

Heston Stochastic Local Volatility

J. Göttker-Schnetmann* K. Spanderen†

September 18, 2015

Abstract

This report describes the implementation of the Heston Stochastic Local Volatility Model in QuantLib.

1 Introduction

The local volatility model is widely used to price exotic equity derivatives. On the other hand it is criticized for an unrealistic volatility dynamics. Some derivatives, especially those containing forward starting features as cliques, will thus not be priced realistically.

Stochastic volatility models (SLV) have been introduced to model the dynamics better and one of the most widely used of those models is the Heston model, although its dynamics can again be criticised for being unrealistic for typical choices of parameters. We nevertheless use this model as a starting point, since an implementation is already available in the QuantLib.

A stochastic local volatility model can combine the desirable features of both models. Vanilla options can be priced exactly and the volatility dynamics can be inherited from the stochastic volatility model.

1.1 Overview

This article is organized as follows: In the second and third section we introduce the mathematics of the model. Since a calibration to real world data often yields Heston parameters which violate the Feller constraint, coordinate transforms are discussed in a large part of section three. The fourth section discusses the general discretization of the model on non-uniform grids. Special care must be given to the zero-flux boundary conditions when evolving probability densities using the Fokker-Planck equation. Because of the singular start values when evolving probability densities we found it helpful to use non-uniform and adaptive grids. For the former we had to

*jschnetm@gmail.com

†klaus@spanderen.de

extend the finite difference framework of the QuantLib, which was up to now could handle only one concentrating point.

The fifth section contains a short discussion of the calibration of the SLV model. The sixth section discusses the results of the newly introduced test cases.

2 Stochastic Local Volatility

2.1 Local Volatility Model

In the Local Volatility Model the volatility $\sigma_{LV}(S, t)$ is function of spot level S_t and time t . The dynamics of the spot price is given by:

$$d \ln S_t = \left(r_t - q_t - \frac{1}{2} \sigma_{LV}^2(S, t) \right) dt + \sigma_{LV}(S, t) dW_t \quad (1)$$

$$\sigma_{LV}^2(S, t) = \left. \frac{\frac{\partial C}{\partial T} + (r_t - q_t) K \frac{\partial C}{\partial K} + q_t C}{\frac{K^2}{2} \frac{\partial^2 C}{\partial K^2}} \right|_{K=S, T=t} \quad (2)$$

The model can be calibrated to yield prices consistent with option market prices. It is often criticized for its unrealistic volatility dynamics. The Dupire formula is mathematically appealing but also unstable.

2.2 Heston Model

$$\begin{aligned} d \ln S_t &= \left(r_t - q_t - \frac{1}{2} \nu_t \right) dt + \sqrt{\nu_t} dW_t^S \\ d\nu_t &= \kappa (\theta - \nu_t) dt + \sigma \sqrt{\nu_t} dW_t^\nu \\ \rho dt &= dW_t^\nu dW_t^S \end{aligned}$$

The Heston model allows for a semi-analytical solution for European call option prices:

$$\begin{aligned} C(S_0, K, \nu_0, T) &= SP_1 - K e^{-(r_t - q_t)T} P_2 \\ P_j &= \frac{1}{2} + \frac{1}{\pi} \int_0^\infty \Re \left[\frac{e^{-iu \ln K} \phi_j(S_0, K, \nu_0, T, u)}{iu} \right] du \end{aligned}$$

Its volatility dynamics is more realistic, but can still be criticized when the Feller constraint is violated. The volatility will then exhibit paths which stay near zero for prolonged periods of time. Additionally it does often not exhibit enough skew for short dated expiries.

2.3 Heston Stochastic Local Volatility

We now add a leverage function $L(S_t, t)$ and mixing factor η :

$$\begin{aligned} d \ln S_t &= \left(r_t - q_t - \frac{1}{2} L(S_t, t)^2 \nu_t \right) dt + L(S_t, t) \sqrt{\nu_t} dW_t^S \\ d\nu_t &= \kappa (\theta - \nu_t) dt + \eta \sigma \sqrt{\nu_t} dW_t^\nu \\ \rho dt &= dW_t^\nu dW_t^S \end{aligned}$$

The leverage function $L(x_t, t)$ is given by probability density $p(S_t, \nu, t)$ and

$$L(S_t, t) = \frac{\sigma_{LV}(S_t, t)}{\sqrt{\mathbb{E}[\nu_t | S = S_t]}} = \sigma_{LV}(S_t, t) \sqrt{\frac{\int_{\mathbb{R}^+} p(S_t, \nu, t) d\nu}{\int_{\mathbb{R}^+} \nu p(S_t, \nu, t) d\nu}} \quad (3)$$

The mixing factor η tunes between stochastic and local volatility.

3 Link between SDE and PDE

The link between SDE and PDE is given by the Feynman-Kac formula. Starting point is a multidimensional SDE of the form:

$$d\mathbf{x}_t = \mu(\mathbf{x}_t, t)dt + \sigma(\mathbf{x}_t, t)d\mathbf{W}_t$$

The Feynman-Kac formula is a PDE for the price of a derivative $u(\mathbf{x}_t, t)$ with boundary condition $u(\mathbf{x}_T, T)$ at maturity T is given by:

$$\partial_t u + \sum_{k=1}^n \mu_k \partial_{x_k} u + \frac{1}{2} \sum_{k,l=1}^n (\sigma \sigma^T)_{kl} \partial_{x_k} \partial_{x_l} u - ru = 0 \quad (4)$$

The Fokker-Planck equation gives the time evolution of the probability density function $p(\mathbf{x}_t, t)$ with the initial condition $p(\mathbf{x}, t=0) = \delta(\mathbf{x} - \mathbf{x}_0)$ by:

$$\partial_t p = - \sum_{k=1}^n \partial_{x_k} [\mu_k p] + \frac{1}{2} \sum_{k,l=1}^n \partial_{x_k} \partial_{x_l} [(\sigma \sigma^T)_{kl} p] \quad (5)$$

3.1 Backward Equation: Feynman-Kac

The SLV model leads to following Feynman-Kac equation for a function $u : \mathbb{R} \times \mathbb{R}_{\geq 0} \times \mathbb{R}_{\geq 0} \rightarrow \mathbb{R}$, $(x, \nu, t) \mapsto u(x, \nu, t)$:

$$\begin{aligned} 0 &= \partial_t u + \frac{1}{2} L^2 \nu \partial_x^2 u + \frac{1}{2} \eta^2 \sigma^2 \nu \partial_\nu^2 u + \eta \sigma \nu \rho L \partial_x \partial_\nu u \\ &\quad + \left(r - q - \frac{1}{2} L^2 \nu \right) \partial_x u + \kappa (\theta - \nu) \partial_\nu u - ru \end{aligned}$$

This PDE can be solved using either Implicit scheme (slow) or more advanced *operator splitting schemes* like modified Craig-Sneyd or Hundsdorfer-Verwer in conjunction with damping steps (fast). The Implementation is mostly harmless by extending `FdmHestonOp`. The PDE can be efficiently solved using operator splitting schemes, preferable the modified Craig-Sneyd scheme.

3.2 Forward Equation: Fokker-Planck

The corresponding Fokker-Planck equation for the probability density $p : \mathbb{R} \times \mathbb{R}_{\geq 0} \times \mathbb{R}_{\geq 0} \rightarrow \mathbb{R}_{\geq 0}, (x, \nu, t) \mapsto p(x, \nu, t)$ is:

$$\begin{aligned} \partial_t p &= \frac{1}{2} \partial_x^2 [L^2 \nu p] + \frac{1}{2} \eta^2 \sigma^2 \partial_\nu^2 [\nu p] + \eta \sigma \rho \partial_x \partial_\nu [L \nu p] \\ &\quad - \partial_x \left[\left(r - q - \frac{1}{2} L^2 \nu \right) p \right] - \partial_\nu [\kappa (\theta - \nu) p] \end{aligned} \quad (6)$$

The numerical solution of the PDE is cumbersome due to difficult boundary conditions and the Dirac delta distribution as the initial condition. Especially the square root process for the variance introduces difficult to handle boundary conditions if the so called Feller constraint is violated.

3.3 Feller Constraint

The square root process

$$d\nu_t = \kappa(\theta - \nu_t)dt + \sigma\sqrt{\nu_t}dW$$

has the following forward Fokker-Planck equation for the probability density $p : \mathbb{R}_{\geq 0} \times \mathbb{R}_{\geq 0} \rightarrow \mathbb{R}_{\geq 0}, (\nu, t) \mapsto p(\nu, t)$

$$\partial_t p = \frac{\sigma^2}{2} \partial_\nu^2 (\nu p) - \partial_\nu (\kappa(\theta - \nu) p) \quad (7)$$

Alternatively this can be written as

$$\partial_t p = \frac{\sigma^2}{2} \nu \partial_\nu^2 p + (\sigma^2 - \kappa(\theta - \nu)) \partial_\nu p + \kappa p \quad (8)$$

so that the derivative operators act on the probability distribution p . A general discussion of the properties of equations of this type can already be found in Feller's paper [1]. For reflecting boundary conditions at $\nu = 0$ this equation has the stationary solution $\partial_t \hat{p} = 0$

$$\hat{p}(\nu) = \beta^\alpha \nu^{\alpha-1} \exp(-\beta\nu) \Gamma(\alpha)^{-1}$$

with $\alpha = \frac{2\kappa\theta}{\sigma^2}, \beta = \frac{\alpha}{\theta}$. Observe that

$$\partial_\nu^2 (\nu \hat{p}) = \partial_\nu (\alpha - \nu\beta) \hat{p} = -\beta \hat{p} + (\alpha - \nu\beta) \partial_\nu \hat{p}$$

with

$$\alpha - \nu\beta = \frac{2}{\sigma^2} (\kappa\theta - \kappa\nu)$$

therefore

$$\frac{\sigma^2}{2} \partial_\nu^2 (\nu \hat{p}) = -\kappa \hat{p} + \kappa(\theta - \nu) \partial_\nu \hat{p}$$

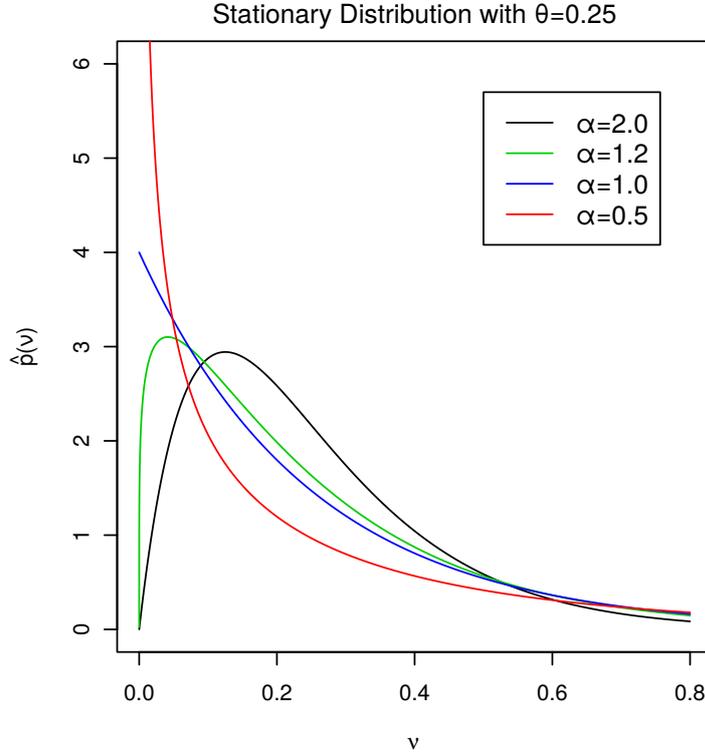


Figure 1: Stationary probability distribution.

and

$$\partial_\nu(\kappa(\theta - \nu)\hat{p}) = -\kappa p + \kappa(\theta - \nu)\partial_\nu\hat{p}$$

which completes the proof that $\partial_t\hat{p} = 0$.

Note that \hat{p} diverges as $\nu \rightarrow 0$ when the Feller constraint $\sigma^2 \leq 2\kappa\theta$ is violated, since then $\alpha - 1 = \sigma^{-2}(2\kappa\theta - \sigma^2) < 0$. The Feller constraint is often violated when calibrating the Heston model to real market data. This is usually not an issue for the backward equation because this equation describes the evolution of the derivative's price backwards in time and these prices usually stay regular on the boundary. For the calibration we have to solve the forward equation and the probability distribution in time. This distribution diverges at the lower boundary if the Feller constraint is violated, therefore standard boundary conditions like Dirichlet and von Neumann do not work.

In this case it is reasonable to apply coordinate transformation to mitigate the issues.

3.3.1 Transformed Square Root Process

The divergence as $\nu \rightarrow 0$ suggests to look for an equation for $q = \nu^{1-\alpha}p$. Using

$$\begin{aligned}\partial_\nu(\nu p) &= \partial_\nu(\nu^\alpha q) = \alpha\nu^{\alpha-1}q + \nu^\alpha\partial_\nu q \\ \partial_\nu^2(\nu p) &= \alpha(\alpha-1)\nu^{\alpha-2}q + 2\alpha\nu^{\alpha-1}\partial_\nu q + \nu^\alpha\partial_\nu^2 q \\ \partial_\nu(p) &= \partial_\nu(\nu^{\alpha-1}q) = (\alpha-1)\nu^{\alpha-2}q + \nu^{\alpha-1}\partial_\nu q\end{aligned}$$

we find

$$\begin{aligned}\partial_t p &= \left(\frac{\sigma^2}{2}\alpha(\alpha-1)\nu^{\alpha-2}q + \sigma^2\alpha\nu^{\alpha-1}\partial_\nu q + \frac{\sigma^2}{2}\nu^\alpha\partial_\nu^2 q \right) + \\ &\quad (\kappa\alpha\nu^{\alpha-1}q + \kappa\nu^\alpha\partial_\nu q) - (\kappa\theta(\alpha-1)\nu^{\alpha-2}q + \kappa\theta\nu^{\alpha-1}\partial_\nu q) \\ &= \nu^{\alpha-1}\frac{\sigma^2}{2}\nu\partial_\nu^2 q + \nu^{\alpha-1}(\sigma^2\alpha + \kappa\nu - \kappa\theta)\partial_\nu q + \nu^{\alpha-1}\frac{2\kappa^2\theta}{\sigma^2}q \\ &\Rightarrow \partial_t q = \frac{\sigma^2}{2}\nu\partial_\nu^2 q + \kappa(\nu + \theta)\partial_\nu q + \frac{2\kappa^2\theta}{\sigma^2}q\end{aligned}\tag{9}$$

This equation has the stationary solution

$$\hat{q}(\nu) = \beta^\alpha \exp(-\beta\nu)\Gamma(\alpha)^{-1}$$

which converges to $\beta^\alpha\Gamma(\alpha)^{-1}$ as $\nu \rightarrow 0$.

3.3.2 Log Square Root Process

Now look at the stochastic process for $z = \log \nu$, which by Itô's lemma is given by

$$dz = \left((\kappa\theta - \frac{\sigma^2}{2})\frac{1}{\nu} - \kappa \right) dt + \sigma \frac{1}{\sqrt{\nu}} dW$$

The forward Fokker-Planck equation for the probability distribution $f : \mathbb{R} \times \mathbb{R}_{\geq 0} \rightarrow \mathbb{R}_{\geq 0}$, $(z, t) \mapsto f(z, t)$ of this process is given by

$$\partial_t f(z, t) = -\partial_z \left((\kappa\theta - \frac{\sigma^2}{2})\frac{1}{\nu} - \kappa \right) f + \partial_z^2 \left(\frac{\sigma^2}{2} \frac{1}{\nu} f \right)\tag{10}$$

with $\nu = \exp(z)$. This can be rewritten as

$$\partial_t f = \frac{\sigma^2}{2} \frac{1}{\nu} \partial_z^2 f + \left((-\kappa\theta - \frac{\sigma^2}{2})\frac{1}{\nu} + \kappa \right) \partial_z f + \kappa\theta \frac{1}{\nu} f\tag{11}$$

for the use as a starting point for a finite difference implementation. This equation has the stationary solution

$$\hat{f}(z) = \beta^\alpha \exp(z\alpha) \exp(-\beta \exp(z))\Gamma(\alpha)^{-1} = \nu\hat{p}(\nu)$$

In a similar fashion as above this can be checked by observing

$$\begin{aligned}\partial_z^2(e^{-z}\hat{f}) &= \partial_z((\alpha-1)e^{-z}-\beta)\hat{f} \\ &= -(\alpha-1)e^{-z}\hat{f} + ((\alpha-1)e^{-z}-\beta)\partial_z\hat{f}\end{aligned}$$

and

$$\partial_z((\kappa\theta - \frac{\sigma^2}{2})e^{-z} - \kappa)\hat{f} = -(\kappa\theta - \frac{\sigma^2}{2})e^{-z}\hat{f} + ((\kappa\theta - \frac{\sigma^2}{2})e^{-z} - \kappa)\partial_z\hat{f}$$

Since

$$\frac{\sigma^2}{2}(\alpha-1) = \kappa\theta - \frac{\sigma^2}{2}$$

and

$$\frac{\sigma^2}{2}\beta = \kappa$$

we see that $\partial_t\hat{f} = 0$. Normation of \hat{p} and \hat{f} can be easily checked via the definition of the Γ function.

Because of the extra factor of ν \hat{f} converges to 0 as $z \rightarrow -\infty$ (corresponding to $\nu \rightarrow 0$) if $\alpha > 0$, even if the Feller constraint is violated. The same applies for general (non-stationary) solutions of both equations.

$f(z) = \nu p(\nu)$ is actually a more general result, when considering the cumulative distribution functions and changing variables $z = \log \nu, dz = \nu^{-1}d\nu$:

$$F(x) = \int_{-\infty}^x f(z)dz = \int_0^{e^x} f(\log \nu)\nu^{-1}d\nu = \int_0^{e^x} p(\nu)d\nu = P(e^x)$$

4 Discretization of the Fokker-Planck Equation

It is a standard technique to use non-uniform grids to improve convergence at critical points without having to introduce too many grid points.

The class `Concentrating1dMesher` had already offered constructing meshes with one concentrating point. This class was extended to allow multiple concentrating points according to [3].

On a non-uniform grid $\{z_1, \dots, z_n\}$ the two-sided approximation of $\partial_z f$ is done by

$$\begin{aligned}\partial_z f(z_i) &\approx \frac{h_{i-1}^2 f_{i+1} + (h_i^2 - h_{i-1}^2) f_i - h_i^2 f_{i-1}}{h_{i-1} h_i (h_{i-1} + h_i)} \\ &= \frac{h_{i-1}}{h_{i-1} + h_i} \frac{f_{i+1} - f_i}{h_i} + \frac{h_i}{h_{i-1} + h_i} \frac{f_i - f_{i-1}}{h_{i-1}}\end{aligned}$$

where $h_i := z_{i+1} - z_i$ and $f_i := f(z_i)$. This can be interpreted as passing a parabola through f_{i-1}, f_i, f_{i+1} or alternatively as a weighted average of

the forward and backward derivative at z_i . The formula collapses to the usual two-sided difference on a uniform grid, but has a lower error on the non-uniform grid.

The second derivative is approximated by

$$\partial_z^2 f(z_i) \approx \frac{h_{i-i}f_{i+1} - (h_{i-1} + h_i)f_i + h_i f_{i-1}}{\frac{1}{2}h_{i-1}h_i(h_{i-1} + h_i)}$$

We finally sort by factors of f_i , which is typically needed within a finite difference implementation. Setting

$$\begin{aligned}\zeta_i &:= h_i h_{i-1} \\ \zeta_i^p &:= h_i(h_{i-1} + h_i) \\ \zeta_i^m &:= h_{i-1}(h_{i-1} + h_i)\end{aligned}$$

we find

$$\begin{aligned}\partial_z f(z_i) &\approx \frac{h_{i-i}}{\zeta_i^p} f_{i+1} + \frac{(h_i - h_{i-1})}{\zeta_i} f_i - \frac{h_i}{\zeta_i^m} f_{i-1} \\ \partial_z^2 f(z_i) &\approx \frac{2}{\zeta_i^p} f_{i+1} - \frac{2}{\zeta_i} f_i + \frac{2}{\zeta_i^m} f_{i-1}\end{aligned}$$

4.1 General Fokker-Planck Equation in 1D

The general Fokker-Planck equation (5) in one dimension has the form

$$\partial_t f = \partial_z^2 A(z) f + \partial_z \hat{B}(z) f \quad (12)$$

When solving the equation on a grid, it is generally more convenient to have the derivative operator act on the function f . The equation now takes the form

$$\partial_t f = A(z) \partial_z^2 f + B(z) \partial_z f + C(z) f \quad (13)$$

with $B(z) = 2\partial_z A(z) + \hat{B}(z)$ and $C(z) = \partial_z \hat{B}(z) + \partial_z^2 A(z)$. Equation (13) has the spacial discretization

$$\begin{aligned}\partial_t f(z_i) &= \frac{2A_i + B_i h_{i-i}}{\zeta_i^p} f_{i+1} + \left(\frac{-2A_i + B_i(h_i - h_{i-1})}{\zeta_i} + C_i \right) f_i \\ &\quad + \frac{2A_i - B_i h_i}{\zeta_i^m} f_{i-1} \\ &=: \gamma_i f_{i+1} + \beta_i f_i + \alpha_i f_{i-1}\end{aligned}$$

This is interpreted as a tridiagonal transfer matrix T with diagonal β_i , upper diagonal γ_i , and lower diagonal α_i :

$$T := \begin{pmatrix} \beta_1 & \gamma_1 & 0 & \dots & & \\ \alpha_2 & \beta_2 & \gamma_2 & 0 & \dots & \\ 0 & \alpha_3 & \beta_3 & \gamma_3 & 0 & \dots \\ \vdots & 0 & \ddots & \ddots & \ddots & 0 \\ & \vdots & & \alpha_{n-1} & \beta_{n-1} & \gamma_{n-1} \\ & & & & \alpha_n & \beta_n \end{pmatrix}$$

Then the time evolution of the discretized probability distribution function

$$\mathbf{p} := \begin{pmatrix} p_1 \\ \vdots \\ p_n \end{pmatrix}$$

on the lattice $\{z_i\}$ is given by

$$\partial_t \mathbf{p} = T \mathbf{p}$$

The Fokker-Planck equations for the square root process (7), the transformed square root process (9), and the square root process in log coordinates (10) are of the general form (13).

4.2 Boundary Condition in one Dimension

The discretization at the boundary cannot use either f_0 or f_{n+1} , because it is not an element of the discretization. Therefore some boundary condition needs to be applied. While often open boundary conditions are used by setting $\partial_z^2 f = 0$ at the boundary and using a one sided first order derivative for $\partial_z f$, the zero flux condition at the boundary is a natural choice for this sort of problem. These conditions avoid leaking of probability mass via the boundaries.

Zero Flux Condition Heuristically these boundary conditions can be derived as follows. No probability weight shall flow from the boundaries of our discretization, i.e. we want the weight within our boundaries to be constant. We add an artificial z_0 below the lower boundary and z_{n+1} above the upper boundary to the grid. Then

$$\partial_t \int_{z_0}^{z_{n+1}} f(z) dz \stackrel{!}{=} 0 .$$

From equation (12) we have

$$\begin{aligned} \partial_t \int_{z_0}^{z_{n+1}} f(z) dz &= \int_{z_0}^{z_{n+1}} \left(\partial_z^2 A(z) f + \partial_z \hat{B}(z) f \right) dz \\ &= \left(\partial_z A(z) f + \hat{B}(z) f \right) \Big|_{z=z_0, z_{n+1}} \end{aligned}$$

The zero flux condition therefore takes the general form

$$[A(z)\partial_z f + B'(z)f] \Big|_{z=z_0, z_{n+1}} = 0$$

with $B' = \hat{B} + \partial_z A$. For a more rigorous derivation of the zero flux boundary condition see [2].

Lower Boundary Let us now look into the discretization of the zero flux boundary conditions. The partial derivative is discretized by a second order forward differentiation, so that all terms are given by grid points

$$\begin{aligned} \partial_z f(z_0) &\approx \frac{-h_0^2 f_2 + (h_1 + h_0)^2 f_1 - ((h_1 + h_0)^2 - h_0^2) f_0}{h_0 h_1 (h_1 + h_0)} \\ &= -\frac{h_0}{\zeta_1^p} f_2 + \frac{(h_0 + h_1)}{\zeta_1} f_1 - \frac{(2h_0 + h_1)}{\zeta_1^m} f_0 \end{aligned}$$

The general zero-flux boundary condition is therefore discretized at the lower boundary as

$$\begin{aligned} 0 &= -\frac{h_0}{\zeta_1^p} A_0 f_2 + \frac{(h_0 + h_1)}{\zeta_1} A_0 f_1 + \left(-\frac{(2h_0 + h_1)}{\zeta_1^m} A_0 + B'_0 \right) f_0 \\ &=: c_1 f_2 + b_1 f_1 + a_1 f_0 \end{aligned}$$

Now

$$\begin{aligned} f_0 &= -\frac{c_1}{a_1} f_2 - \frac{b_1}{a_1} f_1 \\ &= \frac{h_0}{\zeta_1^p} \left(-\frac{(2h_0 + h_1)}{\zeta_1^m} + \frac{B'_0}{A_0} \right)^{-1} f_2 - \frac{h_0 + h_1}{\zeta_1} \left(-\frac{(2h_0 + h_1)}{\zeta_1^m} + \frac{B'_0}{A_0} \right)^{-1} f_1 \end{aligned}$$

and this enables us to remove f_0 from the discretization. Note that the result has a nicely concentrated form of the specifics of the boundary condition in the form of $\frac{B'_0}{A_0}$, the rest is generic. This might not be the best way to compute the result, whenever A_0 is small.

The transfer matrix is modified by

$$\partial_t f_1 = \gamma_1 f_2 + \beta_1 f_1 + \alpha_1 f_0 \quad (14)$$

$$= (\gamma_1 - \alpha_1 \frac{c_1}{a_1}) f_2 + (\beta_1 - \alpha_1 \frac{b_1}{a_1}) f_1 \quad (15)$$

$$=: (\gamma_1 + \Delta\gamma_1) f_2 + (\beta_1 + \Delta\beta_1) f_1 \quad (16)$$

with

$$\begin{aligned} \Delta\gamma_1 &= \alpha_1 \frac{h_0}{\zeta_1^p} \frac{1}{-\frac{(2h_0+h_1)}{\zeta_1^m} + \frac{B'_0}{A_0}} = \frac{h_0}{\zeta_1^p} \frac{2A_1 - B_1 h_1}{\zeta_1^m} \frac{1}{-\frac{(2h_0+h_1)}{\zeta_1^m} + \frac{B'_0}{A_0}} \\ \Delta\beta_1 &= -\alpha_1 \frac{h_0 + h_1}{\zeta_1} \frac{1}{-\frac{(2h_0+h_1)}{\zeta_1^m} + \frac{B'_0}{A_0}} = -\frac{h_0 + h_1}{\zeta_1} \frac{2A_1 - B_1 h_1}{\zeta_1^m} \frac{1}{-\frac{(2h_0+h_1)}{\zeta_1^m} + \frac{B'_0}{A_0}} \end{aligned}$$

This corrects the factors of f_1 and f_2 , i.e. the diagonal and the upper off diagonal.

Upper Boundary The second order backward differentiation at the upper boundary reads

$$\begin{aligned}\partial_z f(z_{n+1}) &\approx -\frac{-h_n^2 f_{n-1} + (h_n + h_{n-1})^2 f_n - ((h_n + h_{n-1})^2 - h_n^2) f_{n+1}}{h_n h_{n-1} (h_n + h_{n-1})} \\ &= \frac{h_n}{\zeta_n^m} f_{n-1} - \frac{h_{n-1} + h_n}{\zeta_n} f_n + \frac{2h_n + h_{n-1}}{\zeta_n^p} f_{n+1}\end{aligned}$$

The general zero-flux boundary condition is therefore discretized at the upper boundary z_{n+1} as

$$\begin{aligned}0 &= \frac{h_n}{\zeta_n^m} A_{n+1} f_{n-1} - \frac{(h_n + h_{n-1})}{\zeta_n} A_{n+1} f_n + \left(\frac{(2h_n + h_{n-1})}{\zeta_n^p} A_{n+1} + B'_{n+1} \right) f_{n+1} \\ &=: c_n f_{n-1} + b_n f_n + a_n f_{n+1}\end{aligned}$$

Now

$$\begin{aligned}f_{n+1} &= -\frac{c_n}{a_n} f_{n-1} - \frac{b_n}{a_n} f_n \\ &= -\frac{h_n}{\zeta_n^m} \left(\frac{(2h_n + h_{n-1})}{\zeta_n^p} + \frac{B'_{n+1}}{A_{n+1}} \right)^{-1} f_{n-1} + \\ &\quad \frac{h_n + h_{n-1}}{\zeta_n} \left(\frac{(2h_n + h_{n-1})}{\zeta_n^p} + \frac{B'_{n+1}}{A_{n+1}} \right)^{-1} f_n\end{aligned}$$

and this again enables us to remove f_{n+1} .

$$\begin{aligned}\partial_t f_n &= \gamma_n f_{n+1} + \beta_n f_n + \alpha_n f_{n-1} \\ &= \left(\beta_n - \gamma_n \frac{c_n}{a_n} \right) f_n + \left(\alpha_n - \gamma_n \frac{b_n}{a_n} \right) f_{n-1} \\ &=: (\beta_n + \Delta\beta_n) f_n + (\alpha_n + \Delta\alpha_n) f_{n-1}\end{aligned}$$

with

$$\begin{aligned}\Delta\alpha_n &= -\gamma_n \frac{h_n}{\zeta_n^m} \frac{1}{\frac{(2h_n + h_{n-1})}{\zeta_n^p} + \frac{B'_{n+1}}{A_{n+1}}} \\ &= -\frac{h_n}{\zeta_n^m} \frac{2A_n + B_n h_{n-1}}{\zeta_n^p} \frac{1}{\frac{(2h_n + h_{n-1})}{\zeta_n^p} + \frac{B'_{n+1}}{A_{n+1}}} \\ \Delta\beta_n &= \gamma_n \frac{h_{n-1} + h_n}{\zeta_n} \frac{1}{\frac{(2h_n + h_{n-1})}{\zeta_n^p} - \frac{B'_{n+1}}{A_{n+1}}} \\ &= \frac{h_{n-1} + h_n}{\zeta_n} \frac{2A_n + B_n h_{n-1}}{\zeta_n^m} \frac{1}{\frac{(2h_n + h_{n-1})}{\zeta_n^p} - \frac{B'_{n+1}}{A_{n+1}}}\end{aligned}$$

This corrects the factors of f_n and f_{n-1} , i.e. the diagonal and the lower off diagonal.

Boundary Factors We define the common factors of the boundary correction as the upper and the lower boundary factors

$$\begin{aligned}\text{bf}^d &:= A_0 \frac{\alpha_1}{a_1} \\ \text{bf}^u &:= A_{n+1} \frac{\gamma_n}{a_n}\end{aligned}$$

then

$$\begin{aligned}\Delta\beta_1 &= -\frac{h_0 + h_1}{\zeta_1} \text{bf}^d \\ \Delta\gamma_1 &= \frac{h_0}{\zeta_1^p} \text{bf}^d \\ \Delta\alpha_n &= -\frac{h_n}{\zeta_n^m} \text{bf}^u \\ \Delta\beta_n &= \frac{h_{n-1} + h_n}{\zeta_n} \text{bf}^u\end{aligned}$$

4.2.1 Square Root Process

From the Fokker-Planck equation (8) we read

$$\begin{aligned}A(\nu) &= \frac{\sigma^2}{2} \nu \\ B(\nu) &= \sigma^2 - \kappa(\theta - \nu) \\ C(\nu) &= \kappa\end{aligned}$$

Therefore

$$\begin{aligned}\mu_i &:= B(\nu_i) = \sigma^2 - \kappa(\theta - \nu_i) \\ \alpha_i &= \frac{2A_i - B_i h_i}{\zeta_i^m} = \frac{\sigma^2 \nu_i - \mu_i h_i}{\zeta_i^m} \\ \beta_i &= \left(\frac{-2A_i + B_i(h_i - h_{i-1})}{\zeta_i} + C_i \right) = \left(\frac{-\sigma^2 \nu_i + \mu_i(h_i - h_{i-1})}{\zeta_i} + \kappa \right) \\ \gamma_i &= \frac{2A_i + B_i h_{i-1}}{\zeta_i^p} = \frac{\sigma^2 \nu_i + \mu_i h_{i-1}}{\zeta_i^p}\end{aligned}$$

The zero-flux boundary condition for the square root process - using equation (7) - is

$$\left[\frac{\sigma^2}{2} \partial_\nu (\nu p) + \kappa(\nu - \theta) p \right] \Big|_{\nu=\nu_0, \nu_{n+1}} = 0$$

i.e.

$$\left[\frac{\sigma^2}{2} \nu \partial_\nu p + \left(\kappa(\nu - \theta) + \frac{\sigma^2}{2} \right) p \right] \Big|_{\nu=\nu_0, \nu_{n+1}} = 0$$

i.e.

$$B'(\nu) = \kappa(\nu - \theta) + \frac{\sigma^2}{2}$$

Therefore

$$\begin{aligned}\beta_1 &\rightarrow \beta_1 + \Delta\beta_1 \\ \gamma_1 &\rightarrow \gamma_1 + \Delta\gamma_1 \\ \alpha_n &\rightarrow \alpha_n + \Delta\alpha_n \\ \beta_n &\rightarrow \beta_n + \Delta\beta_n\end{aligned}$$

with

$$\begin{aligned}\Delta\gamma_1 &= \frac{h_0}{\zeta_1^p} \frac{2A_1 - B_1 h_1}{\zeta_1^m} \frac{1}{-\frac{(2h_1+h_0)}{\zeta_1^m} + \frac{B'_0}{A_0}} \\ &= \frac{h_0}{\zeta_1^p} \frac{\sigma^2 \nu_1 - \mu_1 h_1}{\zeta_1^m} \frac{1}{-\frac{(2h_0+h_1)}{\zeta_1^m} + \frac{\kappa(\nu_0-\theta) + \frac{\sigma^2}{2}}{\frac{\sigma^2}{2}\nu_0}} \\ \Delta\beta_1 &= -\frac{h_0 + h_1}{\zeta_1} \frac{2A_1 - B_1 h_1}{\zeta_1^m} \frac{1}{-\frac{(2h_0+h_1)}{\zeta_1^m} + \frac{B'_0}{A_0}} \\ &= -\frac{h_0 + h_1}{\zeta_1} \frac{\sigma^2 \nu_1 - \mu_1 h_1}{\zeta_1^m} \frac{1}{-\frac{(2h_0+h_1)}{\zeta_1^m} + \frac{\kappa(\nu_0-\theta) + \frac{\sigma^2}{2}}{\frac{\sigma^2}{2}\nu_0}}\end{aligned}$$

Since ν_0 is small it is better to calculate this as:

$$\begin{aligned}\Delta\gamma_1 &= \frac{h_0}{\zeta_1^p} \frac{\sigma^2 \nu_1 - \mu_1 h_1}{\zeta_1^m} \frac{\nu_0}{-\frac{(2h_0+h_1)\nu_0}{\zeta_1^m} + \frac{2\kappa(\nu_0-\theta) + \sigma^2}{\sigma^2}} \\ \Delta\beta_1 &= -\frac{h_0 + h_1}{\zeta_1} \frac{\sigma^2 \nu_1 - \mu_1 h_1}{\zeta_1^m} \frac{\nu_0}{-\frac{(2h_0+h_1)\nu_0}{\zeta_1^m} + \frac{2\kappa(\nu_0-\theta) + \sigma^2}{\sigma^2}}\end{aligned}$$

For the upper boundary

$$\begin{aligned}
\Delta\alpha_n &= -\frac{h_n}{\zeta_n^p} \frac{2A_n + B_n h_{n-1}}{\zeta_n^p} \frac{1}{\frac{(2h_n+h_{n-1})}{\zeta_n^m} + \frac{B'_{n+1}}{A_{n+1}}} \\
&= -\frac{h_n}{\zeta_n^p} \frac{\sigma^2\nu_n + \mu_n h_{n-1}}{\zeta_n^p} \frac{1}{\frac{(2h_n+h_{n-1})}{\zeta_n^m} + \frac{\kappa(\nu_{n+1}-\theta) + \frac{\sigma^2}{2}}{\frac{\sigma^2}{2}\nu_{n+1}}} \\
&= -\frac{h_n}{\zeta_n^p} \frac{\sigma^2\nu_n + \mu_n h_{n-1}}{\zeta_n^p} \frac{1}{\frac{(2h_n+h_{n-1})}{\zeta_n^m} + \frac{2\kappa(\nu_{n+1}-\theta) + \sigma^2}{\sigma^2\nu_{n+1}}} \\
\Delta\beta_n &= \frac{h_{n-1} + h_n}{\zeta_n} \frac{2A_n + B_n h_{n-1}}{\zeta_n^p} \frac{1}{\frac{(2h_n+h_{n-1})}{\zeta_n^m} - \frac{B'_{n+1}}{A_{n+1}}} \\
&= \frac{h_{n-1} + h_n}{\zeta_n} \frac{\sigma^2\nu_n + \mu_n h_{n-1}}{\zeta_n^p} \frac{1}{\frac{(2h_n+h_{n-1})}{\zeta_n^m} + \frac{2\kappa(\nu_{n+1}-\theta) + \sigma^2}{\sigma^2\nu_{n+1}}}
\end{aligned}$$

4.2.2 Transformed Square Root Process

From equation (9) we take

$$\begin{aligned}
A(\nu) &= \frac{\sigma^2}{2}\nu \\
B(\nu) &= \kappa(\nu + \theta) \\
C(\nu) &= \frac{2\kappa^2\theta}{\sigma^2}
\end{aligned}$$

Therefore

$$\begin{aligned}
\mu_i &:= B(\nu_i) = \kappa(\nu_i + \theta) \\
\alpha_i &= \frac{2A_i - B_i h_i}{\zeta_i^m} = \frac{\sigma^2\nu_i - \mu_i h_i}{\zeta_i^m} \\
\beta_i &= \left(\frac{-2A_i + B_i(h_i - h_{i-1})}{\zeta_i} + C_i \right) = \left(\frac{-\sigma^2\nu_i + \mu_i(h_i - h_{i-1})}{\zeta_i} + \frac{2\kappa^2\theta}{\sigma^2} \right) \\
\gamma_i &= \frac{2A_i + B_i h_{i-1}}{\zeta_i^p} = \frac{\sigma^2\nu_i + \mu_i h_{i-1}}{\zeta_i^p}
\end{aligned}$$

The zero-flux boundary condition for the transformed density is derived from the one for p

$$\begin{aligned}
&\left[\frac{\sigma^2}{2}\nu\partial_\nu\nu^{\alpha-1}q + \left(\kappa(\nu - \theta) + \frac{\sigma^2}{2} \right) \nu^{\alpha-1}q \right] \Big|_{\nu=\nu_0, \nu_{n+1}} = 0 \\
&\nu\partial_\nu\nu^{\alpha-1}q = (\alpha - 1)\nu^{\alpha-1}q + \nu\nu^{\alpha-1}\partial_\nu q \\
&\frac{\sigma^2}{2}\nu\partial_\nu q + \frac{\sigma^2}{2}(\alpha - 1)q + \left(\kappa(\nu - \theta) + \frac{\sigma^2}{2} \right) q = \frac{\sigma^2}{2}\nu\partial_\nu q + \kappa\nu q
\end{aligned}$$

i.e.

$$B'(\nu) = \kappa\nu$$

4.2.3 Log Square Root Process

From the Fokker-Planck equation (11) we read

$$\begin{aligned} A(z) &= \frac{\sigma^2}{2} \exp(-z) \\ B(z) &= \left(-\frac{\sigma^2}{2} - \kappa\theta\right) \exp(-z) + \kappa \\ C(z) &= \kappa\theta \exp(-z) \end{aligned}$$

Therefore

$$\begin{aligned} \mu_i &:= B(z_i) = \left(-\frac{\sigma^2}{2} - \kappa\theta\right) \exp(-z_i) + \kappa \\ \alpha_i &= \frac{2A_i - B_i h_i}{\zeta_i^m} = \frac{\sigma^2 \exp(-z_i) - \mu_i h_i}{\zeta_i^m} \\ \beta_i &= \left(\frac{-2A_i + B_i(h_i - h_{i-1})}{\zeta_i} + C_i \right) \\ &= \left(\frac{-\sigma^2 \exp(-z_i) + \mu_i(h_i - h_{i-1})}{\zeta_i} + \kappa\theta \exp(-z_i) \right) \\ \gamma_i &= \frac{2A_i + B_i h_{i-1}}{\zeta_i^p} = \frac{\sigma^2 \exp(-z_i) + \mu_i h_{i-1}}{\zeta_i^p} \end{aligned}$$

Zero Flux Condition As discussed above we impose

$$\partial_t \int_{z_0}^{z_{n+1}} f(z) dz \stackrel{!}{=} 0$$

Now

$$\begin{aligned} \int_{z_0}^{z_{n+1}} \partial_t f(z) dz &= \int_{z_0}^{z_{n+1}} \left(-\partial_z \left(\left(\kappa\theta - \frac{\sigma^2}{2} \right) \frac{1}{\nu} - \kappa \right) f + \partial_z^2 \left(\frac{\sigma^2}{2} \frac{1}{\nu} f \right) \right) dz \\ &= \left(-\left(\left(\kappa\theta - \frac{\sigma^2}{2} \right) \frac{1}{\nu} - \kappa \right) f + \partial_z \left(\frac{\sigma^2}{2} \frac{1}{\nu} f \right) \right) \Big|_{z_0}^{z_{n+1}} \end{aligned}$$

which implies

$$\left(-\left(\left(\kappa\theta - \frac{\sigma^2}{2} \right) \frac{1}{\nu} - \kappa \right) f + \partial_z \left(\frac{\sigma^2}{2} \frac{1}{\nu} f \right) \right) \Big|_{z=z_0, z_{n+1}} \stackrel{!}{=} 0$$

Since

$$\partial_z \left(\frac{\sigma^2}{2} \frac{1}{\nu} f \right) = -\frac{\sigma^2}{2} \frac{1}{\nu} f + \frac{\sigma^2}{2} \frac{1}{\nu} \partial_z f$$

We find for the log process

$$B'(z) = \kappa \left(1 - \frac{\theta}{\nu}\right) = \kappa(1 - \theta \exp(-z))$$

$$\begin{aligned} \Delta\gamma_1 &= \frac{h_0}{\zeta_1^p} \frac{2A_1 - B_1 h_1}{\zeta_1^m} \frac{1}{-\frac{(2h_0+h_1)}{\zeta_1^m} + \frac{B'_0}{A_0}} \\ &= \frac{h_0}{\zeta_1^p} \frac{\sigma^2 \exp(-z_1) - \mu_1 h_1}{\zeta_1^m} \frac{1}{-\frac{(2h_0+h_1)}{\zeta_1^m} + \frac{2\kappa(\exp(z_0)-\theta)}{\sigma^2}} \\ \Delta\beta_1 &= -\frac{h_0 + h_1}{\zeta_1} \frac{2A_1 - B_1 h_1}{\zeta_1^m} \frac{1}{-\frac{(2h_0+h_1)}{\zeta_1^m} + \frac{B'_0}{A_0}} \\ &= -\frac{h_0 + h_1}{\zeta_1} \frac{\sigma^2 \exp(-z_1) - \mu_1 h_1}{\zeta_1^m} \frac{1}{-\frac{(2h_0+h_1)}{\zeta_1^m} + \frac{2\kappa(\exp(z_0)-\theta)}{\sigma^2}} \end{aligned}$$

$$\begin{aligned} \Delta\alpha_n &= -\frac{h_n}{\zeta_n^p} \frac{2A_n + B_n h_{n-1}}{\zeta_n^m} \frac{1}{\frac{(2h_n+h_{n-1})}{\zeta_n^m} + \frac{B'_{n+1}}{A_{n+1}}} \\ &= -\frac{h_n}{\zeta_n^p} \frac{\sigma^2 \exp(-z_n) + \mu_n h_{n-1}}{\zeta_n^m} \frac{\exp(-z_{n+1})}{\frac{(2h_n+h_{n-1}) \exp(-z_{n+1})}{\zeta_n^m} + \frac{2\kappa(1-\theta \exp(-z_{n+1}))}{\sigma^2}} \\ \Delta\beta_n &= \frac{h_{n-1} + h_n}{\zeta_n} \frac{2A_n + B_n h_{n-1}}{\zeta_n^m} \frac{1}{\frac{(2h_n+h_{n-1})}{\zeta_n^m} - \frac{B'_{n+1}}{A_{n+1}}} \\ &= \frac{h_{n-1} + h_n}{\zeta_n} \frac{\sigma^2 \exp(-z_n) + \mu_n h_{n-1}}{\zeta_n^m} \frac{\exp(-z_{n+1})}{\frac{(2h_n+h_{n-1}) \exp(-z_{n+1})}{\zeta_n^m} + \frac{2\kappa(1-\theta \exp(-z_{n+1}))}{\sigma^2}} \end{aligned}$$

4.3 Fokker-Planck Equation of the Heston-Process

The square root process will now play the role of the variance of a stock process $x_t = \log S_t/S_0$

$$\begin{aligned} dx_t &= (r_t - q_t - \frac{\nu_t}{2})dt + \sqrt{\nu_t}dW_t^x \\ d\nu_t &= \kappa(\theta - \nu_t)dt + \sigma\sqrt{\nu_t}dW_t^\nu \\ \rho dt &= dW_t^x dW_t^\nu \end{aligned}$$

It combined process has the following forward Fokker-Planck equation for the probability density $p : \mathbb{R} \times \mathbb{R}_{\geq 0} \times \mathbb{R}_{\geq 0} \rightarrow \mathbb{R}_{\geq 0}$, $(x, \nu, t) \mapsto p(x, \nu, t)$

$$\partial_t p = \frac{1}{2} \partial_x^2 (\nu p) + \left(\frac{\nu}{2} - r_t + q_t\right) \partial_x p + \frac{\sigma^2}{2} \partial_\nu^2 (\nu p) - \partial_\nu (\kappa(\theta - \nu)p) + \partial_x \partial_\nu (\rho \sigma \nu p)$$

Alternatively this can be written as

$$\partial_t p = \frac{\nu}{2} \partial_x^2 p + \left(\frac{\nu}{2} - r_t + q_t + \rho\sigma\right) \partial_x p + \frac{\sigma^2}{2} \nu \partial_\nu^2 p + (\sigma^2 - \kappa(\theta - \nu)) \partial_\nu p + \kappa p + \rho\sigma\nu \partial_x \partial_\nu p$$

which is more suited as a starting point for the implementation of a finite difference scheme.

4.3.1 Transformed Probability Density

Take again $q = \nu^{1-\alpha} p$ in above's Fokker-Planck equation. The derivatives with respect to x do not change, since $\partial_x \nu^{\alpha-1} q = \nu^{\alpha-1} \partial_x q$. Therefore most of the work is done in the preceding section, we only need to look at the term with the mixed derivative:

$$\partial_x \partial_\nu (\nu p) = \partial_x \partial_\nu (\nu^\alpha q) = \alpha \nu^{\alpha-1} \partial_x q + \nu^\alpha \partial_x \partial_\nu q$$

$$\begin{aligned} \partial_t q &= \frac{\nu}{2} \partial_x^2 q + \left(\frac{\nu}{2} - r_t + q_t\right) \partial_x q + \\ &\quad \frac{\sigma^2}{2} \nu \partial_\nu^2 q + \kappa(\nu + \theta) \partial_\nu q + \frac{2\kappa^2\theta}{\sigma^2} q + \\ &\quad \alpha\rho\sigma \partial_x q + \rho\sigma\nu \partial_x \partial_\nu q \\ &= \frac{\nu}{2} \partial_x^2 q + \left(\frac{\nu}{2} - r_t + q_t + \rho\sigma \frac{2\kappa\theta}{\sigma^2}\right) \partial_x q + \\ &\quad \frac{\sigma^2}{2} \nu \partial_\nu^2 q + \kappa(\nu + \theta) \partial_\nu q + \frac{2\kappa^2\theta}{\sigma^2} q + \\ &\quad \rho\sigma\nu \partial_x \partial_\nu q \end{aligned}$$

The zero flux condition takes the form $\forall x$:¹

$$\left[\frac{\sigma^2}{2} \nu \partial_\nu q + \kappa\nu q + \rho\nu\sigma \partial_x q \right] \Big|_{z=z_0, z_{n+1}} = 0$$

4.3.2 Heston Process in $z = \log \nu$

The spot process is the same as above, $x_t = \log S_t/S_0$, instead of the ν process we use the square root process for $z = \log \nu$

$$\begin{aligned} dx_t &= \left(r_t - q_t - \frac{\nu_t}{2}\right) dt + \sqrt{\nu_t} dW_t^x \\ dz_t &= \left(\left(\kappa\theta - \frac{\sigma^2}{2}\right) \frac{1}{\nu} - \kappa\right) dt + \sigma \frac{1}{\sqrt{\nu}} dW_t^\nu \\ \rho dt &= dW_t^x dW_t^\nu \end{aligned}$$

¹See again [2] for a rigorous derivation

The forward Fokker-Planck equation for the probability distribution $f : \mathbb{R} \times \mathbb{R} \times \mathbb{R}_{\geq 0} \rightarrow \mathbb{R}_{\geq 0}$, $(x, z = \log \nu, t) \mapsto f(x, z, t)$ of this process is given by

$$\begin{aligned} \partial_t f(z, t) &= \frac{1}{2} \partial_x^2 (\nu f) + \left(\frac{\nu}{2} - r_t + q_t \right) \partial_x f \\ &\quad - \partial_z \left(\left(\kappa \theta - \frac{\sigma^2}{2} \right) \frac{1}{\nu} - \kappa \right) f + \partial_z^2 \left(\frac{\sigma^2}{2} \frac{1}{\nu} f \right) \\ &\quad + \partial_x \partial_z (\rho \sigma f) \end{aligned}$$

with $\nu = \exp(z)$. This can be rewritten as

$$\begin{aligned} \partial_t f &= \frac{1}{2} \nu \partial_x^2 f + \left(\frac{\nu}{2} - r_t + q_t \right) \partial_x f \\ &\quad + \frac{\sigma^2}{2} \frac{1}{\nu} \partial_z^2 f + \left((-\kappa \theta - \frac{\sigma^2}{2}) \frac{1}{\nu} + \kappa \right) \partial_z f + \kappa \theta \frac{1}{\nu} f \\ &\quad + \rho \sigma \partial_x \partial_z f \end{aligned}$$

The zero-flux boundary condition is

$$\left(\frac{\sigma^2}{2} \frac{1}{\nu} \partial_z f - \kappa \left(1 - \frac{\theta}{\nu} \right) f + \rho \sigma \partial_x f \right) \Big|_{z=z_0, z_{n+1}} \stackrel{!}{=} 0$$

4.4 Discretization of the Heston PDE

The new dimension is discretized in the same way as in the one-dimensional case. Again we do assume a regular discretization. The formulæ from (4) still apply and we end up with a transfer matrix.

$$\begin{aligned} \partial_t f(x_j, z_i, t) &= \gamma_i f_{i+1, j} + \beta_i f_{i, j} + \alpha_i f_{i-1, j} + \\ &\quad c_i f_{i, j+1} + b_i f_{i, j} + a_i f_{i, j-1} + \\ &\quad \text{9-point operator for the mixed derivative} \end{aligned}$$

Note that the coefficients only have labels corresponding to the ν directions, since they have no dependency on x .

The zero flux condition takes the form $\forall x$:

$$\left[\frac{\sigma^2}{2} \nu \partial_z p + \left(\kappa (\nu - \theta) + \frac{\sigma^2}{2} \right) p + \rho \nu \sigma \partial_x p \right] \Big|_{z=z_0, z_{n+1}} = 0$$

which we generalize to

$$\forall x : [A(z) \partial_z f + B'(z) f + C' \partial_x f] \Big|_{z=z_0, z_{n+1}} = 0$$

Using the same three-point approximation for the derivative ∂_z and using the value on the neighboring grid point for the derivative ∂_x , i.e. $C' \partial_x f(x, z_1, t)$

instead of $C'\partial_x f(x, z_0, t)$ and $C'\partial_x f(x, z_n, t)$ instead of $C'\partial_x f(x, z_{n+1}, t)$ the argument of how to integrate the boundary condition into the discretized equation stays the same as for the square root process, we only get the last term as an extra term. Equation (14) is thus modified to

$$\begin{aligned}\partial_t f_1 &= \gamma_1 f_2 + \beta_1 f_1 + \alpha_1 f_0 + \dots \\ &= (\gamma_1 - \alpha_1 \frac{c_1}{a_1}) f_2 + (\beta_1 - \alpha_1 \frac{b_1}{a_1}) f_1 - \alpha_1 \frac{C'}{a_1} \partial_x f_1 + \dots\end{aligned}$$

Because the boundary factors comprise a factor of A_0, A_{n+1} , resp., we have for the last term

$$\alpha_1 \frac{C'}{a_1} = \text{bf}^l \frac{C'}{A_0}$$

and

$$\alpha_n \frac{C'}{a_n} = \text{bf}^u \frac{C'}{A_{n+1}}$$

You will find the implementation of the zero flux boundary condition in `FdmSquareRootOp`.

4.5 Stochastic Local Volatility

As a last extension we add a local leverage function L to the process by replacing $\nu_t \rightarrow \nu_t L^2(x, t)$ in the x_t process. If the ν_t process is constant, this reduces to the well-known local volatility model and the leverage function can be identified with the local volatility surface. A constant ν can be achieved by setting the vol-of-vol $\sigma = 0$ and the start volatility equal to the long-term volatility θ . Without loss of generality we can assume $\nu_t = 1$ in this case. If on the other hand $L \equiv 1$, the model is again the standard Heston model. To make the tuning between the models easier, we will introduce a *mixing parameter* η , by replacing $\sigma \rightarrow \eta\sigma$ later.

$$\begin{aligned}dx_t &= (r_t - q_t - \frac{\nu_t}{2})dt + \sqrt{\nu_t}L(x, t)dW_t^x \\ d\nu_t &= \kappa(\theta - \nu_t)dt + \sigma\sqrt{\nu_t}dW_t^\nu \\ \rho dt &= dW_t^x dW_t^\nu\end{aligned}$$

It has the following forward Fokker-Planck equation for the probability density $p : \mathbb{R} \times \mathbb{R}_{\geq 0} \times \mathbb{R}_{\geq 0} \rightarrow \mathbb{R}_{\geq 0}$, $(x, \nu, t) \mapsto p(x, \nu, t)$

$$\begin{aligned}\partial_t p &= \frac{1}{2}\partial_x^2(\nu L^2 p) + \partial_x(\frac{\nu}{2}L^2 - r_t + q_t)p \\ &\quad + \frac{\sigma^2}{2}\partial_\nu^2(\nu p) - \partial_\nu(\kappa(\theta - \nu)p) \\ &\quad + \partial_x \partial_\nu(\rho \sigma \nu L p)\end{aligned}$$

Now we have an additional choice to make during the implementation. The first is to keep L and p together:

$$\begin{aligned}\partial_t p &= \frac{\nu}{2} \partial_x^2 L^2 p + (-r_t + q_t) \partial_x p + \frac{\nu}{2} \partial_x L^2 p + \rho \sigma \partial_x L p \\ &\quad + \frac{\sigma^2}{2} \nu \partial_\nu^2 p + (\sigma^2 - \kappa(\theta - \nu)) \partial_\nu p + \kappa p \\ &\quad + \rho \sigma \nu \partial_x \partial_\nu L p\end{aligned}\tag{17}$$

The zero flux condition takes the form $\forall x$:

$$\left[\frac{\sigma^2}{2} \nu \partial_z p + \left(\kappa(\nu - \theta) + \frac{\sigma^2}{2} \right) p + \rho \nu \sigma \partial_x L p \right] \Big|_{z=z_0, z_{n+1}} = 0$$

The extension of the zero flux boundary condition to the two dimensional case can be found in `FdmHestonFwdOp`.

4.5.1 Local Volatility Model

As a side-remark and because the local volatility model is a limiting case of the SLV model, we note the equations for this model here. The leverage function L is identified with the local volatility surface $L(x, t) \equiv \sigma(x, t)$:

$$dx_t = (r_t - q_t - \frac{\nu_t}{2}) dt + \sigma(x, t) dW_t^x$$

It has the following forward Fokker-Planck equation for the probability density $p : \mathbb{R} \times \mathbb{R}_{\geq 0} \times \mathbb{R}_{\geq 0} \rightarrow \mathbb{R}_{\geq 0}$, $(x, \nu, t) \mapsto p(x, \nu, t)$

$$\begin{aligned}\partial_t p &= \frac{1}{2} \partial_x^2 (\sigma(x, t)^2 p) + \partial_x \left(\frac{1}{2} \sigma(x, t)^2 - r_t + q_t \right) p \\ &= \frac{1}{2} \partial_x^2 (\sigma(x, t)^2 p) + (-r_t + q_t) \partial_x p + \frac{1}{2} \partial_x (\sigma(x, t)^2 p)\end{aligned}$$

An implementation of the Fokker-Planck operator can be found in the new class `FdmLocalVolFwdOp`.

4.5.2 Transformed Density Function

$$\begin{aligned}\partial_t q &= \frac{\nu}{2} \partial_x^2 L^2 q + (-r_t + q_t) \partial_x q + \partial_x \left(\frac{\nu}{2} L^2 + \rho \sigma \frac{2\kappa\theta}{\sigma^2} L \right) q + \\ &\quad \frac{\sigma^2}{2} \nu \partial_\nu^2 q + \kappa(\nu + \theta) \partial_\nu q + \frac{2\kappa^2\theta}{\sigma^2} q + \\ &\quad \rho \sigma \nu \partial_x \partial_\nu L q\end{aligned}$$

The zero flux condition takes the form $\forall x$:

$$\left[\frac{\sigma^2}{2} \nu \partial_\nu q + \kappa \nu q + \rho \nu \sigma \partial_x L q \right] \Big|_{z=z_0, z_{n+1}} = 0$$

4.5.3 Fokker-Planck Equation in $\log \nu$

$$\begin{aligned}\partial_t f(z, t) &= \frac{1}{2} \partial_x^2 (\nu L^2 f) + \partial_x \left(\frac{\nu}{2} L^2 - r_t + q_t \right) f \\ &\quad - \partial_z \left(\left(\kappa \theta - \frac{\sigma^2}{2} \right) \frac{1}{\nu} - \kappa \right) f + \partial_z^2 \left(\frac{\sigma^2}{2} \frac{1}{\nu} f \right) \\ &\quad + \partial_x \partial_z (\rho \sigma L f)\end{aligned}$$

with $\nu = \exp(z)$. This can be rewritten as

$$\begin{aligned}\partial_t f &= \frac{1}{2} \nu \partial_x^2 L^2 f + (-r_t + q_t) \partial_x f + \frac{\nu}{2} \partial_x L^2 f \\ &\quad + \frac{\sigma^2}{2} \frac{1}{\nu} \partial_z^2 f + \left((-\kappa \theta - \frac{\sigma^2}{2}) \frac{1}{\nu} + \kappa \right) \partial_z f + \kappa \theta \frac{1}{\nu} f \\ &\quad + \rho \sigma \partial_x \partial_z L f\end{aligned}$$

The zero-flux boundary condition is

$$\left(\frac{\sigma^2}{2} \frac{1}{\nu} \partial_z f - \kappa \left(1 - \frac{\theta}{\nu} \right) f + \rho \sigma \partial_x L f \right) \Big|_{z=z_0, z_{n+1}} \stackrel{!}{=} 0$$

4.6 Non-uniform Meshes

Adaptive grids are useful, especially to improve the calibration. The start condition for the Fokker-Planck equation is a highly singular - and concentrated - Greens function. Therefore one might expect, that a grid concentrated at the origin will improve the calculation. Also special care has to be taken for the lower bound if the Feller constraint is violated.

Implementation follows [Tavella & Randall 2000] [3] utilising Peter Casper's Runge-Kutta solver for the ordinary differential equation

$$\begin{aligned}\frac{dY(\epsilon)}{d\epsilon} &= A \left[\sum_{k=1}^n J_k(\epsilon)^{-2} \right]^{-\frac{1}{2}} \\ J_k(\epsilon) &= \sqrt{\beta^2 + (Y(\epsilon) - B_k)^2} \\ Y(0) &= Y_{min} \\ Y(1) &= Y_{max}\end{aligned}$$

based on the coordinate transformation $Y = Y(\epsilon)$ for n critical points B_k with density factors β_k and A chosen such that $Y(1) = Y_{max}$.

4.7 Adaptive Grid Sizes

The probability density is highly concentrated at the beginning and blurs out over a much larger scale later on. The evaluation of the probability

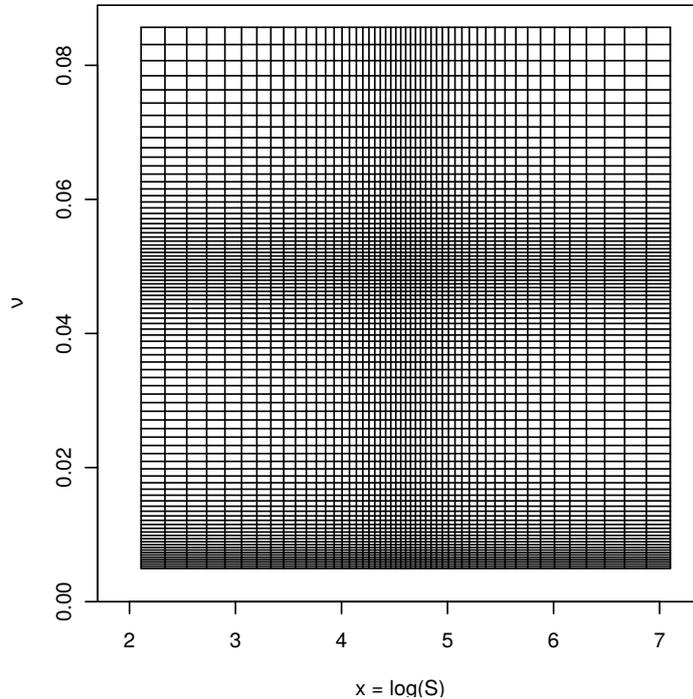


Figure 2: Example for a non-uniform mesh: $x_0 = \ln(100)$, $\nu_0 = 0.05$, Feller constraint is fulfilled

density function is therefore a multi scale problem and best dealt with using an adaptive grid size. The grid should only be defined on areas where the probability density function is not negligibly small or zero. Defining this area is a sort of an chicken and egg problem because the density is what we want to calculate and we do not know in advance where is significant contributions are. Therefore we solve the problem separately for $\ln S$ and ν direction. The ν direction is simple because the Fokker-Planck equation for the square root process has the known closed form solution

$$\nu_t = \frac{\sigma^2 (1 - e^{-\kappa t})}{4\kappa} \chi_d^2 \left(\frac{4\kappa e^{-\kappa t}}{\sigma^2 (1 - e^{-\kappa t})} \nu_0 \right), d = \frac{4\theta\kappa}{\sigma^2}. \quad (18)$$

The high quality implementation of the non-central χ -square distribution in the `boost` library allows to calculate proper quantiles for the distribution density.

It would be difficult to extract the same information for $\ln S$ out of the stochastic local volatility equation (17) without solving the equation. But the same information is already encoded in the implied or local volatility surface of the market. If e.g. the local volatility surface is given then the

one dimensional Fokker-Planck equation can be solved quite easily and the boundaries for the grid size in $\ln S$ can be derived from there.

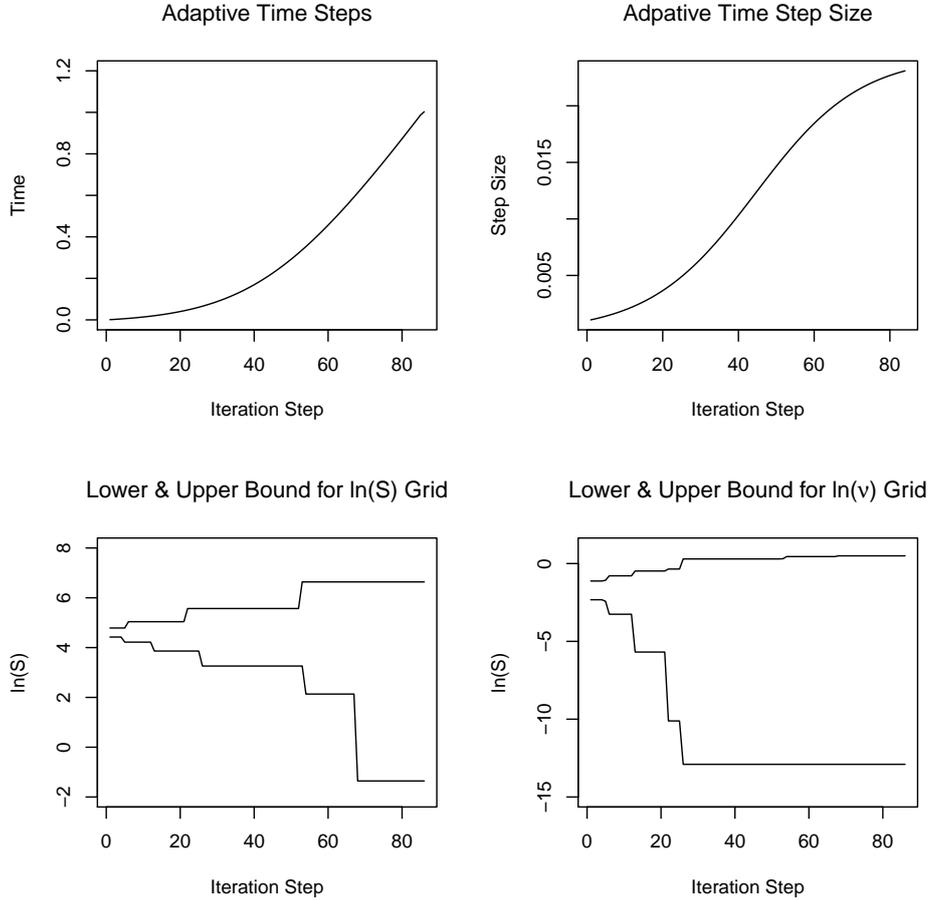


Figure 3: Example for adaptive time and grid meshes used for a realistic Heston local volatility calibration. $x_0 = \ln(100)$, $r = 0.01$, $q = 0.02$, $\kappa = 2.0$, $\theta = 0.074$, $\rho = -0.51$, $\sigma = 0.8$, $\nu_0 = 0.1974$. The market implied volatility surface is shown in figure (4)

The figures (3) show the adaptive time and grid meshes used in a realistic calibration of the Heston stochastic local volatility model.

4.8 Adaptive Time Steps

Clearly the first steps sizes should be small compared with later steps when the density in ν direction is closed to the stationary probability density. The

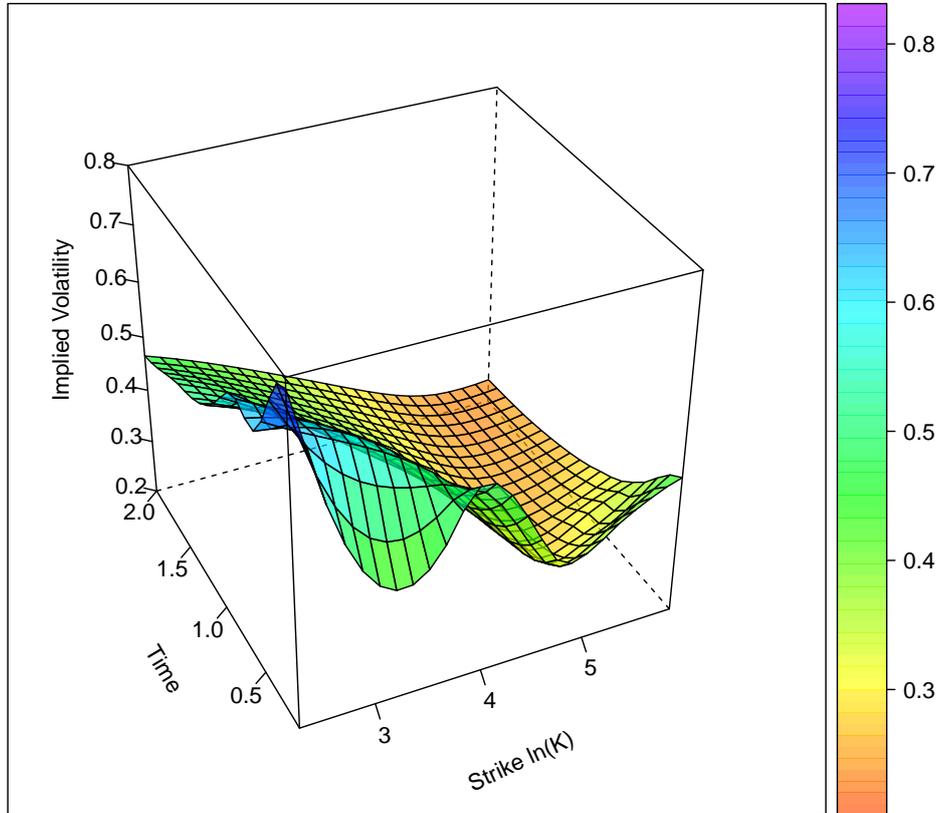


Figure 4: Implied volatility surface for Heston local volatility calibration.

following parametrization of the time steps turned out to be useful

$$\Delta t_i = (1 - e^{-\lambda_i}) \Delta t_m + e^{-\lambda_i} \Delta t_s \quad (19)$$

5 Calibration

The calibration procedure follows Tian et al [4]. A short summary of the calibration procedure is the following. The key to the calibration is equation (3). First we calibrate the Heston model and the Local Volatility model independently. The local volatility model can in principle calibrate exactly to vanilla options, while the Heston model might not produce enough skew at the short end.

Now start with the short dated end of the leverage function by evolving the probability density using the Fokker-Planck equation and integrating the density as in equation (3). The implementation of the calibration can be found in `HestonSLVModel` in method `performCalculation`. This method is called by the `LazyObject` mechanism, if `leverageFunction` is called and a recalculation is necessary. Additionally we have introduced a new local volatility object to store the information of the leverage function in an efficient way, when the grids are non-uniform and adaptive in size. The implementation is found in `FixedLocalVolSurface`.

The mixing parameter could additionally be calibrated by using prices of liquid exotics like barrier options. This is not implemented.

6 Results and Outlook

We have implemented several test cases in the QuantLib test framework. They can be found in the class `HestonSLVModelTest`. The most important one takes a flat local vol surface of 30% and a Heston Model with parameters $S_0 = 100$, $\sqrt{\nu_0} = 24.5\%$, $\kappa = 1$, $\theta = \nu_0$, $\sigma^2 = 0.2$, $\rho = -75\%$. The corresponding implied vol surface is shown in figure 5. It then calibrates the leverage function, which has to flatten the implied vol surface given by the Heston model. The resulting leverage function is shown in figure 6. Afterwards we price a set of vanilla option using this calibrated SLV model to test for the round trip error. We found that using suitable parameters the error in implied volatilities is less than 5 bp.

The mixing parameter is not used during calibration at the moment. Calibrating to exotics is left to futuer extensions of the SLV model.

References

- [1] William Feller. Two singular diffusion problems. *The Annals of Mathematics*, 54(1):173–182, 1951.
- [2] Vladimir Lucic. Boundary conditions for computing densities in hybrid models via PDE methods. *Stochastics*, 84(5-6):705–718, 2012.
- [3] D. Tavella and C. Randall. *Pricing Financial Instruments: The Finite Difference Method*. Wiley, 2000.
- [4] Yu Tian, Zili Zhu, Geoffrey Lee, Fima Klebaner, and Kais Hamza. Calibrating and Pricing with a Stochastic-Local Volatility Model., 2014. <http://ssrn.com/abstract=2182411>.

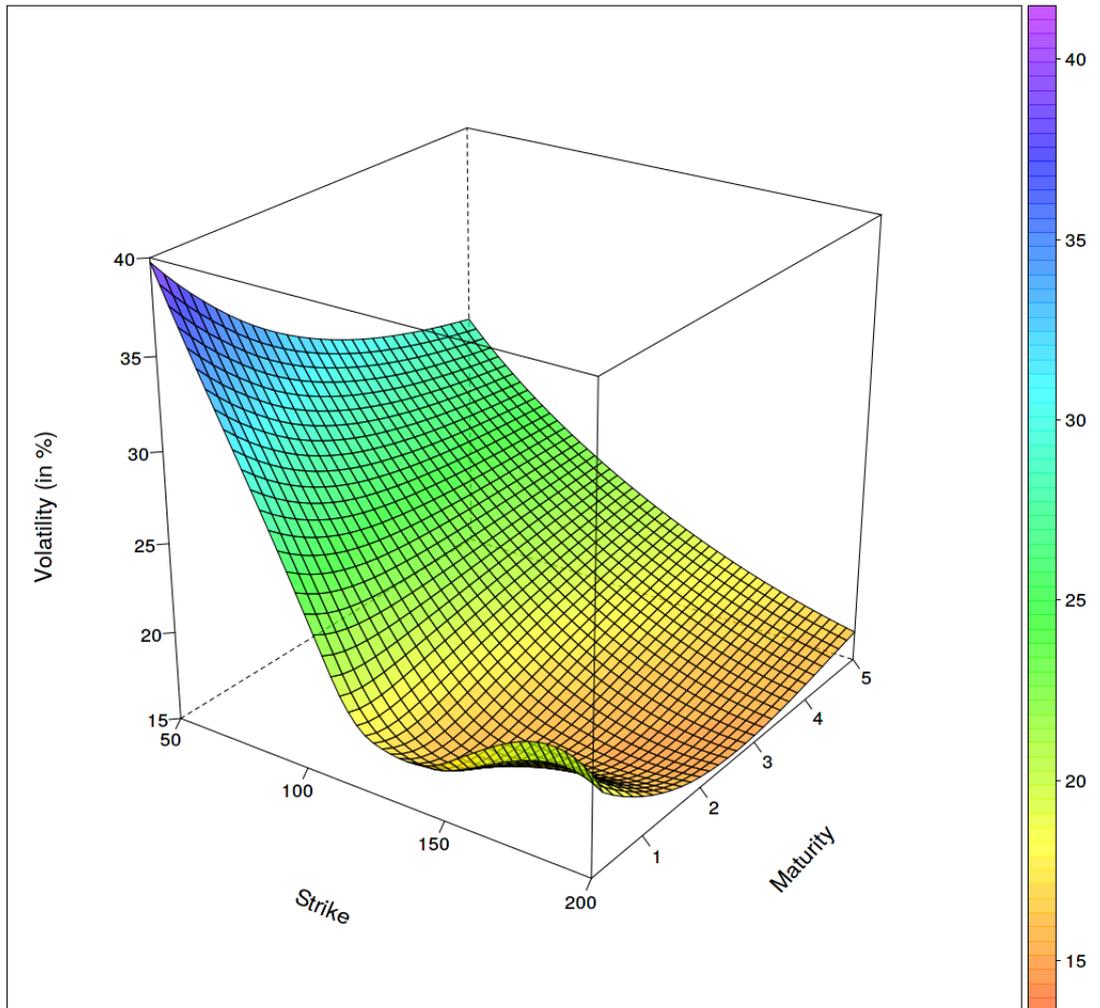


Figure 5: Heston implied volatility surface. $S_0 = 100$, $\sqrt{\nu_0} = 24.5\%$, $\kappa = 1$, $\theta = \nu_0$, $\sigma^2 = 0.2$, $\rho = -75\%$

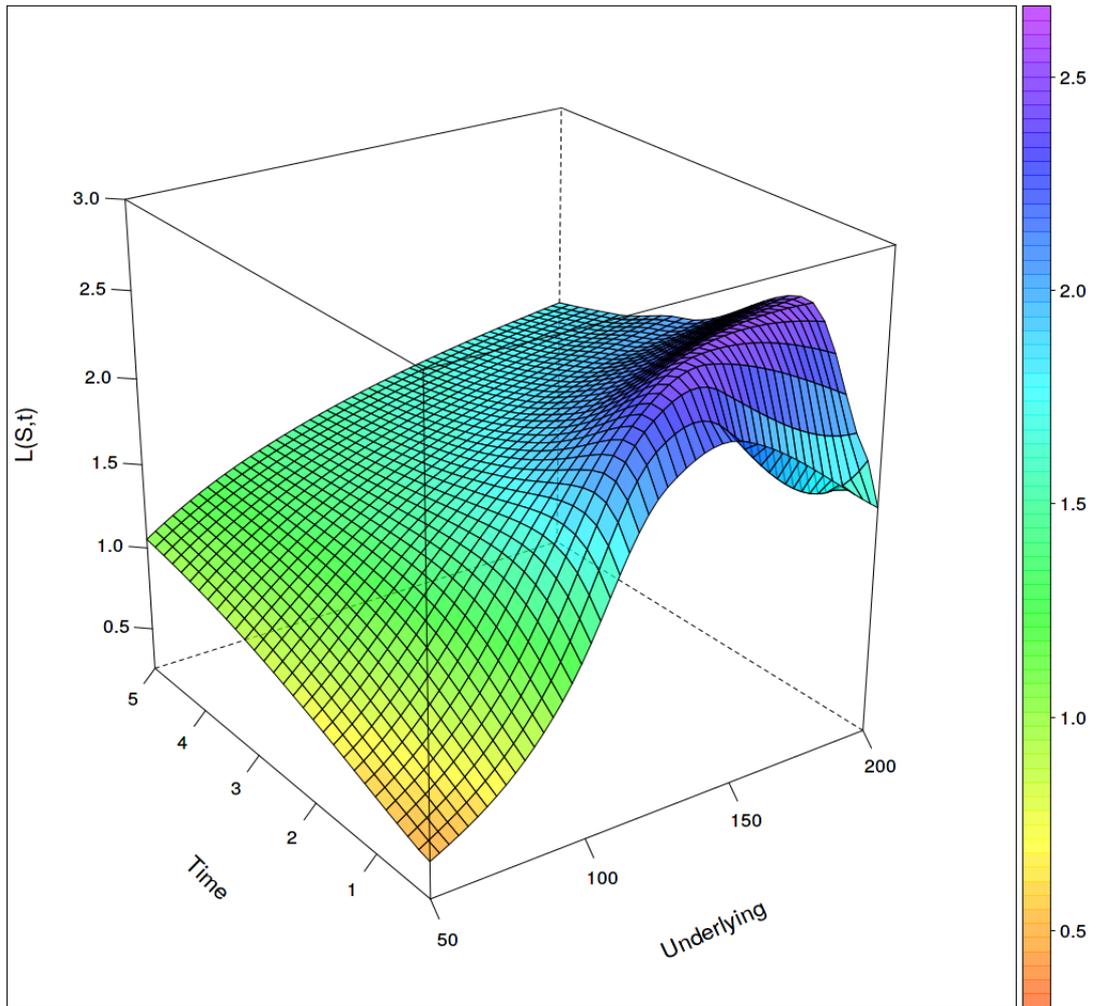


Figure 6: Leverage function after round trip calibration