

Peter Jäckel

# STOCHASTIC VOLATILITY MODELS: PAST, PRESENT AND FUTURE

## Abstract

There are many models for the uncertainty in future instantaneous volatility. When it comes to an actual implementation of a stochastic volatility model for the purpose of the management of *exotic* derivatives, the choice of model is rarely made to capture the particular dynamical features relevant for the specific contract structure at hand. Instead, more often than not, the model is chosen that provides the greatest ease with respect to market calibration by virtue of (semi-)closed form solutions for the prices of *plain vanilla* options. In this presentation, the further implications of various stochastic volatility models are reviewed with particular emphasis on both the dynamic replication of exotic derivatives and on the implementation of the model. Also, a new class of models is suggested that not only allows for the level of volatility, but also for the observed *skew* to vary stochastically over time.



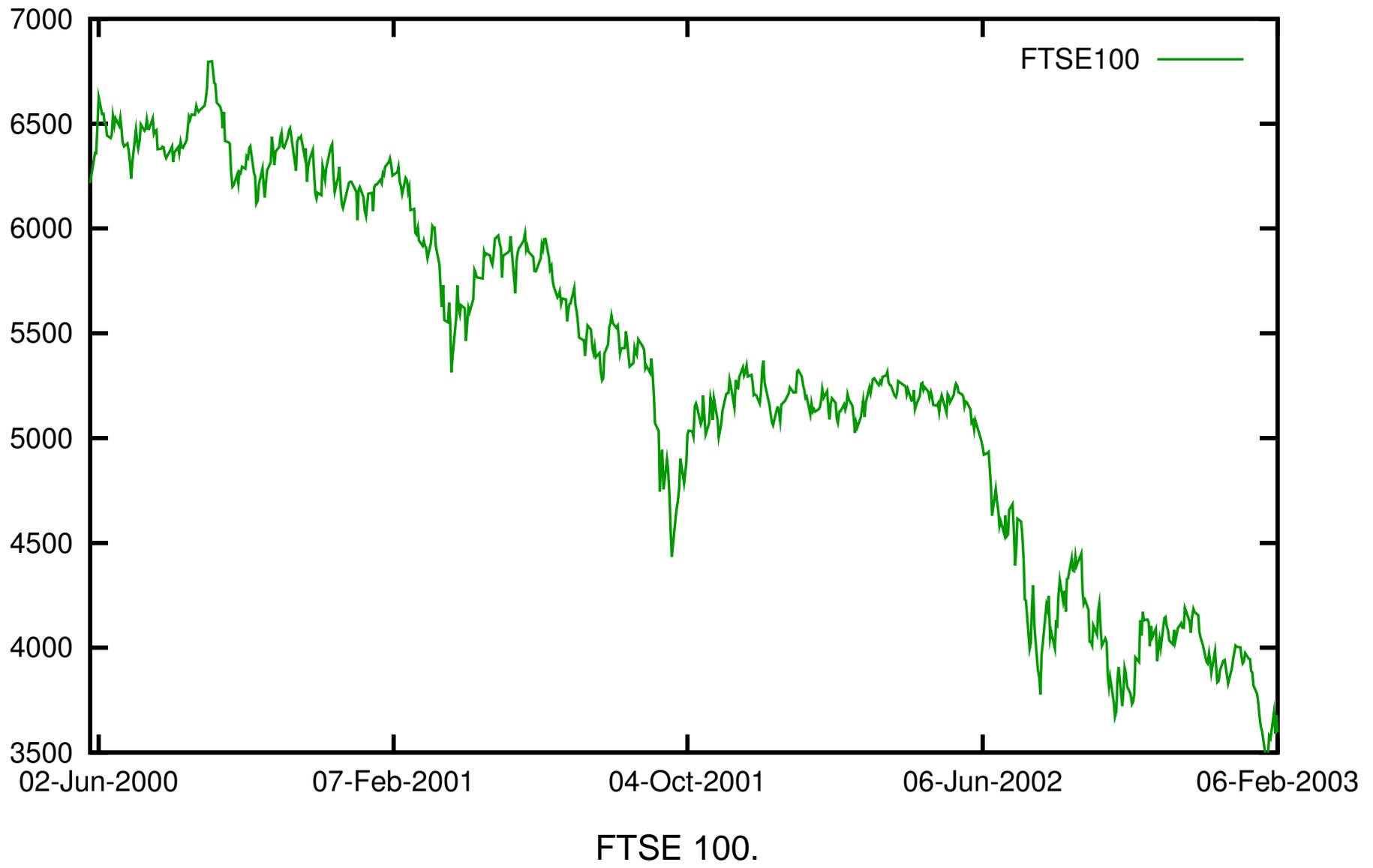
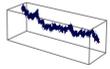
## Overview and Introduction

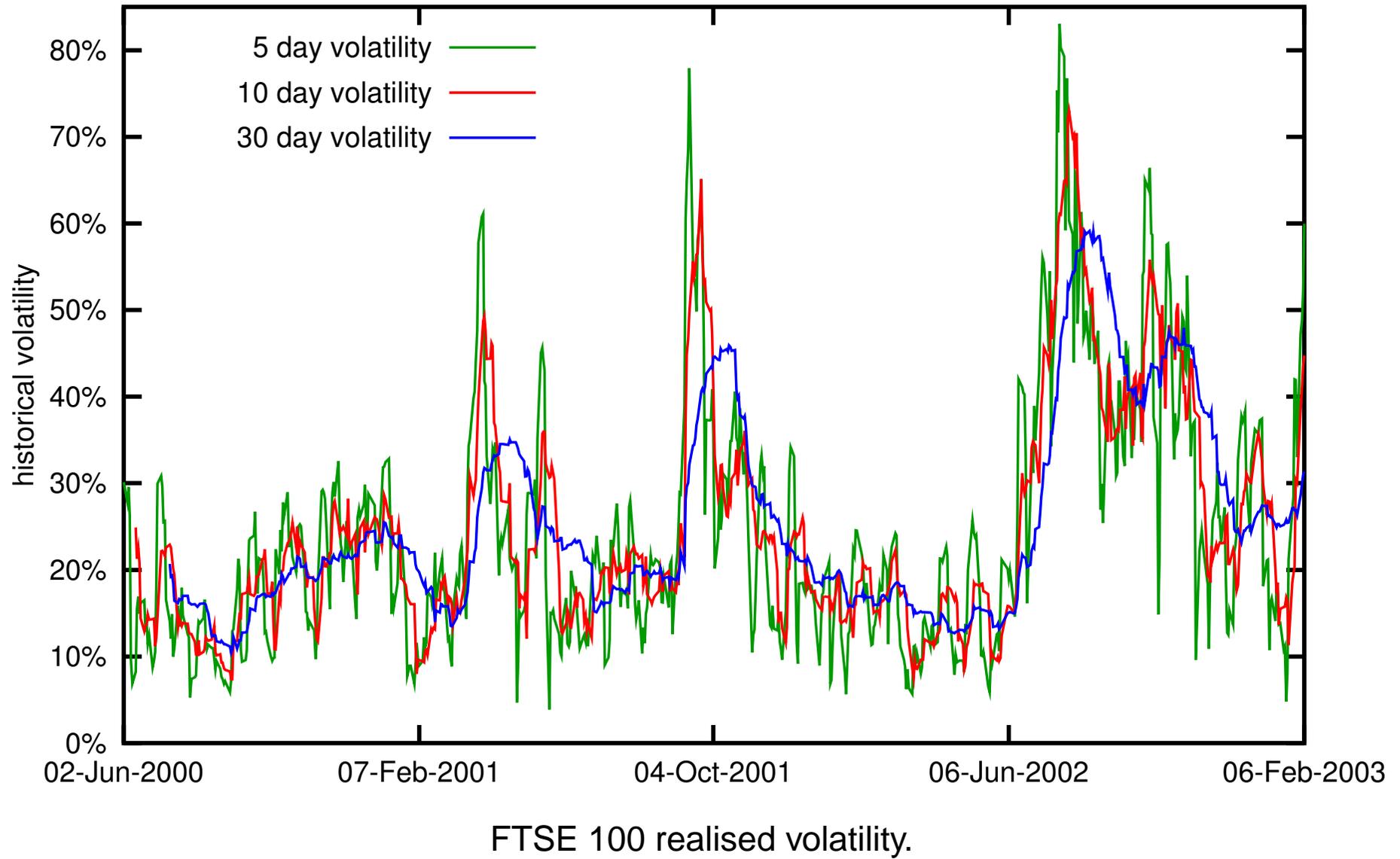
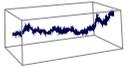
- Why stochastic volatility?
- What stochastic volatility?
- One model to rule them all?
- Mathematical features of stochastic volatility models
- A stochastic skew model
- Monte Carlo methods and stochastic volatility models
- Finite differencing methods and stochastic volatility models



## Why stochastic volatility?

- Realised volatility of traded assets displays significant variability. It would only seem natural that any model used for the hedging of derivative contracts on such assets should take into account that volatility is subject to fluctuations.
- More and more derivatives are explicitly sensitive to future (both implied and instantaneous) volatility levels. Examples are cliquets, globally floored and/or capped cliquets, and many more.
- Some (apparently) comparatively straightforward exotic derivatives such as double barrier options are being re-examined for their sensitivity to uncertainty in volatility.
- New trading ideas such as *exotic volatility options* and *skew swaps*, however, give rise to the need for a new kind of stochastic volatility model: the stochastic skew model.



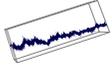




## What stochastic volatility?

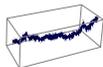
The concept of stochastic volatility, or rather the idea of a second source of risk affecting the level of instantaneous volatility, should not be seen in isolation from the nature of the underlying asset or deliverable contract.

For the three most developed modelling domains of equity, FX, and interest rate derivatives, different effects are considered to be at least partially responsible for the smile or skew observed in the associated option markets.



## Economic effects giving rise to an equity skew

- Leverage effects [Ges77, GJ84, Rub83]. A firm's value of equity can be seen as the net present value of all its future income plus its assets minus its debt. These constituents have very different relative volatilities which gives rise to a leverage related skew.
- Supply and demand. Equivalently, downwards risk insurance is more desired due to the intrinsic asymmetry of positions in equity: by their financial purpose it is more natural for equity to be held long than short, which makes downwards protection more important.
- Declining stock prices are more likely to give rise to massive portfolio rebalancing (and thus volatility) than increasing stock prices. This asymmetry arises naturally from the existence of thresholds below which positions must be cut unconditionally for regulatory reasons.



## Economic effects giving rise to an FX skew and smile

- Anticipated government intervention to stabilise FX rates.
- Government changes that are expected to change policy on trade deficits, interest rates, and other economic factors that would give rise to a market bias.
- Foreign investor FX rate protection.

## Economic effects giving rise to an interest rate skew and smile

- Elasticity of variance and/or mean reversion. In other words, interest rates are for economic reasons linked to a certain band. Unlike equity or FX, interest rates cannot be *split*, *bought back* or *re-valued* and it is this intrinsic difference that connects volatilities to absolute levels of interest rates.
- Anticipated central bank action.



**None of these effects are well described by strong *correlation* between the asset's own driving factor and a second factor governing the uncertainty in *volatility* since they are all based on deterministic relationships.**

**Still, most stochastic volatility models incorporate a skew by virtue of strong correlation of volatility and stock. The strong correlation is usually needed to match the pronounced skew of short-dated plain vanilla options.**

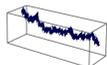
**In this context, one might wonder if it wouldn't be more appropriate to let the stochasticity of volatility explain the market-observed features related to or associated with uncertainty in volatility, and use other mechanisms to account for the skew.**



## One model to rule them all?

An important question that must be asked when a stochastic volatility model is considered is: what is it to be used for?

- Single underlying moderate exotics with strong dependence on forward volatility? Forward starting options? Cliquets?
- Single underlying exotics with strong dependence on forward skew? Globally floored and/or capped cliquets and friends?
- Single underlying exotics with strong path dependence? Barriers of all natures (single, double, layered, range accruals).

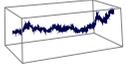


- Multiple underlying moderate exotics with strong dependence on forward volatility? Options on baskets. Cliquets on baskets.
- Multiple underlying moderate exotics with strong dependence on forward skew? Mountain range options.
- Multiple underlying moderate exotics with strong dependence on correlation? Mountain range options.

Not all of these applications would necessarily suggest the use of the same model!

A stochastic volatility model that can be perfectly adequate to capture the risk in one of the above categories may completely miss the exposures in other products.

Example: consider the use of a conventional stochastic volatility model for the management of options on variance swaps versus the use of the same model for options on future market skew in the plain vanilla option market.



## Mathematical features of stochastic volatility models

Heston [Hes93]:  $V[\sigma_S^2] \sim \mathcal{O}(\sigma_S^2)$  (mean reverting)

$$dS = \mu S dt + \sqrt{v} S dW_S \quad (1)$$

$$dv = \kappa(\theta - v)dt + \alpha\sqrt{v} dW_v \quad (2)$$

$$E[dW_S \cdot dW_v] = \rho dt \quad (3)$$

In order to achieve calibration to the market given skew, almost always:

- $0.7 < |\rho| \lesssim 1$  is required.
- $\kappa$  must be very small (*kappa kills the skew*).
- $\alpha$  must be sizeable.
- $\theta$  is by order of magnitude not too far away from the implied volatility of the longest dated option calibrated to.



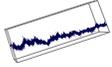
The volatility process can reach zero unless [Fei51, RW00]

$$\kappa\theta > \frac{1}{2}\alpha^2 \quad (4)$$

which is hardly ever given in a set of parameters calibrated to market!

This means the Heston model achieves calibration to today's observed plain vanilla option prices by balancing the probabilities of very high volatility scenarios against those where future instantaneous volatility drops to very low levels.

The average time volatility stays at high or low levels is measured by the mean reversion scale  $1/\kappa$ .

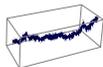


Even when  $\kappa\theta > \frac{1}{2}\alpha^2$ , the long-term distribution of  $\int_t^{t+\tau} \sigma(t)^2 dt$  is sharply peaked at low values of volatility as a result of calibration<sup>1</sup>.

***The dynamics of the calibrated Heston model predict that:  
volatility can reach zero,  
stay at zero for some time,  
or stay extremely low or very high for long periods of time.***

---

<sup>1</sup>see <http://www.dbconvertibles.com/dbquant/Presentations/LondonDec2002RiskTrainingVolatility.pdf>, slides 33–35, for diagrams on this feature.



Stein and Stein / Schöbl and Zhu [SS91, SZ99]:  $V[\sigma_S] \sim \mathcal{O}(1)$  (mean reverting)

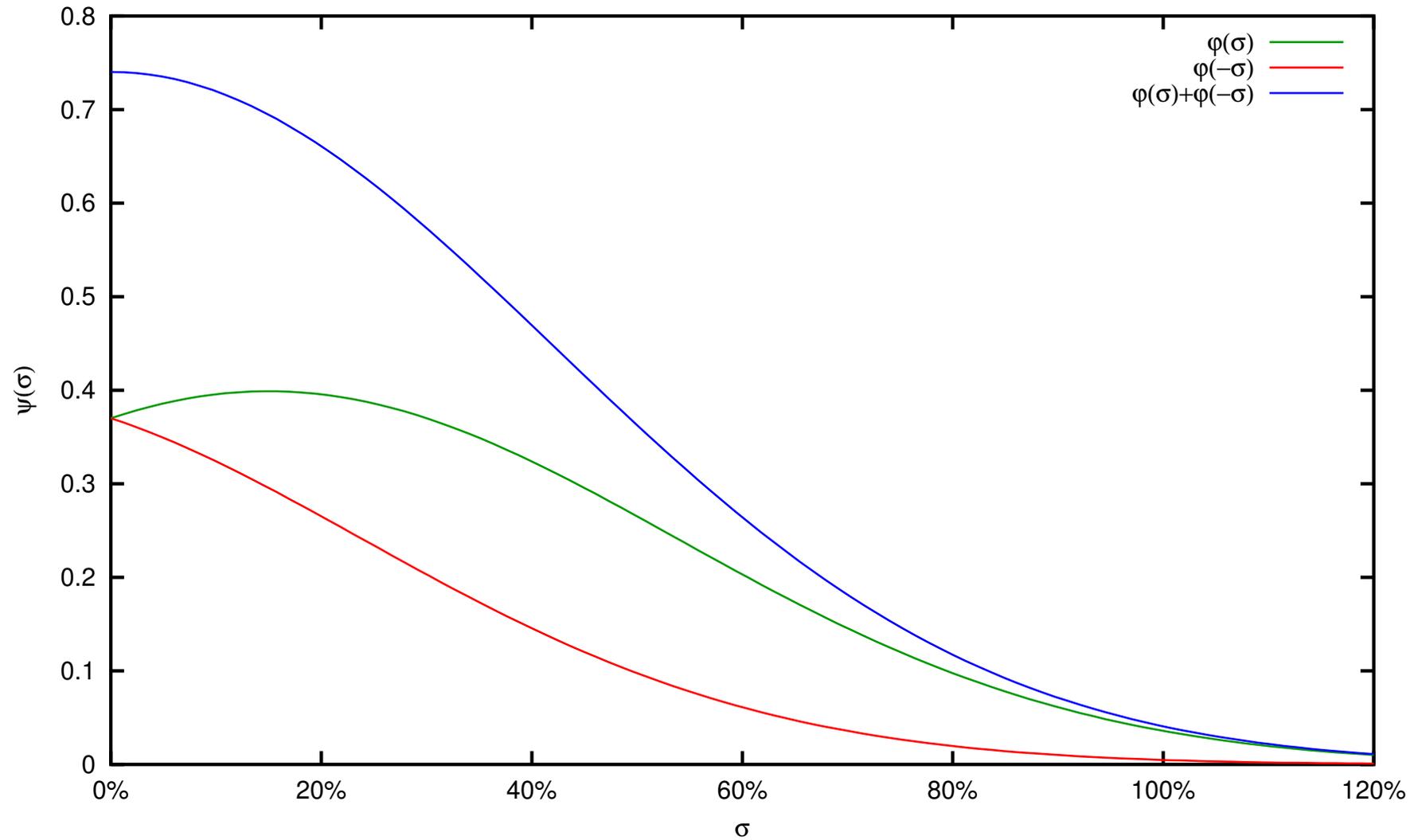
$$dS = \mu S dt + \sigma S dW_S \quad (5)$$

$$d\sigma = \kappa(\theta - \sigma)dt + \alpha dW_\sigma \quad (6)$$

$$E[dW_S \cdot dW_\sigma] = \rho dt \quad (7)$$

The distribution of volatility converges to a Gaussian distribution with mean  $\theta$  and variance  $\frac{\alpha^2}{2\kappa}$ . Since the sign of  $\sigma$  bears meaning only as a sign modifier of the correlation, we have the following two consequences:

- The sign of correlation between movements of the underlying and volatility can suddenly switch.
- The level of volatility has its most likely value at zero.



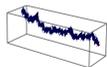
Stationary Stein & Stein volatility distribution for  $\alpha = 0.3$ ,  $\kappa = 0.3$ , and  $\theta = 0.25$ .



***The dynamics of the Stein and Stein / Schöbl and Zhu model  
predict that:***

***volatility is very likely to be near zero,***

***and that the sign of correlation with the spot movement  
driver can switch.***



Hull-White [HW87]:  $V[\sigma_S^2] \sim \mathcal{O}(\sigma_S^4)$  (zero reverting for  $\mu_v < 0$ )

$$dS = \mu_S S dt + \sqrt{v} S dW_S \quad (8)$$

$$dv = \mu_v v dt + \xi v dW_\sigma \quad (9)$$

$$E[dW_S \cdot dW_\sigma] = \rho dt \quad (10)$$

Since  $v$  is lognormally distributed in this model, and since  $\sigma = \sqrt{v}$ , we have

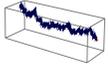
$$E[\sigma(t)] = \sigma(0) \cdot e^{\frac{1}{2}\mu_v t - \frac{1}{8}\xi^2 t} \quad (11)$$

$$V[\sigma(t)] = \sigma(0)^2 \cdot e^{\mu_v t} \cdot \left(1 - e^{-\frac{1}{4}\xi^2 t}\right) \quad (12)$$

$$M[\sigma(T)] = \sigma(0) \cdot e^{\frac{1}{2}(\mu_v - \xi^2)t} \quad (13)$$

where  $M[\cdot]$  is defined as the most likely value.

This means, for  $\mu_v < \frac{1}{4}\xi^2$ , the expectation of volatility converges to the mean-reversion level at zero. For  $\mu_v > \frac{1}{4}\xi^2$ , the expectation diverges.



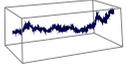
Further, unless  $\mu_v < 0$ , the variance of volatility grows unbounded. In contrast to that, if  $\mu_v < 0$ , the variance of variance diminishes over time. And finally, the most likely value for volatility converges to zero unless  $\mu_v > \xi^2$ .

For the particular case of  $\mu_v = 0$ , we have the special combination of features that the expectation and most likely value of volatility converges to zero, whilst the variance of volatility converges to  $\sigma^2$ .

Any choice of parameters that provides a reasonable match of market given implied volatilities is extremely likely to lead to  $\mu_v < 0$  in which case we have:

***The dynamics of the Hull-White stochastic volatility model predict that:***

***both expectation and most likely value of instantaneous volatility converge to zero.***



Hagan [HKL02]:  $V[\sigma_S] \sim \mathcal{O}(\sigma_S^2)$  (*not* mean reverting)

$$dS = \mu S dt + \sigma S dW_S \quad (14)$$

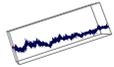
$$d\sigma = \alpha \sigma dW_\sigma \quad (15)$$

$$E[dW_S \cdot dW_\sigma] = \rho dt \quad (16)$$

This model is equivalent to the Hull-White stochastic volatility model for the special case of  $\mu_v = \alpha^2$  and  $\xi = 2\alpha$ . In this model, instantaneous volatility is a martingale but the variance of volatility grows unbounded. At the same time, the *most likely* value for volatility converges to zero.



***The dynamics of the Hagan model predict that:  
the expectation of volatility is constant over time,  
that variance of instantaneous volatility grows without limit,  
and that the most likely value of instantaneous volatility  
converges to zero.***



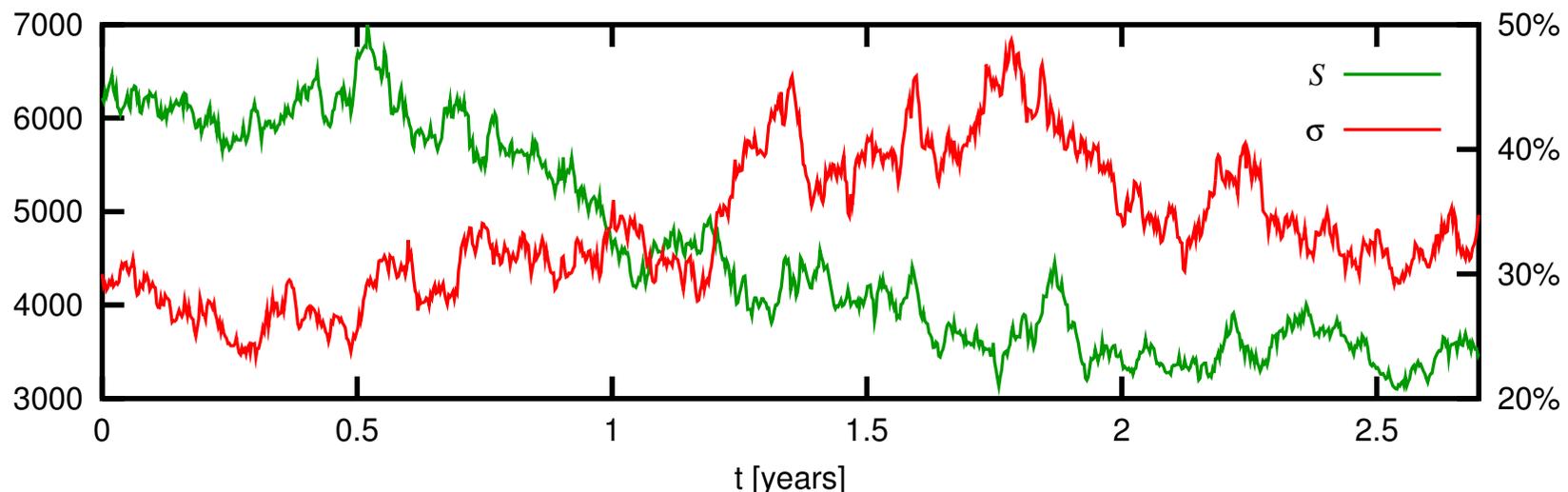
Scott and Scott-Chesney [Sco87, CS89]:  $V[\sigma_S] \sim \mathcal{O}(\sigma_S^2)$  (mean reverting)

$$dS = \mu S dt + e^y S dW_S \quad (17)$$

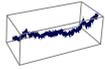
$$dy = \varkappa(\theta - y) dt + \alpha dW_y \quad (18)$$

$$E[dW_S \cdot dW_y] = \rho dt \quad (19)$$

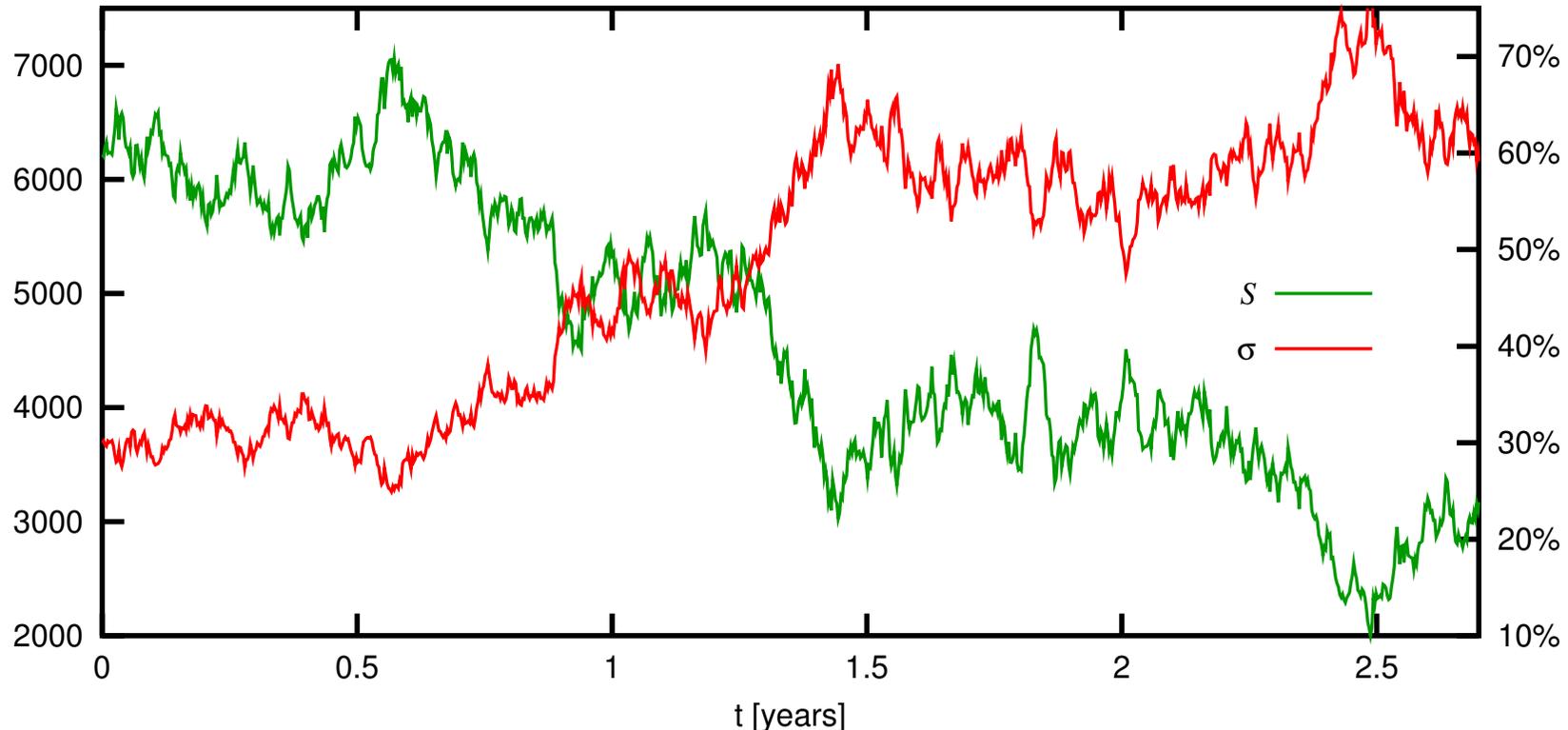
Volatility cannot reach zero, nor does its most likely value converge there.



Sample path for Scott-Chesney model with  $S_0 = 6216$ ,  $r = 5\%$ ,  $d = 1\%$ ,  $\sigma_0 = 30\%$ ,  $\theta = \ln 30\%$ ,  $\varkappa = 0.1$ ,  $\alpha = 40\%$ ,  $\frac{\alpha^2}{2\varkappa} = 2$ ,  $\rho = 0$ . Euler integration with  $\Delta t = 1/365$ .



The market-observable skew of implied volatilities would require a strong negative correlation for this model to be calibrated.

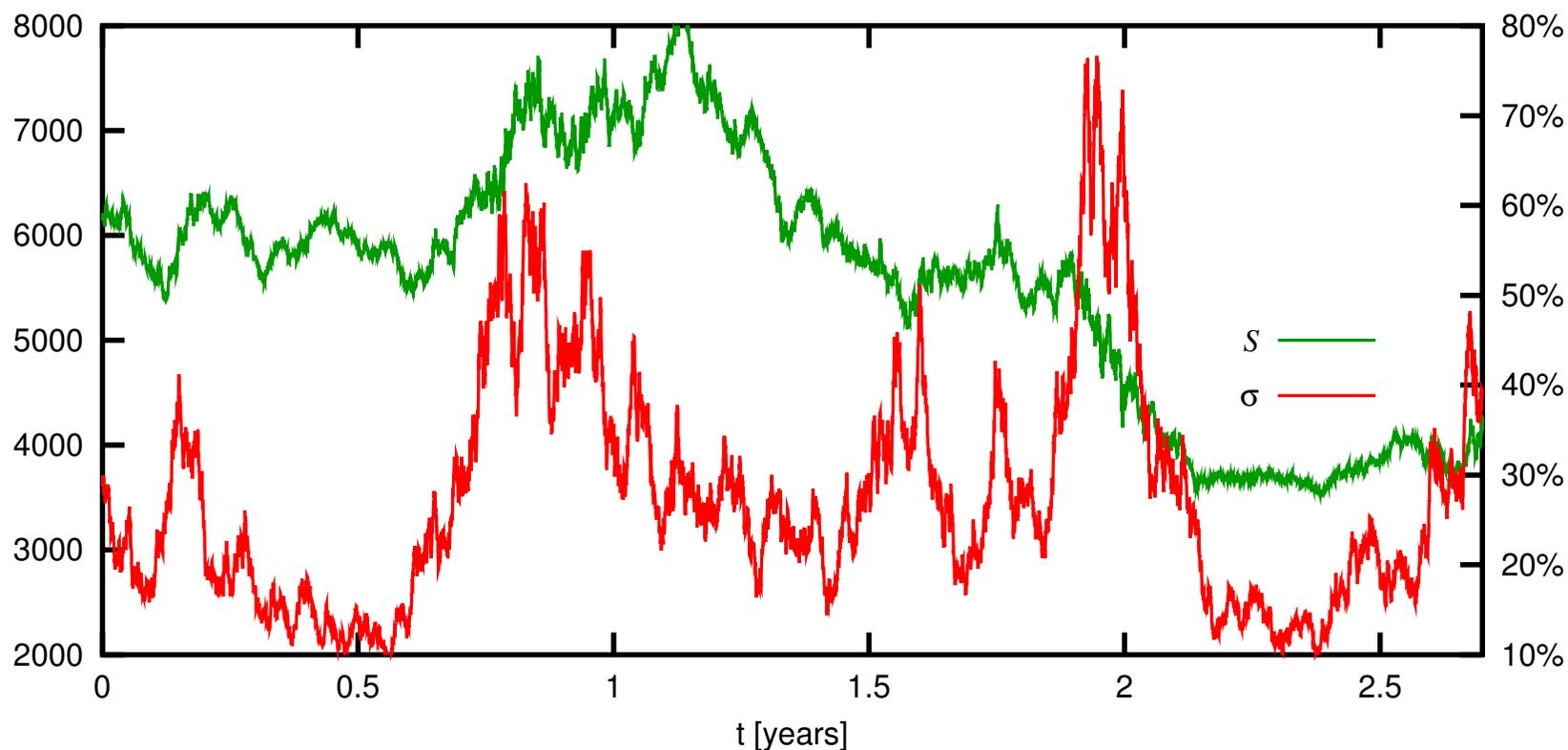


Sample path for Scott-Chesney model with  $S_0 = 6216$ ,  $r = 5\%$ ,  $d = 1\%$ ,  $\sigma_0 = 30\%$ ,  $\theta = \ln 30\%$ ,  $\varkappa = 0.1$ ,  $\alpha = 40\%$ ,  $\frac{\alpha^2}{2\varkappa} = 2$ ,  $\rho = -0.9$ . Euler integration with  $\Delta t = 1/365$ .

However, the required strong correlation between volatility and spot is not supported by any econometric analysis.



Nonetheless, it is possible to reproduce the burstiness of real volatility returns by increasing the mean reversion.



Sample path for Scott-Chesney model with  $S_0 = 6216$ ,  $r = 5\%$ ,  $d = 1\%$ ,  $\sigma_0 = 30\%$ ,  $\theta = \ln 30\%$ ,  $\varkappa = 6$ ,  $\alpha = 1.5$ ,  $\frac{\alpha^2}{2\varkappa} = 0.125$ ,  $\rho = 0$ . Euler integration with  $\Delta t = 1/2920$ .



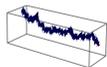
Fouquet *et alii* compare strong mean reversion dynamics with real data and find that it captures the apparent burstiness of realised volatilities very well [FPS00]:

- The larger  $\kappa$ , the more rapidly the volatility distribution converges to its stationary state.
- $1/\kappa$  is the time scale for volatility auto-decorrelation.
- The right measure for uncertainty in volatility is

$$\frac{\alpha^2}{2\kappa},$$

*not*  $\alpha$  on its own.

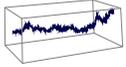
Large mean reversion causes volatility to approach its stationary distribution quickly.



The problem with future volatility being likely to hover near zero for models such as the Heston and the Stein & Stein model goes away when mean reversion is strong.

However, if mean reversion is large, correlation between volatility and spot does not suffice to generate a significant skew.

To achieve market calibration, a different mechanism is needed. This could be independent jumps of the stock itself, or a stock-dependent volatility scaling function.



***The main drawback of the Scott-Chesney model is that:  
it requires very high correlation between the spot and the  
volatility process to calibrate to a pronounced skew,  
and that the skew is fully deterministic.***

***These features are also shared by all of the above discussed  
models.***



## A stochastic skew model

$$dS = \mu S dt + \sigma f(S; \gamma) S dW_S \quad (20)$$

$$d \ln \sigma = \kappa_\sigma (\ln \sigma_\infty - \ln \sigma) dt + \alpha_\sigma dW_\sigma \quad (21)$$

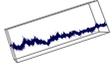
$$d\gamma = \kappa_\gamma (\gamma_\infty - \gamma) dt + \alpha_\gamma dW_\gamma \quad (22)$$

with

$$f(S; \gamma) = e^{\gamma \cdot (\frac{S}{H} - 1)} \quad (23)$$

and

$$\mathbf{E}[dW_\sigma dW_\gamma] = \mathbf{E}[dW_\sigma dW_S] = \mathbf{E}[dW_\gamma dW_S] = 0 \quad (24)$$

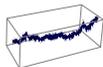


This scaling ensures that:-

- For negative  $\gamma$ , the local volatility scaling factor decays from  $e^{-\gamma}$  for  $S \rightarrow 0$  to 0 for  $S \rightarrow \infty$ .
- The local volatility scaling factor  $f$  at spot level  $H$  is exactly 1.
- The local volatility scaling factor  $f$  change for a spot move of  $\delta \cdot H$  near  $H$  is given by

$$\Delta f = \left. \frac{\partial f}{\partial S} \right|_{S=H} \cdot \delta \cdot H = \frac{\gamma}{H} \cdot \delta \cdot H = \delta \cdot \gamma. \quad (25)$$

In other words,  $\gamma$  is a measure for the local volatility skew at  $H$ .



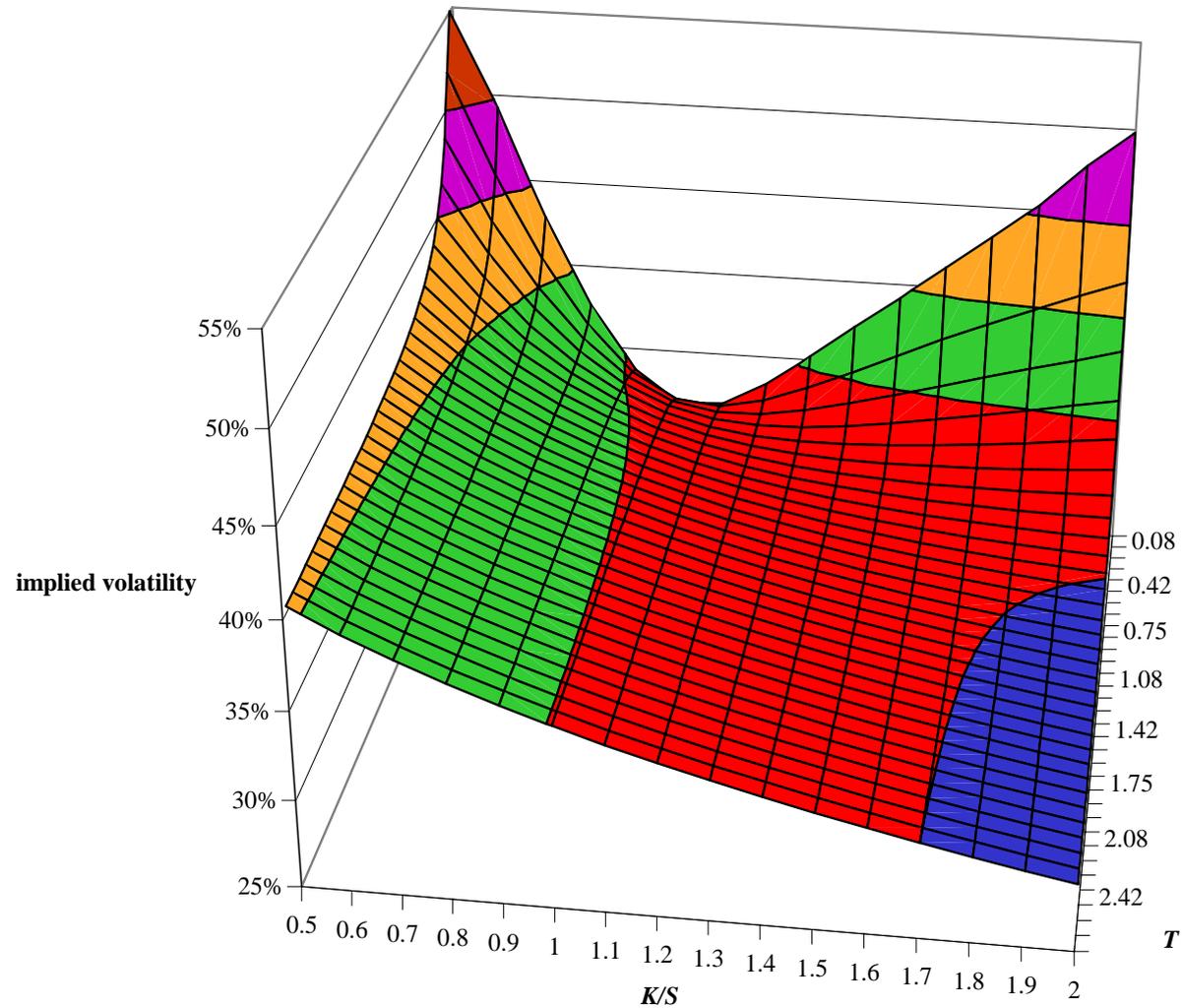
Maintenance of correlation matrices is greatly simplified by the assumption of independence of the individual factors.

The associated partial differential equation governing the boundary value problem of derivatives prices is

$$\begin{aligned} V_t + \underbrace{\left(\mu - \frac{1}{2}e^{2y} f^2(e^x; \gamma)\right)}_{\widehat{\mu}_x} V_x + \underbrace{\kappa_\sigma (\ln \sigma_\infty - y)}_{\widehat{\mu}_y} V_y + \underbrace{\kappa_\gamma (\gamma_\infty - \gamma)}_{\widehat{\mu}_\gamma} V_\gamma & \quad (26) \\ + \frac{1}{2}e^{2y} f^2(e^x; \gamma) V_{xx} + \frac{1}{2}\alpha_\sigma^2 V_{yy} + \frac{1}{2}\alpha_\gamma^2 V_{\gamma\gamma} & = r \cdot V \end{aligned}$$

with

$$x = \ln S \quad \text{and} \quad y = \ln \sigma . \quad (27)$$

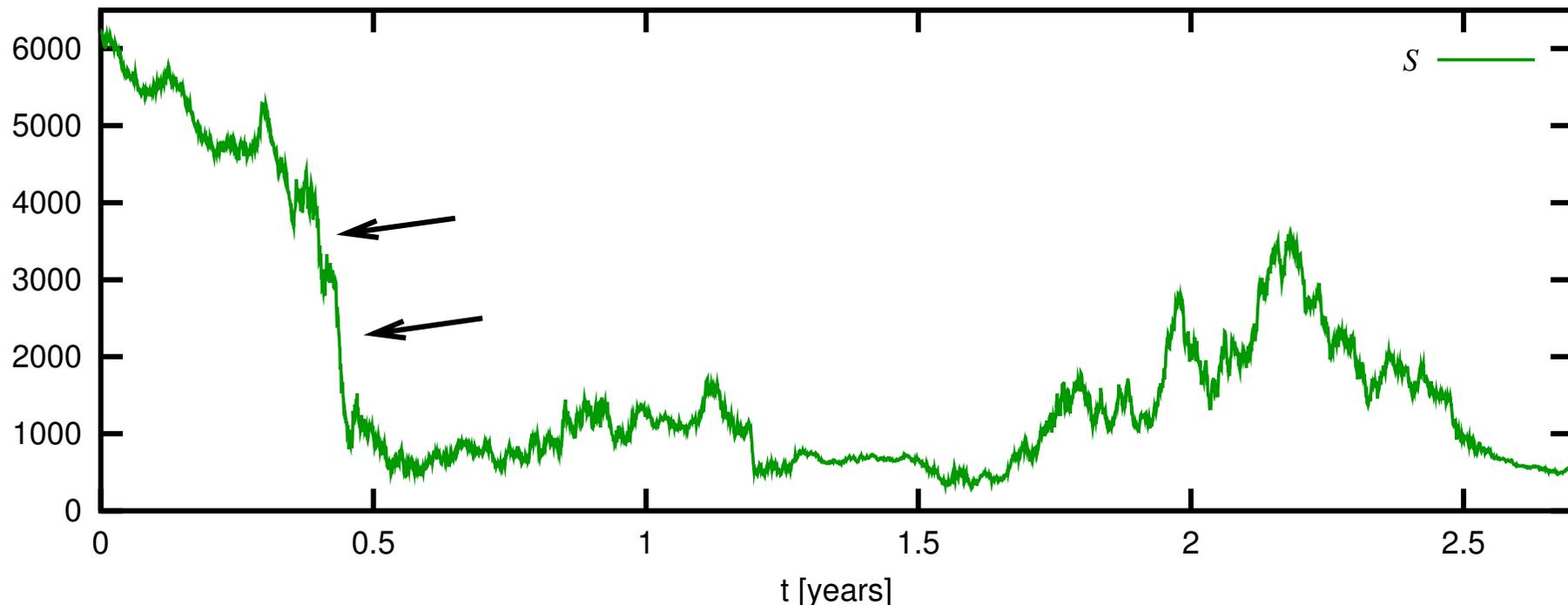


Implied volatility surface for stochastic skew model with  $S_0 = H = 6216$ ,  $r = 5\%$ ,  $d = 1\%$ ,  
 $\sigma_0 = \sigma_\infty = 30\%$ ,  $\kappa_\sigma = 12$ ,  $\alpha_\sigma = 2$ ,  $\sqrt{\frac{\alpha_\sigma^2}{2\kappa_\sigma}} = 41\%$ ,  $\gamma_0 = \gamma_\infty = -0.5$ ,  $\kappa_\gamma = 4$ ,  $\alpha_\gamma = 0.5$ ,  $\sqrt{\frac{\alpha_\gamma^2}{2\kappa_\gamma}} = 0.18$ .

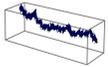


## *Jumps without jumps*

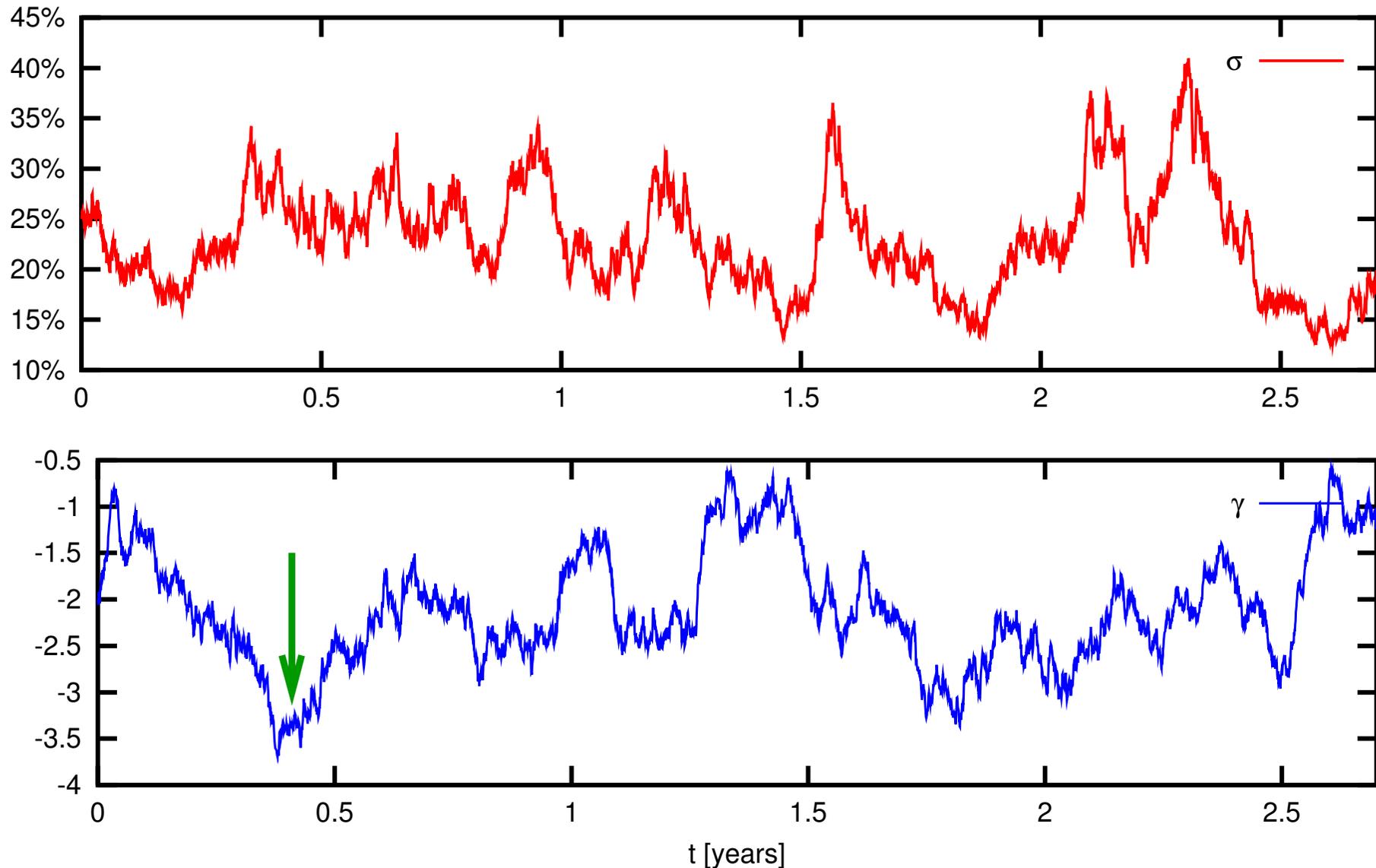
The exponential dependence of the volatility scaling function  $f$  on the spot level  $S$  can lead to jump-like upward (for  $\gamma > 0$ ) or downward (for  $\gamma < 0$ ) rallies when  $|\gamma|$  is of significant size.

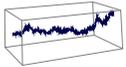


Jump-like almost instantaneous downwards corrections of the spot for  $S_0 = H = 6216$ ,  $r = 5\%$ ,  $d = 1\%$ ,  
 $\sigma_0 = \sigma_\infty = 25\%$ ,  $\kappa_\sigma = 6$ ,  $\alpha_\sigma = 1$ ,  $\sqrt{\frac{\alpha_\sigma^2}{2\kappa_\sigma}} = 29\%$ ,  $\gamma_0 = \gamma_\infty = -2$ ,  $\kappa_\gamma = 3$ ,  $\alpha_\gamma = 2$ ,  $\sqrt{\frac{\alpha_\gamma^2}{2\kappa_\gamma}} = 0.82$ .

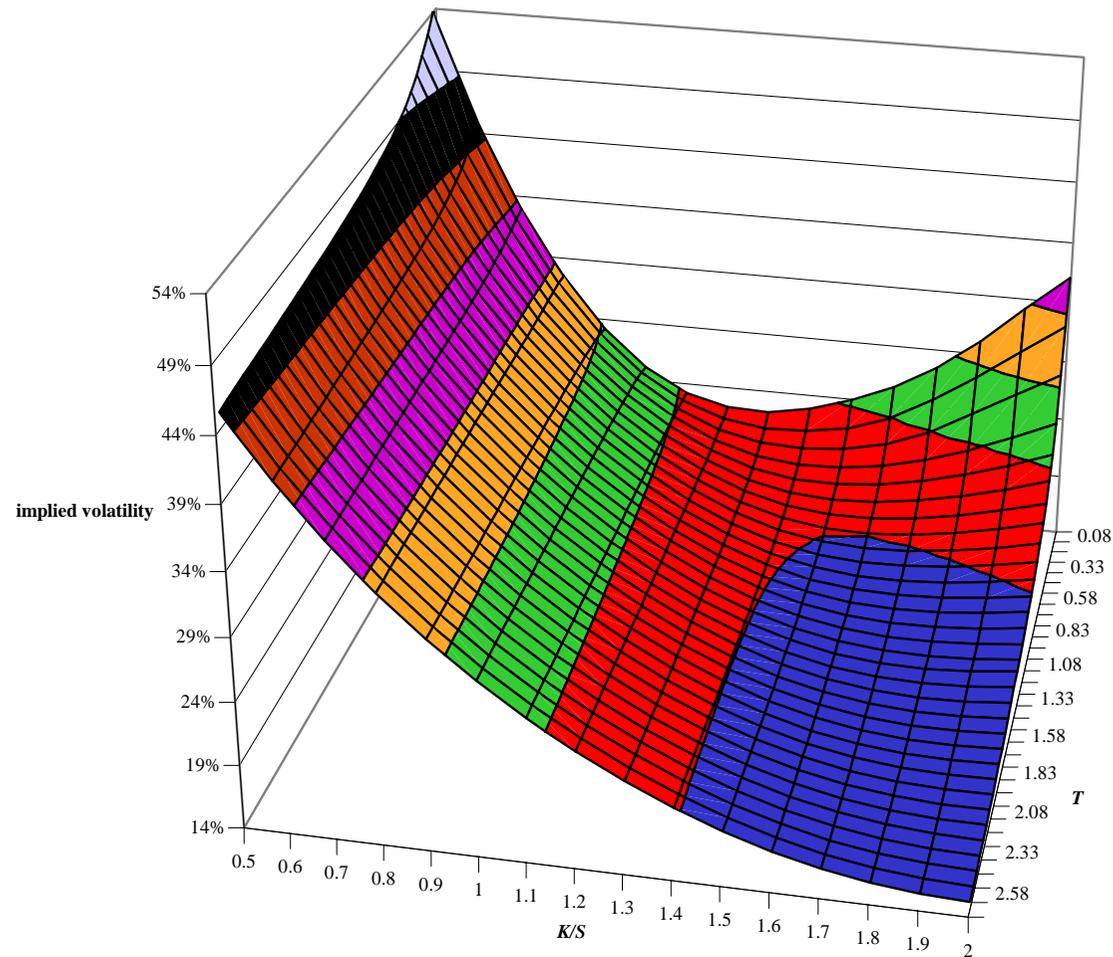


This can happen due to the exponential nature of the scaling function  $f$ , especially during periods of increased  $|\gamma|$ .





These events only occur when the skew is very pronounced:



Implied volatility surface for stochastic skew model with  $S_0 = H = 6216$ ,  $r = 5\%$ ,  $d = 1\%$ ,

$$\sigma_0 = \sigma_\infty = 25\%, \kappa_\sigma = 6, \alpha_\sigma = 1, \sqrt{\frac{\alpha_\sigma^2}{2\kappa_\sigma}} = 29\%, \gamma_0 = \gamma_\infty = -2, \kappa_\gamma = 3, \alpha_\gamma = 2, \sqrt{\frac{\alpha_\gamma^2}{2\kappa_\gamma}} = 0.82.$$

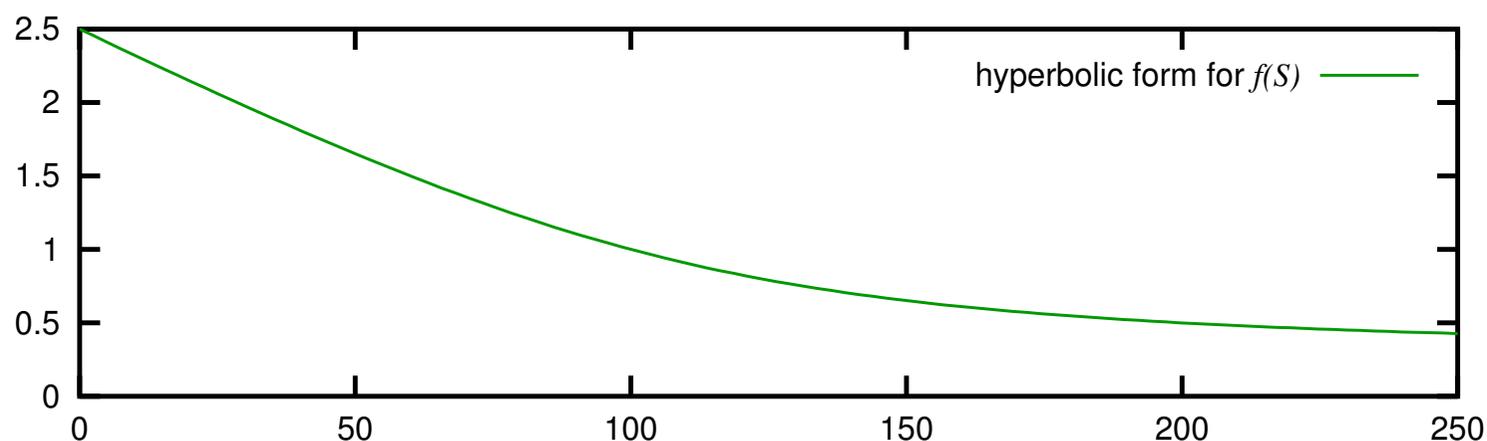


## A hyperbolic alternative

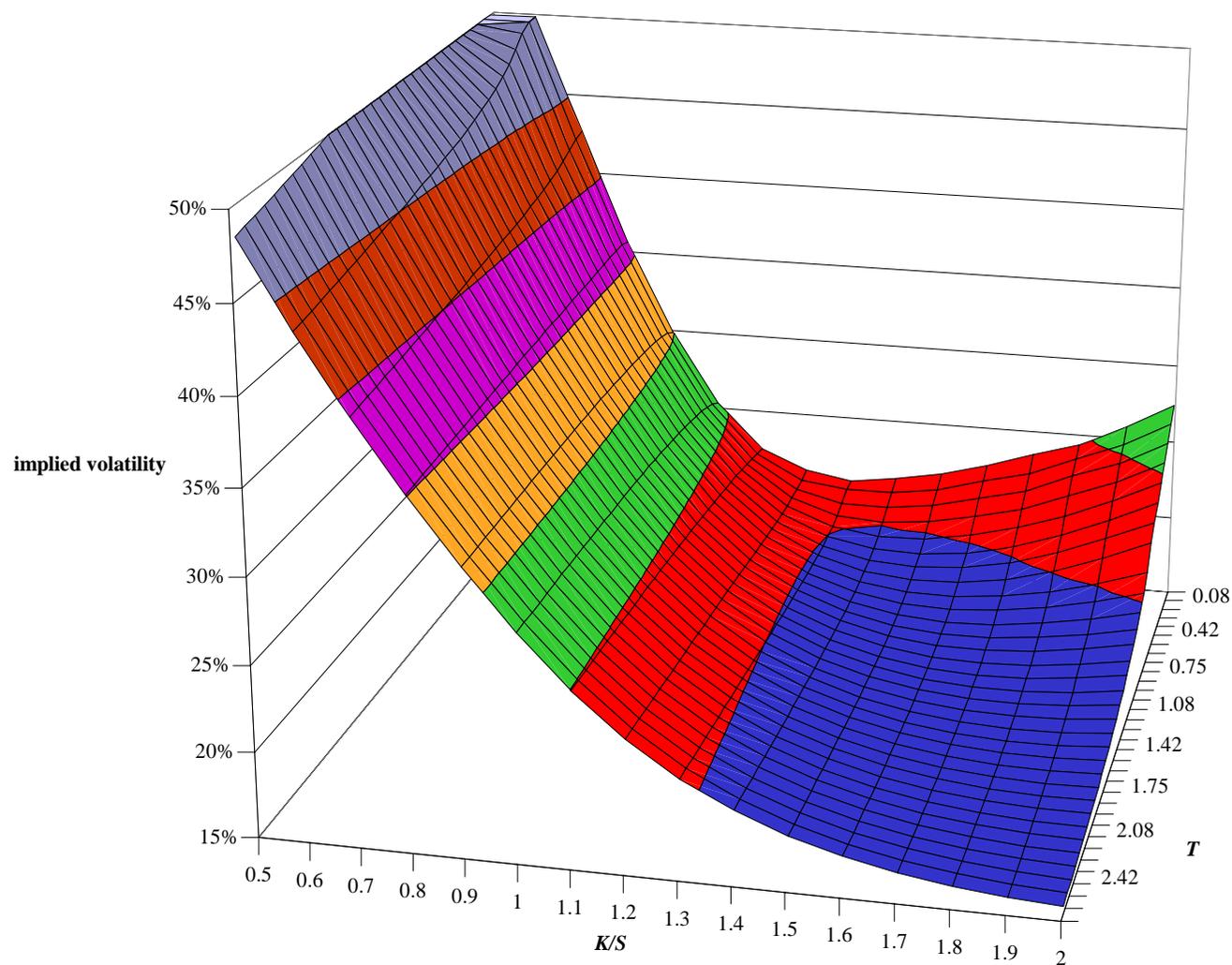
The shown implosions of the spot are caused by the exponential form of the scaling function  $f$  and are technically akin to process explosions seen also for the short rate in a lognormal HJM setting and other equations involving a locally exponential scaling of volatility. Naturally, it is straightforward to use other scaling functions that avoid the spot implosions, should they be undesirable.

An alternative to the exponential scaling is the hyperbolic function

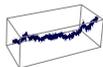
$$f = \gamma \left( \frac{S}{H} - 1 \right) + \sqrt{\gamma^2 \left( \frac{S}{H} - 1 \right)^2 + (1 - \eta)^2} + \eta \quad (28)$$



Hyperbolic example for the scaling function  $f(S)$  with  $\gamma = -1$ ,  $H = 100$ , and  $\eta = 1/4$ .



Implied volatility surface for stochastic skew model with a hyperbolic scaling function  $f$  and  $S_0 = H = 6216$ ,  $r = 5\%$ ,  $d = 1\%$ ,  $\sigma_0 = \sigma_\infty = 25\%$ ,  $\kappa_\sigma = 6$ ,  $\alpha_\sigma = 1$ ,  $\sqrt{\frac{\alpha_\sigma^2}{2\kappa_\sigma}} = 28.87\%$ ,  $\gamma_0 = \gamma_\infty = -3$ ,  $\kappa_\gamma = 3$ ,  $\alpha_\gamma = 3$ ,  $\sqrt{\frac{\alpha_\gamma^2}{2\kappa_\gamma}} = 1.22$ , and  $\eta = 1/4$ .



## Monte Carlo methods and stochastic volatility models

The Heston model is often used to parametrise the observed market volatilities since there are semi-analytical solutions for plain vanilla options under this model.

However, when multi-asset derivatives are priced, we often need to resort to numerical integration of the governing stochastic differential equations.

Euler discretisation of the Heston variance process:

$$\Delta v = \kappa(\theta - v)\Delta t + \alpha\sqrt{v}\sqrt{\Delta t} \cdot z \quad (29)$$

with  $z \sim \mathcal{N}(0, 1)$ . This means for  $z < z^*$  with

$$z^* = -\frac{v + \kappa(\theta - v)\Delta t}{\alpha\sqrt{v\Delta t}} \quad (30)$$

the Euler step causes variance to cross over to the negative domain!



A popular method of choice to avoid this artifact of Euler integration is to use Itô's lemma to transform to coordinates where the Euler step remains in the domain of validity for all possibly drawn Gaussian variates. For the Heston variance process, the coordinate we have to transform to is volatility itself:

$$d\sigma = \frac{\varkappa}{2} \left[ \frac{1}{\sigma} \left( \theta - \frac{\alpha^2}{4\varkappa} \right) - \sigma \right] dt + \frac{1}{2}\alpha dW \quad (31)$$

Alas, it seems we have transformed ourselves from the pan into the fire: whilst equation (2) would always show a positive drift term for all  $\theta > 0$  no matter how close variance came to zero, and only the diffusion component could make it reach zero, the drift term in equation (31) diverges to negative infinity if  $\theta < \frac{\alpha^2}{4\varkappa}$  irrespective of the path taken by the diffusion component.

This means, the transformed equation shows strong (drift-dominated) absorption into zero near zero, whilst the original stochastic differential equation for the variance only exhibits zero as an attainable boundary due to the diffusion component being able to overcome the mean reversion effect (i.e. the positive drift) for  $2\theta\varkappa < \alpha^2$ .



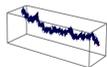
The apparently contradictory behaviour near zero has a simple explanation:

***In an infinitesimal neighbourhood of zero, Itô's lemma cannot be applied to the variance process (2).***

***The transformation of the variance process to a volatility formulation results in a structurally different process !***

Naturally, this feature raises its ugly head in any numerical implementation where we may prefer to use a transformed version of the original equations!

An alternative, when suitable transformations are not available, is to use *implicit* or *mixed* Euler schemes [KP99] in order to ensure that the stepping algorithm does not cause the state variable to leave the domain of the governing equations, possibly in conjunction with Doss's [Dos77] method of constructing pathwise solutions.



An example for such an approach is as follows.

First, let us assume that we have discretised the evolution of time into a sequence of time intervals  $[t_n, t_{n+1}]$ , and that we have drawn an independent Wiener path over those points in time, i.e. that we know  $W(t_n)$  for all  $n$  for one specific path.

Then, approximate  $W(t)$  as a piecewise linear function in between the known values at  $t_n$  and  $t_{n+1}$ , i.e.

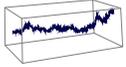
$$W(t) \simeq \gamma_n + \delta_n t \quad \text{for } t \in [t_n, t_{n+1}] \quad (32)$$

with

$$\gamma_n = W(t_n) - \delta_n t_n \quad \text{and} \quad \delta_n = \frac{W(t_n) - W(t_{n+1})}{t_n - t_{n+1}}.$$

Using the resulting dependency  $dW = \delta_n dt$ , this gives us the approximate ordinary differential equation

$$\frac{dv}{dt} \simeq \kappa(\theta - v) + \alpha \delta_n \sqrt{v} \quad (33)$$



which has the implicit solution

$$t - t_n = T(v(t)) - T(v(t_n)) \quad (34)$$

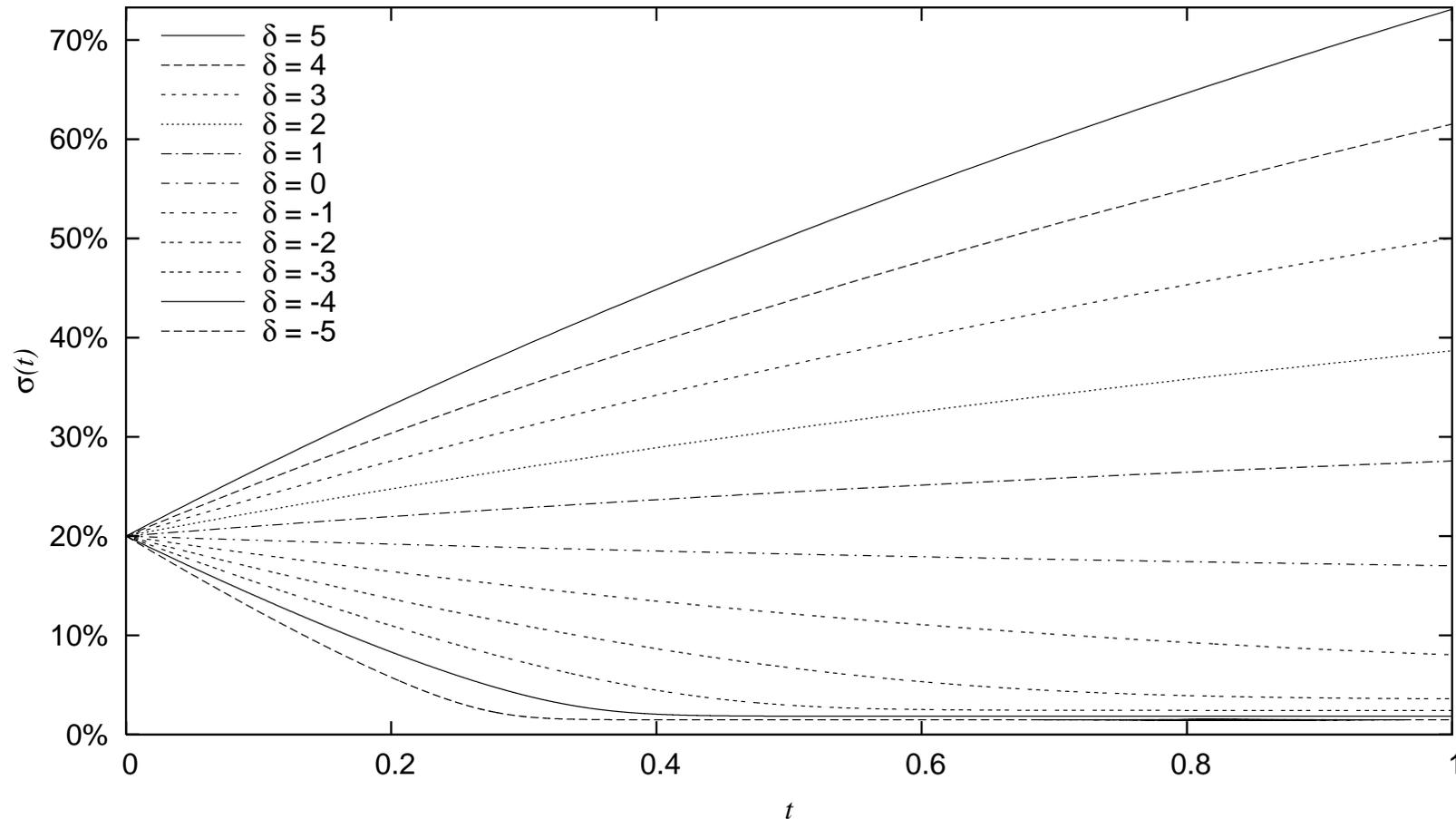
with

$$T(v) = \frac{2\alpha\delta_n}{\kappa\sqrt{\alpha^2\delta_n^2 + 4\theta\kappa^2}} \operatorname{atanh} \left( \frac{2\kappa\sqrt{v} - \alpha\delta_n}{\sqrt{\alpha^2\delta_n^2 + 4\theta\kappa^2}} \right) - \frac{1}{\kappa} \ln(\kappa(v - \theta) - \alpha\delta_n\sqrt{v}) \quad (35)$$

The above equation can be solved numerically comparatively readily since we know that, given  $\delta_n$ , over the time step from  $t_n$  to  $t_{n+1}$ ,  $v$  will move monotonically, and that in the limit of  $\Delta t_n := (t_{n+1} - t_n) \rightarrow \infty$ , for fixed  $\delta_n$ , we have

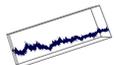
$$\lim_{\Delta t_n \rightarrow \infty} v_{n+1} = \left( \frac{\alpha\delta}{2\kappa} + \sqrt{\left( \frac{\alpha\delta}{2\kappa} \right)^2 + \theta} \right)^2 \quad (36)$$

which can be computed by setting the argument of the logarithm in the right hand side of equation (35) to zero.



Sample paths for the Heston volatility process for  $\sigma(0) = \sqrt{v(0)} = 20\%$ ,  $\sqrt{\theta} = 15\%$ ,  $\alpha = 30\%$ ,  $\kappa = 1$  over a unit time step for different levels of the variate  $\delta = W_v(1) - W_v(0)$  when  $W_v(t)$  is approximated linearly over the time step.

Putting all of the above together enables us to construct paths for the stochastic variance without the need for very small time steps.

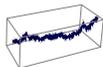


The explicit knowledge of the functional form of the volatility, or variance, path has another advantage:

Fouquet *et alii* [FPS00] explain how we can directly draw the logarithm of the spot level at the end of a large time step  $(t_{n+1} - t_n)$  if we can explicitly compute, for the given volatility or variance path, the quantities

$$\Delta \hat{v}_n := \int_{t_n}^{t_{n+1}} \sigma^2(t) dt \quad (37)$$

$$\Delta \hat{w}_n := \int_{t_n}^{t_{n+1}} \sigma(t) dW(t) \quad (38)$$



The first term poses no difficulty since the primitive of  $T(v)$  can be computed analytically and

$$\int_{t_n}^{t_{n+1}} \sigma^2(t) dt = v(t_{n+1}) \cdot [T(v(t_n)) + \Delta t_n] - v(t_n) \cdot T(v(t_n)) - \int_{v_n}^{v_{n+1}} T(v) dv . \quad (39)$$

The second term requires another approximate numerical scheme which will also be no major obstacle since

$$\begin{aligned} \int_{t_n}^{t_{n+1}} \sigma(t) dW(t) &\simeq \delta_n \cdot \int_{t_n}^{t_{n+1}} \sigma(t) dt \\ &= \delta_n \cdot \left( \sigma(t_{n+1}) \cdot [T(v(t_n)) + \Delta t_n] - \sigma(t_n) \cdot T(v(t_n)) \right. \\ &\quad \left. - \int_{v_n}^{v_{n+1}} T(v) / (2\sqrt{v}) dv \right) . \end{aligned} \quad (40)$$



The numerical approximation is needed for the calculation of the integral on the right hand side of equation (40).

A simple Simpson scheme or Legendre quadrature will produce excellent results given that the function is guaranteed to be monotonic and smooth.

Using all of the above quantities, the draw for the logarithm of the spot level can be constructed as

$$\ln S_{n+1} = \ln S_n + \mu \Delta t_n - \frac{1}{2} \Delta \hat{v}_n + \rho \Delta \hat{\omega}_n + \sqrt{1 - \rho^2} \cdot \sqrt{\Delta \hat{v}_n} \cdot z \quad (41)$$

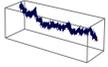
where  $z$  is a standard normal variate that is independent from the variate used to construct the variance step  $v_n \rightarrow v_{n+1}$ .



The above scheme is essentially an extension of the root-mean-square volatility lemma given in [HW87] beyond the case of  $\rho = 0$ .

In comparison, the Stein and Stein / Schöbl and Zhu, Hull-White, Hagan and Scott/Scott-Chesney model can be simulated much more easily since the stochastic differential equation for the volatility component has simple analytical solutions.

Naturally, similar techniques to the one elaborated above for the Heston model can be used to obviate the need for very small time steps.

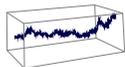


## Finite differencing methods and stochastic volatility models

Non-zero correlation between the different factors makes it impossible to use Alternating Direction Implicit methods (unless we transform away the correlation term which is usually very bad for the handling of boundary conditions, or we combine it with an explicit method for the cross terms which makes the scheme effectively explicit). Explicit methods require rather small time steps in order to avoid explosions due to numerical instabilities.

The multi-dimensional equivalent of the Crank-Nicolson method (also denoted as Peaceman-Rachford-Douglas method [[PR55](#), [DR56](#)]) can be implemented efficiently using iterative solver algorithms such as the stabilised biconjugate gradient method [[GL96](#)] that don't require the explicit specification of the matrix at all.

*All that is needed is a function that carries out the same calculations that would be done in an explicit method. A useful collection of utilities for this purpose is the Iterative Template Library [[LLS](#)].*



For the stochastic skew model, the independence of the three factors makes it possible to use a three factor Alternating Direction Implicit version of the Crank Nicolson method. This means it is possible to have large time steps in a fast finite differencing scheme.

Since we have zero correlation in the stochastic skew model, the number of discretisation layers in both the volatility and the skew factor can be kept small ( $\sim 20$ – $30$ ).

Also, boundary conditions can be kept simple in all directions and in the corners:  $V_{ii} = 0$  for  $i = x, y, \gamma$ .

The speed of three factor ADI implementations is compatible with that of any *safe* implementation involving numerical contour integrals or Fourier inversions of characteristic functions etc.



The generalisation of Alternating Direction Implicit (or Alternating Direction Crank-Nicolson) to multiple spatial dimensions is based on the idea of an *operator split* [PR55, DR56, Mar89].

Take the equation

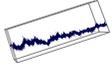
$$V_t + \sum_i \hat{\mu}_i(t, \mathbf{x}) V_{x_i} + \frac{1}{2} \sum_i \sigma_i^2(t, \mathbf{x}) V_{x_i x_i} = r \cdot V, \quad (42)$$

transform away the source term by setting  $u := V e^{-rt}$  (which almost certainly changes your boundary conditions)

$$u_t + \sum_i \underbrace{\left( \hat{\mu}_i \partial_{x_i} + \frac{1}{2} \sigma_i^2 \partial_{x_i}^2 \right)}_{L_i} \cdot u = 0. \quad (43)$$

i.e.

$$(\partial_t + L) \cdot u = 0 \quad \text{with} \quad L = \sum_i L_i. \quad (44)$$



Finite differencing of accuracy order  $\mathcal{O}(\Delta t^2)$ :

$$\partial_t \cdot u \rightarrow \frac{1}{\Delta t} [u(t, \mathbf{x}) - u(t - \Delta t, \mathbf{x})] \quad (45)$$

$$\partial_{x_i} \cdot u \rightarrow \frac{1}{2\Delta x_i} [u(t, \dots, x_i + \Delta x_i, \dots) - u(t, \dots, x_i - \Delta x_i, \dots)] \quad (46)$$

$$\partial_{x_i}^2 \cdot u \rightarrow \frac{1}{\Delta x_i^2} [u(t, \dots, x_i + \Delta x_i, \dots) - 2u(t, \dots, x_i, \dots) + u(t, \dots, x_i - \Delta x_i, \dots)] \quad (47)$$

Discretisation of the differential operators yields  $L_i \rightarrow D_i$  with

$$\begin{aligned} D_i \cdot u(t, \mathbf{x}) &= \hat{\mu}_i(t, \mathbf{x}) \frac{1}{2\Delta x_i} [u(t, \dots, x_i + \Delta x_i, \dots) - u(t, \dots, x_i - \Delta x_i, \dots)] \\ &\quad + \frac{1}{2} \sigma_i^2(t, \mathbf{x}) \frac{1}{\Delta x_i^2} [u(t, \dots, x_i + \Delta x_i, \dots) - 2u(t, \mathbf{x}) + u(t, \dots, x_i - \Delta x_i, \dots)] \end{aligned} \quad (48)$$

$$\begin{aligned} D_i \cdot u(\mathbf{x}) &= \frac{1}{2\Delta x_i^2} \left[ \left( \sigma_i^2(t, \mathbf{x}) + \hat{\mu}_i(t, \mathbf{x}) \Delta x_i \right) u(\dots, x_i + \Delta x_i, \dots) - 2\sigma_i^2(t, \mathbf{x}) u(\mathbf{x}) \right. \\ &\quad \left. + \left( \sigma_i^2(t, \mathbf{x}) - \hat{\mu}_i(t, \mathbf{x}) \Delta x_i \right) u(\dots, x_i - \Delta x_i, \dots) \right] \end{aligned} \quad (49)$$



Crank-Nicolson:

$$(\partial_t + L) \cdot u(t, \mathbf{x}) = 0 \quad (50)$$

is to be approximated by

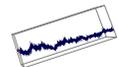
$$\frac{1}{\Delta t} \left[ u(t, \mathbf{x}) - u(t - \Delta t, \mathbf{x}) \right] + \frac{1}{2} D \cdot \left[ u(t, \mathbf{x}) + u(t - \Delta t, \mathbf{x}) \right] = 0 \quad (51)$$

This means, a single step in the Crank-Nicolson scheme is given by solving

$$\left( \mathbf{1} - \frac{1}{2} \Delta t D \right) \cdot u(t - \Delta t, \mathbf{x}) = \left( \mathbf{1} + \frac{1}{2} \Delta t D \right) \cdot u(t, \mathbf{x}) \quad (52)$$

for  $u(t - \Delta t, \mathbf{x})$ .

The *operator split* of the discretised operator  $D = \sum_i D_i$  is to split  $D$  into its commuting components  $\{D_i\}$ , and to solve (52) for each of the  $D_i$  individually in sequence.

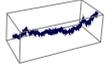


A single time step in the  $n$ -dimensional operator-split finite differencing scheme is thus given by a sequence of  $n$  one-dimensional finite differencing steps. Solve:

$$\begin{aligned}(\mathbf{1} - \frac{1}{2}\Delta t D_1) \cdot \tilde{u}^{(1)}(\mathbf{x}) &= (\mathbf{1} + \frac{1}{2}\Delta t D_1) \cdot u(t, \mathbf{x}) \\(\mathbf{1} - \frac{1}{2}\Delta t D_2) \cdot \tilde{u}^{(2)}(\mathbf{x}) &= (\mathbf{1} + \frac{1}{2}\Delta t D_2) \cdot \tilde{u}^{(1)}(\mathbf{x}) \\&\quad \vdots \quad \quad \quad \vdots \\(\mathbf{1} - \frac{1}{2}\Delta t D_n) \cdot \tilde{u}^{(n)}(\mathbf{x}) &= (\mathbf{1} + \frac{1}{2}\Delta t D_n) \cdot \tilde{u}^{(n-1)}(\mathbf{x})\end{aligned}$$

and set

$$u(t - \Delta t, \mathbf{x}) := \tilde{u}^{(n)}(\mathbf{x}) .$$



For commuting  $D_i$  and  $D_j$ , i.e.  $D_i D_j = D_j D_i$ , this scheme is, like the one-dimensional Crank-Nicolson method, of convergence order  $\mathcal{O}(\Delta t^2)$ :

$$\begin{aligned}
 \tilde{u}^{(j)}(\mathbf{x}) &= \left(\mathbf{1} - \frac{1}{2}\Delta t D_i\right)^{-1} \cdot \left(\mathbf{1} + \frac{1}{2}\Delta t D_i\right) \cdot \tilde{u}^{(j-1)}(\mathbf{x}) \\
 &= \left(\mathbf{1} + \frac{1}{2}\Delta t D_i + \frac{1}{4}\Delta t^2 D_i^2\right) \cdot \left(\mathbf{1} + \frac{1}{2}\Delta t D_i\right) \cdot \tilde{u}^{(j-1)}(\mathbf{x}) + \mathcal{O}(\Delta t^3) \\
 &= \left(\mathbf{1} + \Delta t D_i + \frac{1}{2}\Delta t^2 D_i^2\right) \cdot \tilde{u}^{(j-1)}(\mathbf{x}) + \mathcal{O}(\Delta t^3) \tag{53}
 \end{aligned}$$

$\implies$

$$\begin{aligned}
 u(t - \Delta t) &= \left[ \prod_i \left(\mathbf{1} + \Delta t D_i + \frac{1}{2}\Delta t^2 D_i^2\right) \right] \cdot u(t) + \mathcal{O}(\Delta t^3) \\
 &= \left[ \mathbf{1} + \Delta t \sum_i D_i + \frac{1}{2}\Delta t^2 \sum_{i,j} D_i D_j \right] \cdot u(t) + \mathcal{O}(\Delta t^3) \tag{54}
 \end{aligned}$$

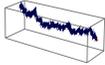


Equation (54) is of precisely the same form as the one we obtain for  $u$  in  $t$  from the continuous equation  $(\partial_t + L) \cdot u = 0$  :

$$u(t - \Delta t) = \left[ \mathbf{1} + \Delta t \sum_i L_i + \frac{1}{2} \Delta t^2 \sum_{i,j} L_i L_j \right] \cdot u(t) + \mathcal{O}(\Delta t^3) \quad (55)$$

In order to avoid a building up of lower order error terms due to the fact that  $D_i$  and  $D_j$  don't always commute perfectly (primarily due to the boundary conditions, but also due to round-off), the ordering of the scheme can be permuted.

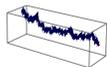
For a three-factor model, this means there are  $3! = 6$  permutations that we can cycle through as shown in the following example.



```
//  
// Schematic sample code for the control block and main loop of a three-dimensional operator split Crank-Nicolson method.  
//  
// This code does not contain examples for the implementation of the actual Crank-Nicolson steps that need to be  
// carried out for each of the three components, nor the incorporation of the lateral boundary conditions.  
//  
const unsigned long n1 = 200, n2 = 30, n3 = 30; // Sample values for the number of spatial levels in each direction.  
  
//  
// There are 6 possible permutations of a sequence of three elements. We therefore adjust the number of steps to be a  
// multiple of 6. When product related event dates are to be considered, this ought to be done for each time interval.  
//  
const unsigned long numberOfSteps = 200, adjustedNumberOfSteps = ((numberOfSteps+5)/6)*6;  
  
//  
// Each scheme consists of three steps. The set of all possible schemes is given by all possible permutations. We sort  
// them such that the last step of any one scheme is different from the first step of the next scheme in the sequence.  
//  
const unsigned long schemes[6][3] = {  
    { 0, 1, 2 }, // D1, D2, D3  
    { 0, 2, 1 }, // D1, D3, D2  
    { 2, 0, 1 }, // D3, D1, D2  
    { 2, 1, 0 }, // D3, D2, D1  
    { 1, 2, 0 }, // D2, D3, D1  
    { 1, 0, 2 }, // D2, D1, D3  
};  
  
//  
// The class ThreeDimensionalContainer is a user-written container for the solution values at the grid nodes.  
// Keep it simple and fast.  
//  
ThreeDimensionalContainer terminalBoundaryCondition(n1,n2,n3), workspace;  
//  
// Here, the terminal boundary conditions should be evaluated to populate the known lattice values at the final point  
// in time which is the starting point for the backwards induction algorithm. The evaluation of the terminal boundary  
// conditions will normally involve the layout of the grid in all three coordinates taking into account potential  
// discontinuities of the terminal boundary condition (effectively the initial values) or its derivative (you should  
// always have a grid level at the strike of plain vanilla options), the precomputation of any coefficient  
// combinations that will be constant for each spatial node through time, etc.  
//  
//  
ThreeDimensionalContainer * threeDimensionalContainers[2] = { &terminalBoundaryCondition, &workspace };  
ThreeDimensionalContainer * knownValues = &terminalBoundaryCondition, * unknownValues;  
unsigned long i, j, k, schemeindex=5, stepInSchemeIndex, containerIndicator=0;
```

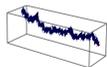


```
//  
// The main loop of backward induction.  
//  
for (i=0;i<adjustedNumberOfSteps;++i){  
  ++schemeindex %= 6;  
  for (stepInSchemeIndex=0;stepInSchemeIndex<3;++stepInSchemeIndex){  
    ++containerIndicator %= 2;  
    unknownValues = threeDimensionalContainers[containerIndicator];  
    switch (schemes[schemeindex][stepInSchemeIndex]){  
      case 0 : // Crank-Nicolson step in D1 to be placed here.  
              //  
              break;  
      case 1 : // Crank-Nicolson step in D2 to be placed here.  
              //  
              break;  
      case 2 : // Crank-Nicolson step in D3 to be placed here.  
              //  
              break;  
    }  
    knownValues = threeDimensionalContainers[containerIndicator];  
  }  
}  
  
//  
// Assuming that the grid levels are stored in the three one-dimensional vectors x1Values[], x2Values[], and  
// x3Values[], and that the spot coordinates are given by x1, x2, and x3, and that we have already asserted that  
// (x1,x2,x3) is inside the grid, we interpolate the solution at (x1,x2,x3) from the grid values.  
//  
for (i=0;x1Values[i]<x1;++i); for (j=0;x2Values[j]<x2;++j); for (k=0;x3Values[k]<x3;++k);  
  
const double p1 = (x1-x1Values[i-1])/(x1Values[i]-x1Values[i-1]), q1 = 1 - p1;  
const double p2 = (x2-x2Values[j-1])/(x2Values[j]-x2Values[j-1]), q2 = 1 - p2;  
const double p3 = (x3-x3Values[k-1])/(x3Values[k]-x3Values[k-1]), q3 = 1 - p3;  
  
//  
// Below, we assume that an object v of class ThreeDimensionalContainer allows you to retrieve  
// the value at the (i,j,k) grid coordinates by the use of the notation v(i,j,k).  
//  
const ThreeDimensionalContainer &v = *knownValues;  
//  
// Trilinear interpolation.  
//  
const double solution = p1*p2*p3*v(i,j,k) + p2*q1*p3*v(i-1,j,k) + p1*q2*p3*v(i,j-1,k) + q1*q2*p3*v(i-1,j-1,k)  
+ p1*p2*q3*v(i,j,k-1) + p2*q1*q3*v(i-1,j,k-1) + p1*q2*q3*v(i,j-1,k-1) + q1*q2*q3*v(i-1,j-1,k-1);
```



# References

- [Baa97] B. E. Baaquie. A Path Integral Approach to Option Pricing with Stochastic Volatility : Some Exact Results. *Journal de Physique*, 1(7):1733–1753, 1997.
- [CN47] J. Crank and P. Nicolson. A practical method for numerical evaluation of solutions of partial differential equations of the heat-conduction type. *Proc. Camb. Philos. Soc.*, 43:50–67, 1947.
- [CS89] M. Chesney and L. Scott. Pricing European Currency Options: A comparison of the modified Black-Scholes model and a random variance model. *Journal of Financial and Quantitative Analysis*, 24:267–284, September 1989.
- [Dos77] H. Doss. Liens entre équations différentielles stochastiques ordinaires. *Annales de l'Institut Henri Poincaré. Probabilités et Statistiques*, 13:99–125, 1977.
- [DR56] J. Douglas and H. H. Rachford. On the numerical solution of heat conduction problems in two and three space variables. *Transactions of the American Mathematical Society*, 82:421–439, 1956.
- [Fel51] W. Feller. Two Singular Diffusion Problems. *Annals of Mathematics*, 54:173–182, 1951.
- [FPS00] J.-P. Fouque, G. Papanicolaou, and K. R. Sircar. *Derivatives in Financial Markets with Stochastic Volatility*. Cambridge University Press, September 2000. ISBN 0521791634.
- [Ges77] R. Geske. The Valuation of Corporate Liabilities as Compound Options. *Journal of Financial and Quantitative Analysis*, 12:541–552, 1977.
- [GJ84] R. Geske and H. E. Johnson. The Valuation of Corporate Liabilities as Compound Options: A Correction. *Journal of Financial and Quantitative Analysis*, 19:231–232, 1984.
- [GL96] G. H. Golub and C. F. Van Loan. *Matrix Computations*. The John Hopkins University Press, 1983,1989,1996.
- [Hes93] S. L. Heston. A closed-form solution for options with stochastic volatility with applications to bond and currency options. *The Review of Financial Studies*, 6:327–343, 1993.
- [HKL02] P. Hagan, D. Kumar, and A. S. Lesniewski. Managing smile risk. *Wilmott*, pages 84–108, September 2002.
- [HW87] J. Hull and A. White. The Pricing of Options on Assets with Stochastic Volatilities. *The Journal of Finance*, 42(2):281–300, June 1987. [faculty.baruch.cuny.edu/lwu/890/HullWhite87.pdf](http://faculty.baruch.cuny.edu/lwu/890/HullWhite87.pdf).
- [HW88] J. Hull and A. White. An Analysis of the Bias in Option Pricing Caused by a Stochastic Volatility. *Advances in Futures and Options Research*, 3:27–61, 1988.
- [Jäc02] P. Jäckel. *Monte Carlo methods in finance*. John Wiley and Sons, February 2002.
- [KP99] P. E. Kloeden and E. Platen. *Numerical Solution of Stochastic Differential Equations*. Springer Verlag, 1992, 1995, 1999.
- [LLS] A. Lumsdaine, L. Lee, and J. Siek. The iterative template library. [www.osl.iu.edu/research/itl](http://www.osl.iu.edu/research/itl).
- [Mar89] G. I. Marchuk. Splitting and alternating direction methods. In J. Lions and P. Ciarlet, editors, *Handbook of Numerical Analysis, Volume I*, volume I, pages 197–462. Elsevier Science Publishers, 1989.
- [Mer73] R. C. Merton. Theory of Rational Option Pricing. *Bell Journal of Economics and Management Science*, 4:141–183, Spring 1973.
- [Mer76] R. C. Merton. Option Pricing When Underlying Stock Returns are Discontinuous. *Journal of Financial Economics*, 3:125–144, 1976.
- [Mer90] R. C. Merton. *Continuous-Time Finance*. Blackwell Publishers Ltd., 1990.
- [MM94] K. W. Morton and D. F. Mayers. *Numerical Solution of Partial Differential Equations*. Cambridge University Press, 1994. ISBN 0521429226.
- [PR55] D. W. Peaceman and H. H. Rachford. The numerical solution of parabolic and elliptic differential equations. *Journal of the Society for Industrial and Applied Mathematics*, 3:28–41, 1955.
- [PTVF92] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery. *Numerical Recipes in C*. Cambridge University Press, 1992.
- [Rub83] M. Rubinstein. Displaced diffusion option pricing. *Journal of Finance*, 38:213–217, March 1983.
- [RW00] L. C. G. Rogers and D. Williams. *Diffusions, Markov Processes and Martingales: Volume 2, Ito Calculus*. Cambridge University Press, September 2000.



- [Sco87] L. Scott. Option Pricing When the Variance Changes Randomly: Theory, Estimation and An Application. *Journal of Financial and Quantitative Analysis*, 22:419–438, December 1987.
- [SS91] E. M. Stein and J. C. Stein. Stock Price Distribution with Stochastic Volatility : An Analytic Approach. *Review of Financial Studies*, 4:727–752, 1991.
- [SZ99] R. Schöbel and J. Zhu. Stochastic Volatility With an Ornstein Uhlenbeck Process: An Extension. *European Finance Review*, 3:23–46, 1999.
- [TR00] D. Tavella and C. Randall. *Pricing Financial Instruments: The Finite Difference Method*. John Wiley and Sons, April 2000. ISBN 0471197602.
- [Wil00] P. Wilmott. *Quantitative Finance*. John Wiley and Sons, 2000.