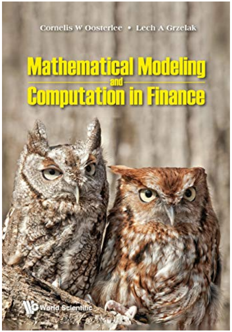
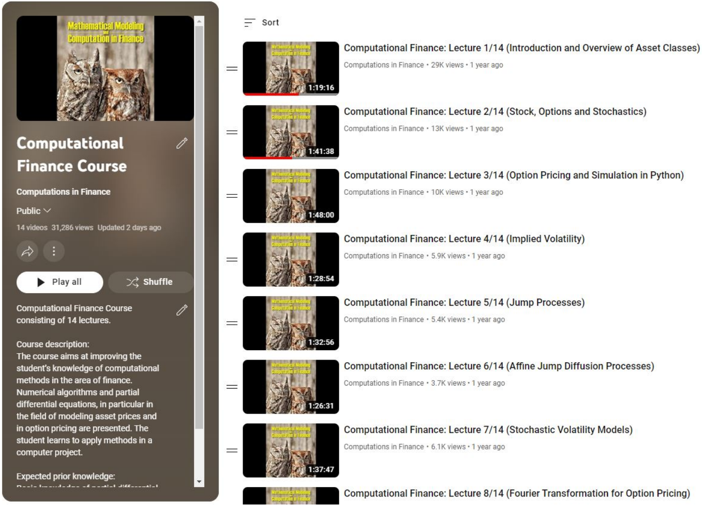
Materials for the course

The course is based on book
"*Mathematical Modeling and Computation in Finance: With Exercises and Python and MATLAB Computer Codes*",
by C.W. Oosterlee and L.A. Grzelak, WSP Europe Ltd, 2019.



- For more details regarding the book go here.
- YouTube Channel with courses can be found here.
- Slides and the codes can be found here.

Materials for the course



- The playlist with the Computational Finance course can be found here.

Questions and Answers, Lectures 1-14, Volume 1

1. *Lecture 1:* Can we use the same pricing models for different asset classes?
2. *Lecture 1:* How is the money savings account related to a zero-coupon bond?
3. *Lecture 2:* What are the challenges in the calculation of implied volatilities?
4. *Lecture 2:* Can you price options using Arithmetic Brownian motion?
5. *Lecture 2:* What is the difference between a stochastic process and a random variable?
6. *Lecture 2:* What are the advantages and disadvantages of using ABM/GBM for modelling a stock process?
7. *Lecture 3:* What sanity checks can you perform for a simulated stock process?
8. *Lecture 3:* What is the Feynman-Kac formula?
9. *Lecture 4:* What is the implied volatility term structure?
10. *Lecture 4:* What are the deficiencies of the Black-Scholes model? Why is the BS model still used?

Questions and Answers, Lectures 1-14, Volume 1

11. *Lecture 5:* How does the so-called Itô's table look like if we include the Poisson jump process?
12. *Lecture 5:* What is the impact of jumps on implied volatility?
13. *Lecture 5:* How to derive a characteristic function for a model with jumps?
14. *Lecture 6:* Is the Heston model with time-dependent parameters affine?
15. *Lecture 6:* Why is adding more and more factors to the pricing models not the best idea?
16. *Lecture 7:* Can you interpret the Heston model parameters and their impact on the volatility surface?
17. *Lecture 7:* Can we model volatility with the Arithmetic Brownian Motion process?
18. *Lecture 8:* What are the benefits of FFT compared to a "brute force" integration?
19. *Lecture 8:* What to do if the FFT/COS method does not converge for an increasing number of expansion terms?
20. *Lecture 9:* What is a standard error? How to interpret it?

Questions and Answers, Lectures 1-14, Volume 1

21. *Lecture 9:* What is weak and strong convergence in Monte Carlo pricing?
22. *Lecture 10:* What are the challenges of discretizing the CIR process using the Euler method?
23. *Lecture 10:* Why do we need Monte Carlo if we have FFT methods for pricing?
24. *Lecture 11:* How to hedge Jumps?
25. *Lecture 11:* What is pathwise sensitivity?
26. *Lecture 12:* What is the Bates model, and how can it be used for pricing?
27. *Lecture 12:* What is the relation between the European options and the Forward-start options?
28. *Lecture 12:* What instruments to choose to calibrate your pricing model?
29. *Lecture 13:* How to calibrate a pricing model? How to choose the objective function?
30. *Lecture 13:* What are the Chooser options?

Question 1/30 (Lecture 1):
Can we use the same pricing models for different asset classes?

- Stock market (selling and buying stocks)
- Option exchanges (trading financial options)
- Interest rates market
- FX market
- Credit market
- Commodity market (gold, metals, corn, meat, futures, ...)
- Energy market (oil, gas, ...)
- OTC market (over the counter, not traded on a regulated exchange)
- Crypto currencies

$\Rightarrow$ Each market gives us mathematical questions

Question 1/30 (Lecture 1):
Can we use the same pricing models for different asset classes?

- When choosing a stochastic process, we need to establish the physical properties of the asset
  - is it always positive?
  - how can we estimate the model parameters?
  - is there an options market?
  - what is the market practice for modelling?

A few examples of processes and their applications:

- Geometric Brownian Motion (GBM) (stocks, FX, etc.)

$$\mathrm{d}X(t) = rX(t)\mathrm{d}t + \sigma X(t)\mathrm{d}W(t),$$

- Ornstein-Uhlenbeck mean reverting processes (interest rates, commodities)

$$\mathrm{d}X(t) = \kappa(\theta - X(t))\mathrm{d}t + \sigma\mathrm{d}W(t),$$

  where: μ, σ, κ and θ are known constants, and $W(t)$ is a Wiener process.
- the Heston model, the Bates model, the local-volatility model etc.

Question 2/30 (Lecture 1): **How is the money savings account related to a zero-coupon bond?**

The simplest concept in finance is the time value of money.

- €1 today is worth more than €1 in a year's time.
- Invest €1 a discrete interest rate of r (assumed to be constant), paid once per year. After one year your bank account contains $1 \times (1 + r)$.
- There are several types of interest:
  - There is simple and compound interest. Simple interest is when the interest you receive is based only on the amount you initially invest, compound interest is when you get interest on your interest.
  - Interest typically comes in two forms, discretely compounded and continuously compounded.

Question 2/30 (Lecture 1):
How is the money savings account related to a zero-coupon bond?

- With $M(t)$ in the bank at time t, how does this increase with time?

  If you check your account a short period later, time $t + \mathrm{d}t$, the amount will have increased by

$$M(t + \mathrm{d}t) - M(t) \approx \frac{\mathrm{d}M}{\mathrm{d}t}\mathrm{d}t + \dots, \quad \text{(Taylor series expansion).}$$

  The interest you receive must be proportional to the amount you have, M, the interest rate r and the time-step, dt. Thus,

$$\frac{\mathrm{d}M}{\mathrm{d}t}\mathrm{d}t = rM(t)\mathrm{d}t \;\text{ giving }\; \frac{\mathrm{d}M}{\mathrm{d}t} = rM(t) \;.$$

  If you have €M(0) initially, then the solution is $M(t) = M(0)\mathrm{e}^{rt}$.

  Conversely, if you know you will get 1€ at time T in the future, its value at an earlier time t is simply

$$\mathrm{e}^{-r(T-t)} \;.$$

Question 2/30 (Lecture 1): **How is the money savings account related to a zero-coupon bond?**

- A zero-coupon bond is a contract with price $P(t,T)$, at time $t < T$, to deliver at time T, $P(T,T) = €1$.

Figure: Cash flow for a zero-coupon bond, $P(t,T)$, with the payment at time T.

Question 2/30 (Lecture 1):
How is the money savings account related to a zero-coupon bond?

- A basic interest rate product is the zero-coupon bond, $P(t, T)$, which pays 1 currency unit at maturity time T, i.e. $P(T, T) = 1$. We are interested in its value at a time $t < T$.
- The fundamental theorem of asset pricing states that the price at time t of any contingent claim with payoff, $H(T)$, is given by:

$$V(t) = \mathbb{E}^{\mathbb{Q}}\left[\mathrm{e}^{-\int_t^T r(z)dz} H(T) \middle| \mathcal{F}(t)\right],$$

where the expectation is taken under the risk-neutral measure $\mathbb{Q}$.

- The price of a zero-coupon bond at time t with maturity T is thus given by:

$$P(t, T) = \mathbb{E}^{\mathbb{Q}}\left[\mathrm{e}^{-\int_t^T r(z)\mathrm{d}z} \middle| \mathcal{F}(t)\right],$$

since $H(T) = V(T) = P(T, T) \equiv 1$.

Question 2/30 (Lecture 1):
How is the money savings account related to a zero-coupon bond?

To summarize:

- Money savings account, $M(t)$:

$$M(T) = M(t)\mathrm{e}^{\int_t^t r(s)\mathrm{d}s},$$

- Zero-coupon bond, $P(t, T)$:

$$P(t, T) = \mathbb{E}\left[\frac{M(t)}{M(T)}\middle|\mathcal{F}(t)\right] = \mathbb{E}\left[\mathrm{e}^{-\int_t^T r(s)\mathrm{d}s}\middle|\mathcal{F}(t)\right],$$

Question 3/30 (Lecture 2 & 4):
What are the challenges in calculating implied volatilities?

Implied Volatility: "the wrong number in the wrong formula to get the right price". [Rebonato 1999]

Mathematically we have:

$$V_c(t, S) = V_c(\sigma, t, S(t), K, T)$$

where *BS* is monotonically increasing in σ (higher volatility corresponds to higher prices). Now, assume the existence of some inverse function

$$g_\sigma(\cdot) = V_c^{-1}(\cdot)$$

so that

$$\sigma_{impl} = g_\sigma(V_c^{mkt}, r, T, K, S_0).$$

By computing the implied volatility for traded options with different strikes and maturities, we can test the Black-Scholes model.

Question 3/30 (Lecture 2 & 4):
What are the challenges in calculating implied volatilities?

The price of a European call option with a strike price of K and maturity T is given by the formula:

$$
\begin{aligned}
V_c(\sigma, t, S(t), K, T) &= S(t) \cdot \mathcal{N}(d_1(t, S)) - \mathrm{e}^{-r(T-t)} K \cdot \mathcal{N}(d_2(t, S(t))), \text{ with} \\
d_1(t, S(t)) &= \frac{1}{\sigma\sqrt{T-t}} \left(\log \frac{S(t)}{K} + \left(r + \frac{1}{2}\sigma^2 \right)(T-t) \right), \\
d_2(t, S(t)) &= d_1(t, S(t)) - \sigma\sqrt{T-t},
\end{aligned}
$$

where $\mathcal{N}$ is the cumulative distribution function for standard normal distribution i.e., $\mathcal{N}(0,1)$.

Question 3/30 (Lecture 2 & 4):
What are the challenges in calculating implied volatilities?

The BS pricing function BS does not have a closed-form solution for its inverse $g_\sigma(\cdot)$. Instead, a root-finding technique is used to solve the equation:

$$V_c(\sigma_{imp}, t, S(t), K, T) - V_c^{mkt} = 0.$$

There are many ways to solve this equation, and one of the most popular method are methods of:

- "Newton"
- "Brent"
- and ideas presented by P.Jackel "Let's be rational". [1]

Since the options prices can move very quickly, it is often important to use the most efficient method when calculating implied volatilities.

[1] http://en.wikipedia.org/wiki/Brent's_method
http://www.jaeckel.org/LetsBeRational.pdf

Question 4/30 (Lecture 2): **Can you price options using Arithmetic Brownian motion?**

We compare the behaviour of the following SDEs:

- Arithmetic Brownian motion (ABM)

$$\mathrm{d}X(t) = \mu \mathrm{d}t + \sigma \mathrm{d}W(t),$$

- Geometric Brownian Motion (GBM)

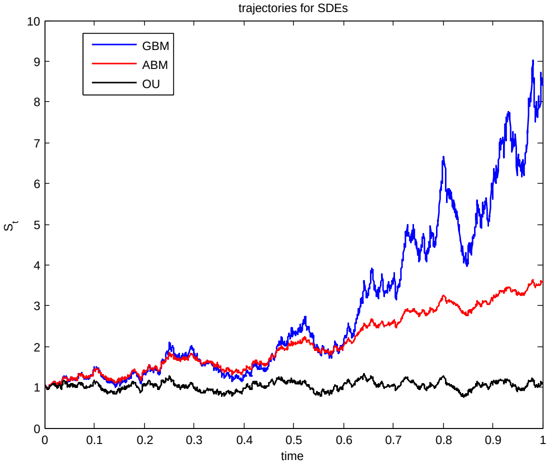
$$\mathrm{d}X(t) = \mu X(t)\mathrm{d}t + \sigma X(t)\mathrm{d}W(t),$$

- Ornstein-Uhlenbeck mean reverting processes (OU)

$$\mathrm{d}X(t) = \kappa(\theta - X(t))\mathrm{d}t + \sigma \mathrm{d}W(t),$$

where: μ, σ, κ and θ are known constants, and $W(t)$ is a Wiener process.

Question 4/30 (Lecture 2):
Can you price options using Arithmetic Brownian motion?



Question 5/30 (Lecture 2): What is the difference between a stochastic process and a random variable?

- We can interpret the observed stock values as a realization $X(t)(\omega)$ of the random variable $X(t)$.
- We look for a model which takes into account almost continuous realizations of the stock prices.

Definition (Stochastic Process)

A stochastic process $X(t)$ is a collection of random variables

$$(X(t), t \in \mathcal{T}) = (X(t)(\omega), t \in \mathcal{T}, \omega \in \Omega)$$

A stochastic process $X(t)$ is a function of two variables:

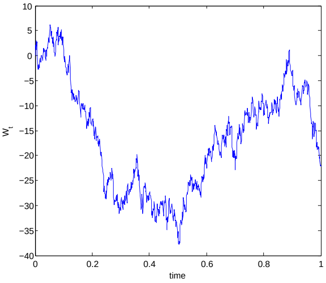
- for a fixed instant t it's a random variable:

$$X(t) = X(t)(\omega) \text{ for } \omega \in \Omega,$$

- for a fixed random outcome $\omega \in \Omega$, it's a function of time

$$X(t) = X(t)(\omega), \text{ for } t \in \mathcal{T}.$$

Process trajectories



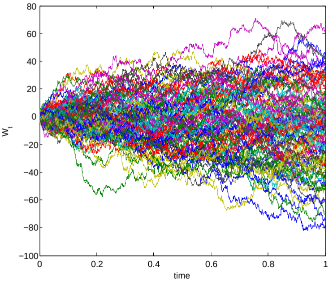


Figure: Sample paths of Brownian motion on $[0, 1]$. Left: 1 path, Right: 100 paths.

Question 6/30 (Lecture 2): **What are the advantages and disadvantages of using ABM/GBM for modelling a stock process?**

We compare the behaviour of the following SDEs:

- Arithmetic Brownian motion (ABM)

$$\mathrm{d}X(t) = \mu \mathrm{d}t + \sigma \mathrm{d}W(t),$$

- Geometric Brownian Motion (GBM)

$$\mathrm{d}X(t) = \mu X(t)\mathrm{d}t + \sigma X(t)\mathrm{d}W(t).$$

When choosing a stochastic process we need to consider the following features:

- Do the stochastic paths reflect physical features of the asset?
- Can the parameters be estimated? and how?
- Can we price derivatives efficiently and adequately?

Question 7/30 (Lecture 3): **What sanity checks can you perform for a simulated stock process?**

There are several simple checks you can perform to check the quality of the simulated process:

- Is the discounted asset a martingale? $\mathbb{E}[S(T)/M(T)] = \frac{S(t_0)}{M(t_0)}$.
- Can you simplify a derivative under investigation? e.g., $K = 0$ for the Call option?
- Stability for the increasing number of the simulated paths.
- Stability for the change of the random seed?
- How do results vary when changing the size of the discretization time step?
- Can the model price back the market instruments?
- Depending on the pricing contract, many other checks may exist.

Question 8/30 (Lecture 3):
What is the Feynman-Kac formula?

Feynman-Kac established a link between partial differential equations (PDEs) and stochastic processes. It offers a method for solving certain PDEs by simulating random paths of a stochastic process. Suppose we are given the PDE:

$$\frac{\partial V}{\partial t} + \bar{\mu}(x,t)\frac{\partial V}{\partial x} + \frac{1}{2}\bar{\sigma}^2(x,t)\frac{\partial^2 V}{\partial x^2} = 0,$$

subject to the final condition $V(x,T) = \eta(x)$, then the Feynman-Kac formula reads:

$$V(x,t) = \mathbb{E}\left(\eta(X(T))|\mathcal{F}(t)\right)$$

where: $X(t)$ is an Ito process driven by the equation:

$$\mathrm{d}X(t) = \bar{\mu}(X(t),t)\mathrm{d}t + \bar{\sigma}(X(t),t)\mathrm{d}W(t),$$

with $W(t)$ is a Wiener process, with initial for $X(t) = x$.

Question 8/30 (Lecture 3):
What is the Feynman-Kac formula?

Solve the following PDE

$$\begin{aligned} \frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 \frac{\partial^2 V}{\partial x^2} &= 0 \\ V(x, T) &= x^2 \end{aligned}$$

where σ is a constant.

Answer: From Feynman-Kac we have:

$$V(x,t) = \mathbb{E}\left(\eta(X(T))|\mathcal{F}(t)\right) = \mathbb{E}\left(X(T)^2|\mathcal{F}(t)\right)$$

where:

$$\begin{aligned} \mathrm{d}X(s) &= 0 \cdot \mathrm{d}t + \sigma \mathrm{d}W(s) \\ X(t) &= x \end{aligned}$$

So we have: $X(T) = x + \sigma\left(W(T) - W(t)\right)$, and $X(T)$ has the distribution $N(x, \sigma\sqrt{T-t})$. Finally $V(t,x) = \sigma^2(T-t) + x^2$.

Question 9/30 (Lecture 4):
What is the implied volatility term structure?

- The BS pricing function BS does not have a closed-form solution for its inverse $g_\sigma(\cdot)$. Instead, a root finding technique is used to solve the equation:

$$BS(\sigma_{impl}, r, T, K, S_0) - V_C^{mkt} = 0.$$

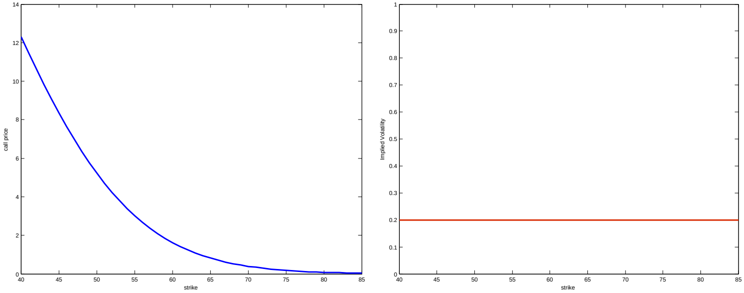
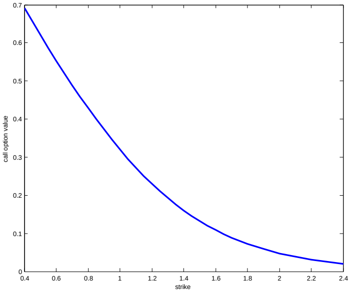


Figure: MODEL–LEFT: BS Call Prices, RIGHT: Implied Volatilities.

Question 9/30 (Lecture 4):
What is the implied volatility term structure?



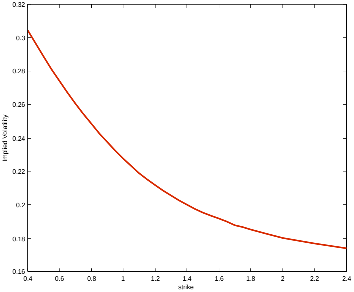


Figure: MARKET DATA– LEFT: Market Call Prices, RIGHT: Implied Volatilities.

Question 9/30 (Lecture 4):
What is the implied volatility term structure?

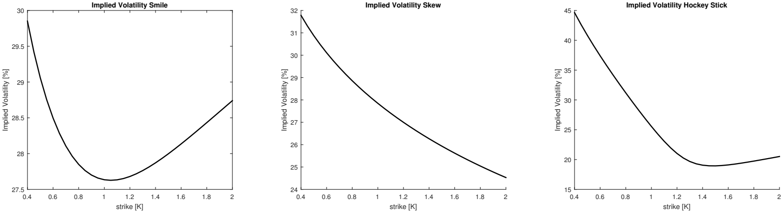


Figure: Typical implied volatility shapes: a smile, a skew and the so-called hockey stick. The hockey stick can be seen as a combination of the implied volatility smile and the skew.

Question 9/30 (Lecture 4):
What is the implied volatility term structure?

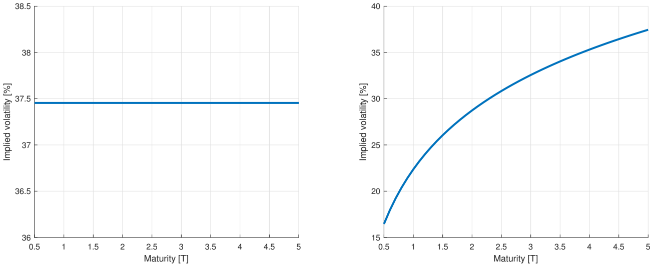


Figure: Comparison of the volatility term structure (for ATM volatilities) for Black-Scholes model with constant volatility σ_* (left-side picture) versus a model with time-dependent volatility $\sigma(t)$ (right-side picture).

Question 10/30 (Lecture 4):
What are the deficiencies of the Black-Scholes model? Why is the BS model still used?

- The Black-Scholes model, and its notion of hedging option contracts by stocks and money, forms the foundation of modern finance. However:
- Delta hedging is supposed to be a continuous process, but in practice it is a discrete process (a hedged portfolio is typically updated once a week or so).
- Empirical studies of financial time series have revealed that the normality assumption for asset prices cannot capture *heavy tails* and *asymmetries*, present in log-asset returns in practice.
- The volatility is supposed to be a known deterministic function of time, which is inconsistent since numerical inversion of the Black-Scholes equation based on market prices from different strikes and fixed maturity, produces a so-called *volatility skew or smile*.

Question 10/30 (Lecture 4):
What are the deficiencies of the Black-Scholes model? Why is the BS model still used?

- $\Rightarrow$ The idea of implied volatility does not fit to the Black-Scholes model
- Look for market consistent asset price models.
- $\Rightarrow$ Use a local volatility, a model with jumps, or stochastic volatility to better fit market data, and incorporate smile effects

Question 11/30 (Lecture 5):
How does the so-called Itô's table look like if we include the Poisson jump process?

- An intuitive explanation for Itô's formula in the case of jumps is that when a jump takes place, i.e. $\mathrm{d}X_{\mathcal{P}}(t) = 1$, the process "jumps" from $X(t_-)$ to $X(t)$, with the jump size determined by function $\bar{J}(t, X(t))$, resulting in the following relation:
$$g(t, X(t)) = g\left(t, X(t_-) + \bar{J}(t, X(t_-)\right).$$
- After the jump at time t, the function $g(\cdot)$ is adjusted with the jump size, which was determined at time t_-.

Question 11/30 (Lecture 5):
How does the so-called Itô's table look like if we include the Poisson jump process?

Lemma (Itô's lemma for Poisson process)

We consider a cádlág process, $X(t)$, defined as:

$$\mathrm{d}X(t) = \bar{\mu}(t, X(t))\mathrm{d}t + \bar{J}(t, X(t_-))\mathrm{d}X_{\mathcal{P}}(t), \quad \textit{with} \quad X(t_0) \in \mathbb{R},$$

where $\bar{\mu}, \bar{J} : [0, \infty) \times \mathbb{R} \to \mathbb{R}$ are deterministic, continuous functions and $X_{\mathcal{P}}(t)$ is a Poisson process, starting at $t_0 = 0$.
For a differentiable function, $g : [0, \infty) \times \mathbb{R} \to \mathbb{R}$, the Itô differential is given by:

$$\begin{aligned}
\mathrm{d}g(t, X(t)) \quad &= \quad \left[\frac{\partial g(t, X(t))}{\partial t} + \bar{\mu}(t, X(t))\frac{\partial g(t, X(t))}{\partial X}\right]\mathrm{d}t \\
&+ \quad \Big[g\left(t, X(t_-) + \bar{J}(t, X(t_-))\right) - g(t, X(t_-))\Big]\mathrm{d}X_{\mathcal{P}}(t),
\end{aligned}$$

where the left limit is denoted by $X(t_-) := \lim_{s \to t} X(s), s < t$, so that, by the continuity of $\bar{J}(\cdot)$, its left limit equals $\bar{J}(t, X(t_-))$.

Question 11/30 (Lecture 5):

How does the so-called Itô's table look like if we include the Poisson jump process?

- We made use of the Itô multiplication table, where the cross terms involving the Poisson process are also handled.
- An intuitive way to understand the Poisson process rule in the table is found in the notion that the term $\mathrm{d}X_{\mathcal{P}} = 1$ with probability $\xi_p \mathrm{d}t$, and $\mathrm{d}X_{\mathcal{P}} = 0$ with probability $(1 - \xi_p \mathrm{d}t)$, which implies that

$$
\begin{aligned}
(\mathrm{d}X_{\mathcal{P}})^2 &= \begin{cases} 1^2 & \text{with probability } \xi_p \mathrm{d}t, \\ 0^2 & \text{with probability } (1-\xi_p \mathrm{d}t) \end{cases} \\
&= \mathrm{d}X_{\mathcal{P}}.
\end{aligned}
$$

	$\mathrm{d}t$	$\mathrm{d}W(t)$	$\mathrm{d}X_{\mathcal{P}}(t)$
$\mathrm{d}t$	0	0	0
$\mathrm{d}W(t)$	0	$\mathrm{d}t$	0
$\mathrm{d}X_{\mathcal{P}}(t)$	0	0	$\mathrm{d}X_{\mathcal{P}}(t)$

Table: Itô multiplication table for Poisson process.

Question 11/30 (Lecture 5):
How does the so-called Itô's table look like if we include the Poisson jump process?

- The dynamics of the combined process are given by:

$$\mathrm{d}X(t) = \bar{\mu}(t, X(t))\mathrm{d}t + \bar{J}(t, X(t_-))\mathrm{d}X_{\mathcal{P}}(t) + \bar{\sigma}(t, X(t))\mathrm{d}W(t),$$

- Assuming that the Poisson process $X_{\mathcal{P}}(t)$ is independent of the Brownian motion $W(t)$, the dynamics of $g(t, X(t))$ are given by:

$$\begin{aligned}
\mathrm{d}g(t, X(t)) &= \left[\frac{\partial g(t, X(t))}{\partial t} + \bar{\mu}(t, X(t))\frac{\partial g(t, X(t))}{\partial X}\right. \\
&+ \left.\frac{1}{2}\bar{\sigma}^2(t, X(t))\frac{\partial^2 g(t, X(t))}{\partial X^2}\right]\mathrm{d}t \\
&+ \Big[g\big(t, X(t_-) + \bar{J}(t, X(t_-))\big) - g\big(t, X(t_-)\big)\Big]\mathrm{d}X_{\mathcal{P}}(t) \\
&+ \bar{\sigma}(t, X(t))\frac{\partial g(t, X(t))}{\partial X}\mathrm{d}W(t),
\end{aligned}$$

Question 12/30 (Lecture 5):
What is the impact of jumps on implied volatility?

- The dynamics for stock $S(t)$ under the risk-neutral measure $\mathbb{Q}$ are therefore given by:

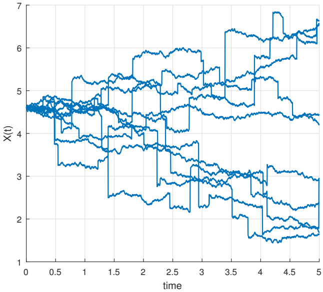
$$\boxed{\frac{\mathrm{d}S(t)}{S(t)} = \left(r - \xi_p \mathbb{E}\left[\mathrm{e}^J - 1\right]\right)\mathrm{d}t + \sigma \mathrm{d}W^{\mathbb{Q}}(t) + \left(\mathrm{e}^J - 1\right)\mathrm{d}X_{\mathcal{P}}^{\mathbb{Q}}(t).}$$

- The process is often presented in the literature as the *standard jump diffusion model*. The standard jump-diffusion model is directly connected to the following $\mathrm{d}X(t)$ dynamics:

$$\mathrm{d}X(t) = \left(r - \xi_p \mathbb{E}\left[\mathrm{e}^J - 1\right] - \frac{1}{2}\sigma^2\right)\mathrm{d}t + \sigma \mathrm{d}W^{\mathbb{Q}}(t) + J\mathrm{d}X_{\mathcal{P}}^{\mathbb{Q}}(t).$$

Question 12/30 (Lecture 5):
What is the impact of jumps on implied volatility?

- In Figure below examples of paths for $X(t)$ and $S(t) = \mathrm{e}^{X(t)}$ are presented. Here, the classical Merton model is used, $J \sim \mathcal{N}(\mu_J, \sigma_J^2)$, where the jumps are symmetric.



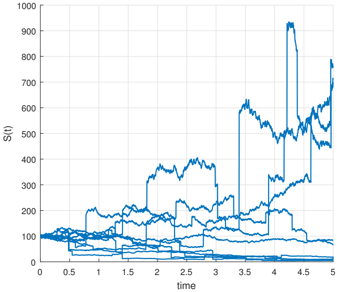
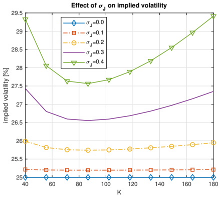
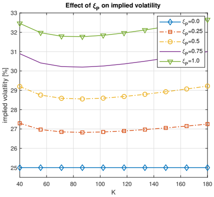


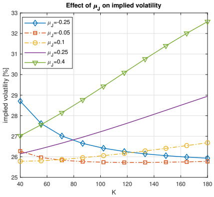
Figure: Left side: Paths of process $X(t)$; Right side: $S(t)$ with $S(t_0) = 100$, $r = 0.05$, $\sigma = 0.2$, $\sigma_J = 0.5$, $\mu_J = 0$, $\xi_p = 1$ and $T = 5$.

Question 12/30 (Lecture 5):
What is the impact of jumps on implied volatility?

- Illustrate the impact of the jump parameters in the Merton jump-diffusion model on the implied volatilities. $J \sim \mathcal{N}(\mu_J, \sigma_J^2)$.
- The influence of ξ_p, σ_J and μ_J. Each parameter is varied individually, while the other parameters are fixed.
- σ_J has a significant impact on the curvature, ξ_p controls the overall level of the implied volatility, whereas μ_J influences the implied volatility slope (the skew).







Question 13/30 (Lecture 5):
How to derive a characteristic function for a model with jumps?

- Merton's jump-diffusion model under $\mathbb{Q}$ consists of a Brownian motion and a compound Poisson process, with $t_0 = 0$,

$$X(t) = X(t_0) + \bar{\mu}t + \sigma W(t) + \sum_{k=1}^{X_{\mathcal{P}}(t)} J_k.$$

  where, $\bar{\mu} = r - \frac{1}{2}\sigma^2 - \xi_p \mathbb{E}\left[\mathrm{e}^J - 1\right]$, $\sigma > 0$, BM $W(t)$, Poisson process $X_{\mathcal{P}}(t), t \geq 0$ with ξ_p and $\mathbb{E}[X_{\mathcal{P}}(t)|\mathcal{F}(0)] = \xi_p t$.
- In the Poisson process setting, the arrival of a jump is independent of the arrival of previous jumps, and the probability of two simultaneous jumps is equal to zero.
- Variable J_k, $k \geq 1$, is an i.i.d. sequence of random variables with a jump-size probability distribution F_J, so $\mathbb{E}[J_k] = \mu_J < \infty$.

Question 13/30 (Lecture 5):
How to derive a characteristic function for a model with jumps?

- With all sources of randomness mutually independent, the characteristic function of $X(t)$ looks,

$$\phi_X(u) := \mathbb{E}\left[\mathrm{e}^{iuX(t)}\right] = \mathrm{e}^{iuX(0)}\mathrm{e}^{iu\bar{\mu}t}\mathbb{E}\left[\mathrm{e}^{iu\sigma W(t)}\right] \cdot \mathbb{E}\left[\exp\left(iu\sum_{k=1}^{X_{\mathcal{P}}(t)} J_k\right)\right].$$

- As $W(t) \sim \mathcal{N}(0,t)$, it follows $\mathbb{E}\left[\mathrm{e}^{iu\sigma W(t)}\right] = \mathrm{e}^{-\frac{1}{2}\sigma^2 u^2 t}$. For the second expectation, consider the summation:

$$\mathbb{E}\left[\exp\left(iu\sum_{k=1}^{X_{\mathcal{P}}(t)} J_k\right)\right] = \sum_{n\geq 0}\mathbb{E}\left[\exp\left(iu\sum_{k=1}^{X_{\mathcal{P}}(t)} J_k\right)\Bigg| X_{\mathcal{P}}(t) = n\right]\mathbb{P}[X_{\mathcal{P}}(t) = n],$$

which results from the tower property of expectations. We have,

$$\begin{aligned}\mathbb{E}\left[\exp\left(iu\sum_{k=1}^{X_{\mathcal{P}}(t)} J_k\right)\right] &= \sum_{n\geq 0}\mathbb{E}\left[\exp\left(iu\sum_{k=1}^{n} J_k\right)\right]\frac{\mathrm{e}^{-\xi_p t}(\xi_p t)^n}{n!}\\ &= \sum_{n\geq 0}\frac{\mathrm{e}^{-\xi_p t}(\xi_p t)^n}{n!}\left(\int_{\mathbb{R}}\mathrm{e}^{iux}f_J(x)\mathrm{d}x\right)^n.\end{aligned}$$

Question 13/30 (Lecture 5):
How to derive a characteristic function for a model with jumps?

- The two nth powers at the right-hand side are seen as a Taylor expansion of an exponential. So,

$$
\begin{aligned}
\mathbb{E}\left[\exp\left(iu\sum_{k=1}^{X_{\mathcal{P}}(t)} J_k\right)\right] &= \mathrm{e}^{-\xi_p t}\sum_{n\geq 0}\frac{1}{n!}\left(\xi_p t\int_{\mathbb{R}}\mathrm{e}^{iux}f_J(x)\mathrm{d}x\right)^n \\
&= \exp\left(\xi_p t\int_{\mathbb{R}}\left(\mathrm{e}^{iux}f_J(x)\mathrm{d}x-1\right)\right) \\
&= \exp\left(\xi_p t\int_{\mathbb{R}}\left(\mathrm{e}^{iux}-1\right)f_J(x)\mathrm{d}x\right) \\
&= \exp\left(\xi_p t\mathbb{E}[\mathrm{e}^{iuJ}-1]\right),
\end{aligned}
$$

using $\int_{\mathbb{R}} f_J(x)\mathrm{d}x = 1$, and $J = J_k$ an i.i.d. sequence of RVs with CDF $F_J(x)$ and PDF $f_J(x)$.

Question 13/30 (Lecture 5):
How to derive a characteristic function for a model with jumps?

- The ChF can thus be written as:

$$\begin{aligned}
\phi_X(u) &= \mathbb{E}\left[\mathrm{e}^{iuX(t)}\right] \\
&= \exp\left(iu(X(0)+\bar{\mu}t)-\frac{1}{2}\sigma^2u^2t\right)\exp\left(\xi_p t\left(\mathbb{E}[\mathrm{e}^{iuJ}-1]\right)\right) \\
&= \exp\left(iu(X(0)+\bar{\mu}t)-\frac{1}{2}\sigma^2u^2t+\xi_p t\int_{\mathbb{R}}\left(\mathrm{e}^{iux}-1\right)f_J(x)\mathrm{d}x\right),
\end{aligned}$$

  with $\bar{\mu}=r-\frac{1}{2}\sigma^2-\xi_p\mathbb{E}\left[\mathrm{e}^J-1\right]$.

- For the Merton model, we have,

$$\mathbb{E}[\mathrm{e}^{iuJ}-1]=\mathrm{e}^{iu\mu_J-\frac{1}{2}u^2\sigma_J^2}-1,\ \mathbb{E}[\mathrm{e}^J-1]=\mathrm{e}^{\mu_J+\frac{1}{2}\sigma_J^2}-1.$$

Question 14/30 (Lecture 6&7):
Is the Heston model with time-dependent parameters affine?

- Suppose we have the following system of SDEs:

$$\mathrm{d}\mathbf{X}(t) = \bar{\mu}(\mathbf{X}(t))\mathrm{d}t + \bar{\sigma}(\mathbf{X}(t))\mathrm{d}\widetilde{\mathbf{W}}(t),$$

with independent Brownian motions $\widetilde{\mathbf{W}}(\mathbf{t})$. For processes in the affine diffusion (AD) class it is assumed that drift, volatility, and interest rate components are of the affine form, i.e.

$$\begin{aligned}
\bar{\mu}(\mathbf{X}(t)) &= a_0 + a_1\mathbf{X}(t) \text{ for } (a_0, a_1) \in \mathbb{R}^n \times \mathbb{R}^{n\times n},\\
\bar{\sigma}(\mathbf{X}(t))\bar{\sigma}(\mathbf{X}(t))^T &= (c_0)_{ij} + (c_1)_{ij}^T\mathbf{X}(t), (c_0, c_1) \in \mathbb{R}^{n\times n} \times \mathbb{R}^{n\times n\times n},\\
r(\mathbf{X}(t)) &= r_0 + r_1^T\mathbf{X}(t), \text{ for } (r_0, r_1) \in \mathbb{R} \times \mathbb{R}^n.
\end{aligned}$$

Question 14/30 (Lecture 6&7): **Is the Heston model with time-dependent parameters affine?**

- Duffie, Pan and Singleton (2000) have shown that for affine diffusion processes the discounted characteristic function, defined as:

$$\phi(\mathbf{X(t)},\mathbf{t},\mathbf{T},\mathbf{u}) \equiv \mathbb{E}^{\mathbb{Q}}\left[\mathrm{e}^{-\int_t^T r(\mathbf{X_s})ds}\mathrm{e}^{i\mathbf{u}\mathbf{X_T}}|\mathcal{F}(t)\right] \text{ for } \mathbf{u}\in\mathbb{C}^{\mathbf{n}},$$

  with boundary condition:

$$\phi(\mathbf{X_T},\mathbf{T},\mathbf{T},\mathbf{u}) = \mathrm{e}^{i\mathbf{u}^T\mathbf{X_T}},$$

  has a solution of the following form:

$$\phi(\mathbf{X(t)},\mathbf{t},\mathbf{T},\mathbf{u}) = \mathrm{e}^{A(\mathbf{u},t,T)+\mathbf{B}(\mathbf{u},t,T)^T\mathbf{X(t)}},$$

  How to find the coefficients $A(\mathbf{u},t,T)$ and $\mathbf{B(u,t,T)^T}$?

- The coefficients $A(\mathbf{u},t,T)$ and $\mathbf{B(u,t,T)^T}$ have to satisfy the following system of Riccati-type ODEs[2]:

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}\tau}A(\mathbf{u},\tau) &= -r_0+\mathbf{B}^Ta_0+\frac{1}{2}\mathbf{B}^Tc_0\mathbf{B},\\ \frac{\mathrm{d}}{\mathrm{d}\tau}\mathbf{B}(\mathbf{u},\tau) &= -r_1+a_1^T\mathbf{B}+\frac{1}{2}\mathbf{B}^Tc_1\mathbf{B}.\end{aligned}$$

[2] Note that we do not consider jumps.

Question 14/30 (Lecture 6&7): **Is the Heston model with time-dependent parameters affine?**

- The Heston model consists of two stochastic differential equations, for the underlying asset price, $S(t)$, and the variance process, $v(t)$, described under the risk-neutral measure, $\mathbb{Q}$, by

$$\begin{aligned} \mathrm{d}S(t) &= rS(t)\mathrm{d}t + \sqrt{v(t)}S(t)\mathrm{d}W_x(t), \\ \mathrm{d}v(t) &= \kappa(\bar{v} - v(t))\mathrm{d}t + \gamma\sqrt{v(t)}\mathrm{d}W_v(t). \end{aligned}$$

  Parameter interpretation.

- A correlation is defined between the underlying Brownian motions, $\mathrm{d}W_v(t)\mathrm{d}W_x(t) = \rho_{x,v}\mathrm{d}t$. Parameters $\kappa \geq 0$, $\bar{v} \geq 0$ and $\gamma > 0$ are called the speed of mean reversion, the long-term mean of the variance process and the volatility of the volatility, respectively.
  - r is the rate of the return,
  - $\bar{v}$ is the long vol, or long run average price volatility ($\lim_{t\to\infty} \mathbb{E}v(t) = \bar{v}$)
  - κ is the rate at which $v(t)$ reverts to $\bar{v}$,
  - γ is the vol- vol, or volatility of the volatility; as the name suggests, this determines the variance of $v(t)$.

Question 14/30 (Lecture 6&7): **Is the Heston model with time-dependent parameters affine?**

- From the definition of the Heston model we have:

$$\left\{\begin{array}{lcl} \mathrm{d}S(t) & = & rS(t)\mathrm{d}t + \sqrt{v(t)}S(t)\mathrm{d}W_x(t) \\ \mathrm{d}v(t) & = & \kappa\left(\bar{v} - v(t)\right)\mathrm{d}t + \gamma\sqrt{v(t)}\mathrm{d}W_v(t) \end{array}\right.$$

Is it affine?

$$\sigma(\mathbf{X(t)})\sigma(\mathbf{X(t)})^T = \left[\begin{array}{ll} v(t)S(t)^2 & S(t)v(t)\gamma\rho_{x,v} \\ S(t)v(t)\gamma\rho_{x,v} & \gamma^2 v(t) \end{array}\right]$$

IT IS NOT AFFINE!

Question 14/30 (Lecture 6&7):
Is the Heston model with time-dependent parameters affine?

- Let us define the log transform: $X(t) = \log S(t)$,

$$\begin{cases} \mathrm{d}X(t) &= \left(r - \frac{1}{2}v(t)\right)\mathrm{d}t + \sqrt{v(t)}\mathrm{d}W_x(t), \\ \mathrm{d}v(t) &= \kappa\left(\bar{v} - v(t)\right)\mathrm{d}t + \gamma\sqrt{v(t)}\mathrm{d}W_v(t). \end{cases}$$

Express the model in two independent Brownian motions

$$\begin{bmatrix} \mathrm{d}X(t) \\ \mathrm{d}v(t) \end{bmatrix} = \begin{bmatrix} r - \dfrac{1}{2}v(t) \\ \kappa(\bar{v} - v(t)) \end{bmatrix} dt + \begin{bmatrix} \sqrt{v(t)} & 0 \\ \rho_{x,v}\gamma\sqrt{v(t)} & \gamma\sqrt{(1-\rho_{x,v}^2)v(t)} \end{bmatrix} \begin{bmatrix} \mathrm{d}\widetilde{W}_x(t) \\ \mathrm{d}\widetilde{W}_v(t) \end{bmatrix}$$

where Brownian motions $\widetilde{W}_x$ and $\widetilde{W}_v$ are independent.
The instantaneous covariance matrix:

$$\bar{\sigma}(\mathbf{X}(\mathbf{t}))\bar{\sigma}(\mathbf{X}(\mathbf{t}))^T = \begin{bmatrix} v(t) & v(t)\gamma\rho_{x,v} \\ v(t)\gamma\rho_{x,v} & \gamma^2 v(t) \end{bmatrix}$$

AFFINE!

Question 14/30 (Lecture 6&7): **Is the Heston model with time-dependent parameters affine?**

- The instantaneous covariance matrix for time-dependent case:

$$\bar{\sigma}(\mathbf{X}(\mathbf{t}))\bar{\sigma}(\mathbf{X}(\mathbf{t}))^T = \begin{bmatrix} v(t) & v(t)\gamma(t)\rho_{x,v}(t) \\ v(t)\gamma(t)\rho_{x,v}(t) & \gamma^2(t)v(t) \end{bmatrix}$$

AFFINE!

- The difficulty lies in the calculation of the corresponding ODEs:

$$\begin{aligned} A(\mathbf{u},\tau) &= \int_{\mathbb{R}^+}\left[-r_0 + \mathbf{B}^T a_0 + \frac{1}{2}\mathbf{B}^T c_0 \mathbf{B}\right]\mathrm{d}\tau, \\ \mathbf{B}(\mathbf{u},\tau) &= \int_{\mathbb{R}^+}\left[-r_1 + a_1^T\mathbf{B} + \frac{1}{2}\mathbf{B}^T c_1 \mathbf{B}\right]\mathrm{d}\tau. \end{aligned}$$

- A possible solution that will offer a closed-form solution for ODEs is based on the piece-wise constant parameters.

Question 15/30 (Lecture 6):
Why is adding more and more factors to the pricing models not the best idea?

- When choosing a model, a few critical points need to be considered:
  - the over-fitting problem;
  - can the corresponding ChF be determined analytically?
  - is the model affine?
  - how to hedge the parameters?
  - can a model be simulated efficiently?
- Always focus on the pricing objective!

Question 16/30 (Lecture 7):

Can you interpret the Heston model parameters and their impact on the volatility surface?

- For $\rho_{x,v} = 0\%$ a higher value of γ gives a more pronounced implied volatility *smile*. A higher volatility-of-volatility parameter increases the implied volatility curvature.
- As the correlation between stock and variance process, $\rho_{x,v}$, gets increasingly negative the slope of the skew in the implied volatility curve increases.

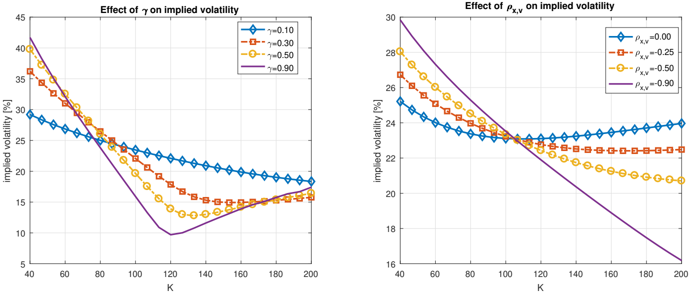


Figure: Impact of variation of the Heston vol-vol parameter γ (left side), and correlation parameter $\rho_{x,v}$ (right side), on the implied volatility as a

Question 16/30 (Lecture 7):

Can you interpret the Heston model parameters and their impact on the volatility surface?

- Speed of mean reversion κ has a limited effect on the implied volatility smile or skew, up to $1\% - 2\%$. κ determines the speed at which the volatility converges to the long-term volatility $\bar{v}$, see the RHS graph, which shows the at-the-money (ATM) implied volatility for different κ.
- With $\bar{v} = 10\%$ ($\sqrt{\bar{v}} \approx 31.62\%$) a large κ-value implies fast convergence of the implied volatility to $\sqrt{\bar{v}}$.

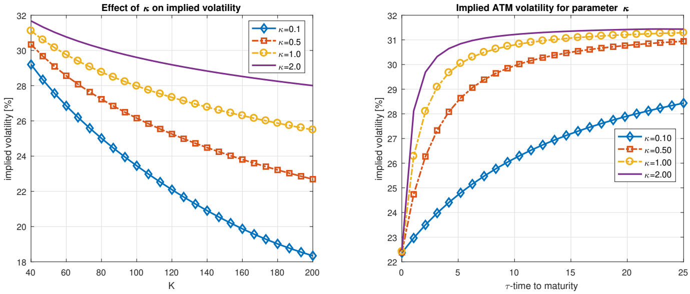


Figure: Impact of variation of the Heston parameter κ on the implied

Question 16/30 (Lecture 7):
Can you interpret the Heston model parameters and their impact on the volatility surface?

- v_0, the initial variance and $\bar{v}$, the variance level, have a similar effect on the implied volatility curve.
- The effect of these two parameters seems to depend on the value of κ, controlling the speed at which the implied volatility converges from $\sqrt{v_0}$ to $\sqrt{\bar{v}}$ (or v_0 to $\bar{v}$).

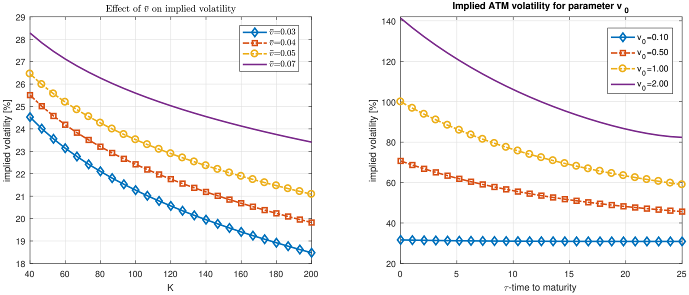


Figure: Impact of changing v_0 and $\bar{v}$ on the Heston implied volatility; left side: $\bar{v}$ as a function of the strike K, right side: v_0 as a function of time

Question 16/30 (Lecture 7):

Can you interpret the Heston model parameters and their impact on the volatility surface?

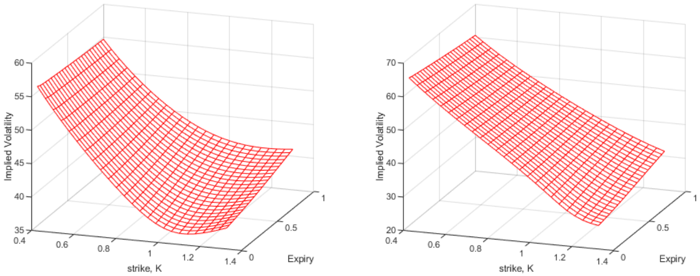


Figure: Implied volatility surface for the Heston model. Left: $\rho = -0.2$. Right: $\rho = -0.8$.

Question 17/30 (Lecture 7): **Can we model volatility with the Arithmetic Brownian Motion process?**

Can we consider a ABM or the OU process to drive volatility, just like under the Heston model?

- The Heston model consists of two stochastic differential equations:

$$\begin{aligned} \mathrm{d}S(t) &= rS(t)\mathrm{d}t + \sqrt{v(t)}S(t)\mathrm{d}W_x(t), \\ \mathrm{d}v(t) &= \kappa(\bar{v} - v(t))\mathrm{d}t + \gamma\sqrt{v(t)}\mathrm{d}W_v(t). \end{aligned}$$

- Volatility, the CIR, the process is non-negative!
- The Schöbel-Zhu model is defined as:

$$\begin{aligned} \mathrm{d}S(t) &= rS(t)\mathrm{d}t + \sigma(t)S(t)\mathrm{d}W_x(t), \\ \mathrm{d}\sigma(t) &= \kappa(\bar{\sigma} - \sigma(t))\mathrm{d}t + \gamma\mathrm{d}W_\sigma(t). \end{aligned}$$

- It is important to notice that under the Heston model we deal with $\sqrt{v(t)}S(t)\mathrm{d}W_x(t)$ where $\mathrm{d}W_x(t)$ can take either positive or negative values. Therefore, it should not matter if we consider a volatility process that takes negative values.
- The bottom line question regarding such a model is: Can the model represent realistic shapes of implied volatilities?

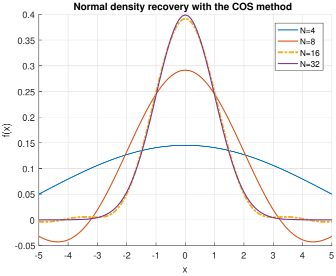
Question 18/30 (Lecture 8):
What are the benefits of FFT compared to a “brute force” integration?

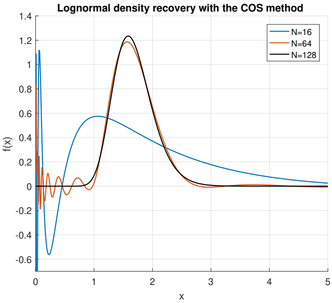
- Let us consider only models for which we don't have an explicit solution (the Black-Scholes model), and only the corresponding ChF is known.
- Assuming a constant interest rate, the following pricing equation for a call option holds:

$$C(t_0, S_0, T, K) = \mathrm{e}^{-rT}\mathbb{E}\left[\max(S(T)-K,0)\right] = \mathrm{e}^{-rT}\int_0^\infty \max(x-K,0) f_{S(T)}(x)\mathrm{d}x,$$

- Once the PDF $f_{S(T)}(x)$ is given, then for pricing of a single option, we only need a single integration.
- The problem is twofold:
  - Computation of the density $f_{S(T)}(x)$ is often unknown and requires multiple integrations over the Fourier space;
  - The calibration routine will typically require evaluations of option prices for multiple strikes, $K_1, K_2, \ldots, K_N$.

Question 19/30 (Lecture 8):
What to do if the FFT/COS method does not converge for increasing expansion terms?



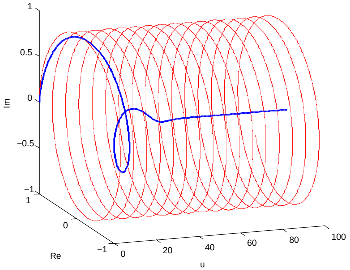


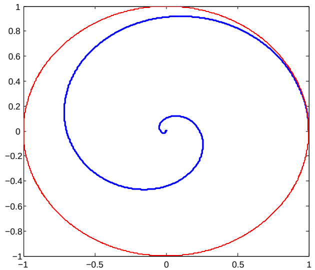
What if convergence does not occur for an increasing number of expansion terms?

Question 19/30 (Lecture 8):
What to do if the FFT/COS method does not converge for increasing expansion terms?

- One possibility is to look at the corresponding ChF and see whether the COS/FFT methods well cover the integration domain:
- The characteristic function for the Black-Scholes asset price is given by:

$$\phi(u,\tau) = \exp\left(i(\log(S_0) + (r - \frac{1}{2}\sigma^2)\tau)u - \frac{1}{2}\sigma^2\tau u^2\right).$$





Question 19/30 (Lecture 8): **What to do if the FFT/COS method does not converge for increasing expansion terms?**

Possible ways to improve Fourier-based methods:

- Compute cumulants for the COS method.
- The COS method: vary L, a and b parameter (sometimes L needs to be very small).
- The COS method: use optimal points a and b based on "*Precise option pricing by the COS method–How to choose the truncation range*" *by Gero Junike, Konstantin Pankrashkin.*
- *The FFT method (the Carr-Madan method): change the upper limit of the Fourier domain.*
- *The FFT method: Modify the so-called dumping factor.*

If these modifications don't help, perform sanity checks like the pricing option for $K->0$ and re-visit derivations for ChF (the moments).

Question 20/30 (Lecture 9):
What is a standard error? How to interpret it?

1. Partition the time interval $[0, T]$, $0 = t_0 < t_1 < \cdots < t_m = T$.
2. Generate asset values, $s_{i,j}$, taking the risk-neutral dynamics of the underlying model. $s_{i,j}$ has two indices, the time points and the Monte Carlo path.
3. Compute the N payoff values, H_j. In the case of European options, $H_j = H(T, s_{m,j})$, in the case of path-dependent options, $H_j = H(T, s_{i,j}),\ i = 1, \dots m$.
4. Compute the average, $\mathbb{E}^{\mathbb{Q}}\left[H(T,S)\middle|\mathcal{F}(t_0)\right] \approx \frac{1}{N}\sum_{j=1}^{N} H_j =: \bar{H}_N$.
5. Calculate the option value as $V(t_0, S) \approx \mathrm{e}^{-r(T-t_0)}\frac{1}{N}\sum_{j=1}^{N} H_j$.
6. Determine the standard error related to the obtained prices in Step 5.

Question 20/30 (Lecture 9):
What is a standard error? How to interpret it?

- By the strong law of large numbers, we know that, for $N \to \infty$,

$$\lim_{N\to\infty} \bar{H}_N(T,S) = \mathbb{E}^{\mathbb{Q}}[H(T,S)], \quad \text{with probability} \quad 1.$$

- To estimate the error due to a finite number of paths, we compute,

$$\begin{aligned}\mathbb{V}\mathrm{ar}^{\mathbb{Q}}\left[\bar{H}_N(T,S)\right] &= \mathbb{V}\mathrm{ar}^{\mathbb{Q}}\left[\frac{1}{N}\sum_{j=1}^{N} H(T,s_{m,j})\right]\\ &= \frac{1}{N^2}\sum_{j=1}^{N}\mathbb{V}\mathrm{ar}^{\mathbb{Q}}[H(T,s_{m,j})],\end{aligned}$$

  given that samples $s_{m,j}$ are drawn independently.
- The unknown variance is approximated by the sample variance,

$$\bar{\upsilon}_N^2 := \frac{1}{N-1}\sum_{j=1}^{N}\left(H(T,s_{m,j}) - \bar{H}_N(T,S)\right)^2.$$

- Standard error ϵ_N is defined as $\epsilon_N := \frac{\bar{\upsilon}_N}{\sqrt{N}}$. When the number of samples increases by a factor 4, the error reduces by a factor 2.

Question 21/30 (Lecture 9):
What is weak and strong convergence in Monte Carlo pricing?

- Denote by x_m the approximation for $X(T)$, where Δt is the time step size, and m corresponds to the last term in the time discretization, $t_i = i \cdot T/m,\ i = 0, \ldots, m$.
- Then, the approximation x_m converges in a strong sense to $X(T)$, with order $\alpha > 0$, if
$$\epsilon^s(\Delta t) := \mathbb{E}^{\mathbb{Q}}\Big[|x_m - X(T)|\Big] = O(\Delta t^{\alpha}).$$
For a sufficiently smooth function $g(\cdot)$, the approximation x_m converges in a weak sense to $X(T)$, with respect to $g(\cdot)$, with order $\beta > 0$, if
$$\epsilon^w(\Delta t) := \Big|\mathbb{E}^{\mathbb{Q}}[g(x_m)] - \mathbb{E}^{\mathbb{Q}}[g(X(T))]\Big| = O(\Delta t^{\beta}).$$
- In other words, a numerical integration method converges in a strong sense, if the asset prices converge, and weak convergence implies a convergent approximation of the probability distribution of $X(T)$, for a given time T. The convergence then concerns only the marginal distribution of $X(T)$.

Question 21/30 (Lecture 9):
What is weak and strong convergence in Monte Carlo pricing?

- For different Δt-values, and also the weak convergence error:

$$
\begin{aligned}
\epsilon^{w}(\Delta t) &= \left| \frac{1}{N}\sum_{j=1}^{N} S_j(T) - \frac{1}{N}\sum_{j=1}^{N} s_{m,j} \right| \\
&= \left| S(t_0)\frac{1}{N}\sum_{j=1}^{N} \mathrm{e}^{\left(r-\frac{1}{2}\sigma^2\right)T+\sigma W_{m,j}} - \frac{1}{N}\sum_{j=1}^{N} s_{m,j} \right|,
\end{aligned}
$$

  where m stands for time $t_m \equiv T$ and the index j indicates the path number at which solutions and are evaluated.

- Note that in this experiment it is crucial to use the same Brownian motion for both equations. If the Brownian motions wouldn't be the same, we wouldn't be able to measure the strong convergence, as the random paths would be different.

Question 21/30 (Lecture 9):
What is weak and strong convergence in Monte Carlo pricing?

- We postulate that

$$\epsilon^s(\Delta t) \leq C \cdot (\Delta t)^{\frac{1}{2}} = \mathcal{O}((\Delta t)^{\frac{1}{2}}).$$

- For different mesh widths Δt, the results are presented in Figure below.

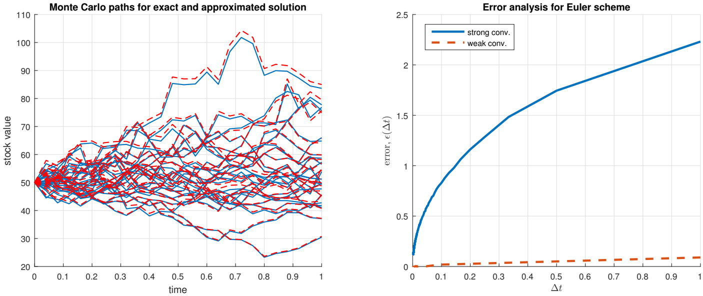


Figure: Left: generated paths with exact simulation versus Euler approximation; Right: error against value of time step Δt for the Euler discretization.

Question 22/30 (Lecture 10): **What are the challenges of discretizing the CIR process using the Euler method?**

- A typical example of a process with probability mass around zero is the CIR process, which was discussed to model the variance for the Heston stochastic volatility model with the following dynamics,

$$\boxed{\mathrm{d}v(t) = \kappa(\bar{v} - v(t))\mathrm{d}t + \gamma\sqrt{v(t)}\mathrm{d}W(t), \quad v(t_0) > 0.}$$

- It is well-known that if the Feller condition, $2\kappa\bar{v} > \gamma^2$, is satisfied, the process $v(t)$ cannot reach zero, and if this condition does not hold, the origin is accessible and strongly reflecting.
- In both cases, the $v(t)$ process cannot become negative.

Question 22/30 (Lecture 10):
What are the challenges of discretizing the CIR process using the Euler method?

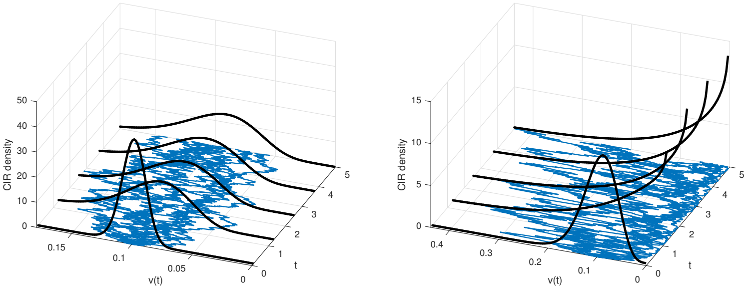


Figure: Paths and the corresponding PDF for the CIR process in the cases where the Feller condition is satisfied and is not satisfied. Simulations were performed with $\kappa = 0.5$, $v_0 = 0.1$, $\bar{v} = 0.1$. Left: $\gamma = 0.1$; Right: $\gamma = 0.35$.

Question 22/30 (Lecture 10): What are the challenges of discretizing the CIR process using the Euler method?

- The nonnegativity problem becomes apparent when a standard discretization is employed. If we apply, for example, the Euler discretization to the process, i.e.,

$$v_{i+1} = v_i + \kappa(\bar{v} - v_i)\Delta t + \gamma\sqrt{v_i}\sqrt{\Delta t}Z.$$

and assume that $v_i > 0$, we may calculate the probability that a next realization, v_{i+1}, becomes negative, i.e. $\mathbb{P}[v_{i+1} < 0]$.

$$\begin{aligned}\mathbb{P}[v_{i+1} < 0 | v_i > 0] &= \mathbb{P}[v_i + \kappa(\bar{v} - v_i)\Delta t + \gamma\sqrt{v_i \Delta t}Z < 0 | v_i > 0] \\ &= \mathbb{P}[\gamma\sqrt{v_i \Delta t}Z < -v_i - \kappa(\bar{v} - v_i)\Delta t | v_i > 0],\end{aligned}$$

which equals,

$$\mathbb{P}[v_{i+1} < 0 | v_i > 0] = \mathbb{P}\left[Z < -\frac{v_i + \kappa(\bar{v} - v_i)\Delta t}{\gamma\sqrt{v_i \Delta t}} \middle| v_i > 0\right] > 0.$$

- Since Z is a normally distributed random variable, it is unbounded. Therefore the probability of v_i becoming negative, is positive under the Euler discretization, implying $\mathbb{P}[v_{i+1} < 0 | v_i > 0] > 0$.

Question 23/30 (Lecture 10): **Why do we need Monte Carlo if we have FFT methods for pricing?**

- Monte Carlo methods are mainly used for pricing, especially of exotic (callable derivatives).
- Fourier-based methods offer a great speed for pricing European options and facilitating fast calibration of the model.
- The pricing speed is crucial in the calibration phase, where many iterations occur.
- Fourier methods also allow for pricing of exotic derivatives (e.g., Bermudas), but their implementation is not generic- therefore, each payoff needs to be re-derived, while Monte Carlo offers flexibility.

Question 24/30 (Lecture 11):
How to hedge jumps?

- In the three graphs $\Delta(t_i)$ (green line) behaves like the stock process $S(t)$, however when the stock $S(t)$ give a call price (pink line) deep in or out of the money, $\Delta(t_i)$ is either 0 or 1.

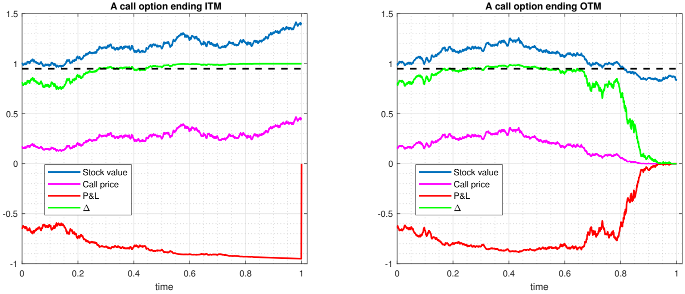


Figure: Delta hedging a call option. Blue: the stock path, pink: the value of a call option, red: P&L(t) portfolio, and green: Δ.

Question 24/30 (Lecture 11):
How to hedge jumps?

- The impact of the frequency of updating the hedge portfolio on the distribution of the P&L(T) is presented. Two simulations have been performed, one with 10 re-balancing actions during the option's lifetime and one with 2000 actions. It is clear that frequent re-balancing brings the variance of the portfolio P&L(T) down to almost 0.

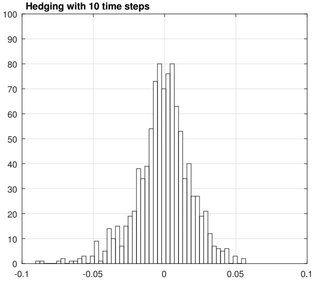


Figure: The impact of the re-balancing frequency on the variance of the P&L(T) portfolio. Left: 10 re-balancing times, Right: 2000 re-balancing times.

Question 24/30 (Lecture 11):
How to hedge jumps?

- Instead of the Black-Scholes model, we consider a jump-diffusion model:

$$\mathrm{d}S(t)/S(t) = \left(r - \xi_p \mathbb{E}[\mathrm{e}^J - 1]\right)\mathrm{d}t + \sigma \mathrm{d}W^{\mathbb{Q}}(t) + (\mathrm{e}^J - 1)\mathrm{d}X_{\mathcal{P}}^{\mathbb{Q}}(t).$$

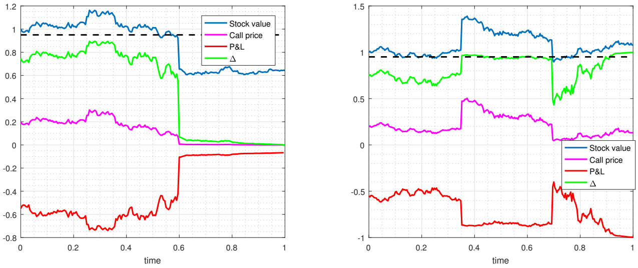


Figure: Delta hedging a call option for a stock with jumps, where we however work with the Black-Scholes delta. Blue: the stock path, pink: value of the call option, red: P&L(t) portfolio, and green: the Δ. Left: a path with one jump time, right: two occurring jumps.

Question 24/30 (Lecture 11):
How to hedge jumps?

- Let us not observe the impact of increased hedging frequency.

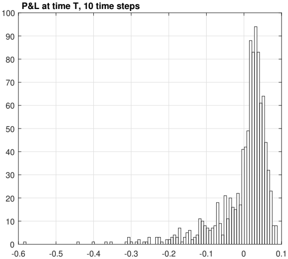


Figure: The impact of the hedging frequency on the variance of the P&L(T) portfolio, with the stock following a jump diffusion process. Left: 10 times hedging during the option lifetime, Right: 2000 times hedging.

Question 24/30 (Lecture 11): **How to hedge jumps?**

- Let us now look at the hedging argument used in the Heston model.
- Suppose we wish to hedge a call option $C(T, K_1)$, then the hedging portfolio will be given by:

$$\Pi = C(T, K_1) + \Delta S(T) + \Delta_2 C(T, K_2).$$

- Thus to hedge the volatility under the stochastic volatility, we need to buy a "volatility sensitive" derivative.
- The same strategy will hold when hedging jumps- we need to build a portfolio of derivatives that also are driven by the jump process $S(T)$.

Question 25/30 (Lecture 11): **What is pathwise sensitivity?**

- To efficiently estimate the sensitivities at time t_0, like with respect to $S(t_0)$ and also to other model parameters.
- The pathwise sensitivity method is applicable with continuous functions of the parameter of interest, and based on interchanging the differentiation and expectation operators,

$$\frac{\partial V}{\partial \theta} = \frac{\partial}{\partial \theta}\mathbb{E}^{\mathbb{Q}}\left[\frac{V(T,S;\theta)}{M(T)}\Big|\mathcal{F}(t_0)\right] = \mathbb{E}^{\mathbb{Q}}\left[\frac{\partial}{\partial \theta}\frac{V(T,S;\theta)}{M(T)}\Big|\mathcal{F}(t_0)\right].$$

- Assuming a constant interest rate, we find,

$$\frac{\partial V}{\partial \theta} = \mathrm{e}^{-r(T-t_0)}\mathbb{E}^{\mathbb{Q}}\left[\frac{\partial V(T,S;\theta)}{\partial S}\frac{\partial S}{\partial \theta}\Big|\mathcal{F}(t_0)\right].$$

Question 25/30 (Lecture 11):
What is pathwise sensitivity?

- As an example, we apply the pathwise sensitivity methodology to a call option under the Black-Scholes model, i.e.,

$$V(T,S;\theta)=\max(S(T)-K,0),\text{with},S(T)=S(t_0)\mathrm{e}^{(r-\frac{1}{2}\sigma^2)(T-t_0)+\sigma(W(T)-W(t_0))}.$$

- The derivative of the payoff with respect to $S(T)$ is given by,

$$\frac{\partial V}{\partial S(T)}=1_{S(T)>K},$$

and the necessary derivatives with respect to $S(t_0)$ and σ are as follows,

$$\begin{aligned}\frac{\partial S(T)}{\partial S(t_0)} &= \mathrm{e}^{(r-\frac{1}{2}\sigma^2)(T-t_0)+\sigma(W(T)-W(t_0))},\\ \frac{\partial S(T)}{\partial \sigma} &= S(T)\left(-\sigma(T-t_0)+W(T)-W(t_0)\right).\end{aligned}$$

Question 25/30 (Lecture 11):
What is pathwise sensitivity?

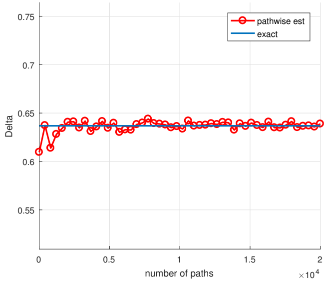
- So, for the estimates of delta and vega we obtain,

$$
\begin{aligned}
\frac{\partial V}{\partial S(t_0)} &= \mathrm{e}^{-r(T-t_0)}\mathbb{E}^{\mathbb{Q}}\left[1_{S(T)>K}\mathrm{e}^{(r-\frac{1}{2}\sigma^2)(T-t_0)+\sigma(W(T)-W(t_0))}\right]\\
&= \mathrm{e}^{-r(T-t_0)}\mathbb{E}^{\mathbb{Q}}\left[\frac{S(T)}{S(t_0)}1_{S(T)>K}\right],
\end{aligned}
$$

and

$$
\begin{aligned}
\frac{\partial V}{\partial \sigma} &= \mathrm{e}^{-r(T-t_0)}\mathbb{E}^{\mathbb{Q}}\left[1_{S(T)>K}S(T)\left(-\sigma(T-t_0)+W(T)-W(t_0)\right)\right]\\
&= \frac{\mathrm{e}^{-r(T-t_0)}}{\sigma}\mathbb{E}^{\mathbb{Q}}\left[S(T)\left(\log\left(\frac{S(T)}{S(t_0)}\right)-\left(r+\frac{1}{2}\sigma^2\right)(T-t_0)\right)1_{S(T)>K}\right].
\end{aligned}
$$

Question 25/30 (Lecture 11): **What is pathwise sensitivity?**



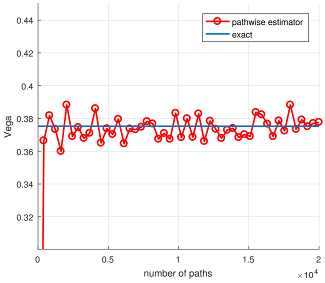


Figure: The Black-Scholes delta (left) and vega (right) estimated by the pathwise sensitivity method. The parameters are $S(t_0) = 1$, $r = 0.06$, $\sigma = 0.3$, $T = 1$, $K = S(t_0)$.

More examples, with the application to the Heston model can be found in the lecture notes.

Question 26/30 (Lecture 12):
What is the Bates model, and how can it be used for pricing?

- Empirical studies have shown that the Heston model is not able to calibrate well to short-term implied volatilities.
- The Bates model also generalizes the Heston model by adding jumps to the Heston stock price process. The model is described by the following system of SDEs:

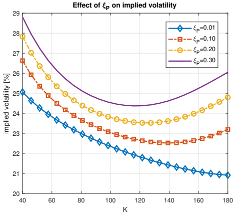
$$\begin{cases} \frac{\mathrm{d}S(t)}{S(t)} = \left(r - \xi_p \mathbb{E}[\mathrm{e}^J - 1]\right)\mathrm{d}t + \sqrt{v(t)}\mathrm{d}W_x(t) + \boxed{\left(\mathrm{e}^J - 1\right)\mathrm{d}X_{\mathcal{P}}(t),} \\ \mathrm{d}v(t) = \kappa\left(\bar{v} - v(t)\right)\mathrm{d}t + \gamma\sqrt{v(t)}\mathrm{d}W_v(t), \end{cases}$$

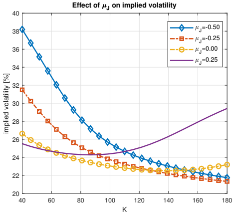
with Poisson process $X_{\mathcal{P}}(t)$ with intensity ξ_p, and normally distributed jump sizes, J, with expectation μ_J and variance σ_J^2, i.e. $J \sim \mathcal{N}(\mu_j, \sigma_j^2)$. $X_{\mathcal{P}}(t)$ is assumed to be independent of the Brownian motions and of the jump sizes.

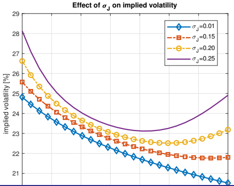
Question 26/30 (Lecture 12):

What is the Bates model, and how can it be used for pricing?

- In the experiments we use $T = 1$, $S_0 = 100$; the Heston model parameters have been set to $\kappa = 1.2$, $\bar{v} = 0.05$, $\gamma = 0.05$, $\rho_{x,v} = -0.75$, $v_0 = 0.05$ and $r = 0$. The three jump parameters, i.e. the intensity ξ_p, the jump mean μ_J and jump volatility σ_J, have been varied.







Question 26/30 (Lecture 12):
What is the Bates model, and how can it be used for pricing?

- Figures shows quite similar patterns for the implied volatilities regarding parameter changes in ξ_p and σ_J. An increase of the parameter values increases the curvature of the implied volatilities. The plots also indicate that the implied volatilities from the Bates model can be higher than those obtained by the Heston model.
- The explanation lies in the fact that the additional, uncorrelated jump component is present in the dynamics of the Bates model, which may increase the *realized* stock variance. An increase of volatility of the stock increases the option price as the option is more likely to end in the money.
- In the case of parameter μ_J, which represents the *expected* location of the jumps, the implied volatilities reveal an irregular behavior. Parameter μ_J can take either positive or negative values, as it determines whether the stock paths will have upwards or downwards jumps.

Question 27/30 (Lecture 12): **What is the relation between European and Forward-start options?**

- Forward-starting options, also called the performance options. The value of an option does not depend on the value a stock but on stocks performance.
- Forward starting options may be considered as European options with a future starting time.
- Whereas in European options the initial stock value, S_0, is known at initial time t_0, in the case of forward-starting options the initial stock value is unknown as it will be determined at some future time T_1.
- Forward-starting options do not depend on today's value of the underlying asset but on a value which is based on the performance of the asset over some future time period $[T_1, T_2]$.

Question 27/30 (Lecture 12): **What is the relation between European and Forward-start options?**

- Forward-starting options are often considered as building blocks for other financial derivatives, like cliquet (also called ratchet) options, consisting of a series of consecutive forward-start options. Cliquet options are encountered in the equity and interest rate markets.
- A useful feature of these Cliquet option contracts is that they allow for a periodic locking in of profits. If a certain asset performed well a period $[T_i, T_{i+1}]$, but performed poorly in $[T_{i+1}, T_{i+2}]$, an investor will receive the profit accumulated in the first period and does not loose money in the second period. As the profits can be limited as well, i.e. in the contract it can be specified that the profit cannot be larger than, say 20%, cliquet options may be a relatively cheap alternative for plain-vanilla European options.

Question 27/30 (Lecture 12): **What is the relation between European and Forward-start options?**

- We start with some general results regarding the pricing of forward-start options that do not depend on any particular model. In the subsections to follow different alternatives for the underlying asset dynamics will be discussed. For two maturities T_1 and T_2, so that $t_0 < T_1 < T_2$, we define a forward-starting option, as follows:

$$V(T_2; T_1, T_2) = \max\left(\frac{S(T_2) - S(T_1)}{S(T_1)} - K, 0\right),$$

  with a constant strike K.

- The contract value is determined by the percentage performance of underlying asset $S(t)$, measured at the future times T_1 and T_2. Obviously, for $t_0 = T_1$ the payoff collapses to a standard "scaled" European option with maturity at time T_2, and payoff equal to:

$$V(T_2; T_1, T_2) = \frac{1}{S_0}\max\left(S(T_2) - K^*, 0\right), \quad K^* = S_0(K+1).$$

Question 27/30 (Lecture 12):
What is the relation between European and Forward-start options?

- The payoff can also be expressed as:

$$V(T_2, S; T_1, T_2) = \max\left(\frac{S(T_2)}{S(T_1)} - (K+1), 0\right).$$

As the contract will pay at time T_2, today's value is given by the following pricing equation:

$$V(t_0, S; T_1, T_2) = M(t_0)\mathbb{E}^{\mathbb{Q}}\left[\frac{1}{M(T_2)}\max\left(\frac{S(T_2)}{S(T_1)} - k, 0\right)\Big|\mathcal{F}(t_0)\right],$$

with $k = K + 1$, and $M(t)$ the money-savings account.

Question 27/30 (Lecture 12):
What is the relation between European and Forward-start options?

- Under constant interest rates the pricing equation reads

$$\begin{aligned} V(t_0, S; T_1, T_2) &= \frac{M(t_0)}{M(T_2)}\mathbb{E}^{\mathbb{Q}}\left[\max\left(\frac{S(T_2)}{S(T_1)} - k, 0\right)\Big|\mathcal{F}(t_0)\right] \\ &= \frac{M(t_0)}{M(T_2)}\mathbb{E}^{\mathbb{Q}}\left[\max\left(\mathrm{e}^{x(T_1,T_2)} - k, 0\right)\Big|\mathcal{F}(t_0)\right] \end{aligned}$$

where

$$x(T_1, T_2) := \log X(T_1, T_2) = \log \frac{S(T_2)}{S(T_1)} = \log S(T_2) - \log S(T_1).$$

- Note that we haven't made any assumptions regarding dynamics of the stock process.

Question 28/30 (Lecture 12): **What instruments to choose to calibrate your pricing model?**

- Pricing approach:
  1. Start with some financial product
  2. Model asset prices involved (SDEs)
  3. Calibrate the model to market data (numerics, optimization)
  4. Model product price correspondingly (P(I)DE or integral)
  5. Price the product of interest (numerics, MC)
  6. Set up a hedge to remove the risk to the product (optimization)

Question 28/30 (Lecture 12):
What instruments to choose to calibrate your pricing model?

- European option- an option that may be only exercised on expiration day;
- American option - an option that may be exercised on any trading day (also on the expiration);
- Bermudan option - an option that may be exercised only on specified dates;
- Exotic or path-dependent options (value does not only depend on the current stock value but, for example, also on the average value of the stock price development).

Question 29/30 (Lecture 13):
How to calibrate a pricing model? How to choose the objective function?

- Typically, the pricing model is calibrated to European options. Many calibration variants are possible.
- The crucial point of model calibration is that the model should be calibrated to hedging instruments, e.g., a model that prices exotic derivatives with 10y expiry should not be calibrated to 20y options.
- Possible choices regarding the objective function are:
  - Standard, weighted target function:

$$\epsilon = \sum_{i=1}^{N_T}\sum_{j=1}^{N_K}\omega_{i,j}\left(C(T_i, K_j) - C(\theta, T_i, K_j)\right)^2.$$

  - Implied volatility-based target function

$$\epsilon = \sum_{i=1}^{N_T}\sum_{j=1}^{N_K}\omega_{i,j}\left(\sigma(T_i, K_j) - \sigma(\theta, T_i, K_j)\right)^2.$$

  - Calibration based only on OTM options.

Question 30/30 (Lecture 13):
What are the Chooser options?

- Binary: Cash or Nothing: Pays out Q at expiry T if option is in the money $S > K$, otherwise expires worthless. Payoff:

$$V(S,T) = Q1_{S \geq K}$$

- Compound option: Call on a call: the right to buy a 'call with maturity T and strike K' at time T_0 for the price K_0. Payoff:

$$V^{CC}(S,T_0,K_0,K,T) = \max\left[V^C(S,K,T) - K_0, 0\right]$$

- Chooser option: Gives the holder the right to choose whether the underlying option at time T_0 is a Call or a Put with the same strike K and maturity T. The payoff of a chooser option is

$$V^{CH}(S,K,T_0,T) = \max\left[V^C(S,K,T), V^P(S,K,T)\right]$$