
- W_X and W_V are scalar Brownian motions such that $dW_X(t) * dW_Y(t) = \rho dt$, where the correlation $\rho \in [-1, 1]$

- $X(t)$ represents an asset price process
- $V(t)$ represents the instantaneous variance of relative changes to $X(t)$, in the sense that the quadratic variation of $dX(t)/X(t)$ over $[t, t + dt]$ is $V(t)dt$

An application of Ito's Lemma shows that the previous system is equivalent to

$$d\log(X(t)) = -\frac{1}{2}V(t)dt + \sqrt{V(t)}dW_X(t)$$
$$dV(t) = k(\theta - V(t))dt + \epsilon\sqrt{V(t)}dW_V(t)$$

Risk-Neutral

- Here the process X is considered drift-less
- In the case in which X is the price of an asset, if we want to compute the option price with MC method, then we must consider the risk-neutral version

$$d\log(X(t)) = \mathbf{r\,dt} \quad -\frac{1}{2}V(t)dt + \sqrt{V(t)}dW_X(t)$$

$$dV(t) = k(\theta - V(t))dt + \epsilon\sqrt{V(t)}dW_V(t)$$

- $e^{-r\,dt + X(t)}$, and its discretization, must be martingale

Euler Scheme

$$dY(t) = \mu(Y(t))dt + \sigma(Y(t))dW(t)$$

can be approximated by

- Define $t_i = i\Delta t$
- If $t \in (t_i, t_{i+1})$, then
  - $dY(t) \approx Y(t_{i+1}) - Y(t_i)$
  - $dW(t) \approx \sqrt{\Delta t}Z$, $Z \sim N(0,1)$
  - $\mu(Y(t)) \approx \mu(Y(t_i))$ and $\sigma(Y(t)) \approx \sigma(Y(t_i))$

If we denote $Y_i = Y(t_i)$, then

$$Y_{i+1} = Y_i + \mu(Y_i)\Delta t + \sigma(Y_i)\sqrt{\Delta t}Z$$

$$\log(X_{i+1}) = log(X_i) - \frac{1}{2}V_i\Delta t + \sqrt{V_i}Z_X\sqrt{\Delta t}$$

$$V_{i+1} = V_i + k(\theta - V_i)\Delta t + \epsilon\sqrt{V_i}Z_V\sqrt{\Delta t}$$

where Z_X and Z_Y are standard normal variables with correlation ρ

How to generate Z_X and Z_V?

1. Generate $U_1, U_2 \approx^{i.i.d.} U[0,1]$
2. $Z_V := \phi^{-1}(U_1)$, where ϕ is the c.d.f.
3. $Z_X := \rho Z_V + \sqrt{1-\rho^2}\,\phi^{-1}(U_2)$

A Matlab Code

```
clear all
T=2; N=20; dt=T/N; t=[0:dt:T];
rho=0.5; k=1; epsilon=0.4; theta=1; x0=1; v0=0.3;

X(1)=x0;  V(1)=v0;
for i=1:N
  ZV=randn; ZX=randn;
  ZX=rho*ZV+sqrt(1-rho^2)*ZX;
  X(i+1)=exp(log(X(i))-0.5*V(i)*dt+...
                    sqrt(V(i))*ZX*sqrt(dt));
  V(i+1)=V(i)+k*(theta-V(i))*dt+...
                    epsilon*sqrt(V(i))*ZV*sqrt(dt);
end
plot(t,X,'*');
```

Problem

The simulated V can be negative with non-zero probability

First Solution

Replace $V(i)$ with $\max(V(i), 0)$

However $V(i) = 0 \Rightarrow \quad V(i+1) = k\theta\Delta t$, i.e., $V(i+1)$ is deterministic

Advanced Scheme

Leif Andersen, *Efficient Simulation of the Heston Stochastic Volatility Model*, 2007
http://ssrn.com/abstract=946405

The author introduces two advanced schemes to simulate V

- TG Scheme (Truncated Gaussian)
- QE Scheme (Quadratic-Exponential)

and an advanced scheme to simulate X

About the Euler Scheme...

One immediate problem with the scheme above is that the discrete process for V can become negative with non-zero probability, which in turn would make computation of $\sqrt{\hat{V}}$ impossible and cause the time-stepping scheme to fail. To get around this problem, several remedies have been proposed in the literature; see [LKD] for a review of various "fixes". The scheme that appears to produce the smallest discretization bias can be written on the form

$$\ln \hat{X}(t+\Delta) = \ln \hat{X}(t) - \frac{1}{2}\hat{V}(t)^{+}\Delta + \sqrt{\hat{V}(t)^{+}}Z_X\sqrt{\Delta}, \quad (6)$$

$$\hat{V}(t+\Delta) = \hat{V}(t) + \kappa\left(\theta - \hat{V}(t)^{+}\right)\Delta + \varepsilon\sqrt{\hat{V}(t)^{+}}Z_V\sqrt{\Delta}, \quad (7)$$

where we use the notation $x^{+} = \max(x, 0)$. In [LKD] this scheme is denoted "full truncation"; its main characteristic is that the process for V is allowed to go below zero, at which point the process for V becomes deterministic with an upward drift of $\kappa\theta$.

Andersen-pag.6

Is the problem of touching zero real?

Proposition 2 *Assume that* $V(0) > 0$. *If* $2\kappa\theta \geq \varepsilon^2$

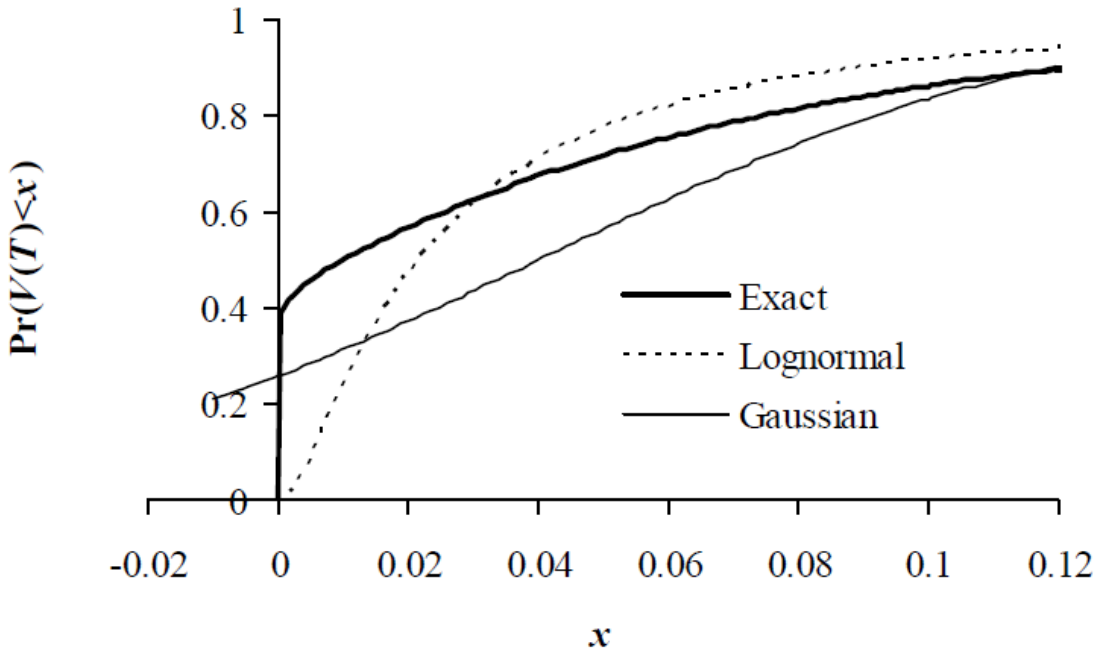
then the process for V *can never reach zero.*

If $2\kappa\theta < \varepsilon^2$, *the origin is accessible and strongly reflecting.*

Andersen-pag.4

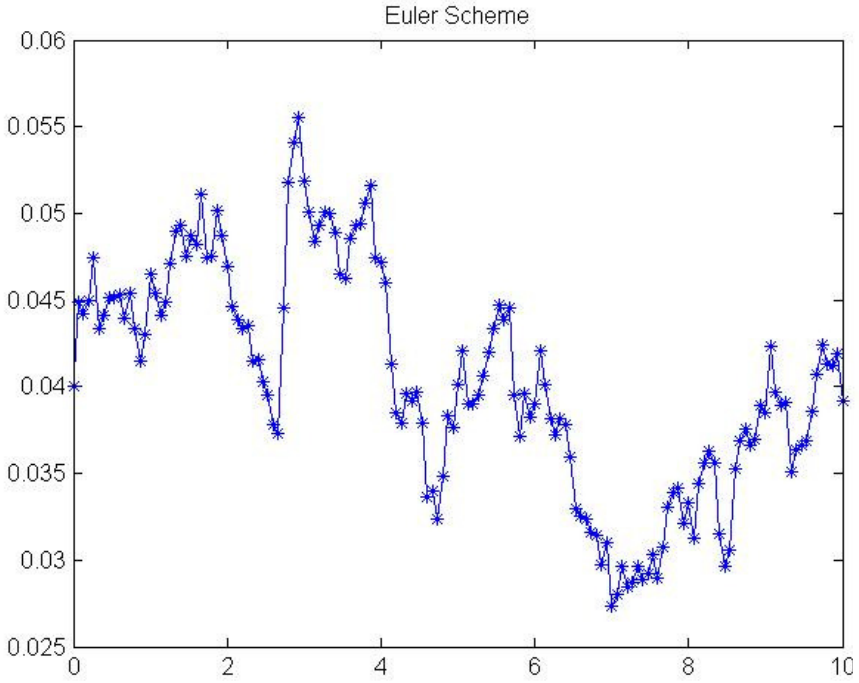
In typical applications $2\kappa\theta$ is often significantly below ε^2, so the likelihood of hitting zero is often quite significant. Indeed, the process for V often has a strong affinity for the area around the origin, as is evident from the distribution graph in Figure 1. For comparison, we have superimposed Gaussian and lognormal distributions matched to the first two moments of V; evidently neither of these distributions are particularly good proxies for the true distribution of V.

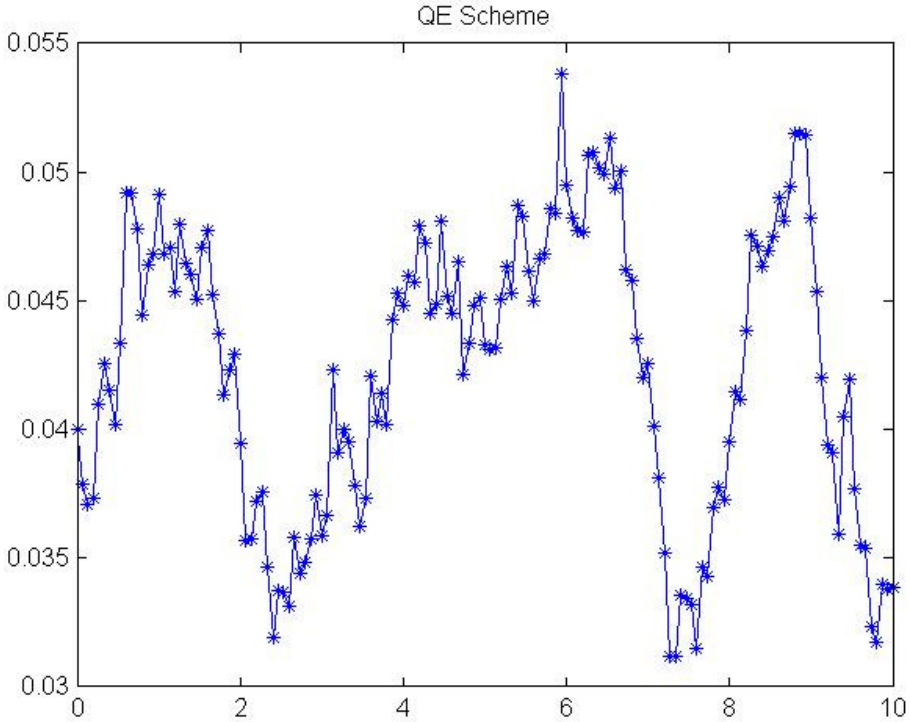
Figure 1: Cumulative Distribution of V



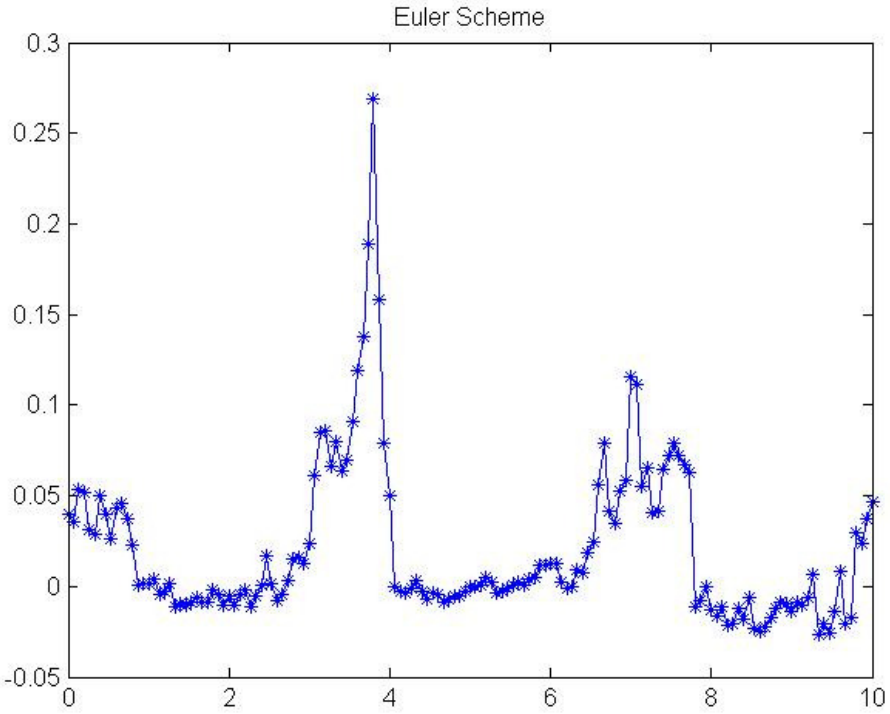
Notes: The figure shows the cumulative distribution function for $V(T)$ given $V(0)$, with $T = 0.1$. Model parameters were $V(0) = \theta = 4\%$, $\kappa = 50\%$, and $\varepsilon = 100\%$. The lognormal and Gaussian distributions in the graph were parameterized by matching mean and variances to the exact distribution of $V(T)$.

k=0.5, theta=0.04, epsilon=0.04

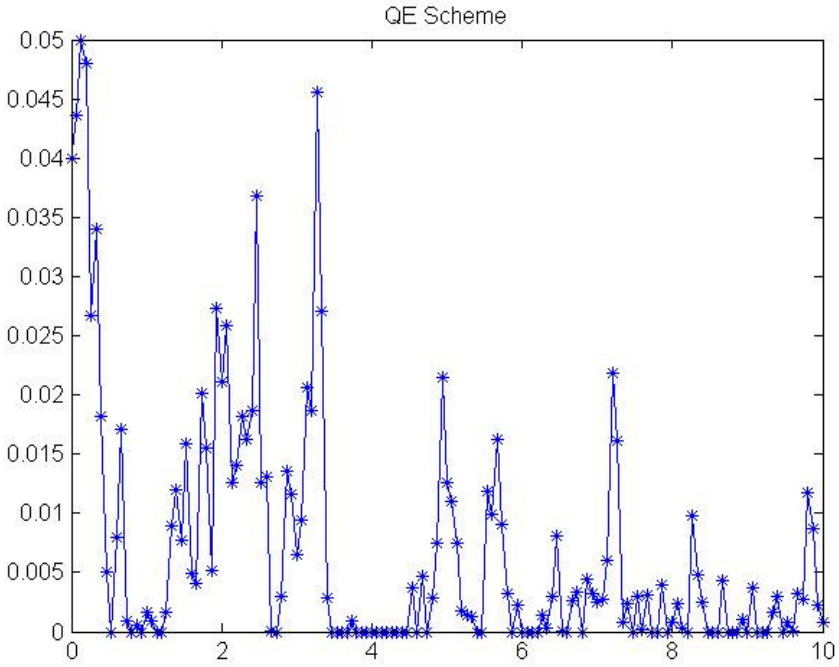


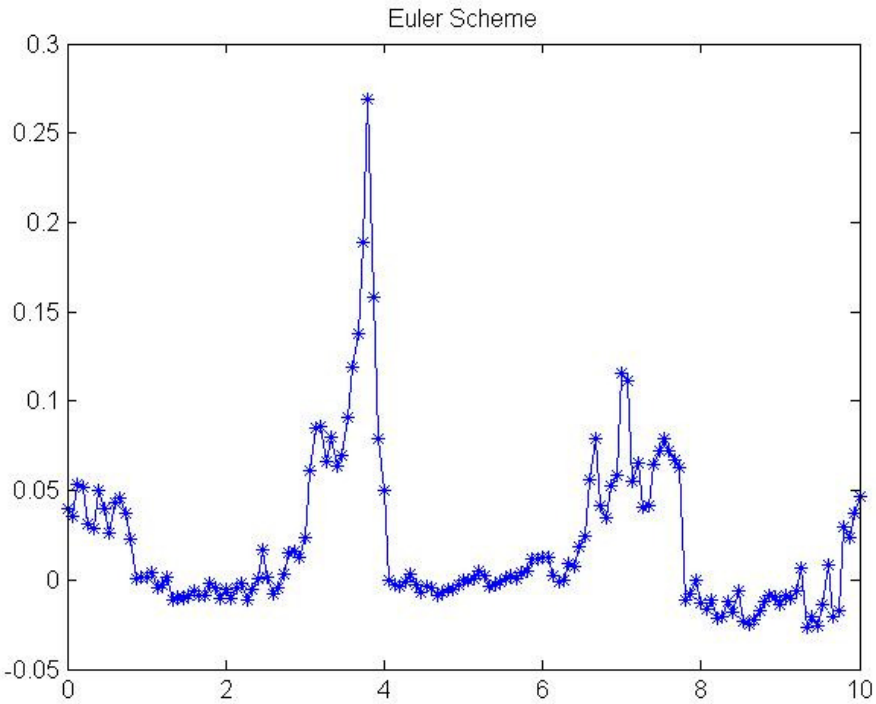


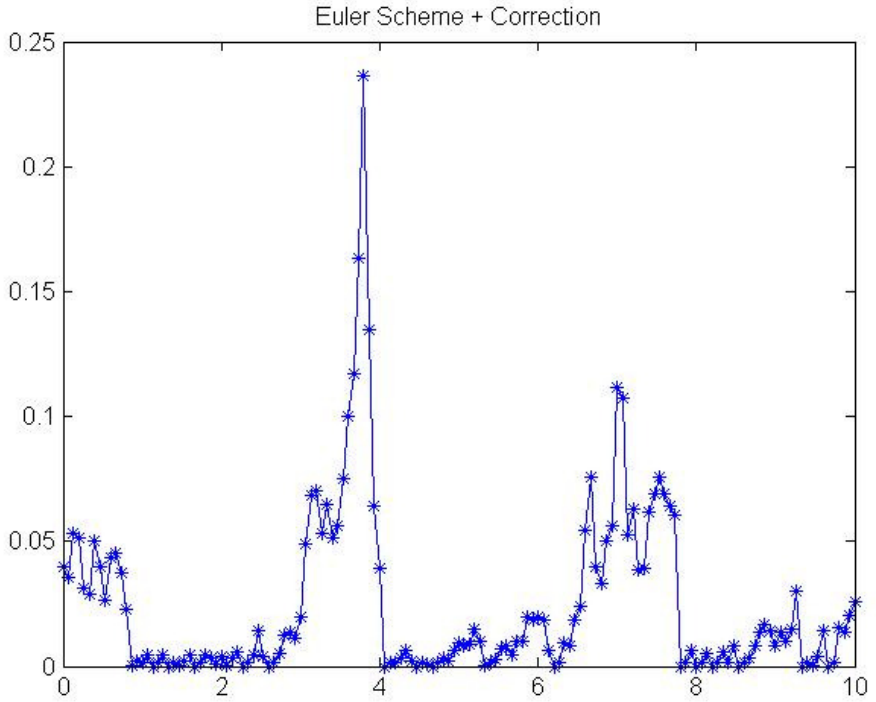
k=0.5, theta=0.04, epsilon=0.4



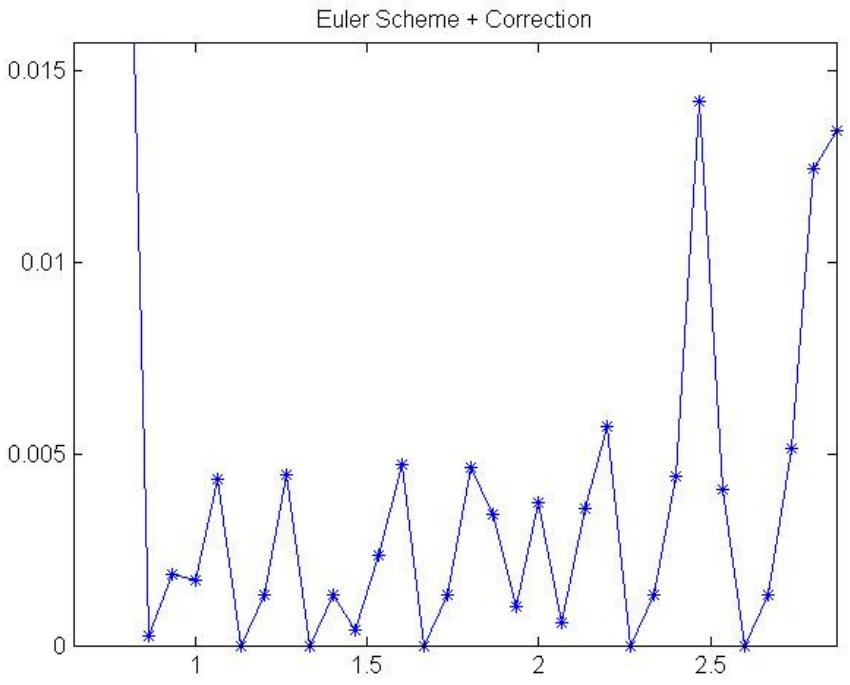
→ *Warning: Imaginary parts of complex X and/or Y arguments ignored*

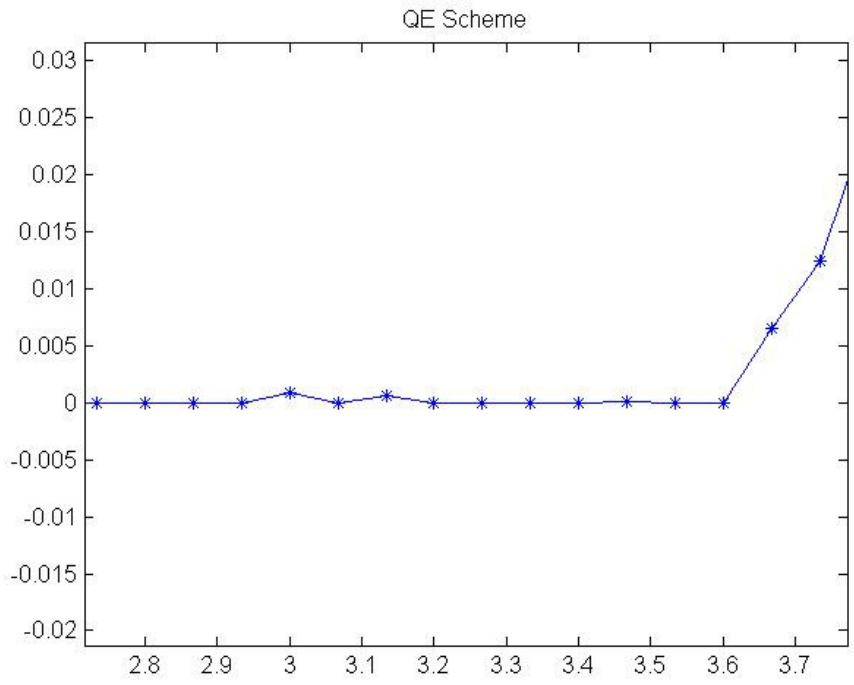



Euler Scheme
0.3
0.25
0.2
0.15
0.1
0.05
0
-0.05
0
2
4
6
8
10


Euler Scheme + Correction
0.25
0.2
0.15
0.1
0.05
0
0
2
4
6
8
10

Zooming...





(left): if zero is touched, the increase is deterministic

(right): if zero is touched, the process can remain near zero

QE Scheme

In the Heston model the volatility follows a CIR process, here denoted as process (2)

Proposition 1 *Let $F_{\chi'^2}(y;\nu,\lambda)$ be the cumulative distribution function for the non-central chi-square distribution with ν degrees of freedom and non-centrality parameter λ:*

$$F_{\chi'^2}(z;\nu,\lambda) = e^{-\lambda/2}\sum_{j=0}^{\infty}\frac{(\lambda/2)^j}{j!2^{\nu/2+j}\Gamma(\nu/2+j)}\int_0^z z^{\nu/2+j-1}e^{-x/2}\,dx.$$

For the process (2) define

$$d = 4\kappa\theta/\varepsilon^2; \quad n(t,T) = \frac{4\kappa e^{-\kappa(T-t)}}{\varepsilon^2\left(1-e^{-\kappa(T-t)}\right)}, T > t.$$

Let $T > t$. Conditional on $V(t)$, $V(T)$ is distributed as $e^{-\kappa(T-t)}/n(t,T)$ times a non-central chi-square distribution with d degrees of freedom and non-centrality parameter $V(t)n(t,T)$. That is,

$$\Pr\left(V(T) < x|V(t)\right) = F_{\chi'^2}\left(\frac{x\cdot n(t,T)}{e^{-\kappa(T-t)}}; d, V(t)\cdot n(t,T)\right).$$

Andersen-pag.3

Moments

Corollary 1 *Let $T > t$. Conditional on $V(t)$, $V(T)$ has the following first two moments:*

$$\begin{aligned} \mathrm{E}\left(V(T)|V(t)\right) &= \theta + (V(t) - \theta)\, e^{-\kappa(T-t)}; \\ \mathrm{Var}\left(V(T)|V(t)\right) &= \frac{V(t)\varepsilon^2 e^{-\kappa(T-t)}}{\kappa}\left(1 - e^{-\kappa(T-t)}\right) + \frac{\theta\varepsilon^2}{2\kappa}\left(1 - e^{-\kappa(T-t)}\right)^2. \end{aligned}$$

Andersen-pag.3

QE Scheme

Q=Quadratic

We derive our new scheme in steps. The first step is based on an observation that a non-central chi-square with moderate or high non-centrality parameter can be well-represented by a power-function applied to a Gaussian variable. See [Pat], [Pe], and [Pit0], as well as the survey in [JK]. While there is evidence that a cubic transformation of a Gaussian variable is preferable, such a scheme could not preserve non-negative values for the V process and we abandon it in favor of a quadratic representation, along the lines of [Pat]. Specifically, for sufficiently large[7] values of $\hat{V}(t)$, we write

$$\hat{V}(t+\Delta) = a\,(b+Z_V)^2 \tag{23}$$

where Z_V is a standard Gaussian random variable, and a and b are certain constants, to be determined by moment-matching. a and b will depend on the the time-step Δ and $\hat{V}(t)$, as well as the parameters in the SDE for V.

Andersen-pag.31

QE Scheme

E=Exponential

The scheme (23) does not work well for low values of $V(t)$ – in fact the moment-matching exercise fails to work – so we supplement it with a scheme to be used for low values of $\hat{V}(t)$. For this, we take inspiration from the asymptotic density in (12) and use an approximated density for $\hat{V}(t+\Delta)$ of the form

$$\Pr\left(\hat{V}(t+\Delta)\in[x,x+dx]\right)\approx\left(p\delta(0)+\beta(1-p)e^{-\beta x}\right)dx,\quad x\geq 0, \tag{24}$$

where δ is a Dirac delta-function, and p and β are non-negative constants to be determined. As in the TG scheme, we have a probability mass at the origin, but now the strength of this mass (p) is explicitly specified, rather than implied from other parameters. The mass at the origin is supplemented with an exponential tail, qualitatively similar to that of the density (12). It can be verified that if $p\in[0,1]$ and $\beta\geq 0$, then (24) constitutes a valid density function.

Sampling according to (24) is straightforward and efficient. To see this, first we integrate (24) to generate a cumulative distribution function

$$\Psi(x) = \Pr\left(\hat{V}(t+\Delta) \leq x\right) = p + (1-p)\left(1 - e^{-\beta x}\right), \quad x \geq 0.$$

We notice that the inverse of Ψ is readily computable:

$$\Psi^{-1}(u) = \Psi^{-1}(u; p, \beta) = \begin{cases} 0, & 0 \leq u \leq p, \\ \beta^{-1} \ln\left(\frac{1-p}{1-u}\right), & p < u \leq 1. \end{cases} \tag{25}$$

By the standard inverse distribution function method, we thus get the simple sampling scheme

$$\hat{V}(t+\Delta) = \Psi^{-1}(U_V; p, \beta) \tag{26}$$

where U_V is a draw from a uniform distribution. Note that this scheme is extremely fast to execute.

What are the distributions' parameters?

$$m = \theta + \left(\hat{V}(t) - \theta\right) e^{-\kappa\Delta}; \tag{17}$$

$$s^2 = \frac{\hat{V}(t)\varepsilon^2 e^{-\kappa\Delta}}{\kappa}\left(1 - e^{-\kappa\Delta}\right) + \frac{\theta\varepsilon^2}{2\kappa}\left(1 - e^{-\kappa\Delta}\right)^2, \tag{18}$$

such that

$$\psi = \frac{s^2}{m^2} = \frac{\frac{\hat{V}(t)\varepsilon^2 e^{-\kappa\Delta}}{\kappa}\left(1 - e^{-\kappa\Delta}\right) + \frac{\theta\varepsilon^2}{2\kappa}\left(1 - e^{-\kappa\Delta}\right)^2}{\left(\theta + \left(\hat{V}(t) - \theta\right) e^{-\kappa\Delta}\right)^2}. \tag{19}$$

Differentiating this expression with respect to $\hat{V}(t)$ shows that $\partial\psi/\partial\hat{V}(t) < 0$ for all $\hat{V}(t) \geq 0$, such that the largest possible value for ψ is obtained for $\hat{V}(t) = 0$, and the smallest possible value for $\hat{V}(t) = \infty$. Inserting these values for $\hat{V}(t)$ into (19 shows that $\psi \in (0, \varepsilon^2/(2\kappa\theta)]$.

Andersen-pag.11

ψ is used to determine if the value of V is large (Q scheme) or small (E scheme)

Proposition 5 *Let m and s be as defined in Proposition 4 (equations (17) and (18)), and set $\psi = s^2/m^2$. Provided that $\psi \leq 2$, set*

$$b^2 = 2\psi^{-1} - 1 + \sqrt{2\psi^{-1}}\sqrt{2\psi^{-1} - 1} \geq 0 \tag{27}$$

and

$$a = \frac{m}{1 + b^2}. \tag{28}$$

Let $\hat{V}(t + \Delta)$ be as defined in (23); then $\mathrm{E}\left(\hat{V}(t + \Delta)\right) = m$ and $\mathrm{Var}\left(\hat{V}(t + \Delta)\right) = s^2$.

Proposition 6 *Let m, s, and ψ be as defined in Proposition 4. Assume that $\psi \geq 1$ and set*

$$p = \frac{\psi - 1}{\psi + 1} \in [0, 1), \tag{29}$$

and

$$\beta = \frac{1 - p}{m} = \frac{2}{m(\psi + 1)} > 0. \tag{30}$$

Let $\hat{V}(t + \Delta)$ be as defined in (26); then $\mathrm{E}\left(\hat{V}(t + \Delta)\right) = m$ and $\mathrm{Var}\left(\hat{V}(t + \Delta)\right) = s^2$.

Andersen-pag.14-15

3.2.4 *Summary of QE algorithm*

Assume that some arbitrary level $\psi_c \in [1,2]$ has been selected. The detailed algorithm for the QE simulation step from $\hat{V}(t)$ to $\hat{V}(t+\Delta)$ is then:

1. Given $\hat{V}(t)$, compute m and s^2 from equations (17) and (18).
2. Compute $\psi = s^2/m^2$
3. Draw a uniform random number U_V rand
4. **If** $\psi \leq \psi_c$:
   (a) Compute a and b from equations (28) and (27)
   (b) Compute $Z_V = \Phi^{-1}(U_V)$, e.g. using the algorithm in [Moro] randn
   (c) Use (23), i.e. set $\hat{V}(t+\Delta) = a(b+Z_V)^2$
5. **Otherwise**, if $\psi > \psi_c$:
   (a) Compute β and p according to equations (30) and (29)
   (b) Use (26), i.e. set $\hat{V}(t+\Delta) = \Psi^{-1}(U_V; p, \beta)$, where Ψ^{-1} is given in (25).

Andersen-pag.16-17

```
%%%-- Andersen - QE scheme -------------------------
Psi_c = 1.5;
V(1)=x0;
for i=1:N
    % STEP 1 & 2
    % Calculate m,s,Psi
    m=theta+(V(i)-theta)*exp(-k*dt);
    s2=V(i)*epsilon^2*exp(-k*dt)*(1-exp(-k*dt))/k+...
              theta*epsilon^2*(1-exp(-k*dt))^2/(2*k);
    s=sqrt(s2);
    Psi = (s2)/(m^2);
```

```
    % STEP 3,4,5
    if (Psi>Psi_c)
        % Exponential approximation
        p = (Psi-1)/(Psi+1);%Prob. mass in 0
        beta = (1-p)/m;%exponent of density tail
        % gets V from inverse CDF applied to uniform U
        U = rand;
        V(i+1)=(log((1-p)/(1-U))/beta)*(U>p);
    else
        % Quadratic approx
        invPsi = 1/Psi;
        b2=2*invPsi-1+sqrt(2*invPsi)*sqrt(2*invPsi-1);
        a=m/(1+b2);
        V(i+1)=a*(sqrt(b2)+randn)^2;
    end
end
```

What about X?

$$\ln\hat{X}(t+\Delta) = \ln\hat{X}(t) + \frac{\rho}{\varepsilon}\left(\hat{V}(t+\Delta) - \hat{V}(t) - \kappa\theta\Delta\right) + \Delta\left(\frac{\kappa\rho}{\varepsilon} - \frac{1}{2}\right)\left(\gamma_1\hat{V}(t) + \gamma_2\hat{V}(t+\Delta)\right)$$
$$+ \sqrt{\Delta}\sqrt{1-\rho^2}\sqrt{\gamma_1\hat{V}(t) + \gamma_2\hat{V}(t+\Delta)}\cdot Z$$
$$= \ln\hat{X}(t) + K_0 + K_1\hat{V}(t) + K_2\hat{V}(t+\Delta) + \sqrt{K_3\hat{V}(t) + K_4\hat{V}(t+\Delta)}\cdot Z, \qquad (33)$$

where Z is a standard Gaussian random variable, independent of $\hat{V}$, and $K_0, \ldots, K_4$ are given by

$$K_0 = -\frac{\rho\kappa\theta}{\varepsilon}\Delta, \quad K_1 = \gamma_1\Delta\left(\frac{\kappa\rho}{\varepsilon} - \frac{1}{2}\right) - \frac{\rho}{\varepsilon},$$
$$K_2 = \gamma_2\Delta\left(\frac{\kappa\rho}{\varepsilon} - \frac{1}{2}\right) + \frac{\rho}{\varepsilon}, \quad K_3 = \gamma_1\Delta(1-\rho^2), \quad K_4 = \gamma_2\Delta(1-\rho^2).$$

Note that the K_i, $i = 0, \ldots, 4$ depend on the time-step as well as on the constants γ_1 and γ_2.

Andersen-pag.19

For given values of γ_1 and γ_2, the scheme constitutes our proposed discretization scheme for $\ln X$. It is to be combined with a simulation scheme for V, in the following fashion:

1. Given $\hat{V}(t)$, generate $\hat{V}(t+\Delta)$ using one of the time-stepping schemes in Section 3
2. Draw a uniform random number U, independent of all random numbers used for $\hat{V}(t+\Delta)$
3. Set $Z = \Phi^{-1}(U)$, e.g. using the algorithm in [Moro]
4. Given $\ln\hat{X}(t)$, $\hat{V}(t)$, and the value for $\hat{V}(t+\Delta)$ computed in Step 1, compute $\ln\hat{X}(t+\Delta)$ from (33).

Andersen-pag.19

- γ_1 and γ_2 can be taken equal to 0.5
- step 2 and 3 can be substituted by randn

```
% Implementation of QE Scheme for Heston model
clear; close all;

% PARAMETERS
epsilon=1; k=0.5; rho=-0.9; theta=0.04;
T=10; V0=0.04; X0=100;
N=150; dt=T/N; t=[0:dt:T];
V = zeros(N,1); X = zeros(N,1);

% Random changes for volatility are sampled from one of
% two distributions, depending on the ratio Psi = s^2/m^2,
% where m & s are mean and variance of next volatility
% value, conditioned to current one.
% Scheme 1 (exponential approx.) selected when Psi>Psi_cutoff
% Scheme 2 (quadratic approx.) selected when 0<Psi<Psi_cutoff

% choose Psi_cutoff = 1.5 as suggested in article
Psi_cutoff = 1.5;
```

```
%--- DISCRETIZE V ----------------------------------
V(1)=V0;
for i=1:N
    % STEP 1 & 2: Calculate m,s,Psi
    m=theta+(V(i)-theta)*exp(-k*dt);
    m2=m^2;
    s2=V(i)*epsilon^2*exp(-k*dt)*(1-exp(-k*dt))/k+...
        theta*epsilon^2*(1-exp(-k*dt))^2/(2*k);
    s=sqrt(s2);
    Psi = (s2)/(m2); invPsi = 1/Psi;
    % STEP 3,4,5: use exp or quad scheme
    if (Psi>Psi_cutoff) % Exponential approximation
      p_exp = (Psi-1)/(Psi+1);
      beta_exp = (1-p_exp)/m;
      U = rand;
      V(i+1)=(log((1-p_exp)/(1-U))/beta_exp)*(U>p_exp);
```

```
    else  % Quadratic approx
      b2_quad=2*invPsi-1+sqrt(2*invPsi)*sqrt(2*invPsi-1);
      a_quad = m/(1+b2_quad);
      V(i+1) = a_quad*(sqrt(b2_quad)+randn)^2;
    end
end
%-- DISCRETIZE X ------------------------------------
% No martingale correction applied
X(1)=X0;
gamma1=0.5; gamma2=0.5; %central discretization scheme
k0=-rho*k*theta*dt/epsilon;
k1=gamma1*dt*(k*rho/epsilon-0.5)-rho/epsilon;
k2=gamma2*dt*(k*rho/epsilon-0.5)+rho/epsilon;
k3=gamma1*dt*(1-rho^2); k4=gamma2*dt*(1-rho^2);
for i=1:N
      X(i+1)=exp(log(X(i))+k0+k1*V(i)+k2*V(i+1)+...
    sqrt(k3*V(i)+k4*V(i+1))*icdf('norm',rand,0,1));
end
```

In the risk-neutral case…

- Add *r* to the drift
- The numerical discretization must be itself a martingale --> replace K_0 with

$$K_0^* = -\ln M - \left(K_1 + \frac{1}{2}K_3\right)\hat{V}(t).$$

Define

$$M = \mathrm{E}\left(e^{A\hat{V}(t+\Delta)}|\hat{V}(t)\right) > 0, \quad A = K_2 + \frac{1}{2}K_4 = \frac{\rho}{\varepsilon}(1+\kappa\gamma_2\Delta) - \frac{1}{2}\gamma_2\Delta\rho^2.$$

If $\psi \leq \psi_c$, *then*

$$\mathrm{E}\left(e^{A\hat{V}(t+\Delta)}|\hat{V}(t)\right) = \frac{\exp\left(\frac{Ab^2a}{1-2Aa}\right)}{\sqrt{1-2Aa}},$$

where A must satisfy

$$A < \frac{1}{2a}.$$

If, on the other hand, $\psi > \psi_c$, *then*

$$\mathrm{E}\left(e^{A\hat{V}(t+\Delta)}|\hat{V}(t)\right) = p + \frac{\beta(1-p)}{\beta - A},$$

provided that

$$A < \beta.$$

Andersen-pag.20-22