INSTITUTO SUPERIOR DAS CIÊNCIAS, DO TRABALHO E DA EMPRESA
DEPARTAMENTO DE FINANÇAS

ISCTE Business School
Instituto Universitário de Lisboa

UNIVERSIDADE DE LISBOA
FACULDADE DE CIÊNCIAS
DEPARTAMENTO DE MATEMÁTICA

Ciências ULisboa

Option Pricing under jump-diffusion models

Ricardo Filipe Machado Venâncio

Mestrado em Matemática Financeira

Dissertação orientada por:
João Pedro Vidal Nunes

2016

Acknowledgments

This thesis is the final result after enrolling in the Master of Mathematics applied to Finance. It lead me to achieve great knowledge about Finance, particularly on how to approach the question of computing a fair value for financial contracts.

In first place, I thank to all the students and professors I had contact with in FCT-UNL, where I got my bachelor's degree in Mathematics. Long were the hours I spent in teachers' offices clarifying my doubts which were promptly and enthusiastically answered. Their patience and dedication on my and my colleagues formation was crucial to become the person I am today.

In second place, I thank my supervisor, João Nunes, who patiently taught me the main topics related to the Pricing theory and who helped me with many questions I had, without any hesitation. He gave me important tools so that I could design the work I am presenting here and without which I could not develop this dissertation.

I also appreciate the help of all the people who I worked with in BNP Paribas, who encouraged me during my learning process while I was working as an intern. They proposed me challenges and posed questions who helped me to better understand how Financial markets work and contributed to my professional and personal growth. Thanks to this working experience, I was able to establish a link between what is developed in theory and relate it to the practice.

Finally, I thank my family for all the support, specially in the moments I was absent living and working abroad.

I hope the work I developed is useful for who studies matters related to Option Pricing and the reader find it interesting to explore.

Resumo

Nesta tese, apresentam-se métodos para resolver numericamente equações diferenciais por forma a obter preços de contractos financeiros. Em particular, dá-se ênfase a opções vanilla de estilo europeu e americano cujo activo subjacente segue um modelo de difusão com saltos. Quanto à distribuição destes útltimos, destacam-se o modelo de Merton, que considera que eles têm uma distribuição Normal, e o de Kou, onde é assumida uma dupla distribuição exponencial. Este tipo de modelos representa uma extensão dos clássicos modelos de difusão, como o famoso modelo de Black-Scholes-Merton, e tem como objectivo superar algumas das falhas inerentes a este último, tal como caudas muito curtas e picos baixos da distribuição do logaritmo dos retornos do activo, que não reflectem, em geral, o sentimento dos investores nos mercados financeiros, aliando, ao mesmo tempo, a simplicidade e eficiência dos modelos de difusão.

Para alcançar o nosso objectivo, estabelece-se inicialmente qual é a equação que descreve a dinâmica do valor dos preços das opções referidas em relação a vários parâmetros, tal como o valor do preço do activo subjacente e o tempo até à maturidade. Em seguida, constróem-se partições para a resolução numérica do problema, através da discretização da função que descreve o preço do contracto financeiro por diferenças finitas. Esta abordagem é útil visto que permite obter preços de contractos cujo "payoff" não é tão simples quanto o de opções vanilla e para os quais não existem fórmulas fechadas ou semi-fechadas para o seu valor em cada momento do tempo até à sua maturidade.

No final, expõem-se os resultados encontrados para diferentes resoluções das partições, comparados com referências da literatura, e apresentam-se algumas conclusões.

Abstract

In this dissertation, methods to solve numerically partial differential equations in order to obtain prices for contingent claims are presented. In particular, we highlight European and American style vanilla options, whose underlying asset follows a jump-diffusion model. For the distribution of the jumps, the Merton and Kou models are studied. The former considers these have a Normal distribution, whereas the latter assumes a double-exponential. These type of models represent an extension of the classic diffusion models, such as the famous Black-Scholes-Merton, and has the goal of overcoming its flaws, such as thin tails and low peaks in the distribution of the logarithm of the asset returns, that do not reflect the general investors sentiment in the financial markets, while maintaining the simplicity and tractability inherent to diffusion models.

To accomplish our goal, an equation describing the relation of the value of the referred options on several parameters, such as the time-to-maturity and the spot value of the underlying asset is suggested. We then build partitions in order to numerically solve our problem using finite differences, discretizing the function which provides the price of our contingent claim. This approach is useful, since it allows to obtain prices of contracts whose payoff is not as simple as the vanilla options' and for which it does not exist closed or semi-closed formulae for its value at each point in time until maturity.

Finally, we expose results found for each one of partitions considered, comparing them with values in the literature, and some conclusions are presented.

Keywords

Option Pricing

Lévy processes

Diffusion models

Jump-diffusion models

European Options

American Options

LU decomposition

Tridiagonal matrices

Numerical discretization

Finite differences

Contents

1 Introduction

Although options were created a long time ago, it was only on the 26th April 1973 that the Chicago Board Options Exchange (CBOE) traded the first listed option. Since then, interest in options has excessively grown from only 911 contracts on 16 underlying stocks to millions of contracts with total notional value of several trillion dollars in CBOE and in other stock exchanges around the world.

An option, as the name itself indicates, gives to its holder the right to buy or sell an asset, but not the obligation to do so, unlike other contracts such as futures or forwards. To compensate for this benefit, the holder of the option must pay an up-front fee, called the "premium". The most basic existing options are vanilla calls and puts. Call options give the entity that hold them the right to buy a determined amount of an asset at or until a specified time in the future, defined as the "expiration date" or "maturity date", for a pre-specified price, which we name as the "strike price". Puts, on the other hand, permit the entity to sell the asset. The asset in question, in this instance, can be a simple stock, index, bond or a basket composed of several securities.

Based on the time at which options can be exercised, they can be divided into two categories: European or American options. The former gives the right to its owner to exercise the option at the maturity date T, whereas the latter can be exercised at any time until T.

Let us assume the underlying asset is a stock with price S, using S_t as notation for its value at time t, prior to T. In the case of a call option, at maturity, it will only be exercised if $S_T > K$, with K being the strike price, whereas otherwise it will expire

valuing zero. In the case of a put option, the logic is similar, but this contract will only have non-null value if $S_T < K$.

Like any other contingent claim, options are attractive, because they are highly useful in hedging and for speculation. Hedging refers to the act of using financial instruments to cover the position a market agent has in its portfolio in order to avoid big losses, in other words, reducing risk. As a simple example, a put option may be a good instrument to acquire if one is afraid that the price of a certain stock will fall. This way, a minimum selling price is guaranteed (the strike) and the risk of a loss is minimized.

These kinds of financial products obviously require a price to be quoted in the market. And since nobody has a crystal ball, that reveals what happens to an asset's future price, mathematical models are needed to compute their fair-value. To this end, the work carried out by Robert Merton, Fisher Black and Myron Scholes is of greater importance. They developed a set of formulae based on the Brownian motion which allow us to price many financial derivatives such as calls, puts and barrier options [10]. The Black-Scholes-Merton (BSM) model tries to replicate the dynamics of a financial asset through a geometric Brownian motion, i.e., for a given moment in time t, the expression that gives an asset's price is $S_t = S_0 e^{\mu t + \sigma W_t}$ and the continuous compounded return, $\ln(S_t/S_0)$, has a normal distribution. Here, $\{W_t, t \geq 0\}$ is a Brownian motion, which has a normal distribution with mean 0 and variance t, μ is called the drift, which measures the annualized average return, σ is called the volatility, which corresponds to the annualized standard deviation of the underlying asset price return, S_0 is the initial stock price and "ln" represents the Neperian logarithm.

However, take a look at Figure A.1, in the Appendix. In this figure, you can visualize Standard & Poor's 500 (S&P500 for short) prices through the year of 2015. This is a very liquid American stock market index based on the market capitalization of 500 large companies that have their common stock listed on the NYSE (New York Stock Exchange) or NASDAQ (National Association of Securities Dealer Automated Quotation system). As you can see by the first chart, there seems to be a large fall in the

S&P index price towards the end of August 2015. This is highlighted by the computation of log returns (defined as the continuously compounded returns) and represented in Figure A.2. A.2 also reveals a peak at the end of August. This was due to a fall in China's stock market, which at that time affected the U.S.' markets. On the other hand, Figure A.5, which represents a histogram of the log returns of the S&P, reveals some interesting characteristics: an asymmetric feature, a higher peak and two heavier tails than those of the normal distribution. To illustrate this idea, in this same figure, a normal distribution with the same mean and standard deviation as the price series is plotted in red to compare.

As another example, in January 2015, the Swiss National Bank (SNB) discontinued the minimum EUR/CHF exchange rate, which was fixed with a floor of 1.20 since 2011. This measure taken by the SNB, provoked a huge fluctuation of this rate which is illustrated by Figures A.3 and A.4. These charts, such as in the previous example, stand respectively for the values of the EUR/CHF in 2015 and the corresponding log returns. Looking into the histogram A.6, it is also evident that the log-returns of the underlying are not normally distributed. To sustain this idea, we first present two definitions:

Definition 1.1. *Given a random variable X, the kurtosis and skewness of X are respectively defined as $\mathbb{E}[(\frac{X-\mu}{\sigma})^3]$ and $\mathbb{E}[(\frac{X-\mu}{\sigma})^4]$, where $\mathbb{E}(\cdot)$ is the expected value of the random variable X.*

The price series corresponding to the S&P500 presents a skewness equal to -1.2752 and kurtosis equal to 4.1427. The EUR/CHF log returns have these values respectively equal to 1.8168 and 8.5888. Comparing to the normal distribution, which has a skewness of zero and kurtosis equal to 3, we see that for both series, the values are far from normal. Moreover, we see that these series are "leptokurtic", meaning their kurtosis is above 3.

A Jarque-Bera (see [7] for more details) test was performed to evaluate the normality of the presented series. This test has as null hypothesis that the sample data of a random variable comes from a normal distribution. The alternative hypothesis is that it does not. With a confidence level of 95%, the null hypothesis was rejected, which means the

assumption of a normal distribution for the log-returns of these assets is not plausible. This is even more evident if we look into the histogram of S&P500 because it displays a high peak and asymmetric heavy tails. These characteristics are not exclusive for this index and are also present for almost all financial asset prices like individual stocks, foreign exchange rates and interest rates.

Therefore, as you can observe, the normal distribution does not fit very well with empirical data, since the empirical distribution of asset returns exhibits fat tails and skewness, a behaviour that deviates from normality and is inconsistent with the BSM model. Besides, as also evident in the images referred before, asset price processes have jumps or spikes and must be taken into account when pricing financial contingent claims. This is crucial for the estimation of current market values of portfolios held by companies that deal on a daily basis with products that depend on financial assets, such as banks and hedge funds. More importantly, it is essential to provide prices of contingent claims to clients who take into account these events.

It is then evident that under the geometric Brownian motion variations of great amplitude are not likely to happen as it assumes the log-returns have a normal distribution. This is documented with empirical evidence in [30]. Besides, Brownian motion is almost surely continuous by definition, a feature that is inconsistent with the existence of referred variations, which are commonly called "jumps" in literature. One way to try to get around the problem is to consider Lévy processes in which non-marginal variations are more likely to happen as a consequence of fat-tailed distributions based processes. Lévy based models are therefore much more realistic.

On the other hand an essential requirement for using a pricing model, independently of the considered distribution for the asset returns, is that it should ensure that the prices of actively traded instruments such as Europeans options coincide as much as possible with their market price. However, despite its excellent analytical tractability, this condition is not fulfilled in the case of the BSM model. The BSM model does not track the market's implied volatility well enough. In fact, by inverting the BSM formula with respect to the volatility for a series of options with different strikes and

with the same maturity, one should obtain approximately the same (constant) implied volatility. However, it is an empirical fact that the graph of the implied volatility is not a horizontal line as a function of strike or as a function of time to maturity, but resembles a "smile" instead, as reported by Hull [29]. Thus, using a single value for the volatility in order to price options with different strikes and maturities leads us to get prices that are not in accordance with the ones displayed in the financial markets.

It then makes sense to consider an alternative for the BSM model. Jump-diffusion models and Lévy based models, proposed in the late 1980s and early 1990s, are attractive because they explain the jump patterns exhibited by some stocks. Studies reveal that Lévy models are realistic when pricing options close to maturity [14]. However, they are more difficult to handle numerically. And in contrast to the basic BSM model, Lévy models do not immediately make obvious which hedging strategy leads to an instantaneous risk-free portfolio.

The aim of this thesis is then to introduce the Lévy processes and models based on these in order to incorporate the features discussed above and try to overcome some of the flaws in the BSM model. It is important to come up with a pricing model that traders and other market agents can utilize that captures the behavior of implied volatility smiles more accurately and considers the occurrence of jumps in order to handle the risks of trading. Lévy processes provide us with appropriate tools to adequately and consistently describe all these observations, both in the 'real' and in the 'risk-neutral' world as we shall see through the rest of this thesis.

The main idea is to replace the Normal distribution of the increments by a more general distribution that is able to better reflect the stylized facts such as the skewness and excess kurtosis present in the financial markets. Examples of distributions that attempt to achieve this are the Variance Gamma [40], the Normal Inverse Gaussian [4], the CGMY [11] and the Hyperbolic Model [18]. These admit the possible existence of an infinite number of jumps in any interval and so are called "infinite activity models". Here, the price behavior on small time intervals is modeled by jumps and the Brownian motion is no longer needed. We will, however, deal only with jump diffusions. For a

more interested reader in infinite activity models, we suggest the consultation of [14], [42] or [44]. Pure jump models with finite activity are also possible but, according to [14], they do not lead to a realistic description of price dynamics.

In defiance of the analytical tractability offered by the Lévy processes, the constraints of independence and stationarity of their increments bring some drawbacks. Mandelbrot and Hudson [30] advocate Lévy processes lead to rigid scaling properties for marginal distributions of returns which are not observed in empirical time series. On the other hand, in spite of being able to calibrate the implied volatility patterns for a single maturity under the risk neutral measure, they fail to reproduce correct option prices over a range of different maturities. Both issues are owed to the fact that exponential Lévy models do not allow for time inhomogeneity. It has been observed that the estimated volatilities change stochastically over time and are clustered as reported in [14] or [30].

Local volatility models were proposed as an alternative. For example, Cox [16] introduced the Constant Elasticity Variance (CEV) model which attempted to introduce the leverage effect. It tries to incorporate the fact that in equity markets volatility increases when prices decrease due to investors fear sentiment. The opposite effect occurs in commodity markets. Heston [28] considered a diffusion-based stochastic volatility model in which the price and volatility are correlated and the latter follows a squared-root process.

A jump-diffusion stochastic volatility model was proposed by Bates in [5], which deals with this problem by adding proportional log-normal jumps to the Heston stochastic volatility model.

Despite the flaws inherent in Lévy processes, they are very useful for pricing purposes and provide an analytical tractability that more sophisticated models do not. In this thesis we start by presenting some theory on the important concepts related to Lévy processes in Chapter 2 and then introduce two models in Chapter 3, one which dates back to Merton [41] and considers that the logarithm of the asset price returns follows a diffusion with jumps that have a normal distribution and a more recent one, developed

by Kou [33] where it is assumed that jump sizes have a double exponential distribution. Presented these two jump-diffusion models, we explain, in Chapter 4, the rationale behind the fact that assuming assets are driven by Lévy processes leads to the existence of markets which are not complete. We establish the existence of a partial integro-differential equation (PIDE) using a risk-free measure in Chapter 5, that describes the dynamics of the price of a European put option under the jump-diffusion model. Chapter 6, provides a localization error estimate using the asymptotic behavior of the European put option and Chapter 7 explains how to construct the implicit method by using three time levels.

One can apply several numerical tools to calculate option prices such as Monte Carlo simulation and finite-difference methods or use numerical integration techniques when the closed-form solution of the characteristic function of the log returns is known.

Monte Carlo simulation in its most basic form is probably the simplest numerical method one can implement (see [26] for more details). As long as American features are not required and great accuracy is not necessary, Monte Carlo is a very good method for pricing options. However, these methods may take considerable time to simulate.

The purpose of this dissertation is to develop a finite difference method which avoids iteration at each time step and has a second-order convergence rate to solve the PIDE under the jump-diffusion model. Almendral and Oosterlee [1] suggested finite difference and element methods with the second-order backward differentiation formula (BDF2) in the time variable for pricing European options under the jump-diffusion model. These implicit methods with the BDF2 use iterative techniques to solve linear systems involving dense matrices and implement the FFT for the integral term to reduce the computational complexity. The numerical method in [15] uses two time levels but only has a first-order convergence rate without iteration at each time step. Another numerical method in [22] also uses two time levels according to the Crank-Nicolson scheme and has the second-order convergence rate. However, it must carry out iterations at each time step. For this reason, it is necessary to use three time levels. We particularly focus on the construction of linear systems whose coefficient matrices are not dense but

tridiagonal matrices instead. So the implicit method can be solved easily by using LU decomposition.

Regarding American-style options, d'Halluin-Forsyth-Labahn [23] and d'Halluin-Forsyth-Vetzal [22] proposed an implicit method of the Crank–Nicolson type combined with a penalty method for pricing American options under the Merton model. In [23] the authors showed that the fixed point iteration at each time step converges to the solution of a linear system of discrete penalized equations. In the case of American options under the Kou model, Toivanen [51] developed numerical methods coupled with two alternative ways to solve the LCPs. One is the operator splitting method introduced by Ikonen and Toivanen [31], and the other is the penalty method proposed by d'Halluin, Forsyth, and Labahn [23]. These numerical methods use iterative techniques to solve linear systems involving dense coefficient matrices.

We provide results in Chapter 8 that show our method is stable and consistent with respect to the discrete l^2-norm in the sense of the Von Neumann analysis and prove it has second-order convergence rate.

Chapter 9, presents the numerical results we obtain for vanilla European options for the implicit method with three time levels under the Merton and Kou models. Chapters 10, 11 and 12 represent an extension of the European options to options with American-style features where theory related to this kind of contracts is described as along with the results for the numerical tests performed.

Finally, we draw some conclusion in Chapter 14.

2 Lévy processes

Before defining a Lévy process, a few concepts are introduced which are useful to comprehend what follows next. For a better understanding or a more detailed study, the reading of [32] or [39] is suggested.

Definition 2.1. *A stochastic process is a mathematical model for the occurrence, at each moment after an initial time, of a random phenomenon. The randomness is captured by the introduction of a measurable space called the sample space, on which probability measures can be placed. From now on, we define* Ω *as the probability space, the set which contains all the possible events. Thus, a stochastic process is a collection of random variables* $(X_t)_{t\in\mathbb{R}^+}$ *on* $(\Omega, \mathcal{F})$*, which take values in a second measurable space* $(\Omega', \mathcal{F}')$*, called the state space.* $\mathcal{F}$ *and* $\mathcal{F}'$ *represent sigma-algebras and in this framework, the state space will be the d-dimensional Euclidean space equipped with the sigma-field of the Borel sets. The index* $t \in [0, +\infty)$ *of the random variables* X *is interpreted as the time.*

Definition 2.2. *A filtration in a measurable space* $(\Omega, \mathcal{F})$ *is a family of sub-*σ*-algebras of* $\mathcal{F}$*,* $(\mathcal{F}_t)_{t\in I}$*, with* I *being a set of indexes, such that if* $0 \leq s \leq s'$ *then* $\mathcal{F}_s \subset \mathcal{F}_{s'}$*.*

The σ-algebra $\mathcal{F}_t$ represents the information we have at time t. When, for each $t \geq 0$, the random variable X_t is $\mathcal{F}_t$-measurable, we say that the process $(X_t)_{t\in\mathbb{R}_0^+}$ is adapted to $(\mathcal{F}_t)_{t\in I}$.

Within all the results that will be presented from now on, the following hypotheses shall be implicitly assumed:

- It is possible for short-sales to take place, i.e., we can sell an asset which we do not possess;
- The asset's quantities we are trading may not be integer values, they may be any real number;
- The price for selling a security is the same as the price for buying the same security;
- There are no transacting costs;
- The market is completely liquid, meaning that it is always possible to buy or sell any asset we want.

These hypothesis are not precisely verified in real life, but they allow for a convenient simplification of our models. Moreover, we shall also suppose that no arbitrage opportunities exist. This will be important for pricing derivative products using a "risk-neutral measure". In a nutshell, an arbitrage opportunity is a way to obtain profits, without an initial investment and without any risk. In general, when those appear in the market, they quickly vanish, returning the assets' prices to their "equilibrium value" and that is why they are assumed to be nonexistent.

Definition 2.3. *Suppose* $(\Omega, \mathcal{F}, \mathbb{P})$ *is a probability space, where* Ω *is the set containing all the possible events,* $\mathcal{F}$ *a* $\sigma-$*algebra and* $\mathbb{P}$ *a probability measure, often mentioned as the "physical measure" in the literature.*

A one-dimensional stochastic process $\{L_t\}_{t\geq 0}$ *on a probability space* $(\Omega, \mathcal{F}, \mathbb{P})$ *is a Lévy process if (see [43]):*
(1) $L_0 = 0$ *almost surely (a.s.);*
(2) It has independent increments, that is, for any $0 \leq s < t < s' < t'$ *the random variables* L_t - L_s *and* $L_{t'}$ - $L_{s'}$ *are independent;*
(3) It has stationary increments, that is, for any $0 \leq s \leq t \leq \infty$ *the law of* L_t - L_s *only depends on the increments t - s;*

(4) It is stochastically continuous, that is, $\lim_{t\to 0}\mathbb{P}(|L_{s+t}-L_s|\geq\epsilon)=0, \quad \forall\epsilon\geq 0$*;*

(5) The sample paths are right continuous with left limits a.s.

Examples of Lévy processes are the Brownian motion, the Poisson processes and also its extension to a Compound Poisson Process. These processes are defined and some of their properties are described in the Appendix.

From the definition, we can deduce that the value of a Lévy process in a certain point in time does not depend on its past values, i.e., it evolves independently of what happened before that time instant. Besides, if we look into a specific time interval, that is, to an increment from one point to another with a determined length, the distribution of the difference between the variable in the last point and the one at the first point is the same as the distribution of an increment with the same length, but considering other two moments in time. This may be a flaw when applying to reality, but from an application point of view, it is a simplification which will make the pricing model we are introducing in the next chapters much simpler to analyse. To illustrate this idea, imagine, for instance, that we are modelling the number of cars who cross a certain bridge in one day. Usually, the number of cars who cross it between 2 and 3 pm is much less than the number who pass between 8 am and 9 am or between 6 pm and 7 pm, because the latter usually is the rush hour. Therefore, it might not be suitable that the distribution in the rush hour is the same as the one between 2 and 3 pm.

2.1 Geometric Lévy processes

Once we define a Lévy process, we can model the asset value by what we call the geometric Lévy process, $S_t = S_0 e^{L_t}$, on the filtered probability space $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbb{P})$, where $\mathcal{F}_t$ is the filtration generated by the Lévy process $\{L_t\}_{t\geq 0}$. This process somehow represents an extension to the geometric Brownian motion (GBM). The GBM is the solution of the stochastic differential equation

$$\frac{dS_t}{S_t} = \mu dt + \sigma W_t \tag{2.1}$$

and, as we shall see, the geometric Lévy process (GLP) follows a similar dynamic, but with an added jump component. Here, μ stands for the mean of the asset's return and σ its annualised volatility.

The GBM is a case of a diffusion process, which assumes no jumps take place, contrasting with jump-diffusion processes.

Definition 2.4. *A pure jump process $(J_t)_{t\geq 0}$ is constant between jumps and is adapted and right-continuous.*

As an example, Figure (A.7) in Appendix illustrates a simulation of a Poisson process, which is a particular case of a pure jump process. The expected number of events per unit time was set to 0.1. The horizontal axis represents time and the vertical the value of the random variable, denoted by $N(t)$.

When we combine diffusion processes with pure jump ones, we create what is defined as a jump-diffusion process:

Definition 2.5. *A jump-diffusion or jump process is a process of the form*

$$\begin{aligned} X_t &= X_0 + \int_0^t \gamma_s dW_s + \int_0^t \theta_s ds + J_t \\ &:= X_t^C + J_t, \end{aligned}$$

where $(J_t)_{t\geq 0}$ is a pure jump process and X_t^C is the continuous part of X_t. $(\gamma_s)_{s\geq 0}$ and $(\theta_s)_{s\geq 0}$ are adapted processes and $(W_s)_{s\geq 0}$ is a Brownian motion.

As with most of density distributions, it is often useful to know what is the respective characteristic function. The next very useful theorem, commonly referred as the Lévy–Khintchine representation, allows the description of the characteristic function of the large family of Lévy-process and despite being very technical, its proof can be found in [49].

2.2 Lévy–Khintchine representation

Theorem 2.6. *Let $(L_t)_{t\geq 0}$ be a Lévy process. For all $z \in \mathbb{R}$ and $t \geq 0$,*

$$\mathbb{E}(e^{izL_t}) = \exp\left[t\left(-\frac{a}{2}z^2 + ibz + \int_{\mathbb{R}} (e^{izx} - 1 - izx\mathbb{1}_{|x|\leq 1})d\nu(x)\right)\right], \tag{2.2}$$

where a is a non-negative real number, b is a real number and ν is a measure on $\mathbb{R}$ satisfying $\nu\{0\} = 0$ and $\int_{\mathbb{R}} \min(1, x^2)\, d\nu(x) < \infty$.

The set of three parameters (a, b, ν) is commonly known as the generating Lévy triplet. The first parameter a is called Gaussian variance, since it is associated with the Brownian part of the Lévy process and the third quantity ν is called Lévy measure. If $\nu = 0$ then L_t is a drifted Brownian motion and if $a = 0$ then L_t is said to be purely non-Gaussian.

From the definition of a Lévy process, we know that the sample paths are càdlag ("continue à droite, limité à gauche"), over finite intervals $[0, t], t \leq T$, i.e., they are right-continuous and have left limits almost surely. Thus, any path has only a finite number of jumps with absolute jump size larger than ϵ for any $\epsilon > 0$. As a consequence, the sum of jumps along $[0, t]$ with absolute jump size bigger than 1 is a finite sum for each path. Of course instead of the threshold 1, one could use any number $\epsilon > 0$ here. Contrary to the sum of the big jumps, the sum of the small jumps does not converge in general. There might be too many small jumps to get convergence. This is the reason why we need the term $izx\mathbb{1}_{|x|\leq 1}$ in equation (2.2) in general, so that the integral converges. However, according to Sato [49], it is necessary a Lévy measure satisfies $\int_{\mathbb{R}} (|x| \wedge 1)\nu(dx) < \infty$, (this means that the jump part of the Lévy process is of finite variation), in order for the truncation of small jumps not to be needed in (2.2). The symbol $\wedge$ stands for the minimum between two quantities (in this case, $|x|$ and 1). Furthermore, if the Lévy density decays fast enough as $x \to \pm\infty$, we can replace $izx\mathbb{1}_{|x|\leq 1}$ by izx. If the process is of finite variation, then we do not need this term at all.

Generally speaking, a jump Lévy process can display either finite activity or infinite

activity. In the former case, the aggregate jump arrival rate is finite, whereas in the latter case, an infinite number of jumps can occur in any finite time interval. Within the infinite activity category, the sample path of the jump process can either exhibit finite variation or infinite variation. In the former case, the aggregate absolute distance travelled by the process is finite, while in the latter case, the aggregate absolute distance travelled by the process is infinite over any finite time interval.

$\nu(dx)$ is interpreted as the expected number of jumps per unit of time whose size belongs to the set dx. Therefore, we formally summarize the last paragraphs in the following two results, found in [49]:

Theorem 2.7. *Let L_t be a Lévy process with triplet (a, b, ν).*

(1) If $\nu(\mathbb{R}) < \infty$, then almost all paths of L_t have a finite number of jumps on every compact interval. The Lévy process has finite activity.

(2) If $\nu(\mathbb{R}) = \infty$, then almost all paths of L_t have an infinite number of jumps on every compact interval. The Lévy process has infinite activity.

Theorem 2.8. *Let L_t be a Lévy process with triplet (a, b, ν).*

(1) If $\sigma^2 = 0$ and $\int_{\mathbb{R}} (|x| \wedge 1)\nu(dx) < \infty$, then almost all paths of L_t have finite variation.

(2) If $\sigma^2 \neq 0$ or $\int_{\mathbb{R}} (|x| \wedge 1)\nu(dx) = \infty$, then almost all paths of L_t have infinite variation.

For the particular case of a compound Poisson process, only a finite number of jumps in any finite time interval take place, meaning the truncation function in (2.2) can be omitted when deriving the characteristic function of this process.

2.3 Drifted Brownian motion with a finite number of jumps

Consider the following process, corresponding to the drifted Brownian motion with a finite number of jumps:

$$L_t = \left(\mu - \frac{\sigma^2}{2}\right) t + \sigma W_t + \sum_{i=0}^{N_t} Y_i. \tag{2.3}$$

Y_i, $i \in \{0, \ldots, N_t\}$, is a sequence of independent, identically distributed random variables with a common probability density function p; N_t, $t \geq 0$, is a Poisson process with intensity λ and N_t, Y_i and W_t are mutually independent.

Y_i is a function of the jump absolute size, defined from now on by η_t. Suppose that in an infinitesimal time frame dt, S_t jumps to $\eta_t S_t$. The relative jump size is then: $dS_t = \frac{\eta_t S_t - S_t}{S_t} = \eta_t - 1$. The latter represents the infinitesimal change in S_t per unit of time, after a jump occurred.

2.4 The SDE

As found in [9] or [39], in order to guarantee no "free lunch" takes place (remember that one of our hypothesis is that no arbitrage exists), it is required and sufficient that there exists an equivalent martingale measure $\mathbb{Q}$. By definition, this means the following:

- $\forall A \in \mathcal{F}, \mathbb{Q}(A) > 0 \iff \mathbb{P}(A) > 0$ and
- The discounted process, $(\tilde{S}_t)_{t\in[0,T]} = (e^{-(r-q)t} S_t)_{t\in[0,T]}$ is a martingale in the probability space $(\Omega, \mathcal{F}, \mathbb{Q})$.

The definition of a martingale can, for instance, be found in [9].

As a consequence of the existence of this measure in our arbitrage-free context, the value of a contingent claim at time $t \in [0, T]$ is the discounted value of its expected value under the risk neutral measure. Given we are assuming our asset follows a GLP

and that S_t is $\mathcal{F}_t$-measurable and non-negative, this is equivalent to:

$$\begin{aligned} & \mathbb{E}(S_T|\mathcal{F}_t) = S_t e^{(r-q)(T-t)} \\ \iff & \mathbb{E}(S_t e^{L_T}|\mathcal{F}_t) = S_t e^{(r-q)(T-t)} \\ \iff & \mathbb{E}(e^{L_T}|\mathcal{F}_t) = e^{(r-q)(T-t)} \end{aligned} \tag{2.4}$$

Now, let us compute the characteristic function of (2.3), implicitly assuming it is conditional to $\mathcal{F}_t$:

$$\begin{aligned} \forall z \in \mathbb{C},\, \mathbb{E}(e^{izL_t}) = & \\ & = \mathbb{E}\left[\exp\left(iz \left((\mu - \frac{\sigma^2}{2})t + \sigma W_t + \sum_{i=0}^{N_t} Y_i \right) \right) \right] \\ & = e^{iz(\mu - \frac{\sigma^2}{2})t} \times \mathbb{E}\left[e^{iz\sigma W_t} \times \exp\left(iz \sum_{i=0}^{N_t} Y_i \right) \right] \\ & = e^{iz(\mu - \frac{\sigma^2}{2})t} \times \mathbb{E}[e^{iz\sigma W_t}] \times \mathbb{E}\left[\exp\left(iz \sum_{i=0}^{N_t} Y_i \right) \right] \end{aligned}$$

Here we used the fact that the compound Poisson process and the Brownian motion are independent. Since the exponential function is continuous, then the random variables $e^{\,izW_t}$ and $e^{\,iz\sum_{i=0}^{N_t} Y_i}$ are independent. Consequently, the expected value of their product is equal to the product of the respective expected values. Now, recalling that the Brownian motion follows a normal distribution with mean zero and variance t, $\mathbb{E}(e^{iz\sigma W_t})$ can be easily computed:

$$\mathbb{E}(e^{iz\sigma W_t}) = e^{iz\sigma \times 0 - t\frac{(z\sigma)^2}{2}} = e^{-t\frac{(z\sigma)^2}{2}}$$

On the other hand, the generating moment function of the compound Poisson process $X_t = \sum_{i=0}^{N_t} Y_i$ is given by (see Theorem A.5 in the Appendix):

$$\forall z \in \mathbb{C}, \mathbb{E}(e^{izX_t}) = \exp\left(t\lambda \int_{\mathbb{R}} (e^{izx} - 1) p(x)\, dx \right), \tag{2.5}$$

where p represents the probability density function of the random variables Y_i, $i \in \{0, \ldots, N_t\}$.

The moment generating function (mgf for short) of our Lévy-process (2.3) can be finally written as:

$$\forall\, z \in \mathbb{C}, \mathbb{E}(e^{izL_t}) = \exp\left\{ t\left(iz\Big(\mu - \frac{\sigma^2}{2}\Big) - \frac{(z\sigma)^2}{2} + \lambda \int_{\mathbb{R}} (e^{izx} - 1)p(x)\,dx \right) \right\} \tag{2.6}$$

Under the equivalent martingale measure (EMM from now on) $\mathbb{Q}$, using the fact that the exponential function is injective and evaluating the mgf at the point $z = -i$ and (2.4) at $t = 0$, we get:

$$\begin{aligned}
&\mathbb{E}(e^{L_T}) = e^{(r-q)T} \\
\iff\ & r - q = i\left(\mu - \frac{\sigma^2}{2}\right) \times (-i) - \frac{\sigma^2(-i)^2}{2} + \lambda \int_{\mathbb{R}} (e^{x} - 1)p(x)\,dx \\
\iff\ & r - q = \left(\mu - \frac{\sigma^2}{2}\right) + \frac{\sigma^2}{2} + \lambda \int_{\mathbb{R}} (e^{x} - 1)p(x)\,dx \\
\iff\ & r - q = \mu + \lambda\zeta
\end{aligned}$$

where ζ is defined by: $\zeta := \int_{\mathbb{R}} (e^x - 1)p(x)\,dx$.

Summing and subtracting the exponential of $iz\lambda t \int_{\mathbb{R}} xp(x)\,dx$ to the moment generating function we got, it becomes:

$$\begin{aligned}
\forall z \in \mathbb{C}, \mathbb{E}(e^{izL_t}) = \exp\Bigg\{ t\Bigg(& iz\Big(\mu - \frac{\sigma^2}{2}\Big) - \frac{(z\sigma)^2}{2} + \lambda \int_{\mathbb{R}} (e^{izx} - 1)p(x)\,dx \Bigg) \\
& + iz\lambda t \int_{\mathbb{R}} xp(x)dx \; - iz\lambda t \int_{\mathbb{R}} xp(x)dx \Bigg\}.
\end{aligned} \tag{2.7}$$

$\int_{\mathbb{R}} xp(x)dx$ is the expected value of each of the variables Y_i, ($\mathbb{E}(Y)$ for short, since they follow the same distribution), so the last equation is equivalent to:

$$\begin{aligned}
\forall z \in \mathbb{C}, \mathbb{E}(e^{izL_t}) = \exp\Bigg\{ t\Bigg(& iz[r - q - \lambda\zeta - \frac{\sigma^2}{2} + \lambda\mathbb{E}(Y)] \\
& - \frac{\sigma^2}{2}z^2 + \lambda \int_{\mathbb{R}} [e^{izx} - 1 - izx]p(x)dx \Bigg) \Bigg\}
\end{aligned} \tag{2.8}$$

We can now identify the Lévy triplet of this process as being equal to

$$(\sigma, r - q - \lambda\tilde{\zeta} - \frac{\sigma^2}{2}, \lambda P)$$

where $\tilde{\zeta} := \zeta - \mathbb{E}(Y)$ and $P(x) = p(x)dx$.

In Chapter 3 of [42], the author proves that the exponential of the Lévy process (2.3) is the solution of the stochastic differential equation (SDE):

$$\frac{dS_t}{S_t} = (\mu - \lambda\tilde{\zeta})dt + \sigma W_t + (\eta_t - 1)dN_t. \tag{2.9}$$

Combining a Brownian motion with drift and a compound Poisson process, we obtain the simplest case of a jump-diffusion — a process which sometimes jumps and has a continuous but random evolution between the jump times and that is why we are trying to model our asset dynamics using (2.3). Intuitively speaking, the diffusion part takes into account the normal fluctuations in the risky asset's price caused by economic factors such as changes in capitalization rates or a temporary imbalance between supply and demand. These are considered not to be very significant, since they only cause marginal fluctuations in the price. The jump component is added to the diffusion process to reflect the fact that non-marginal variations happen in discrete points in time. These appear as a consequence of certain events such as a new product which is going to be launched and is expected to generate significant profits to the company or the expectation of a significant change in the political and economic regime like the Brexit referendum which took place the last 23rd June 2016.

Nonetheless, it is now necessary to choose the density function for the jumps. Two of the most common jump distributions are the Normal density, which is the one introduced by Merton in [41] and the double-exponential, presented by Kou in [33]. These are described in the next chapter and we shall then see how we can apply them in option pricing.

3 Jump-diffusion models

In this thesis, we confine ourselves to just two models: the Merton's and the Kou's model.

A problem with jump diffusion models is that, in general, they do not yield a closed-form solution for option prices; instead one has to solve them numerically. This will be done in the next chapters, where we describe a numerical method in order to price vanilla options. However, in the specific cases of the Merton's and Kou's model, it is possible to express the solution in terms of an infinite series. This enables us, in particular, to compare the numerical results to the analytical solution, in order to get an idea of the error in our method.

3.1 Merton's model

As the first to explore jump diffusion models, Merton in [41] assumes the Neperian logarithm of the sequence of jumps Y_i, $i \in \{0, \ldots, N_t\}$ in (2.3) follows a normal distribution with mean μ_J and variance σ_J^2, i.e., a distribution with following associated density function:

$$p(x) := \frac{1}{\sigma_J\sqrt{2\pi}} \exp\left[-\frac{(x-\mu_J)^2}{2\sigma_J^2}\right], \quad \forall x \in \mathbb{R}. \tag{3.1}$$

Figure 3.1 represents an example of a simulation of the Merton's model for illustration purposes.

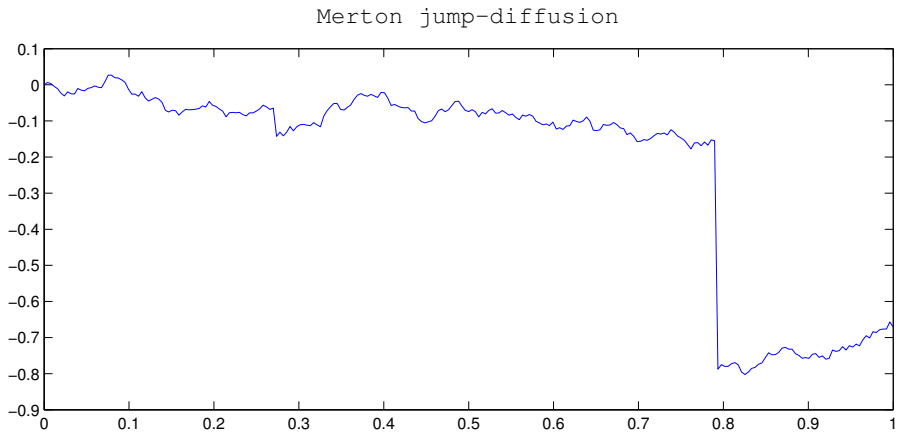

Figure 3.1: Simulation of a jump-diffusion path of the Merton's model

In the horizontal axis, we represent the time in years, ranging from 0 (today) and 1. The values in the vertical axis correspond to the log-returns.

Some of the properties of this model are for example studied in [42] and we are summarizing some of them here.

Let $H(S_T)$ be the value of a contingent claim at its maturity ($H(S_T) = (S_T - K)^+$ in the case of a European call and $H(S_T) = (K - S_T)^+$ for a European put), $H(S_t)$ its intrinsic value at time $t \in [0, T]$ and

$$H^{BS}(S_t, \sigma, \tau, r - q) := e^{-r\tau} E^{\mathbb{Q}}[H(S_t e^{(r-q-\frac{\sigma^2}{2})\tau} + \sigma W_\tau^{\mathbb{Q}})]. \tag{3.2}$$

The latter corresponds to the pricing formula developed by Black-Scholes-Merton in [10] for a European-style plain vanilla option, where it is supposed that the underlying follows a GBM. Here, S_t, σ, r and q represent, respectively, the price of the underlying asset at time t, the diffusion part associated to the Brownian motion in (2.3), the risk-free rate and the continuous dividend yield inherent to the underlying asset. $\tau := T - t$ is the time until the maturity.

Let j stand for the number of jumps occurring during the time-period τ, which follows a Poisson distribution. According to [42], the pricing formula for the Merton model is:

$$H_t^M(S_t) = \sum_{j \geq 0} \frac{e^{-\lambda\tau}(\lambda\tau)^j}{j!} H^{BS}(S_j \equiv S_t e^{j\mu_J + \frac{j\sigma_J^2}{2} - \lambda(e^{\mu_J + \frac{\sigma_J^2}{2}} - 1)\tau}, \sqrt{\sigma^2 + \frac{j\sigma_J^2}{\tau}}, \tau, r - q) \tag{3.3}$$

As we can see, it is possible for the jump-diffusion to represent the price of a vanilla call or a put as a weighted average of the standard Black-Scholes-Merton prices.

Define d_1 and d_2 as:

$$d_1 = \frac{\ln(S_t/K) + (r - q + \sigma^2/2)(T-t)}{\sigma\sqrt{T-t}} \quad \text{and}$$
$$d_2 = d_1 - \sigma\sqrt{T-t}.$$

The prices of a call and put options at time t, under the BSM model, are respectively equal to:

$$Call_t^{BSM} = e^{-q(T-t)} S_t \mathcal{N}(d_1) - K e^{-r(T-t)} \mathcal{N}(d_2), \tag{3.4}$$
$$Put_t^{BSM} = K e^{-r(T-t)} \mathcal{N}(-d_2) - e^{-q(T-t)} S_t \mathcal{N}(-d_1), \tag{3.5}$$

where $\mathcal{N}$ represents the cumulative distribution function of the standard normal law.

Using the equations above, we plot the Figures 3.2 and 3.3 for illustrative purposes. They represent the price of vanilla options with strike equal to 100, for each spot price, in a range from 80 to 120.

As we can observe, the value of the options under the Merton's model in comparison with the value under the Black-Scholes model is larger as the rate arrival of jumps increases. It is also possible to show that this also happens when the variance of the jump distribution increases. This was expected, since more jumps with greater variance represent more uncertainty in the expected final payoff and lead to more potential earnings for the option's owner.

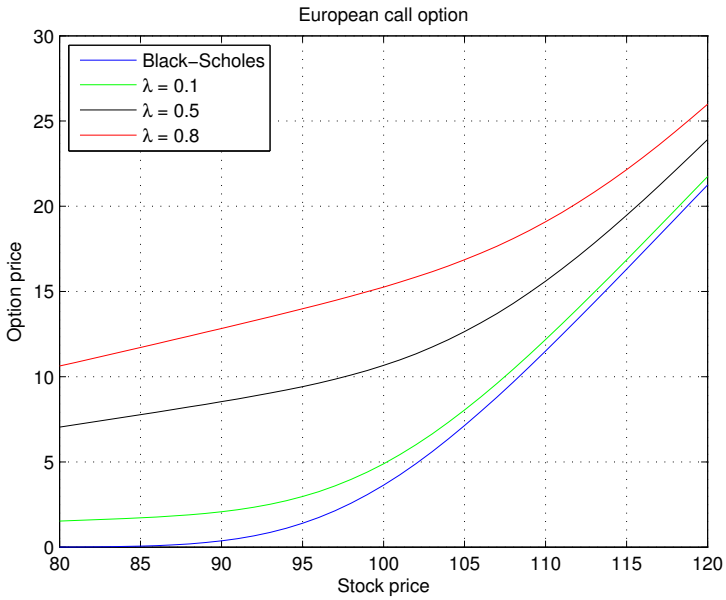

Figure 3.2: Call option prices with
$r = 0.05, \sigma = 0.15, q = 0, \mu_J = -0.9, \sigma_J = 0.45, T = 0.25$.

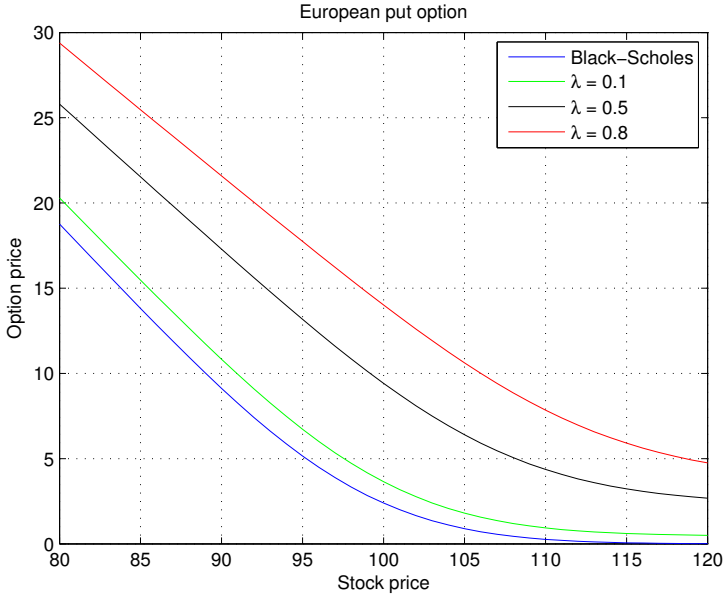

Figure 3.3: Put option prices with
$r = 0.05, \sigma = 0.15, q = 0, \mu_J = -0.9, \sigma_J = 0.45, T = 0.25$.

3.2 Kou's model

In [33], Kou suggests a jump-diffusion model based in a simple double exponential jump-diffusion. This means the jumps Y_i, $i \in \{0, \ldots, N_t\}$ in (2.3) follow a distribution described as follows:

$$p(x) := p\eta_1 e^{-\eta_1 x} \mathbb{1}_{\{x \geq 0\}} + (1-p)\eta_2 e^{\eta_2 x} \mathbb{1}_{\{x<0\}}, \quad \forall x \in \mathbb{R}, \tag{3.6}$$

with $p, \eta_1, \eta_2 \geq 0$.

For $x \geq 0$, the expected value of a variable with distribution (3.6) is $p \times \frac{1}{\eta_1}$ and for $x < 0$, it is equal to $(1-p) \times \frac{1}{\eta_2}$ (when the respective denominators are not null). From this observation, we interpret p as the probability of occurring an upward jump and $(1-p)$ the probability of occurring a downward one. Remark that the computation of the first expected value is only possible if $\eta_1 \geq 1$, so that its value is finite. In practical terms, this means that $\frac{1}{\eta_1} < 1$, i.e, the average jump size cannot exceed 100%, which is quite an acceptable hypothesis that we shall take in consideration from now on.

Figure 3.4 represents a simulation of the path of the log returns of an underlying, following the dynamics described by (2.3), but this time the measure of the jumps is given by (3.6).

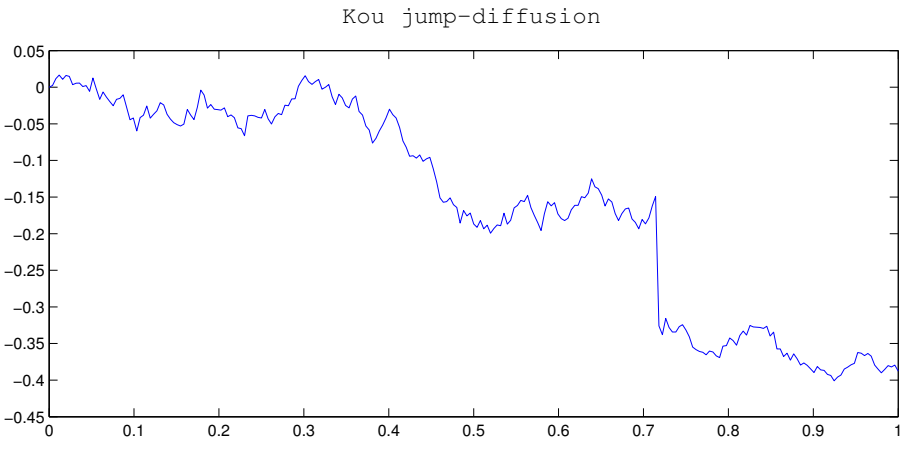

Figure 3.4: Jump-diffusion path simulation for the Kou model

3.2.1 Incremental returns

Quoting [33], "the empirical tests performed in Ramezani and Zeng (2002) suggest that the double exponential jump-diffusion model fits stock data better than the normal jump-diffusion model, and both of them fit the data better than the classical geometric Brownian motion model.". This statement corroborates the fact that the Kou model is suitable for our purpose and one of the reasons it was chosen to price financial derivatives. To better understand why is this true, let us first compute the incremental return of an underlying whose price at time t is S_t. As stated in the previous chapter, we are assuming S_t is the exponential of the Lévy process (2.3). Denoting by Δ an incremental variation, the incremental return becomes:

$$\begin{aligned}
\frac{\Delta S_t}{S_t} &= \frac{S_{t+\Delta t}}{S_t} - 1 \\
&= \frac{\exp\left\{\left(\mu - \frac{1}{2}\sigma^2\right)(\Delta t + t) + \sigma W_{t+\Delta t} + \sum_{i=0}^{N_{t+\Delta t}} Y_i\right\}}{\exp\left\{\left(\mu - \frac{1}{2}\sigma^2\right)t + \sigma W_t + \sum_{i=0}^{N_t} Y_i\right\}} - 1 \\
&= \exp\left\{\left(\mu - \frac{1}{2}\sigma^2\right)\Delta t + \sigma(W_{t+\Delta t} - W_t) + \sum_{i=N_t+1}^{N_{t+\Delta t}} Y_i\right\} - 1.
\end{aligned} \tag{3.7}$$

If Δt is sufficiently small (for instance, as when we consider daily observations), the return can be approximated in distribution, ignoring the terms with orders higher than Δt. Considering the second order approximation in a MacLaurin series of the exponential function ($e^x \approx 1 + x + \frac{1}{2}x^2$), the return of the asset is approximately equal to:

$$\begin{aligned}
\frac{\Delta S_t}{S_t} \approx \left(\mu - \frac{\sigma^2}{2}\right)\Delta t + \sigma(W_{t+\Delta t} - W_t) + \sum_{i=N_t+1}^{N_{t+\Delta t}} Y_i + \frac{1}{2}\Bigg(\left[\left(\mu - \frac{\sigma^2}{2}\right)\Delta t\right]^2 \\
+2\left(\mu - \frac{\sigma^2}{2}\right)\Delta t \times \left[\sigma(W_{t+\Delta t} - W_t) + \sum_{i=N_t+1}^{N_{t+\Delta t}} Y_i\right] + \left[\sigma(W_{t+\Delta t} - W_t) + \sum_{i=N_t+1}^{N_{t+\Delta t}} Y_i\right]^2\Bigg)
\end{aligned} \tag{3.8}$$

Now, to simplify the last equation, note the following facts:

- The random variable $(W_{t+\Delta t} - W_t)$ has expected value equal to 0 and variance Δt. Concerning the compound Poisson process $\sum_{i=N_t+1}^{N_{t+\Delta t}} Y_i$, its mean is $\lambda \Delta t \mathbb{E}(Y_i)$ and variance $\lambda \Delta t \mathbb{E}(Y_i^2)$, as proven in Theorem A.11, found in the Appendix;

- The Brownian motion and the Compound Poisson process are assumed to be independent, so the variance of its sum is equal to sum of the respective variances.

Using the last statements, we can imply the following:

- $\Big[\sigma(W_{t+\Delta t} - W_t) + \sum_{i=N_t}^{N_{t+\Delta t}} Y_i\Big]$ is a random variable with mean $\lambda \Delta t \mathbb{E}(Y_i)$ and variance $\lambda \Delta t \mathbb{E}(Y_i^2) + \sigma^2 \Delta t$. The multiplication of this variable with $2\Big(\mu - \frac{\sigma^2}{2}\Big)\Delta t$ makes this term of order $(\Delta t)^{\frac{3}{2}}$;

- The expected value of the product of $\sigma(W_{t+\Delta t} - W_t)$ and $\sum_{i=N_t}^{N_{t+\Delta t}} Y_i$ is 0 and its variance is of order $(\Delta t)^2$ (it is equal to $\sigma^2 \Delta t \times \lambda \Delta t \mathbb{E}(Y_i^2)$);

- $\Big[\sum_{i=N_t}^{N_{t+\Delta t}} Y_i\Big]^2$ has expected value and variance of at least order $o((\Delta t)^2)$. The calculus involved is not so straightforward as the previous ones, but using the independence of the variables Y_i, $i \in \{N_t + 1, ..., N_{t+\Delta t}\}$, this fact can be proven.

Gathering all this information together, the increment returns become:

$$\frac{\Delta S_t}{S_t} \approx \mu \Delta t + \sigma \sqrt{\Delta t} Z + \sum_{i=N_t+1}^{N_{t+\Delta t}} Y_i \tag{3.9}$$

with $Z \sim \mathcal{N}(0, 1)$.

Appealing to Theorem A.8, found in the Appendix, for $\lambda \Delta t$ small, the probability the Poisson process N_t having one jump is approximately $\lambda \Delta t$, having none is $1 - \lambda \Delta t$ and more or equal to two is 0. So, we can approximate the summation in the previous SDE as:

$$\sum_{i=N_t+1}^{N_{t+\Delta t}} Y_i \approx \begin{cases} Y_i, & \text{with probability } \lambda \Delta t \\ 0, & \text{with probability } 1 - \lambda \Delta t \end{cases}$$

Incorporating this approximation into (3.9), we can write:

$$\frac{\Delta S_t}{S_t} \approx \mu\Delta t + \sigma\sqrt{\Delta t}Z + BY, \tag{3.10}$$

where B is a Bernoulli random variable with $P(B = 1) = \lambda\Delta t$ and $P(B = 0) = 1-\lambda\Delta t$ and the distribution of Y is given by (3.6).

Notice that without the last part, BY, the last equation corresponds to the GBM, where the incremental return of the underlying follows a normal distribution.

3.2.2 Incremental returns distribution

We are now computing the density distribution of (3.10). In first place, assume

$$\sum_{i=N_t+1}^{N_{t+\Delta t}} Y_i \approx Y_i.$$

The distribution of the sum of two independent variables is equal to the convolution of the respective density functions (see Theorem A.2) and from that fact, it follows that for any $z \in \mathbb{R}$, the density function of the sum of normal and double-exponential random variables present in (3.10) is equal to:

$$\begin{aligned}
&\int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi\sigma^2\Delta t}} e^{-\frac{(z-x-\mu\Delta t)^2}{2\sigma^2\Delta t}} \times [p\eta_1 e^{-\eta_1 x}\mathbb{1}_{\{x\geq 0\}} + (1-p)\eta_2 e^{\eta_2 x}\mathbb{1}_{\{x<0\}}]dx \\
&= \int_{0}^{+\infty} \frac{1}{\sqrt{2\pi\sigma^2\Delta t}} p\eta_1 e^{\frac{-x^2+2x(z-\mu\Delta t)-(z-\mu\Delta t)^2-2\eta_1 x\sigma^2\Delta t}{2\sigma^2\Delta t}} dx \\
&+ \int_{-\infty}^{0} \frac{1}{\sqrt{2\pi\sigma^2\Delta t}} (1-p)\eta_2 e^{\frac{-x^2+2x(z-\mu\Delta t)-(z-\mu\Delta t)^2+2\eta_2 x\sigma^2\Delta t}{2\sigma^2\Delta t}} dx \\
&= p\eta_1 e^{-\frac{(z-\mu\Delta t)^2}{2\sigma^2\Delta t}} \int_{0}^{+\infty} \frac{1}{\sqrt{2\pi\sigma^2\Delta t}} e^{\frac{-x^2+2x(z-\mu\Delta t-\eta_1\sigma^2\Delta t)}{2\sigma^2\Delta t}} dx \\
&+ (1-p)\eta_2 e^{-\frac{(z-\mu\Delta t)^2}{2\sigma^2\Delta t}} \int_{-\infty}^{0} \frac{1}{\sqrt{2\pi\sigma^2\Delta t}} e^{\frac{-x^2+2x(z-\mu\Delta t+\eta_2\sigma^2\Delta t)}{2\sigma^2\Delta t}} dx \\
&= p\eta_1 e^{-\frac{(z-\mu\Delta t)^2}{2\sigma^2\Delta t}} e^{\frac{(z-\mu\Delta t-\eta_1\sigma^2\Delta t)^2}{2\sigma^2\Delta t}} \int_{0}^{+\infty} \frac{1}{\sqrt{2\pi\sigma^2\Delta t}} e^{-\frac{[x-(z-\mu\Delta t-\eta_1\sigma^2\Delta t)]^2}{2\sigma^2\Delta t}} dx
\end{aligned}$$

$$+(1-p)\eta_2 e^{-\frac{(z-\mu\Delta t)^2}{2\sigma^2\Delta t}} e^{\frac{(z-\mu\Delta t+\eta_2\sigma^2\Delta t)^2}{2\sigma^2\Delta t}} \int_{-\infty}^{0} \frac{1}{\sqrt{2\pi\sigma^2\Delta t}} e^{-\frac{[x-(z-\mu\Delta t+\eta_2\sigma^2\Delta t)]^2}{2\sigma^2\Delta t}} dx$$

$$=p\eta_1 e^{-\frac{(z-\mu\Delta t)^2}{2\sigma^2\Delta t}} e^{\frac{(z-\mu\Delta t)^2}{2\sigma^2\Delta t}} e^{-\frac{2(z-\mu\Delta t)\eta_1\sigma^2\Delta t}{2\sigma^2\Delta t}} e^{\frac{(\sigma^2)^2\eta_1^2(\Delta t)^2}{2\sigma^2\Delta t}} \int_{0}^{\infty} \frac{1}{\sqrt{2\pi\sigma^2\Delta t}} e^{-\frac{[x-(z-\mu\Delta t-\eta_1\sigma^2\Delta t)]^2}{2\sigma^2\Delta t}} dx$$

$$+(1-p)\eta_2 e^{-\frac{(z-\mu\Delta t)^2}{2\sigma^2\Delta t}} e^{\frac{(z-\mu\Delta t)^2}{2\sigma^2\Delta t}} e^{\frac{2(z-\mu\Delta t)\eta_2\sigma^2\Delta t}{2\sigma^2\Delta t}} e^{\frac{(\sigma^2)^2\eta_2^2(\Delta t)^2}{2\sigma^2\Delta t}} \int_{-\infty}^{0} \frac{1}{\sqrt{2\pi\sigma^2\Delta t}} e^{-\frac{[x-(z-\mu\Delta t+\eta_2\sigma^2\Delta t)]^2}{2\sigma^2\Delta t}} dx$$

$$=p\eta_1 e^{-(z-\mu\Delta t)\eta_1} e^{\frac{\sigma^2\eta_1^2\Delta t}{2}} \Phi\Big[\frac{z-\mu\Delta t-\eta_1\sigma^2\Delta t}{\sigma\sqrt{\Delta t}}\Big]$$
$$+(1-p)\eta_2 e^{(z-\mu\Delta t)\eta_2} e^{\frac{\sigma^2\eta_2^2\Delta t}{2}} \Phi\Big[-\frac{z-\mu\Delta t+\eta_2\sigma^2\Delta t}{\sigma\sqrt{\Delta t}}\Big],$$

where $\Phi(\cdot)$ represents the cumulative normal distribution function.

On the other hand, $\mu\Delta t+\sigma\sqrt{\Delta t}Z$, with $Z\sim\mathcal{N}(0,1)$, has distribution $\mathcal{N}(\mu\Delta t,\sigma^2\Delta t)$, so, its density function is given by:

$$\frac{1}{\sigma\sqrt{\Delta t}}\phi\Big(\frac{x-\mu\Delta t}{\sigma\sqrt{\Delta t}}\Big), \forall x\in\mathbb{R},$$

where $\phi(y) := \frac{e^{-\frac{1}{2}y^2}}{\sqrt{2\pi}}, \forall y\in\mathbb{R}$ is the standard normal density function.

Thus, for the case

$$\sum_{i=N_t+1}^{N_{t+\Delta t}} Y_i \approx 0,$$

the density of (3.10) corresponds to $\frac{1}{\sigma\sqrt{\Delta t}}\phi(\frac{x-\mu\Delta t}{\sigma\sqrt{\Delta t}})$.

Summarizing the previous results, the density function of the approximation for the returns as described in (3.10), denominated by g, can now be written as:

$$g(x) = \frac{1-\lambda\Delta t}{\sigma\sqrt{\Delta t}}\phi\Big(\frac{x-\mu\Delta t}{\sigma\sqrt{\Delta t}}\Big) + \lambda\Delta t\Bigg\{p\eta_1 e^{\frac{\sigma^2\eta_1^2\Delta t}{2}} e^{-(x-\mu\Delta t)\eta_1}\Phi\Bigg(\frac{x-\mu\Delta t-\sigma^2\eta_1\Delta t}{\sigma\sqrt{\Delta t}}\Bigg)$$
$$+(1-p)\eta_2 e^{\frac{\sigma^2\eta_2^2\Delta t}{2}} e^{(x-\mu\Delta t)\eta_2}\Phi\Bigg(-\frac{x-\mu\Delta t+\sigma^2\eta_2\Delta t}{\sigma\sqrt{\Delta t}}\Bigg)\Bigg\}.$$

Looking at (3.10) and using expected value's linearity, we can affirm that the mean of the a random variable with distribution g is $\mu\Delta t+\lambda\left(\frac{p}{\eta_1}-\frac{(1-p)}{\eta_2}\right)\Delta t$, since the mean of the normal distribution in (3.10) is $\mu\Delta t$ and the one of the product of the Binomial random variable and the variable Y is $\lambda\Delta t\times\left(\frac{p}{\eta_1}-\frac{(1-p)}{\eta_2}\right)$. The variance is the sum of the

variance of the normal distribution, corresponding to $\sigma^2\Delta t$, and the one of BY, because they are independent. Assuming also that the outcome of the Bernoulli's variable is in nothing correlated to the outcome of Y, the variance of BY is then:

$$\begin{aligned}
Var(BY) &= \mathbb{E}[(BY)^2] - \mathbb{E}^2[BY] \\
&= \mathbb{E}[B^2]\mathbb{E}[Y^2] - \mathbb{E}^2[B]\mathbb{E}^2[Y] \\
&= \mathbb{E}[B^2](Var[Y] + \mathbb{E}^2[Y]) + (Var[B] - \mathbb{E}[B^2])\mathbb{E}^2[Y] \\
&= \mathbb{E}[B^2]Var[Y] + Var[B]\mathbb{E}^2[Y] \\
&= \mathbb{E}[B]Var[Y] + Var[B]\mathbb{E}^2[Y] \\
&= \lambda\Delta t\left\{p(1-p)\left(\frac{1}{\eta_1} + \frac{1}{\eta_2}\right)^2 + \left(\frac{p}{\eta_1^2} + \frac{(1-p)}{\eta_2^2}\right)\right\} \\
&+ \Delta t(1-\Delta t)\left(\frac{p}{\eta_1} - \frac{(1-p)}{\eta_2}\right)^2.
\end{aligned}$$

Therefore, the variance of the density function g is given by:

$$\sigma^2\Delta t + \lambda\Delta t\left\{p(1-p)\left(\frac{1}{\eta_1} + \frac{1}{\eta_2}\right)^2 + \left(\frac{p}{\eta_1^2} + \frac{(1-p)}{\eta_2^2}\right)\right\} + \Delta t(1-\Delta t)\left(\frac{p}{\eta_1} - \frac{(1-p)}{\eta_2}\right)^2.$$

3.2.3 Empirical evidence

Figure 3.5 is a plot of the density function g, along with the one corresponding to the normal distribution, both with the same mean and variance.

Figures 3.6, 3.7 and 3.8 represent "zoom-in" plots of the graph of g, respectively evidencing the contrast between the peaks, left and right side tails of the normal law and g. As we can observe, the double-exponential distribution has an asymmetric leptokurtic feature for the asset returns as well as fatter tails than the normal distribution with the same mean and variance. The parameters used were the same as in [33].

Kou's model fits a commonly observed behaviour in the markets: an overreaction and under-reaction to various good or bad news. More precisely, in the absence of outside news, the asset price simply follows a GBM. Good or bad news arrive according to a Poisson process, and the asset price changes in response, depending on the jump size distribution. Because the double exponential distribution has both a high peak

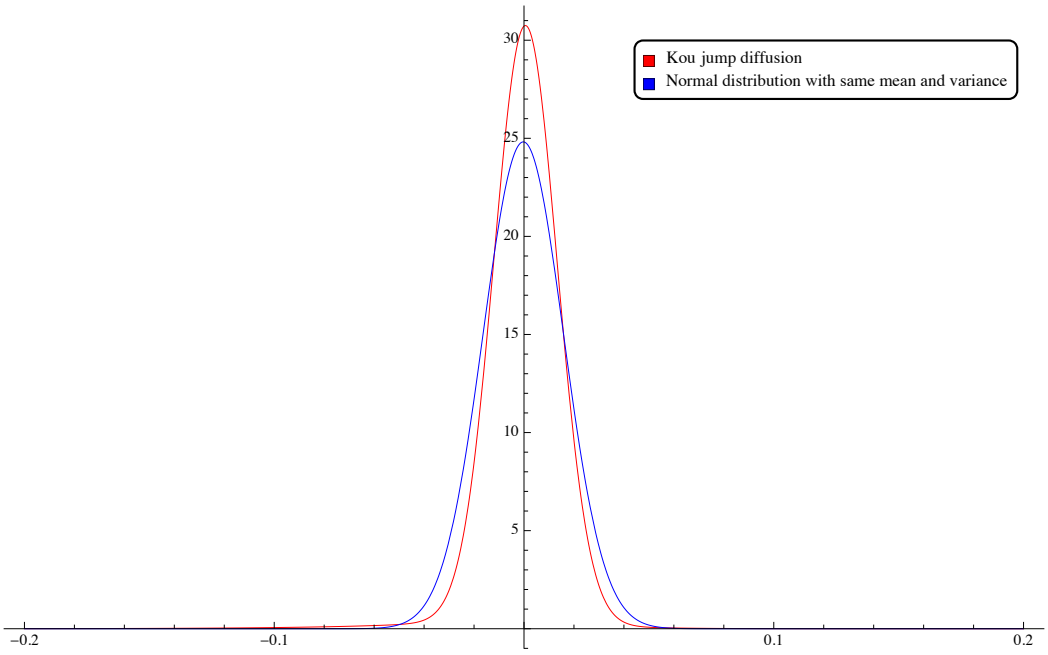

Figure 3.5: Plot of the density function g and the normal distribution. The parameters used were:
$\mu = 0.15, \sigma = 0.20, \lambda = 10, \eta_1 = 50, \eta_2 = 25, p = 0.3.$

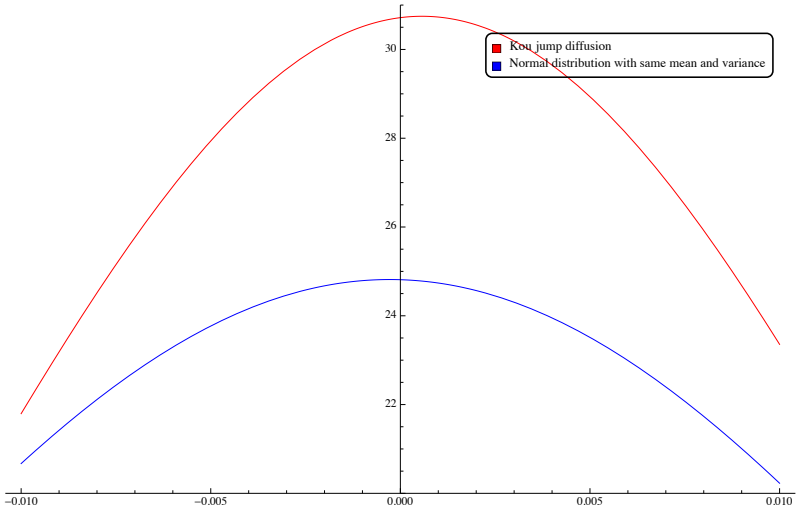

Figure 3.6: Peak

and heavy tails, it can be used to model both the overreaction (attributed to the heavy tails) and under-reaction (attributed to the high peak) to outside news. Therefore, the

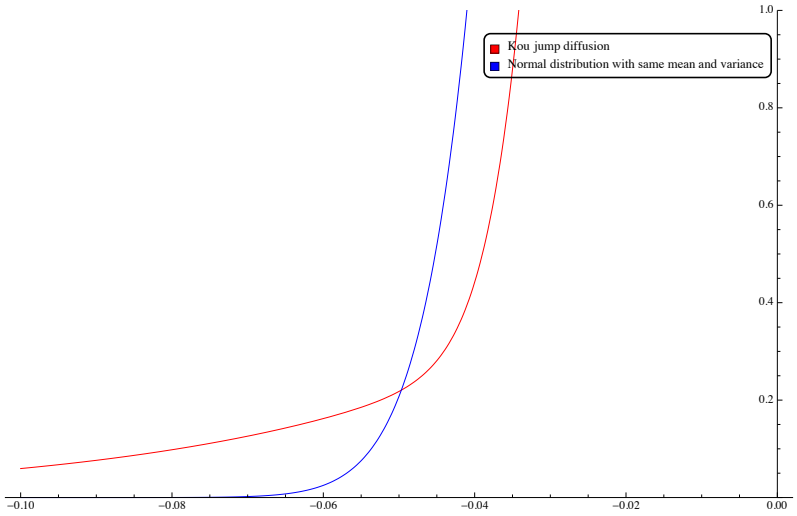

Figure 3.7: Left side tail

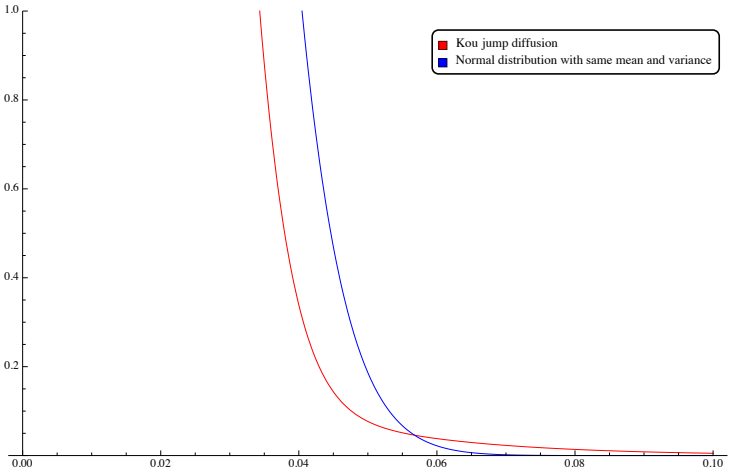

Figure 3.8: Right side tail

double exponential jump-diffusion model can be interpreted as an attempt to build a simple model, within the traditional random walk and efficient market framework to incorporate investors' sentiment. Besides, it also allows for a simple pricing of more complicated and exotic options like path-dependent options as Kou states in [33]. This simplicity makes this model quite popular in the literature, representing an improvement in comparison to the BSM model, while it maintains its analytical tractability.

3.2.4 Option pricing

In order to obtain prices for vanilla options and other financial securities, [33] uses semi-closed formulas which are also used in [42]. However, we are presenting a different approach, based on the framework of Bakshi, Madan, Green and Heston. On page 218 of [3], these authors suggest pricing European vanilla options based on the characteristic function of the random process we consider. This way, the reader has an alternative way for obtaining vanilla option prices which only requires the numerical valuation of integrals. For the Kou's model case, using (2.2), we can compute the characteristic function of a Lévy process L_t, denoted by ϕ_{L_t}, with a double exponential jump size distribution (3.6) in the following way (for more details, see [44]):

$$\phi_{L_t}(u) := \mathbb{E}[e^{iuL_t}] = e^{t\left[iu\mu - \frac{\sigma^2 u^2}{2} + iu\lambda\left(\frac{p}{\eta_1 - iu} - \frac{(1-p)}{\eta_2 + iu}\right)\right]}, \tag{3.11}$$

where $\mu = \left(r - q - \frac{\sigma^2}{2}\right)$.

Using equation (3.11), we can get the characteristic function of the random variable $\ln(S_t)$ which we denote by $\phi_{ln(S_t)}$.

The prices of European-style options are then computed as:

$$Call_t^{Eur} = S_t e^{-q(T-t)}\Pi_1 - Ke^{-r(T-t)}\Pi_2, \tag{3.12}$$

$$Put_t^{Eur} = Ke^{-r(T-t)}(1 - \Pi_2) - S_t e^{-q(T-t)}(1 - \Pi_1), \tag{3.13}$$

where

$$\Pi_1 = \frac{1}{2} + \frac{1}{\pi}\int_0^{+\infty} Re\Big(\frac{e^{-iuln(K)}\phi_{ln(S_T)}(u-i)}{iu\phi_{ln(S_T)}(-i)}\Big)du, \tag{3.14}$$

$$\Pi_2 = \frac{1}{2} + \frac{1}{\pi}\int_0^{+\infty} Re\Big(\frac{e^{-iuln(K)}\phi_{ln(S_T)}(u)}{iu}\Big)du. \tag{3.15}$$

The integrals above can be calculated using, for example, a Gauss-Legendre quadrature, as done in [21].

The same way we did for Merton, we plot the price of vanilla options, for each spot price, in a range from 80 to 120, with strike equal to 100, using the formulas above (see Figures 3.9 and 3.10 below). The same conclusions are reached as for Merton's model:

an increase on the jump intensity leads to an option price increase and for any case, an option price under Kou's model is more expensive than one priced with BSM model.

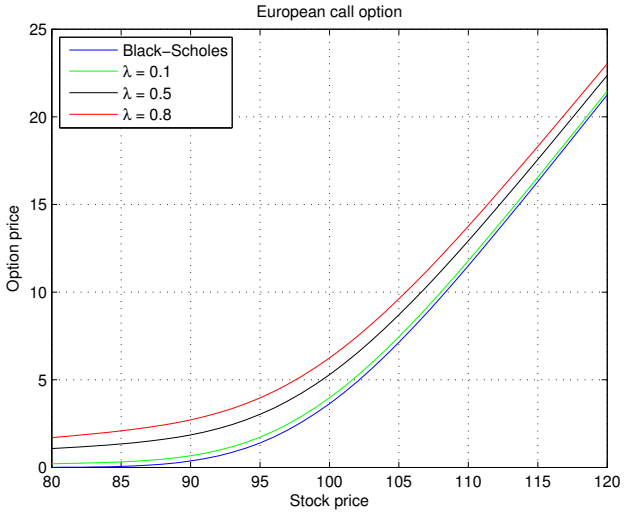

Figure 3.9: Call option prices with
$r = 0.05, \sigma = 0.15, T = 0.25, K = 100,$
$q = 0, \lambda = 0.10, \eta_1 = 3.0465, \eta_2 = 3.0775, p = 0.3445.$

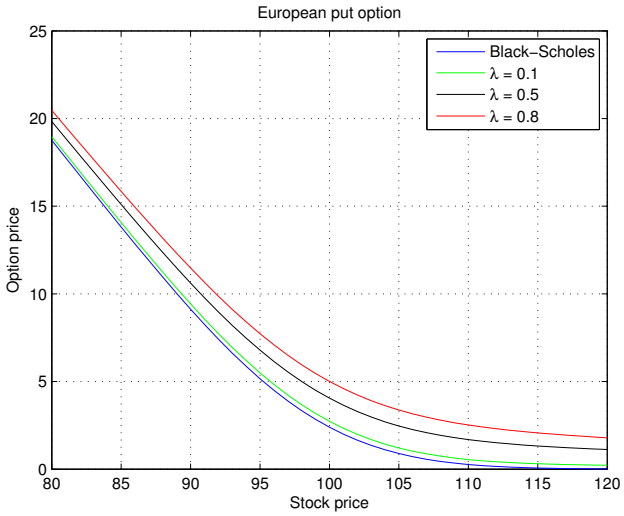

Figure 3.10: Put option prices with
$r = 0.05, \sigma = 0.15, T = 0.25, K = 100,$
$q = 0, \lambda = 0.10, \eta_1 = 3.0465, \eta_2 = 3.0775, p = 0.3445.$

4 Incomplete markets

As mentioned in Chapter 2, the assumption that the trading of assets takes place without arbitrage opportunities leads to the existence of a martingale measure $\mathbb{Q}$ and the discounted process, $(\tilde{S}_t)_{t\geq 0}$, is a martingale in the probability space $(\Omega, \mathcal{F}, \mathbb{Q})$.

Now, a fundamental question naturally arises: is this measure unique? The answer is: No. In the Lévy models context, there might be several measures that turn the discounted process $(\tilde{S}_t)_{t\geq 0}$ a martingale. As a consequence of the Second Fundamental Theorem of Asset Pricing, the market is not complete. This means that it is not possible to construct a self-financing trading portfolio that eliminates the risk of an existing position. In other words, if we acquire a certain financial product, such as a vanilla option, it will not be possible to trade other assets at inception of its acquisition, in such a way we guarantee a perfect hedging. For a better understading of these concepts, we suggest the reading of [9].

We are now giving some insight on why are markets incomplete when assets are driven by Lévy processes, based on [45]. The goal is not at all to present an exhausting prove of this claim, but to provide some results which give an idea of its veracity. Accomplished this goal, we then suggest a measure from among the infinite number of possible choices.

4.1 The Lévy measure

A convenient tool for analyzing the jumps of a Lévy process is the random measure of jumps of the process. Consider a set $A \in \mathcal{B}(\mathbb{R} \setminus \{0\})$, where $\mathcal{B}(\mathbb{R} \setminus \{0\})$ is the Borel σ-algebra on $\mathbb{R} \setminus \{0\}$, such that $0 \notin \bar{A}$. Letting $0 \leq t \leq T$, the random measure of the jumps of the process L is defined by:

$$\forall \omega \in \Omega,\ \mu^L(\omega; t, A) := \#\{0 \leq s \leq t : \Delta L_s(\omega) \in A\} = \sum_{s \leq t} \mathbb{1}_A(\Delta L_s(\omega)). \tag{4.1}$$

Hence, the measure $\mu^L(\omega; t, A)$ counts the jumps of the process L of size A up to time t. It is possible to prove that μ^L is a Poisson random measure and that its intensity, defined by

$$\nu(A) = \mathbb{E}[\mu^L(1, A)] = \mathbb{E}[\sum_{s \leq 1}(\Delta L_s(\omega))] \tag{4.2}$$

is a σ-finite measure on $\mathbb{R} \setminus \{0\}$, In (4.2), the first argument ω was suppressed for simplification purposes.

$\mu^L(1, A)$ measures the number of jumps of size A until $t = 1$. Thus, the indicator function in (4.1) equals 1, when the jump is of size A and 0 otherwise. This is how we obtain the last equality in (4.2).

4.2 Incompleteness evidence

Let $\mathbb{Q}$ be a probability measure on $\mathcal{F}$, equivalent to $\mathbb{P}$. Its Radon-Nikodym derivative $\frac{d\mathbb{Q}}{d\mathbb{P}}$ is a non-negative integrable random variable, so, using tower's law, it is not difficult to see that the process $Z_t = \mathbb{E}^{\mathbb{P}}\left[\frac{d\mathbb{Q}}{d\mathbb{P}}|\mathcal{F}_t\right], \forall\, 0 \leq t \leq T$, is a martingale under the probability $\mathbb{P}$ and its expected value under this measure equals 1.

To formalize the result in last paragraph, we present the following theorem, found in [45].

Theorem 4.1. *Given two equivalent measures $\mathbb{P}$ and $\mathbb{Q}$, there exists a unique, positive, $\mathbb{P}$-martingale $Z = (Z_t)_{0 \leq t \leq T}$ such that $Z_t = \mathbb{E}^{\mathbb{P}}\left[\frac{d\mathbb{Q}}{d\mathbb{P}}|\mathcal{F}_t\right], \forall\, 0 \leq t \leq T$. Z is called the density process of $\mathbb{Q}$ with respect to $\mathbb{P}$.*

Let $L = (L_t)_{0 \leq t \leq T}$ be a Lévy process with triplet (a, b, ν) under $\mathbb{P}$, with finite first moment. Then, L has the canonical decomposition:

$$L_t = bt + \sqrt{a}W_t + \int_0^t \int_{\mathbb{R}} x(\mu^L - \nu^L)(ds, dx).$$

As referred in [47], in the last stochastic integral of the previous expression, ν^L represents the compensator of the Poisson random measure μ^L and it is a product measure of the Lévy with the Lebesgue measures, letting us therefore write $\nu^L(ds, dx)$ as $\nu^L(dx)ds$.

This author also proves that if $\mathbb{P} \sim \mathbb{Q}$ with density process Z, there exists a deterministic process β and a measurable non-negative deterministic process Y, satisfying

$$\int_0^t \int_{\mathbb{R}} |x(Y(s, x) - 1)|\nu(dx)ds < \infty$$

and

$$\int_0^t (a\beta_s^2)ds < \infty,$$

$\mathbb{Q}$-a.s. for $0 \leq t \leq T$.

Under the new measure, $\tilde{W}_t = W_t - \int_0^t (\sqrt{a}\beta_s)ds$ is a Brownian motion, $\tilde{\nu}^L(ds, dx) = Y(s, x)\nu^L(ds, dx)$ is the $\mathbb{Q}$-compensator of μ^L and L has the following canonical decomposition under $\mathbb{Q}$:

$$L_t = \tilde{b}t + \sqrt{a}\tilde{W}_t + \int_0^t \int_{\mathbb{R}} x(\mu^L - \nu^L)(ds, dx),$$

where $\tilde{b}t = bt + \int_0^t a\beta_s ds + \int_0^t \int_{\mathbb{R}} x(Y(s, x) - 1)\nu^L(ds, dx)$.

For what we purposed to show, we can suppose, without loss of generality, that β and Y previously mentioned are deterministic and independent of time. Then, in this case:

$$\tilde{b} = b + a\beta + \int_{\mathbb{R}} x(Y - 1)\nu^L(dx).$$

Under the risk neutral measure $\mathbb{Q}$, the asset price has mean rate of return $\mu = r - q$ and the discounted process $(\tilde{S}_t)_{0 \leq t \leq T}$ is a martingale under $\mathbb{Q}$. Therefore, according to

[47], the drift term $\tilde{b}$ takes the form

$$\tilde{b} = r - q - \frac{\tilde{a}}{2} - \int_{\mathbb{R}} (e^x - 1 - x)\tilde{\nu}^L(dx). \tag{4.3}$$

Using this equation, we can now show why it is not possible to establish the existence of a unique martingale measure. Assume that the price process of a financial asset is modeled as an exponential Lévy process under both the real and the risk-neutral measure. Suppose that these measures, denoted $\mathbb{P}$ and $\mathbb{Q}$, are equivalent and consider the triplets of the Lévy processes under $\mathbb{P}$ and $\mathbb{Q}$ respectively are (a, b, ν) and $(\tilde{a}, \tilde{b}, \tilde{\nu})$. Now, as seen earlier, these triplets are related via $\tilde{a} = a, \tilde{\nu} = Y\nu$ and

$$\tilde{b} = b + a\beta + \int_{\mathbb{R}} x(Y - 1)\nu^L(dx), \tag{4.4}$$

where (β, Y) is the tuple of functions related to the density process. On the other hand, equating (4.3) and (4.4) and using $\tilde{a} = a, \tilde{\nu}^L = Y\nu$, we have that

$$\begin{aligned} 0 &= b + a\beta + \int_{\mathbb{R}} x(Y - 1)\nu^L(dx) - r + q + \frac{\tilde{a}}{2} + \int_{\mathbb{R}} (e^x - 1 - x)\tilde{\nu}(dx) \\ \iff 0 &= b - r + q + a\Big(\beta + \frac{1}{2}\Big) + \int_{\mathbb{R}} \Big((e^x - 1)Y - x\Big)\nu^L(dx). \end{aligned} \tag{4.5}$$

Therefore, we have one equation, but two unknown parameters, β and Y, stemming from the change of measure. Every solution tuple (β, Y) of equation (4.5) corresponds to a different equivalent martingale measure, which explains why the market is not complete.

Now, we are left to choose a martingale measure in order to proceed with pricing.

4.3 The Esscher transform

One of the most referred measures in the literature, as for instance in [25], is the called *Esscher transform*. Gerber and Shiu in [25] were the first to show that it can be used in option pricing and provide an interpretation of this approach in terms of maximal expected utility.

Definition 4.2. *The Esscher transform is any change of* $\mathbb{P}$ *to a locally equivalent*

measure $\mathbb{Q}$ *with a density process* $Z_t = \frac{d\mathbb{Q}}{d\mathbb{P}}\Big|_{\mathcal{F}_t}$ *of the form*

$$Z_t = \frac{e^{\theta L_t}}{mgf(\theta)^t},$$

where $\theta \in \mathbb{R}$*, and* $mgf(\theta)$ *denotes the moment generating function of* L_1*.*

In stock price modeling, the Esscher transform is an useful tool for finding an equivalent probability measure under which discounted stock prices are martingales. According to [47], it is necessary and sufficient for its existence that the following condition is verified, for at least one value θ:

$$mgf(\theta) = mgf(\theta + 1). \tag{4.6}$$

If this equation has a solution, we guarantee that $\mathbb{Q}$ defines an EMM and L_t becomes a Lévy process under $\mathbb{Q}$. This is indeed the fact in particular for the Merton and Kou's models, because the logarithm of the moment generating function, ln[mgf(u)], is strictly convex for a non-degenerate distribution, as proved in the next theorem.

Theorem 4.3. *Let* μ *be a non-degenerate probability measure on* $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ *which possesses a moment generating function. Then, the logarithm of this moment generating function possesses a strictly positive second derivative on the interior of its range of existence. In particular, it is strictly convex on its range of existence.*

Proof. Consider an arbitrary value u from the interior of the range of existence of the moment generating function. The second derivative of the logarithm of the moment generating function, taken at u is:

$$\frac{d^2}{du^2}\Big(\ln(\text{mgf (u)})\Big) = \frac{\text{mgf"(u)mgf(u)} - [\text{mgf'(u)}]^2}{\text{mgf(u)}^2} = \frac{\int x^2 e^{ux}\mu(dx)}{\int e^{ux}\mu(dx)} - \left[\frac{\int x e^{ux}\mu(dx)}{\int e^{ux}\mu(dx)}\right]^2.$$

The last algebraic expression corresponds to the variance of the non-degenerate distribution $\frac{e^{ux}}{\int e^{ux}\mu(dx)}\mu(dx)$, so it is strictly positive.

In order to prove our claim, first remark that since the logarithm function is injective, finding a solution to equation (4.6) is equivalent to solve the following:

$$\ln(\text{mgf}(\theta + 1)) - \ln(\text{mgf}(\theta)) = \ln(1)$$

Applying the mean value theorem to the left-hand side function in the previous equation, we know there exists $\xi \in]\theta, \theta+1[$ such that:

$$\frac{d}{d\theta}\Big(\ln(\text{mgf}(\theta+1)) - \ln(\text{mgf}(\theta))\Big) = [\ln(\text{mgf})]''(\xi)$$

We already proved the second derivative of the last term is strictly positive, so ln(mgf) is then a function with positive first derivative. This in particular means the left hand side function has an inverse and consequently, (4.6) has a unique solution. ■

4.3.1 The relative entropy

The reason for the choice of an Esscher transform is due to the fact that, as shown in [19], for exponential Lévy models, it corresponds to the minimal entropy martingale measure, i.e., the equivalent martingale measure which minimizes the relative entropy and it has also the property of preserving the Lévy structure of the model. For more clarification on this topic, we suggest the reading of [47] and also of [25]. Borrowing notation from [24], we define the relative entropy as follows:

Definition 4.4. *Let $\mathcal{G}$ be a sub-sigma field of $\mathcal{F}$. The relative entropy of $\mathbb{Q}$ with respect to $\mathbb{P}$ is:*

$$\mathcal{H}_{\mathcal{G}}(\mathbb{Q}|\mathbb{P}) := \begin{cases} \int_\Omega ln\big(\frac{d\mathbb{Q}}{d\mathbb{P}}\Big|_{\mathcal{G}}\big) d\mathbb{Q}, & if \quad \mathbb{Q} \ll \mathbb{P} \text{ on } \mathcal{G} \\ +\infty, & if \quad \mathbb{Q} \not\ll \mathbb{P} \text{ on } \mathcal{G} \end{cases}$$

where $\frac{d\mathbb{Q}}{d\mathbb{P}}\Big|_{\mathcal{G}}$ stands for the Radon-Nikodym derivative of $\mathbb{Q}|_{\mathcal{G}}$ with respect to $\mathbb{P}|_{\mathcal{G}}$ and $\mathbb{Q} \ll \mathbb{P}$ means $\mathbb{Q}$ is absolutely continuous with respect to $\mathbb{P}$.

Intuitively speaking, if one thinks of the measure $\mathbb{P}$ as encapsulating some information about how the market behaves, then pricing options by Esscher transform amounts to choosing the equivalent martingale measure which is closest to $\mathbb{P}$ in terms of its information content.

We have now sufficient "ammunition" to build a method for pricing financial derivatives and that is exactly what it will be done in the next Chapter.

5 The PIDE

In this section, we briefly explain a partial integro-differential equation (PIDE) whose solution will be approximated numerically in the next sections. The PIDE we are going to introduce captures the dependence between the time and space variables, having consequently significant importance for our purposes.

Let $\{w(t, S_t)\}_{t\geq 0}$ be the value of a European contingent claim on the asset S and $H(S_T) := w(T, S_T)$ its value at the maturity.

Considering that we are working in a "free-arbitrage scenario", the discounted value process $\{e^{-rt}w(t, S_t)\}_{t\geq 0}$ is a martingale under the risk-neutral measure $\mathbb{Q}$ and thus, $\forall\, t \geq 0$,

$$e^{-rt}w(t, S_t) = \mathbb{E}^{\mathbb{Q}}[e^{-rT}w(T, S_T)|\mathcal{F}_t] \iff e^{r(T-t)}w(t, S_t) = \mathbb{E}^{\mathbb{Q}}[H(S_T)|\mathcal{F}_t]. \tag{5.1}$$

Since $S_T = S_0e^{L_T}$ and $S_t = S_0e^{L_t}$, dividing S_T by the latter and then using the process' stationarity, we have:

$$\frac{S_T}{S_t} = \frac{S_0e^{L_T}}{S_0e^{L_t}} \iff S_T = S_te^{L_{T-t}}.$$

Therefore, equation (5.1) becomes equivalent to:

$$e^{r(T-t)}w(t, S_t) = \mathbb{E}^{\mathbb{Q}}[H(S_te^{L_{T-t}})|\mathcal{F}_t].$$

Generalizing, we finally get:

$$w(t, s) = e^{-r(T-t)}\mathbb{E}^{\mathbb{Q}}[H(se^{L_{T-t}})|\mathcal{F}_t].$$

Let us apply the change of variables x=ln(s) and write $v(t, x) := w(t, e^x)$. Since

$s = e^x$, this modification leads to the equality

$$v(t,x) = e^{-r(T-t)}\mathbb{E}^{\mathbb{Q}}[H(e^{x+L_{T-t}})|\mathcal{F}_t]$$

Theorem 5.1. *Suppose that the function $v(t,x)$ that has just been defined belongs to the set $C^{1,2}([0,T[\times\mathbb{R}) \cap C^0([0,T[\times\mathbb{R})$ and assume $\int_{\mathbb{R}} |x| d\nu(x) < \infty$.*

Then, v(t,x) satisfies the following PIDE:

$$v_t + \frac{a}{2}v_{xx} + bv_x - rv + \int_{\mathbb{R}} (v(t,x+y) - v(t,x) - v_x(t,x)y)d\nu(y) = 0, \tag{5.2}$$

$\forall\,(t,x) \in [0,T[\,\times\mathbb{R}$, *with final condition $v(T,x) = H(e^x), \forall x \in \mathbb{R}$.*

v_x and v_{xx} respectively denote the first and second order derivatives of v in respect to x and v_t the first derivative of v in respect to t.

Proof. See Proposition 1.12 of [47]. ■

In the equation above and from now on, every time we write v, we mean $v(t,x)$.

Recalling that for the process (2.3), we have

$$a = \sigma^2, \quad b = \Big(r - q - \lambda\widetilde{\zeta} - \frac{\sigma^2}{2}\Big), \quad \nu = \lambda P(y),$$

the PIDE (5.2) corresponds to:

$$v_t + \frac{1}{2}\sigma^2 v_{xx} + \Big(r - q - \lambda\widetilde{\zeta} - \frac{\sigma^2}{2}\Big)v_x - rv + \int_{\mathbb{R}} [v(t,x+y) - v(t,x) - v_x(t,x)y]\lambda dP(y) = 0,$$

with terminal condition $v(T,x) = H(e^x)$, $\forall x \in \mathbb{R}$.

Using the definition of $\widetilde{\zeta}$, noticing that $v(t,x)$ and $v_x(t,x)$ do not depend on y and applying the integral's linearity, the last equation becomes equal to:

$$v_t + \frac{1}{2}\sigma^2 v_{xx} + \Big(r - q - \lambda(\zeta - \mathbb{E}[Y]) - \frac{\sigma^2}{2}\Big)v_x - rv + \lambda\Big[\int_{\mathbb{R}} v(t,x+y)\,dP(y) - v(t,x)\int_{\mathbb{R}} dP(y) - v_x(t,x)\int_{\mathbb{R}} y\,dP(y)\Big] = 0.$$

$\int_{\mathbb{R}} dP(y) = 1$, because it is the integral over $\mathbb{R}$ of a probability's density function and $\int_{\mathbb{R}} y\,dP(y) = \mathbb{E}[Y]$. Thus, we get:

$$v_t + \frac{1}{2}\sigma^2 v_{xx} + \Big(r - q - \lambda\zeta + \lambda\mathbb{E}[Y] - \frac{\sigma^2}{2}\Big)v_x - rv$$

$$+\lambda \int_{\mathbb{R}} v(t, x+y)\, dP(y) - \lambda v - \lambda v_x \times \mathbb{E}[Y] = 0$$

$$\iff v_t + \frac{1}{2}\sigma^2 v_{xx} + \Big(r - q - \lambda\zeta - \frac{\sigma^2}{2}\Big)v_x - (r+\lambda)v + \lambda \int_{\mathbb{R}} v(t, x+y)\, dP(y) = 0 \quad (5.3)$$

Therefore, we conclude that v is the solution of the PIDE with constant coefficients (5.3) and terminal condition $v(T, x) = H(e^x)$, $\forall x \in \mathbb{R}$, and $dP(y) = p(y)dy$.

Finally, the time variable is transformed to obtain a forward problem in time. If we set $u(\tau, \cdot) := v(T - \tau, \cdot)$ and thereafter multiply both sides of equation (5.3) by (-1), it becomes:

$$u_\tau = \frac{1}{2}\sigma^2 u_{xx} + \Big(r - q - \frac{1}{2}\sigma^2 - \lambda\zeta\Big)u_x - (r+\lambda)u + \lambda \int_{\mathbb{R}} u(\tau, x+y)\, dP(y),$$

i.e.,

$$u_\tau = \frac{1}{2}\sigma^2 u_{xx} + \Big(r - q - \frac{1}{2}\sigma^2 - \lambda\zeta\Big)u_x - (r+\lambda)u + \lambda \int_{\mathbb{R}} u(\tau, z)p(z-x)\, dz, \qquad (5.4)$$

$\forall\, (\tau, z) \in [0, T[\times \mathbb{R}^+$, with initial condition $u(0, x) = H(e^x)$, $\forall x \in \mathbb{R}$.

This is true, because $\frac{\partial v}{\partial t} = \frac{\partial u}{\partial \tau}\frac{\partial \tau}{\partial t} = -\frac{\partial u}{\partial \tau}$; $\frac{\partial v}{\partial x} = \frac{\partial u}{\partial x}$ and $\frac{\partial^2 v}{\partial x^2} = \frac{\partial^2 u}{\partial x^2}$.

For the Merton model,

$$\begin{aligned}
\zeta &= \int_{\mathbb{R}} (e^x - 1) dF_{Merton}(x) \\
&= \frac{1}{\sqrt{2\pi\sigma_J^2}} \int_{\mathbb{R}} e^{x - \frac{(x-\mu_J)^2}{2\sigma_J^2}}\, dx - 1 \\
&= \frac{1}{\sqrt{2\pi\sigma_J^2}} \int_{\mathbb{R}} e^{\frac{2x\sigma_J^2 - x^2 + 2x\mu_J - \mu_J^2}{2\sigma_J^2}}\, dx - 1 \\
&= \frac{1}{\sqrt{2\pi\sigma_J^2}} \int_{\mathbb{R}} e^{-\frac{x^2 - 2x(\mu_J + \sigma_J^2) + \mu_J^2}{2\sigma_J^2}}\, dx - 1 \\
&= \frac{1}{\sqrt{2\pi\sigma_J^2}} \int_{\mathbb{R}} e^{-\frac{(x-(\mu_J+\sigma_J^2))^2 - (\mu_J+\sigma_J^2)^2 + \mu_J^2}{2\sigma_J^2}}\, dx - 1 \\
&= \frac{1}{\sqrt{2\pi\sigma_J^2}} \int_{\mathbb{R}} e^{-\frac{(x-(\mu_J+\sigma_J^2))^2 - 2\mu_J\sigma_J^2 - (\sigma_J^2)^2}{2\sigma_J^2}}\, dx - 1
\end{aligned}$$

$$= e^{-\frac{-2\mu_J\sigma_J^2-(\sigma_J^2)^2}{2\sigma_J^2}} \frac{1}{\sqrt{2\pi\sigma_J^2}} \int_{\mathbb{R}} e^{-\frac{(x-(\mu_J+\sigma_J^2))^2}{2\sigma_J^2}} dx - 1$$
$$= e^{\mu_J+\frac{\sigma_J^2}{2}} - 1.$$

In Kou's model case:

$$\zeta = \int_{\mathbb{R}} (e^x - 1) dF_{Kou}(x)$$
$$= (1-p)\eta_2 \int_{-\infty}^{0} e^{x(1+\eta_2)} dx + p\eta_1 \int_{0}^{+\infty} e^{x(1-\eta_1)} dx - 1$$
$$= \frac{(1-p)\eta_2}{1+\eta_2} + \frac{p\eta_1}{\eta_1 - 1} - 1.$$

6 Localization error estimates

In order to numerically solve the PIDE (5.4), we first localize the variables and the integral term to bounded domains, i.e., we have to restrict the domain $\mathbb{R}$ of the spatial variable to a bounded interval. In what follows, we will mainly focus on pricing a European-style put, thus particularizing the results for this type of option. Nevertheless, using the put-call parity, it is possible to obtain the price of a European-style call option and most of the results below can be extended to more complex payoffs, with the right adjustments. Most of the results presented from now on are based in the articles of Kwon and Lee [37] and Almendral and Osterlee [1].

It is well-known that the asymptotic behavior of the price of a European put option is described by

$$\lim_{x\to-\infty}\left[u(\tau,x)-(Ke^{-r\tau}-e^{x})\right]=0,\quad \text{and} \tag{6.1}$$

$$\lim_{x\to+\infty}[u(\tau,x)]=0. \tag{6.2}$$

Let us consider $\bar{u}(\tau,x)$ as the solution of the following localized problem with the asymptotic behavior given by the previous equations:

$$\bar{u}_\tau(\tau,x)=\mathcal{L}\bar{u}(\tau,x),\quad (\tau,x)\in\,]0,T]\times\,]-X,X[\tag{6.3}$$

$$\bar{u}(\tau,x)=g(\tau,x),\quad x\in\mathbb{R}\setminus\,]-X,X[\tag{6.4}$$

$$\bar{u}(0,x)=h(x),\quad x\in\,]-X,X[\tag{6.5}$$

where $X>0$, $g(\tau,x):=(Ke^{-r\tau}-e^{x})^{+}$, $h:=H(e^{x})$ and $\mathcal{L}$ is the integro-differential linear operator:

$$\mathcal{L}\bar{u}(\tau,x) := \frac{1}{2}\sigma^2\bar{u}_{xx} + (r - q - \frac{1}{2}\sigma^2 - \lambda\zeta)\bar{u}_x - (r+\lambda)\bar{u} + \lambda \int_{\mathbb{R}} \bar{u}(\tau,x)p(z-x)\,dz. \quad (6.6)$$

We describe the relationship between the solution u of (5.4) and the solution $\bar{u}(\tau,x)$ in the following theorem, as proved in [15].

Theorem 6.1. *Suppose that h is bounded ($\|h\|_\infty < \infty$) and there exists $\alpha > 0$ such that*

$$\int_{\|x\|>1} e^{\alpha\|x\|} d\nu(x) < \infty.$$

Then

$$|u(\tau,x) - \bar{u}(\tau,x)| \leq C\|h\|_\infty e^{-\alpha(X-|x|)}, \quad \forall x \in]-X, X[, \quad (6.7)$$

where C is a constant not depending on x.

In the case of a put option, $\|h\|_\infty < \infty$, because, $\|h\|_\infty$ is at most equal to K and the second hypothesis is also verified due to the martingale property (see [15] for more details on this). Thus, the last theorem applies to the case of a put option. From (6.7), we can also see that the localization error is more pronounced near the boundary.

7 Numerical implementation for European options

In this chapter, we concentrate on a straightforward numerical scheme for solving the initial valued PIDE:

$$u_\tau(\tau, x) = \mathcal{L}u(\tau, x), \quad (\tau, x) \in]0, T] \times]-X, X[\tag{7.1}$$

$$u(\tau, x) = g(\tau, x), \quad x \in \mathbb{R} \setminus]-X, X[\tag{7.2}$$

$$u(0, x) = h(x), \quad x \in]-X, X[\tag{7.3}$$

where $X > 0$, $\mathcal{L}$ is the integro-differential linear operator defined in equation (6.6), $h(x) := H(e^x)$ and $g(\tau, x) := (Ke^{-r\tau} - e^x)$.

Let us split the operator $\mathcal{L}$ into three parts as follows:

$$\mathcal{L}u(\tau, x) = \mathcal{D}u(\tau, x) + \mathcal{I}u(\tau, x) - (r + \lambda)u(\tau, x),$$

where $\mathcal{D}$ is the differential operator of $\mathcal{L}$ and $\mathcal{I}$ is its integral operator such that:

$$\mathcal{D}u(\tau, x) := \frac{\sigma^2}{2}u_{xx}(\tau, x) + (r - q - \frac{1}{2}\sigma^2 - \lambda\zeta)u_x(\tau, x), \quad \text{and} \tag{7.4}$$

$$\mathcal{I}u(\tau, x) := \lambda \int_{\mathbb{R}} u(\tau, z)p(z - x)\, dz. \tag{7.5}$$

In order to numerically approximate the integral term $\mathcal{I}u$ in (7.5), we split it into two regions: $\Omega := [-X, X]$ and $\mathbb{R}\setminus\Omega$. On the complement of the computational domain, we use the boundary conditions

$$u(\tau, x) \approx Ke^{-r\tau} - e^x \quad \text{on }]-\infty, -X[, \quad \text{and}$$

$$u(\tau, x) \approx 0 \quad \text{on }]X, +\infty[.$$

Over $\mathbb{R} \setminus \Omega$, (7.5) is then given by:

$$\epsilon(\tau, x, X) = \int_{-\infty}^{-X} (Ke^{-r\tau} - e^z) p(z - x)\, dz \tag{7.6}$$

In the case of the Merton model, the jump density function p is defined in (3.1) and thus:

$$\begin{aligned}
\epsilon(\tau, x, X) &= \int_{-\infty}^{-X} (Ke^{-r\tau} - e^z) \frac{1}{\sqrt{2\pi\sigma_J^2}} e^{-\frac{(z-x-\mu_J)^2}{2\sigma_J^2}}\, dz \\
&= \underbrace{\int_{-\infty}^{-X} Ke^{-r\tau} \frac{1}{\sqrt{2\pi\sigma_J^2}} e^{-\frac{(z-x-\mu_J)^2}{2\sigma_J^2}}\, dz}_{:=A} - \underbrace{\int_{-\infty}^{-X} \frac{1}{\sqrt{2\pi\sigma_J^2}} e^{-\frac{-2z\sigma_J^2 + (z-x-\mu_J)^2}{2\sigma_J^2}}\, dz}_{:=B},
\end{aligned}$$

where

- $A = Ke^{-r\tau}\Phi\left[-\frac{X + x + \mu_J}{\sigma_J} \right]$ and

- $$\begin{aligned}
B &= \int_{-\infty}^{-X} \frac{1}{\sqrt{2\pi\sigma_J^2}} e^{-\frac{z^2 - 2z(\sigma_J^2 + x + \mu_J) + (x+\mu_J)^2}{2\sigma_J^2}}\, dz \\
&= \int_{-\infty}^{-X} \frac{1}{\sqrt{2\pi\sigma_J^2}} e^{-\frac{[z - (\sigma_J^2 + x + \mu_J)]^2 - (\sigma_J^2)^2 - 2\sigma_J^2(x+\mu_J)}{2\sigma_J^2}}\, dz \\
&= e^{\frac{\sigma_J^2}{2} + \mu_J + x} \int_{-\infty}^{-X} \frac{1}{\sqrt{2\pi\sigma_J^2}} e^{-\frac{[z - (\sigma_J^2 + x + \mu_J)]^2}{2\sigma_J^2}}\, dz \\
&= e^{\frac{\sigma_J^2}{2} + \mu_J + x} \Phi\left[-\frac{X + \sigma_J^2 + \mu_J + x}{\sigma_J} \right].
\end{aligned}$$

Subtracting B to A:

$$\epsilon(\tau, x, X) = Ke^{-r\tau}\Phi\left[-\frac{X + x + \mu_J}{\sigma_J} \right] - e^{\frac{\sigma_J^2}{2} + \mu_J + x} \Phi\left[-\frac{X + \sigma_J^2 + \mu_J + x}{\sigma_J} \right]. \tag{7.7}$$

For the Kou model, its density p is defined in (3.6) and consequently:

$$\epsilon(\tau, x, X) = \int_{-\infty}^{-X} (Ke^{-r\tau} - e^z)[p\eta_1 e^{-\eta_1(z-x)} \mathbb{1}_{\{(z-x)\geq 0\}} + (1-p)\eta_2 e^{\eta_2(z-x)} \mathbb{1}_{\{(z-x)<0\}}]\, dz$$

$$= \int_{-\infty}^{-X} Ke^{-r\tau}(1-p)\eta_2 e^{\eta_2(z-x)} - \int_{-\infty}^{-X} (1-p)\eta_2 e^{z(\eta_2+1)-\eta_2 x}\, dz$$

$$= K(1-p)e^{-r\tau-\eta_2(X+x)} - (1-p)\frac{\eta_2}{\eta_2+1}e^{-\eta_2 x - X(1+\eta_2)}. \tag{7.8}$$

We remark that the first indicator function is equal to zero, since z lies in $]-\infty, -X[$ and $x \in]-X, X[$. This means that $(z-x)$ is always non-positive.

7.1 Time and Space grid

In this section, we present a uniform grid of the time and space domains with the goal of then applying a finite differences discretization to the approximated solution of our problem (7.1)-(7.3).

Consider a uniform mesh in space and another in time, i.e., let us take uniform mesh points on the truncated domain $[0,T]\times[-X,X]$. Choosing positive integers N and M, larger than 1, we define by k the ratio $\frac{T}{M-1}$ and by h the quotient $\frac{2X}{N-1}$, creating the points $x_i = -X + (i-1)h$, $\forall i \in \{1,...,N\}$ and $\tau_m = (m-1)k, \forall m \in \{1,...,M\}$. After considering these partitions, we define an approximation of the value of u in the point τ_m and x_i as u_i^m, that is, $u_i^m \approx u(\tau_m, x_i)$ and $p_{ij} := p(x_j - x_i)$.

For the integral approximation of (7.5), we use the composite trapezoidal rule on $[-X,X]$, which leads to:

$$\int_{\mathbb{R}} u(\tau_m, z)p(z-x_i)\, dz \approx \frac{h}{2}[u_1^m p_{i,1} + u_N^m p_{i,N} + 2\sum_{j=2}^{N-1} u_j^m p_{i,j}] + \epsilon(\tau_m, x_i, X), \tag{7.9}$$

where $i \in \{2,...,N-1\}$.

7.2 Approximation by finite differences

For the time and space variables, we approximate them by the following finite differences:

$$u_\tau(\tau_m, x_i) \approx \frac{u_i^{m+1} - u_i^{m-1}}{2k}, \tag{7.10}$$

$$u_{xx}(\tau_m, x_i) \approx \frac{1}{2}\left(\frac{u_{i+1}^{m+1} - 2u_i^{m+1} + u_{i-1}^{m+1}}{h^2} + \frac{u_{i+1}^{m-1} - 2u_i^{m-1} + u_{i-1}^{m-1}}{h^2}\right), \tag{7.11}$$

$$\text{and } u_x(\tau_m, x_i) \approx \frac{1}{2}\left(\frac{u_{i+1}^{m+1} - u_{i-1}^{m+1}}{2h} + \frac{u_{i+1}^{m-1} - u_{i-1}^{m-1}}{2h}\right), \tag{7.12}$$

$\forall i \in \{2, ..., N-1\}, m \in \{2, ..., M-1\}$.

As mentioned in [37], at a mesh point (τ_m, x_i) we approximate the first-order time derivative by the central difference with the second-order accuracy and the second-order spatial derivative by the average of the second-order central differences on the $(m-1)th$ and the $(m+1)th$ time levels. We also approximate the first-order spatial derivative by the average of the central differences $(m-1)th$ and the $(m+1)th$ time levels. Three time levels are involved in this approximation scheme, which is very similar to the Crank–Nicholson scheme. For an insight on how to get the mentioned approximations, the reading of section A.3 in Appendix is advised.

7.3 Numerical scheme

The operators $\mathcal{D}, \mathcal{I}$, and $\mathcal{L}$ at each mesh point (τ_m, x_i) can be approximated by discrete operators $\mathcal{D}_\Delta$, $\mathcal{I}_\Delta$ and $\mathcal{L}_\Delta$:

$$\mathcal{D}u_i^m \approx \mathcal{D}_\Delta\left(\frac{u_i^{m+1} + u_i^{m-1}}{2}\right), \quad \mathcal{I}u_i^m \approx \mathcal{I}_\Delta u_i^m, \quad \mathcal{L}u_i^m \approx \mathcal{L}_\Delta u_i^m,$$

where the finite difference operator D_Δ and the discrete integral operator I_Δ are defined by:

$$\mathcal{D}_\Delta u_i^m := \frac{\sigma^2}{2}\frac{u_{i+1}^m - 2u_i^m + u_{i-1}^m}{h^2} + \left(r - q - \lambda\zeta - \frac{\sigma^2}{2}\right)\frac{u_{i+1}^m - u_{i-1}^m}{2h}, \tag{7.13}$$

$$\mathcal{I}_\Delta u_i^m := \frac{\lambda h}{2}\left[u_1^m p_{i,1} + u_N^m p_{i,N} + 2\sum_{j=2}^{N-1} u_j^m p_{i,j}\right] + \lambda\epsilon(\tau_m, x_i, X). \tag{7.14}$$

The discrete operator $\mathcal{L}_\Delta$ is defined as:

$$\mathcal{L}_\Delta u_i^m := \begin{cases} \mathcal{D}_\Delta u_i^m + \mathcal{I}_\Delta u_i^m - (r+\lambda)u_i^m, & \text{if } m = 2 \\ \mathcal{D}_\Delta\left(\frac{u_i^{m+1} + u_i^{m-1}}{2}\right) + \mathcal{I}_\Delta u_i^m - (r+\lambda)u_i^m, & \text{if } m \geq 3 \end{cases}$$

As the reader might have noticed, the above discretization method of the integro-

differential operator $\mathcal{L}u$ involves three time levels. This is why the numerical method we are introducing is called as the "implicit method with three time levels". Besides, we need two initial values on the first and second time levels in order to obtain the solution in the following ones. The value u^1 on the first time level is provided by the payoff function, and the value u^2 on the second time level can be obtained by applying the explicit method to the integro-differential operator $\mathcal{L}u$.

Summarizing, as an approximate value of the solution u_i^m, we seek $\tilde{u}_i^m$ which is the solution of the following time-stepping problem:

Initial condition:

$\tilde{u}_i^1 = h(x_i), \quad \text{for } 1 \leq i \leq N.$

Boundary condition:

$\tilde{u}_i^m = g(\tau_m, x_i), \quad \text{for } 2 \leq m \leq M \text{ and } i \in \{1, N\}.$

For $m = 2$:

$$\frac{\tilde{u}_i^m - \tilde{u}_i^{m-1}}{k} = \mathcal{D}_\Delta \tilde{u}_i^{m-1} + \mathcal{I}_\Delta \tilde{u}_i^{m-1} - (r+\lambda)\tilde{u}_i^{m-1}, \quad \text{for } 2 \leq i \leq N-1. \tag{7.15}$$

For $2 < m \leq M$:

$$\frac{\tilde{u}_i^m - \tilde{u}_i^{m-2}}{2k} = \mathcal{D}_\Delta\Big(\frac{\tilde{u}_i^m + \tilde{u}_i^{m-2}}{2}\Big) + \mathcal{I}_\Delta \tilde{u}_i^{m-1} - (r+\lambda)\tilde{u}_i^{m-1}, \quad \text{for } 2 \leq i \leq N-1. \tag{7.16}$$

8 Consistency, stability and convergence

We now proceed with the analysis of consistency, stability, and convergence of the implicit method presented in the end of the last section.

8.1 Consistency

In simple words, we analyse the local (truncation) error of the described numerical method, defined as the error committed by one step of the method. That is, it is the difference between the result given by the method, assuming that no error was made in earlier steps, and the exact solution. The discretization of a PIDE should become exact as the mesh size tends to zero when considering a consistent numerical scheme, i.e., the truncation error should vanish.

Theorem 8.1. *Let $v \in C^\infty(]0,T] \times [-X,X])$ satisfy the initial and boundary conditions (7.2) and (7.3). For any $(\tau_m, x_i) \in]0,T] \times]-X,X[$, the following result is valid when considering sufficiently small h and k:*

For $m = 1$:

$$v_\tau(\tau_m, x_i) - \mathcal{L}v(\tau_m, x_i) - \left(\frac{v(\tau_{m+1}, x_i) - v(\tau_m, x_i)}{k} - \mathcal{L}_\Delta v(\tau_m, x_i) \right) = O(k + h^2). \quad (8.1)$$

For $m \geq 2$:

$$v_\tau(\tau_m, x_i) - \mathcal{L}v(\tau_m, x_i) - \left(\frac{v(\tau_{m+1}, x_i) - v(\tau_{m-1}, x_i)}{2k} - \mathcal{L}_\Delta v(\tau_m, x_i) \right) = O(k^2 + h^2). \quad (8.2)$$

Proof. See [37]. ■

8.2 Stability

In this section, we shall investigate the stability of the implicit method we have been considering by using the Von Neumann analysis (see [50] for more details). In a nutshell, the term stability designates that any numerical errors, introduced at some stage of the calculations, are propagated in a mild fashion – i.e. do not blow up in the subsequent steps of the method.

Define

$$\begin{aligned}\alpha_1 &= \frac{\sigma^2}{h^2} - \frac{1}{h}\Big(r - q - \frac{1}{2}\sigma^2 - \lambda\zeta\Big),\\ \alpha_2 &= \frac{2\sigma^2}{h^2},\\ \alpha_3 &= \frac{\sigma^2}{h^2} + \frac{1}{h}\Big(r - q - \frac{1}{2}\sigma^2 - \lambda\zeta\Big).\end{aligned}$$

The operator $\mathcal{D}_\Delta$ can then be rewritten as:

$$\begin{aligned}\mathcal{D}_\Delta u_i^m &= u_{i-1}^m\left[\frac{\sigma^2}{2h^2} - \frac{\Big(r - q - \lambda\zeta - \frac{\sigma^2}{2}\Big)}{2h}\right] - \frac{1}{2}u_i^m\Big(\frac{2\sigma^2}{h^2}\Big) + u_{i+1}^m\left[\frac{\sigma^2}{2h^2} + \frac{\Big(r - q - \lambda\zeta - \frac{\sigma^2}{2}\Big)}{2h}\right]\\ &= \frac{\alpha_1}{2}u_{i-1}^m - \frac{\alpha_2}{2}u_i^m + \frac{\alpha_3}{2}u_{i+1}^m.\end{aligned}$$

Let us focus our attention to (7.16), that is, consider $i \in \{2, ..., N-1\}$ and $m > 2$. We have:

$$\begin{aligned}\frac{\tilde{u}_i^m - \tilde{u}_i^{m-2}}{2k} &= \frac{1}{2}\frac{\alpha_1\tilde{u}_{i-1}^m - \alpha_2\tilde{u}_i^m + \alpha_3\tilde{u}_{i+1}^m + \alpha_1\tilde{u}_{i-1}^{m-2} - \alpha_2\tilde{u}_i^{m-2} + \alpha_3\tilde{u}_{i+1}^{m-2}}{2} - (r+\lambda)\tilde{u}_i^{m-1}\\ &\quad + \frac{\lambda h}{2}\left[\tilde{u}_1^{m-1}p_{i,1} + 2\sum_{j=2}^{N-1}\tilde{u}_j^{m-1}p_{i,j} + u_N^{m-1}p_{i,N}\right] + \lambda\epsilon(\tau_{m-1}, x_i, X)\end{aligned}$$

$$\iff$$

$$\begin{aligned}&-\frac{\alpha_1 k}{2}\tilde{u}_{i-1}^m + \Big(1 + \frac{\alpha_2 k}{2}\Big)\tilde{u}_i^m - \frac{\alpha_3 k}{2}\tilde{u}_{i+1}^m\\ &= \frac{\alpha_1 k}{2}\tilde{u}_{i-1}^{m-2} + \Big(1 - \frac{\alpha_2 k}{2}\Big)\tilde{u}_i^{m-2} + \frac{\alpha_3 k}{2}\tilde{u}_{i+1}^{m-2} - 2k(r+\lambda)\tilde{u}_i^{m-1}\\ &\quad + \lambda hk\left[\tilde{u}_1^{m-1}p_{i,1} + 2\sum_{j=2}^{N-1}\tilde{u}_j^{m-1}p_{i,j} + u_N^{m-1}p_{i,N}\right] + 2k\lambda\epsilon(\tau_{m-1}, x_i, X) \qquad (8.3)\end{aligned}$$

For stability analysis purposes, we neglect the last term in the right-hand side of the previous equation, because it does not depend on $\tilde{u}$. Replacing u_i^m by $g^m e^{i(i\theta)}$ and subsequently dividing the equation by $g^{m-2}e^{i(i\theta)}$, we obtain:

$$
\begin{aligned}
&- \alpha_1 \frac{k}{2} g^2 e^{-i\theta} + \left(1 + \alpha_2 \frac{k}{2}\right) g^2 - \alpha_3 \frac{k}{2} g^2 e^{i\theta} \\
&= \alpha_1 \frac{k}{2} e^{-i\theta} + \left(1 - \alpha_2 \frac{k}{2}\right) + \alpha_3 \frac{k}{2} e^{i\theta} \\
&+ \lambda k h \Big[g e^{i(3-m)\theta} p_{i,1} + g e^{i(N-(m-2))\theta} p_{i,N} + 2 \sum_{j=2}^{N-1} g e^{i(j-(m-2))\theta} p_{i,j} \Big] - 2k(r+\lambda) g
\end{aligned}
$$

$\Longleftrightarrow$

$$
\begin{aligned}
&g^2 \Big[- \alpha_1 \frac{k}{2} (cos(\theta) - isin(\theta)) + \left(1 + \alpha_2 \frac{k}{2}\right) - \alpha_3 \frac{k}{2} (cos(\theta) + isin(\theta)) \Big] \\
&= \alpha_1 \frac{k}{2} (cos(\theta) - isin(\theta)) + \left(1 - \alpha_2 \frac{k}{2}\right) + \alpha_3 \frac{k}{2} (cos(\theta) + isin(\theta)) \\
&+ \lambda k h \Big[g e^{i(3-m)\theta} p_{i,1} + g e^{i(N-(m-2))\theta} p_{i,N} + 2 \sum_{j=2}^{N-1} g e^{i(j-(m-2))\theta} p_{i,j} \Big] - 2k(r+\lambda) g
\end{aligned}
$$

$\Longleftrightarrow$

$$
\begin{aligned}
&g^2 \Big[1 + \alpha_2 \frac{k}{2} - \frac{k}{2} \Big((\alpha_1 + \alpha_3) cos(\theta) - isin(\theta) \left(\alpha_1 - \alpha_3\right) \Big) \Big] \\
&= \Big(1 - \alpha_2 \frac{k}{2}\Big) + \frac{k}{2} \Big((\alpha_1 + \alpha_3) cos(\theta) - isin(\theta) \left(\alpha_1 - \alpha_3\right) \Big) \\
&+ \lambda k h \Big[g e^{i(3-m)\theta} p_{i,1} + g e^{i(N-(m-2))\theta} p_{i,N} + 2 \sum_{j=2}^{N-1} g e^{i(j-(m-2))\theta} p_{i,j} \Big] - 2k(r+\lambda) g
\end{aligned}
$$

$\Longleftrightarrow$

$$
\begin{aligned}
&g^2 \Big[1 + \frac{k\sigma^2}{h^2} - \frac{k\sigma^2}{h^2} cos(\theta) - isin(\theta) \frac{k}{h} \Big(r - q - \frac{1}{2}\sigma^2 - \lambda\zeta \Big) \Big] \\
&= \lambda k h \Big[g e^{i(3-m)\theta} p_{i,1} + g e^{i(N-(m-2))\theta} p_{i,N} + 2 \sum_{j=2}^{N-1} g e^{i(j-(m-2))\theta} p_{i,j} \Big] - 2k(r+\lambda) g \\
&+ \Big(1 - \frac{k\sigma^2}{h^2}\Big) + \frac{k\sigma^2}{h^2} cos(\theta) + isin(\theta) \frac{k}{h} \Big(r - q - \frac{1}{2}\sigma^2 - \lambda\zeta \Big)
\end{aligned}
$$

$\Longleftrightarrow$

$$
g^2 \Big[1 + \frac{k\sigma^2}{h^2} (1 - cos(\theta)) - i \frac{k}{h} \Big(r - q - \frac{1}{2}\sigma^2 - \lambda\zeta \Big) sin(\theta) \Big]
$$

$$= 1 - \frac{k\sigma^2}{h^2}(1 - cos(\theta)) + i\frac{k}{h}\Big(r - q - \frac{1}{2}\sigma^2 - \lambda\zeta\Big)sin(\theta)$$
$$+ 2\lambda k\frac{h}{2}g\Big[e^{i(3-m)\theta}p_{i,1} + e^{i(N-(m-2))\theta}p_{i,N} + 2\sum_{j=2}^{N-1} e^{i(j-(m-2))\theta}p_{i,j}\Big] - 2k(r+\lambda)g$$

$\Longleftrightarrow$

$$\beta_2 g^2 - 2\beta_1 g - \beta_0 = 0,$$

with

$$\beta_2 := 1 + \frac{k\sigma^2}{h^2}(1 - cos(\theta)) - i\frac{k}{h}\Big(r - q - \frac{1}{2}\sigma^2 - \lambda\zeta\Big)sin(\theta),$$
$$\beta_1 := -k(r+\lambda) + \lambda k(F_R + iF_I),$$
$$\beta_0 := 1 - \frac{k\sigma^2}{h^2}(1 - cos(\theta)) + i\frac{k}{h}\Big(r - q - \frac{1}{2}\sigma^2 - \lambda\zeta\Big)sin(\theta),$$

and where F_R and F_I are real numbers such that

$$F_R + iF_I = \frac{h}{2}\Big[e^{i(3-m)\theta}p_{i,1} + e^{i(N-(m-2))\theta}p_{i,N} + 2\sum_{j=2}^{N-1} e^{i(j-(m-2))\theta}p_{i,j}\Big].$$

To prove the stability, we borrow the following lemma from [50]:

Lemma 8.2. *A finite difference scheme for a scalar equation is stable if and only if all the roots, $g_r(\theta)$, of the amplification polynomial $P(\theta, k, h) := \beta_2 g^2 - 2\beta_1 g - \beta_0$ satisfy the following conditions:*

(i) There is a constant C such that $|g_r| \leq 1 + Ck$;

(ii) There are positive constants c_0 and c_1 such that if $c_0 \leq |g_r| \leq 1 + Ck$, then $|g_r|$ is a simple root and for any other root $|g_s|$ the relation

$$|g_r - g_s| \geq c_1$$

holds for k and h sufficiently small.

Using this lemma, we prove that our implicit method is stable.

Theorem 8.3. *The finite difference method (7.15)–(7.16) is stable in the sense of the Von Neumann analysis if k is smaller than $1/(2r + 4\lambda)$.*

Proof. Let us consider the difference method (7.15)–(7.16). We prove its stability by using Lemma 8.2. Assume g is a root of the amplification polynomial $P(\theta, k, h)$.

$$|g| = \left|\frac{2\beta_1 \pm \sqrt{4\beta_1^2 - 4\beta_2\beta_0}}{2\beta_2}\right| = \left|\frac{\beta_1 \pm \sqrt{\beta_1^2 - \beta_2\beta_0}}{\beta_2}\right| = \left|\frac{\beta_1}{\beta_2} \pm \sqrt{\frac{\beta_1^2 - \beta_2\beta_0}{\beta_2^2}}\right|$$
$$\leq 2\left|\frac{\beta_1}{\beta_2}\right| + \left|\frac{\beta_0}{\beta_2}\right|^{\frac{1}{2}} \leq 1 + 2(r + 2\lambda)k,$$

because $|\beta_2| > 1$, $|\beta_1| < (r + 2\lambda)k$ and $\left|\frac{\beta_2}{\beta_0}\right| = \left|\frac{\beta_2}{2-\beta_2}\right| \leq \frac{|\beta_2|}{2+|\beta_2|} < 1$.

We have just proven (i) in Lemma 8.2.

Now, let g_r and g_s be two roots of $P(\theta, k, h)$. Assume g_r is a root which has modulus larger than 1 (the case where $c_0 = 1$). Then:

$$\begin{aligned}
|2g_r - (g_r + g_s)| &\geq 2|g_r| - |g_r + g_s| \\
|g_r - g_s| &\geq 2|g_r| - |g_r + g_s| \\
|g_r - g_s| &\geq 2|g_r| - (|g_s| - |g_r|) \\
|g_r - g_s| &\geq 2 - 1 - 2(r + 2\lambda)k + 1 \\
|g_r - g_s| &\geq 2 - 2(r + 2\lambda)k
\end{aligned}$$

So, taking k sufficiently small, we can for instance guarantee $2(r + 2\lambda)k$ is smaller than one. Therefore, defining $c_1 := 1$, we can establish the inequality: $|g_r - g_s| \geq c_1$ and thus fulfill the condition (ii) of Lemma 8.2.

In conclusion, we proved the scheme is stable. ■

8.3 Convergence

We are now showing that our numerical scheme is convergent. Generally speaking, convergence is verified when the numerical solution of a problem approaches its exact solution, by converging to it as the mesh size tends to zero. Mathematically, this means the norm of the error we commit tends to zero, when the mesh size does too.

Before formalizing the proof of the convergence property, we introduce a few useful concepts, which are examined in more detail in [34].

8.3.1 Preliminary definitions and results

Let $\mathbb{R}^N$ be the space of column vectors of dimension N with real entries. For a given column vector $x := (x_1, x_2, \ldots, x_N)^T \in \mathbb{R}^N$, using the usual notation $h := \frac{2X}{N-1}$, we define the l^2 and l_∞ norms of the vector x as:

$$\|x\|_{l^2} := \left(h \sum_{j=1}^{N} |x_j|^2 \right)^{\frac{1}{2}} \quad \text{and} \quad \|x\|_{l^\infty} := \max_{1 \le j \le N} |x_j|. \tag{8.4}$$

If A is a matrix with N rows, we define the matrix norms as:

$$\|A\|_2 := \max_{x \neq 0} \frac{\|Ax\|_{l^2}}{\|x\|_{l^2}} \quad \text{and} \quad \|A\|_\infty := \max_{x \neq 0} \frac{\|Ax\|_{l^\infty}}{\|x\|_{l^\infty}}. \tag{8.5}$$

Let A^T be the transpose of a matrix A. If A is a square matrix with real entries, then

$$\|A\|_2 = \sqrt{\lambda_{\max}(A^T A)}, \tag{8.6}$$

where $\lambda_{\max}(A^T A)$ is the greatest eigenvalue of $A^T A$ and

$$\|A\|_\infty = \max_j \left(\sum_{k=1}^{N} |a_{jk}| \right), \tag{8.7}$$

where a_{jk} is the jth row and kth column element of matrix A.

For the respective proofs, we refer the Theorems A.4 and A.5, found in the Appendix.

From the last equality, it is easy to conclude that if A is a square Toeplitz matrix whose diagonals are constant, then

$$\|A\|_\infty = \|A^T\|_\infty. \tag{8.8}$$

To prove the convergence in the discrete l^2-norm, we need the following lemma.

Lemma 8.4. *Let $\{a_n\}_{n \ge 0}$ be a non-negative sequence such that, for $n \ge 2$,*

$$a_n \le a_{n-2} + Cka_{n-1} + D,$$

where C and D are positive constants. If $a_0 = 0$, then, for $n \ge 2$,

$$a_n \le (1 + Ck)^{n-1} a_1 + D \sum_{j=0}^{n-2} (1 + Ck)^j, \tag{8.9}$$

where k is a positive constant.

Proof. Assume $a_0 = 0$. Let $n = 2$.

$$\begin{aligned} a_2 &\leq a_0 + Cka_1 + D \\ &\leq a_1 + Cka_1 + D \\ &= (1 + Ck)a_1 + D \end{aligned}$$

If $n = 3$:

$$\begin{aligned} a_3 &\leq a_1 + Cka_2 + D \\ &\leq a_1 + Ck((1 + Ck)a_1 + D) + D \\ &= (1 + Ck + (Ck)^2)a_1 + D + DCk \\ &\leq (1 + 2Ck + (Ck)^2)a_1 + D + D(1 + Ck) \\ &= (1 + Ck)^2 a_1 + D\sum_{j=0}^{1}(1 + Ck)^j \end{aligned}$$

We are now using mathematical induction to prove our claim. Take $m \geq 4$ and suppose the inequality (8.9) is valid for any $n \leq m$. For $n = m + 1$, we have:

$$\begin{aligned} a_{m+1} &\leq a_{m-1} + Cka_m + D \\ &\leq (1 + Ck)^{m-2} a_1 + D\sum_{j=0}^{m-3}(1 + Ck)^j + Ck\Big[(1 + Ck)^{m-1} a_1 + D\sum_{j=0}^{m-2}(1 + Ck)^j\Big] + D \\ &\leq [(1 + Ck)^{m-2} + Ck(1 + Ck)^{m-1} a_1 + D\left[\sum_{j=0}^{m-3}(1 + Ck)^j + Ck\sum_{j=0}^{m-2}(1 + Ck)^j + 1\right] \\ &= (1 + Ck)^{m-2}[1 + Ck + Ck^2)]a_1 \\ &+ D\left[\sum_{j=0}^{m-3}(1 + Ck)^j + Ck(1 + Ck)^{m-2} + Ck\sum_{j=0}^{m-3}(1 + Ck)^j + 1\right] \\ &\leq (1 + Ck)^{m-2}[1 + 2Ck + Ck^2)]a_1 \\ &+ D\left[(1 + Ck)\sum_{j=0}^{m-3}(1 + Ck)^j + Ck(1 + Ck)^{m-2} + 1\right] \\ &= (1 + Ck)^m a_1 + D\left[\sum_{j=0}^{m-3}(1 + Ck)^{j+1} + Ck(1 + Ck)^{m-2} + 1\right] \\ &\leq (1 + Ck)^m a_1 + D\left[\sum_{j=1}^{m-2}(1 + Ck)^j + (1 + Ck)(1 + Ck)^{m-2} + (1 + Ck)^0\right] \end{aligned}$$

$$\leq (1+Ck)^m a_1 + D\sum_{j=0}^{m-1}(1+Ck)^j.$$

We proved the inequality (8.9) holds for $n = m+1$, given that it is verified for $n \leq m$, so, we are able to conclude it is valid for any $n \geq 2$. ■

8.3.2 Convergence proof

Let us define the error vector $\xi^m := (\xi_2^m, \xi_3^m, \ldots, \xi_{N-1}^m)^T$ on time level m by

$$\xi_i^m = u_i^m - \tilde{u}_i^m, \quad \text{for } 2 \leq i \leq N-1,$$

and the boundary errors by

$$\xi_1^m = u_1^m - \tilde{u}_1^m \quad \text{and} \quad \xi_N^m = u_N^m - \tilde{u}_N^m.$$

With the definition above, we are going to prove second-order convergence in the discrete l^2-norm, as done in the following theorem.

Theorem 8.5. *Let $u(\tau, x)$ be the solution of the initial-valued PIDE in (7.1)-(7.3) and let $\bar{u}_i^m$ be the solution of the finite difference approximation in (7.15)–(7.16). If k and h are sufficiently small, then there exists a positive constant C', independent of h and k, such that for $2 \leq m \leq M$*

$$\|\xi^m\|_{l^2} \leq C'(k^2 + h^2). \tag{8.10}$$

Proof. In (8.2), we can replace v with the classical solution u. Using (8.2) and (8.3), we obtain the error equation for $3 \leq m \leq M$ and $2 \leq i \leq N-1$:

$$\begin{aligned}
&-\alpha_1 \frac{k}{2}\xi_{i-1}^m + \left(1 + \frac{\alpha_2 k}{2}\right)\xi_i^m - \frac{\alpha_3 k}{2}\xi_{i+1}^m \\
&= \alpha_1 \frac{k}{2}\xi_{i-1}^{m-2} + \left(1 - \frac{\alpha_2 k}{2}\right)\xi_i^{m-2} + \frac{\alpha_3 k}{2}\xi_{i+1}^{m-2} \\
&+ \lambda hk\left[\xi_1^{m-1}p_{i,1} + 2\sum_{j=2}^{N-1}\xi_j^{m-1}p_{i,j} + \xi_N^{m-1}p_{i,N}\right] - 2k(r+\lambda)\xi_i^{m-1} + k\,O(k^2+h^2).
\end{aligned}$$

The last equation can be rewritten in matrix form as follows:

$$\Big(I + \frac{k}{2}B\Big)\xi^m = \Big(I - \frac{k}{2}B\Big)\xi^{m-2} + C\xi^{m-1} + d, \tag{8.11}$$

where I is the identity matrix of size $N-2, C=(c_{ij})$ and $B=(b_{ij})$ are square matrices of order $N-2$ with entries

$$b_{i,j}=\begin{cases}-\frac{\sigma^2}{h^2}+\frac{1}{h}(r-q-\lambda\zeta-\frac{\sigma^2}{2}), & \text{if } i=j-1 \text{ and } 2\leq i\leq N-2\\ \frac{2\sigma^2}{h^2}, & \text{if } i=j \text{ and } 1\leq i\leq N-2\\ -\frac{\sigma^2}{h^2}-\frac{1}{h}(r-q-\lambda\zeta-\frac{\sigma^2}{2}), & \text{if } i=j+1 \text{ and } 1\leq i\leq N-3\\ 0, & \text{otherwise}\end{cases}$$

$$c_{i,j}=\begin{cases}2kh\lambda p_{i,j}-2k(r+\lambda), & \text{if } i=j \text{ and } 1\leq i\leq N-2\\ 2kh\lambda p_{i,j}, & \text{if } i\neq j \text{ and } 1\leq i\leq N-2\end{cases}$$

and $d=(d_j)$ is the column vector of size $N-2$ with entries

$$d_j=kO(h^2+k^2), \quad \text{for } j=1\leq j\leq N-2.$$

Now, attempt to the following inequalities:

$$\left|1+\frac{k\sigma^2}{h^2}\right|>\frac{k}{2}\left|-\frac{\sigma^2}{h^2}+\frac{1}{h}(r-q-\lambda\zeta-\frac{\sigma^2}{2})\right|+\frac{k}{2}\left|-\frac{\sigma^2}{h^2}-\frac{1}{h}(r-q-\lambda\zeta-\frac{\sigma^2}{2})\right|$$
$$\iff 2h^2+2k\sigma^2>k\left|\sigma^2-h(r-q-\lambda\zeta-\frac{\sigma^2}{2})\right|+k\left|\sigma^2+h(r-q-\lambda\zeta-\frac{\sigma^2}{2})\right|$$

The inequation above is the necessary and sufficient requirement in order for the matrix $\left(I+\frac{k}{2}B\right)$ to be strictly diagonally dominated. Letting h being sufficiently small, the algebraic expressions inside the absolute values become non-negative, since σ^2 is non-negative. Thus:

$$2h^2+2k\sigma^2>k\left(\sigma^2-h(r-q-\lambda\zeta-\frac{\sigma^2}{2})\right)+k\left(\sigma^2+h(r-q-\lambda\zeta-\frac{\sigma^2}{2})\right)$$

$\iff$

$$2h^2+2k\sigma^2-k\sigma^2-k\sigma^2>0$$

$\iff$

$$2h^2>0.$$

Since the last inequality holds true for any value of h, we can conclude there exists a sufficiently small h such that the matrix referred before is a strictly diagonally dominated matrix. As proven in Theorem A.3, this implies it is invertible and thus $\left(I + \frac{k}{2}B\right)^{-1}$ exists. Multiplying both sides of (8.11) by $\left(I + \frac{k}{2}B\right)^{-1}$ and taking the l^2-norm gives:

$$\begin{aligned}
\|\xi^m\|_{l^2} =& \left\| \left(I + \frac{k}{2}B\right)^{-1}\left(I - \frac{k}{2}B\right)\xi^{m-2} + \left(I + \frac{k}{2}B\right)^{-1} C\xi^{m-1} + \left(I + \frac{k}{2}B\right)^{-1} d \right\|_{l^2} \\
\leq& \left\| \left(I + \frac{k}{2}B\right)^{-1}\left(I - \frac{k}{2}B\right)\xi^{m-2} \right\|_{l^2} \\
&+ \left\| \left(I + \frac{k}{2}B\right)^{-1} C\xi^{m-1} \right\|_{l^2} + \left\| \left(I + \frac{k}{2}B\right)^{-1} d \right\|_{l^2} \\
\leq& \left\| \left(I + \frac{k}{2}B\right)^{-1}\left(I - \frac{k}{2}B\right) \right\|_2 \|\xi^{m-2}\|_{l^2} \\
&+ \left\| \left(I + \frac{k}{2}B\right)^{-1} C \right\|_2 \|\xi^{m-1}\|_{l^2} + \left\| \left(I + \frac{k}{2}B\right)^{-1} \right\|_2 \|d\|_{l^2}.
\end{aligned} \tag{8.12}$$

We shall estimate the three matrix norms in the last inequality. $\left(I + \frac{k}{2}B\right)$ is a tridiagonal Toeplitz matrix, with constant values along all its diagonals. Thus, and as proved in [35], each of its eigenvalues, λ_h, is given by:

$$\lambda_h = a + 2\sqrt{bc}\cos\left(\frac{h\pi}{N-2}\right), \quad h \in \{1, \ldots, N-3\}, \tag{8.13}$$

where a, b and c are respectively the diagonal, lower and upper diagonal constants. The quotient $\frac{h\pi}{N-2}$ is a value between zero and π (exclusive) and therefore, the cosine function in (8.13) always assumes different values (it is injective in $]0, \pi[$). The constants a, b and c are never equal to zero, so we can state that the matrix $\left(I + \frac{k}{2}B\right)$ has $N-2$ non-zero different eigenvalues and consequently, it is diagonizable.

By definition, there is an invertible matrix Q such that $Q^{-1}\Lambda_W Q = \left(I + \frac{k}{2}B\right)$, where Λ_W is a diagonal matrix, whose elements correspond to the eigenvalues of the right-hand side matrix, which we now denote by W. The columns of Q are eigenvectors associated to the referred eigenvalues and they constitute a basis of the $\mathbb{R}^{N-2}$ space. We can analogously prove that $\left(I - \frac{k}{2}B\right)$ is diagonizable (and logically, $\left(I - \frac{k}{2}B\right)^{-1}$ too).

Let us take a basis of eigenvectors of W that is also a basis of eigenvectors for $Z := \left(I - \frac{k}{2}B\right)^{-1}$. Denote the corresponding eigenvalues of W by $\lambda_1, ..., \lambda_{N-2}$ and those of Z by $\mu_1, ..., \mu_{N-2}$. Representing by Λ_Z the diagonal matrix which is similar to Z and whose diagonal components are its eigenvalues, we prove WZ commute:

$$WZ = Q\Lambda_W Q^{-1} Q\Lambda_Z Q^{-1} = Q\Lambda_W \Lambda_Z Q^{-1} = Q\Lambda_Z Q^{-1} Q\Lambda_W Q^{-1} = ZW.$$

We remark we used the fact that diagonal matrices trivially commute in multiplication.

Let $z \in \mathbb{R}^{N-2}$ be an arbitrary column-vector chosen. Using the fact that the matrices above commute, we have:

$$\begin{aligned}
&\left\| \left(I + \frac{k}{2}B\right)^{-1} \left(I - \frac{k}{2}B\right) z \right\|_{l^2}^2 = \left\| \left(I - \frac{k}{2}B\right) \left(I + \frac{k}{2}B\right)^{-1} z \right\|_{l^2}^2 \\
&= h \left(\left(I + \frac{k}{2}B\right)^{-1} z \right)^T \left(I - \frac{k}{2}B\right)^T \left(I - \frac{k}{2}B\right) \left(I + \frac{k}{2}B\right)^{-1} z \\
&= h \left(\left(I + \frac{k}{2}B\right)^{-1} z \right)^T \left[\left(I + \frac{k}{2}B\right)^T \left(I + \frac{k}{2}B\right) - k(B^T + B) \right] \left(I + \frac{k}{2}B\right)^{-1} z \\
&= hz^T \left(\left(I + \frac{k}{2}B\right)^T \right)^{-1} \left(I + \frac{k}{2}B\right)^T \left(I + \frac{k}{2}B\right) \left(I + \frac{k}{2}B\right)^{-1} z \\
&- khz^T \left(\left(I + \frac{k}{2}B\right)^T \right)^{-1} (B^T + B) \left(I + \frac{k}{2}B\right)^{-1} z \\
&= hz^T z - khz^T \left(\left(I + \frac{k}{2}B\right)^{-1} \right)^T (B + B^T) \left(I + \frac{k}{2}B\right)^{-1} z \\
&= \|z\|_{l^2}^2 - hk \left(\left(I + \frac{k}{2}B\right)^{-1} z \right)^T (B + B^T) \left(I + \frac{k}{2}B\right)^{-1} z. \qquad (8.14)
\end{aligned}$$

$B + B^T$ has the form:

$$B + B^T = \frac{2\sigma^2}{k^2} \begin{bmatrix} 2 & -1 & 0 & \cdots & 0 \\ -1 & 2 & -1 & \ddots & \vdots \\ 0 & -1 & 2 & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & -1 \\ 0 & \cdots & 0 & -1 & 2 \end{bmatrix}.$$

Using (8.13), it is clear $B+B^T$ is positive definite, because it is a tridiadiagonal Toeplitz

matrix with constant values along all its diagonals.

Consequently from the definition, for any vector $x \in \mathbb{R}^{N-2}, x^T(B + B^T)x > 0$ and thus, the equality from equation (8.14) is majored by $\|z\|_{l^2}^2$. Having the definition of l^2-norm in mind, we can then conclude that

$$\left\| \Big(I + \frac{k}{2}B\Big)^{-1} \Big(I - \frac{k}{2}B\Big) \right\|_2 \leq 1. \tag{8.15}$$

Now, we prove that

$$\left\| \Big(I + \frac{k}{2}B\Big)^{-1} \right\|_2 \leq 1. \tag{8.16}$$

To accomplish this goal, we are claiming the following inequality is true for any $z \in \mathbb{R}^{N-2}$:

$$\|z\|_{l^2} \leq \left\| \Big(I + \frac{k}{2}B\Big)z \right\|_{l^2}. \tag{8.17}$$

The latter inequality allows us to conclude our claim, because in the case of the l^2-norm, it holds that $\|A^{-1}\|_2 = \frac{1}{\|A\|_2}$. This is easy to see, using equation (8.6).

The matrix B with real entries can be split into two parts such that

$$B = \frac{1}{2}(B + B^T) + \frac{1}{2}(B - B^T).$$

$B + B^T$ is a symmetric matrix and $B - B^T$ is a skew-symmetric. This directly implies that the quadratic form of the latter is zero. In fact, taking $z \in \mathbb{R}^{N-2}$,

$$\begin{aligned}
z^T(B - B^T)z &= (z^T(B - B^T)z)^T \\
z^T(B - B^T)z &= (z^T(B^T - B)z)^T \\
z^T(B - B^T)z &= -(z^T(B - B^T)z)^T \\
2z^T(B - B^T)z &= 0 \\
z^T(B - B^T)z &= 0.
\end{aligned}$$

Using the last result and applying the Cauchy–Schwarz inequality gives:

$$\begin{aligned}
\left\| \Big(I + \frac{k}{2}B\Big)z \right\|_{l^2} \|z\|_{l^2} &\geq hz^T\Big(I + \frac{k}{2}B\Big)z \\
&= hz^Tz + h\frac{k}{2}z^T\left(\frac{1}{2}(B + B^T) + \frac{1}{2}(B - B^T)\right)z
\end{aligned}$$

$$
\begin{aligned}
&= \|z\|_{l^2}^2 + h\frac{k}{4} z^T (B + B^T) z \\
&\geq \|z\|_{l^2}^2 + \frac{k}{4} \lambda_{min}(B + B^T) \|z\|_{l^2}^2 \\
&\geq \|z\|_{l^2}^2,
\end{aligned}
$$

where $\lambda_{min}(B + B^T) > 0$ is the smallest eigenvalue of $(B + B^T)$. Hence, we prove our claim.

To understand how did we get the penultimate inequality as above, we use the following Theorem and the fact that $B + B^T$ is a positive definite matrix.

Theorem 8.6. *Let A be a symmetric positive definite matrix. The following inequality holds for any $x \in \mathbb{R} \setminus \{0\}$:*

$$\lambda_{\min} \|x\|_{l^2}^2 \leq x^T A x \leq \lambda_{\max} \|x\|_{l^2}^2,$$

where $\lambda_{\min}$ and $\lambda_{\max}$ are, respectively, the minimum and the maximum eigenvalues of A.

Proof. By the spectral Theorem, A is similar to a diagonal matrix. Then, consider A can be decomposed as $A = Q^T \Lambda Q$, where Λ is a diagonal matrix with entries $\lambda_1, \ldots, \lambda_N$, corresponding to the positive eigenvalues of A, and Q is an orthogonal matrix.
Define $y := \frac{x}{\sqrt{x^T x}}$ and $z := Qy$. Let $x \in \mathbb{R} \setminus \{0\}$.

$$\frac{x^T A x}{\|x\|_{l^2}^2} = \frac{x^T A x}{x^T x} = y^T A y = y^T Q^T \Lambda Q y = z^T \Lambda z = \sum_{i=1}^{N} \lambda_i |z_i|^2,$$

where $z_i, i \in \{1, \ldots, N\}$ are the entries of the vector z. Now, notice that for a fixed z, the expression above is minimal for $\lambda_i = \lambda_{\min}$ and maximal if $\lambda_i = \lambda_{\max}$. Consequently:

$$\lambda_{\min} \sum_{i=1}^{N} |z_i|^2 \leq \frac{x^T A x}{\|x\|_{l^2}^2} \leq \lambda_{\max} \sum_{i=1}^{N} |z_i|^2.$$

But $\sum_{i=1}^{N} |z_i|^2$ is equal to the square of the norm of z which is equal to 1:
$\|z\|_{l^2}^2 = z^T z = \|Qy\|_{l^2}^2 = (Qy)^T Qy = y^T Q^T Q y = y^T y = \frac{x^T x}{x^T x} = 1$.

Therefore, we conclude that

$$\lambda_{\min} \|x\|_{l^2}^2 \leq x^T A x \leq \lambda_{\max} \|x\|_{l^2}^2.$$

■

We are now estimating the norm $\left\|\left(I+\frac{k}{2}B\right)^{-1}C\right\|_2$, using the fact that C is a Toeplitz matrix, as long as the grid we choose is uniform. But first, we estimate a lower bound for the infinity norm.

Let A be a matrix whose maximum eigenvalue is denoted by $\lambda^A_{\max}$ and $v \in \mathbb{R}^{N-2}\backslash\{0\}$ an eigenvector associated to this eigenvalue. By definition:

$$Av = \lambda^A_{\max} v.$$

This implies:

$$\begin{aligned}
&\|Av\|_{l^\infty} = \|\lambda^A_{\max} v\|_{l^\infty} \\
\iff \\
&\|Av\|_{l^\infty} = |\lambda^A_{\max}|\|v\|_{l^\infty} \\
\iff \\
&\frac{\|Av\|_{l^\infty}}{\|v\|_{l^\infty}} = |\lambda^A_{\max}| \\
\Rightarrow \\
&\max_{v\neq 0} \frac{\|Av\|_{l^\infty}}{\|v\|_{l^\infty}} \geq |\lambda^A_{\max}| \\
\iff \\
&\|A\|_\infty \geq |\lambda^A_{\max}|.
\end{aligned}$$

Applying this result in the following inequalities, we get:

$$\begin{aligned}
\left\|\left(I+\frac{k}{2}B\right)^{-1}C\right\|_2 \leq& \left\|\left(I+\frac{k}{2}B\right)^{-1}\right\|_2 \|C\|_2 \\
\leq& \|C\|_2 = \sqrt{\lambda_{\max}(C^TC)} \\
\leq& \sqrt{\|C^TC\|_\infty} \leq \sqrt{\|C^T\|_\infty \|C\|_\infty} \\
=& \sqrt{\|C\|_\infty \|C\|_\infty} = \|C\|_\infty \\
=& \max_j \left(\sum_{k=1}^{N-2} |c_{jk}|\right)
\end{aligned}$$

$$\leq 2k(r+\lambda) + 2k\lambda \underbrace{h \sum_{k=1}^{N-2} p_j k}_{\leq 1}$$

$$\leq 2k(r+2\lambda).$$

Hence, through (8.12) we have:

$$\|\xi^m\|_{l^2} \leq \|\xi^{m-2}\|_{l^2} + 2k(r+2\lambda)\|\xi^{m-1}\|_{l^2} + \|d\|_{l^2}, \tag{8.18}$$

and we are ready to use Lemma 8.4. With the initial condition $\|\xi^1\|_{l^2} = 0$ and using the formula for the sum of the first $m-1$ elements of a geometric progression, we get:

$$\begin{aligned}\|\xi^m\|_{l^2} &\leq (1+2(r+2\lambda)k)^{m-1}\|\xi^2\|_{l^2} + \|d\|_{l^2} \sum_{j=0}^{m-2} (1+2k(r+2\lambda))^j \\ &\leq e^{2(r+2\lambda)T}\|\xi^2\|_{l^2} + \frac{e^{2(r+2\lambda)T}-1}{2k(r+2\lambda)}\|d\|_{l^2}. \end{aligned} \tag{8.19}$$

Now, let us estimate the two discrete vector norms $\|\xi^2\|_{l^2}$ and $\|d\|_{l^2}$. On time level 2, we can obtain the local truncation error vector $\|\xi^2\|_{l^2}$ in the form

$$\xi_i^2 = kO(h^2+k), \quad \text{for } 1 \leq i \leq N-2.$$

Hence, we have the estimate

$$\|\xi^2\|_{l^2} = kO(k+h^2). \tag{8.20}$$

For the column vector d, we have the estimate

$$\|d\|_{l^2} = kO(k^2+h^2). \tag{8.21}$$

Therefore, combining (8.19), (8.20) and (8.21), we have (8.10) as desired.

■

9 Numerical results for European options

In this section, we implement numerical simulations to price European put options under the Merton and Kou models. We use the implicit method described in Chapter 7 so that it leads to tridiagonal linear systems. Each tridiagonal linear system can be solved by using a LU decomposition with $O(N)$ operations, where N is the number of spatial steps. We apply the FFT algorithm to compute the product of a dense Toeplitz matrix and a column vector in the discrete integral operator in (7.14). The FFT algorithm allows us to reduce the number of multiplications from $O(N^2)$ to $O(N\log_2 N)$, provided that N is typically a power of 2. Thus, the implicit method with three time levels requires totally $O(MN\log_2 N)$ operations, where N is the number of spatial steps and M the number of steps in time. For more details on how to apply a FFT to the matrix multiplication of a Toeplitz matrix and a vector, we refer to Appendix.

All the simulations presented here were computed in MATLAB using a Macbook Pro with processor 2,3 GHz Intel Core i5.

9.1 Results for the Merton model

We start with the Merton model, which has the density function of jump sizes given by (3.1). Parameters used in the simulation are the following:

$$\sigma = 0.15, \quad r = 0.05, \quad \sigma_J = 0.45, \quad \mu_J = -0.90,$$

$$d = 0, \quad \lambda = 0.10, \quad T = 0.25, \quad K = 100.$$

These values are also used by [22] to compute prices of European-style calls and in our experiments, we calculated the fair price of put vanilla options for three different initial values and a unique strike ($K = 100$). In order to keep the accuracy of our method larger than second-order, we build a cubic spline to evaluate the prices of the option at non-mesh points of stock prices. To test our method, we compare the results obtained with reference values, which are calculated using the numerical series (3.3). We could similarly apply the put-call parity to the values presented in [22] to obtain prices for European-style puts. The referred results can be found in Table 1 and as the reader might notice, the pointwise errors are decreasing with the increase of the number of steps.

Table 1: Prices of standard European-style put options under the Merton jump-diffusion model

Steps	Spot	Value	Error	CPU (sec)
$M = 25$	90	9.286528	-	0.012645
$N = 128$	100	3.147057	-	0.012569
	110	1.402729	-	0.012770
$M = 50$	90	9.289424	0.002896	0.032274
$N = 256$	100	3.148688	0.001631	0.029000
	110	1.401531	0.001198	0.032306
$M = 100$	90	9.286416	0.003008	0.095487
$N = 512$	100	3.148942	0.000254	0.092192
	110	1.401161	0.000370	0.099015
$M = 200$	90	9.285666	0.000750	0.257156
$N = 1024$	100	3.149006	0.000064	0.240354
	110	1.401056	0.000105	0.262576
$M = 400$	90	9.285478	0.000188	0.836382
$N = 2048$	100	3.149021	0.000015	0.798052
	110	1.401193	0.000137	0.815349
$M = 800$	90	9.285432	0.000046	3.510451
$N = 4096$	100	3.149025	0.000004	3.495371
	110	1.401188	0.000005	3.406741

The reference values are 9.285418 at $S = 90$, 3.149026 at $S = 100$ and 1.401186 at $S = 110$. The truncated domain is $[-1.5, 1.5]$, for the variable $x = \ln(S/S_0)$. M is the number of time steps and N is the number of spatial steps. The third column represents the computed values for each initial value and the fourth the difference between the value computed for a given spot value and the one calculated in the previous row. The last column shows how much computation time we needed.

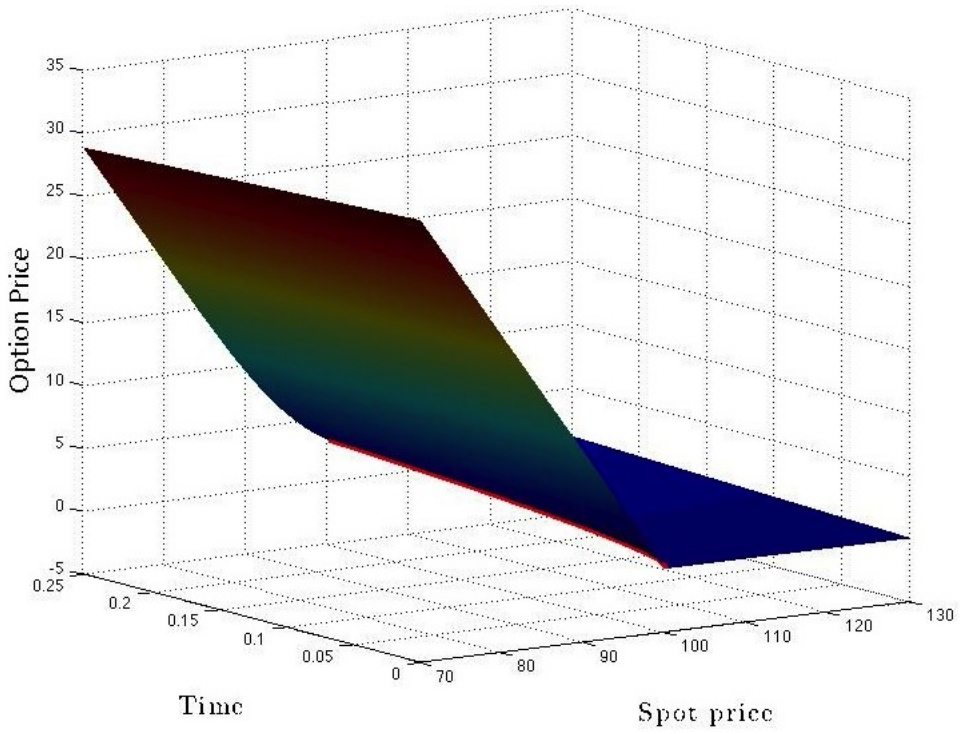

Figure 9.1: Price of a European put option under the Merton model

For illustration purposes, we plot the price of European put options under the Merton model in Figure 9.1.

9.2 Results for the Kou model

For the Kou model, whose density function is given by (3.6), the parameters used are:

$$\sigma = 0.15, \quad r = 0.05, \quad \eta_1 = 3.0465, \quad \eta_2 = 3.0775,$$
$$d = 0, \quad \lambda = 0.10, \quad T = 0.25, \quad K = 100, \quad p = 0.3445.$$

As done in the Merton model case, we compare our results with the ones presented in [22], using the put-call parity. We could analogously use (3.12)-(3.15) to obtain the prices of European-style puts as reported in Table 2.

Table 2: Prices of standard European-style put options under the Kou jump-diffusion model

Steps	Spot	Value	Error	CPU (sec)
$M = 25$	90	9.432506	-	0.009598
$N = 128$	100	2.732566	-	0.012768
	110	0.557687	-	0.011418
$M = 50$	90	9.434658	0.002152	0.030867
$N = 256$	100	2.731925	0.000641	0.024575
	110	0.553610	0.004077	0.025017
$M = 100$	90	9.431838	0.002820	0.061299
$N = 512$	100	2.731519	0.000406	0.060633
	110	0.552541	0.001069	0.064405
$M = 200$	90	9.430967	0.000871	0.252093
$N = 1024$	100	2.731372	0.000147	0.213662
	110	0.552250	0.000301	0.217982
$M = 400$	90	9.430666	0.000061	0.688922
$N = 2048$	100	2.731311	0.000142	0.724175
	110	0.552392	0.000137	0.709273
$M = 800$	90	9.430550	0.000116	3.619007
$N = 4096$	100	2.731284	0.000027	3.292017
	110	0.552372	0.000020	3.369464

The reference values are 9.430457 at $S = 90$, 2.731259 at $S = 100$ and 0.552363 at $S = 110$. The truncated domain is $[-1.5, 1.5]$, for the variable $x = \ln(S/S_0)$. M is the number of time steps and N is the number of spatial steps. The third column represents the computed values for each initial value and the fourth the difference between the value obtained for a given spot and the one calculated in the previous row. The last column shows how much time we needed to get the estimated values.

We plotted in Figure 9.2 the values of European-style put options under the Kou model.

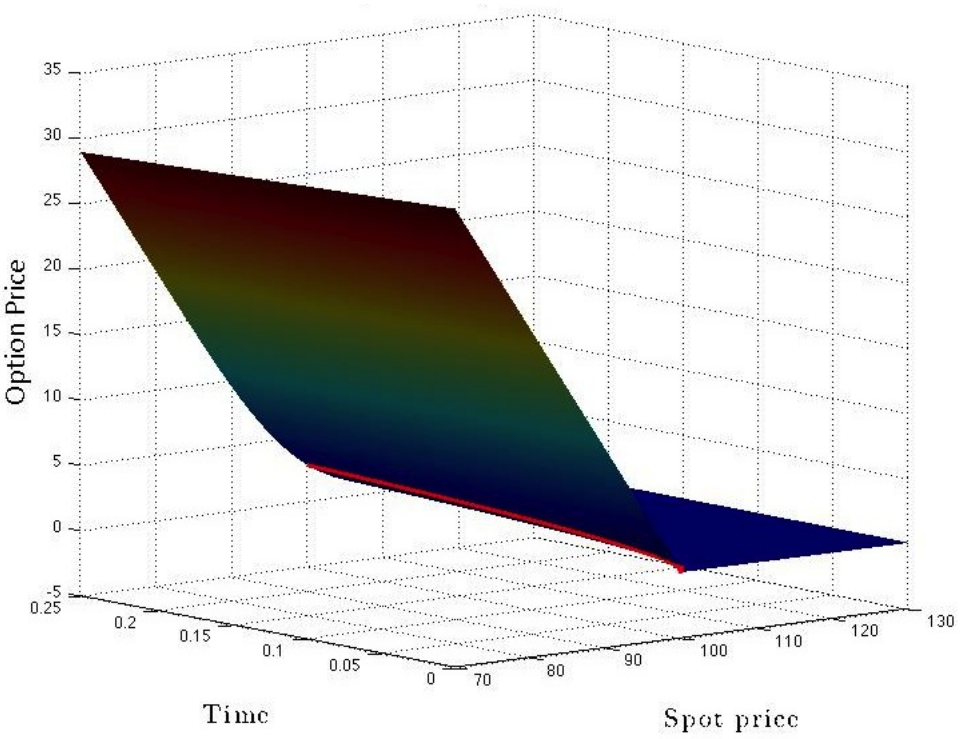

Figure 9.2: Price of a European put option under the Kou model

10 American option pricing

An American option only differs from a European one, because it can be exercised at any time until its maturity. Up to now we have assumed that a contract, like an option, can only be exercised exactly at time T. In real life, a large number of options can in fact be exercised at any time prior to T. The choice of exercise time is thus left to the holder of the claim. A contract with this feature is called an American contract. To put it more formally, let us fix a final exercise date T and a contract function H. The European version of this contract will, as usual, pay the amount $H(S_T)$ at time T to the holder, where H denotes the payoff of the contract. If the contract, on the other hand, is of the American type, then the holder will obtain the amount $H(S_t)$ if he/she chooses to exercise it at time t, prior to T. The situation is complicated further by the fact that the exercise time t does not have to be chosen a priori (i.e. at $t = 0$). It can be selected on the basis of the information generated by the stock price process, and thus, the holder will in fact choose a random exercise time τ. The exercise time (or rather, exercise strategy) τ has to be chosen such that the decision on whether to exercise the contract at time t or not, depends only upon the information generated by the price process up to time t.

American contracts are thus more complicated to analyze than their European counterparts, since the holder of the claim has to decide on an optimal exercise strategy. Problems of this kind are quite hard to solve and analytically they lead to so called "free boundary value problems", instead of the corresponding parabolic PDEs for the European counterparts.

Consider an underlying asset with price S_t at time t. The dynamics of S_t are assumed to be represented by the following SDE:

$$\frac{\partial S_t}{\partial t} = (r - q - \lambda\tilde{\zeta})dt + \sigma dW_t + (\eta_t - 1)dN_t \tag{10.1}$$

as in (2.9), with $\mu = r - q$.

In the above, r is the risk-free rate, assumed to be constant, μ is the instantaneous return per time unit, the constant σ is the instantaneous volatility per unit time, $(W_t)_{t\geq 0}$ a standard Wiener process and $(N_t)_{t\geq 0}$ is a Poisson process whose increments satisfy the same conditions as in (2.9). We assume that S pays a continuous dividend yield of rate q.

Consider an American put option with strike K and expiration date T, written on an underlying asset with price $S_t, t \in [0, T]$, assumed to obey the equation (10.1). The payoff of this option at each point t is simply $(K - S_t)^+$, which symbolically is written as $\max(0, K - S_t)$.

10.1 The Linear Complementary Problem

According to great part of the literature which explains and presents a study on American option pricing theory such as, for example, [14] (Proposition 12.3) and [6], the value of an American option u satisfies the Linear Complementary Problem (LCP):

$$\begin{cases} u_\tau(\tau, x) - \mathcal{L}u(\tau, x) \geq 0, & (10.2) \\ u(\tau, x) \geq h(x), & (10.3) \\ (u_\tau(\tau, x) - \mathcal{L}u(\tau, x))(u(\tau, x)) - h(x)) = 0, & (10.4) \end{cases}$$

for all $(\tau, x) \in]0, T] \times]-\infty, \infty[$. The payoff function of u is here represented by the function h. In section 7 of [17], we are introduced with a intuitive explanation on why an American put option follows the above equations, for the particular case where the asset follows a GBM.

In a nutshell, the second inequality says the option's value must be at least worth

its payoff and is a consequence of the no arbitrage assumption.

The first inequality would be an equality if we were in the European option case. When it is optimal to hold the option the equality is valid. Otherwise, it is optimal to exercise the option and the equality in (10.3) is satisfied. In this case, the payoff is the solution for the LCP and a strict inequality in (10.2) is verified.

The last equation as above translates the idea that either the option's holder decides not to exercise it and, in this case, $u_\tau(\tau, x) = \mathcal{L}u(\tau, x)$, or otherwise it does exercise and we have $u(\tau, x) = h(x)$.

Similarly to the European case, regarding the asymptotic behavior of the American put option, the value of $u(\tau, x)$ is $K - e^x$ when the log price x is a sufficiently small negative number, and the value $u(\tau, x)$ tends to zero as the log price x approaches infinity.

Now that we have our problem set, we are left to construct a pricing model to obtain American option prices and perform numerical experiments based on finite differences methods as done for European options.

11 Numerical Implementation for American options

In this section we formulate a numerical method to solve the LCP (10.2)-(10.4) for the American put option by using the implicit method discussed in the previous sections for the European option, combined with the operator splitting method described in [31] and [51].

As a reference, we suggest the reading of [37] and [38], since our analysis is mostly based on these two articles.

11.1 Operator splitting method

We start with the operator splitting method to solve the LCP (10.2)-(10.4). The splitting method is introduced by Ikonen and Toivanen [31] to evaluate the price of the American put option under the BSM model, and the method is studied by Toivanen [51] under the Kou model. To simplify our problem, we define the auxiliary variable φ as $\varphi := u_\tau - \mathcal{L}u$, implying that the LCP (10.2)-(10.4) for the American put option can be reformulated as:

$$\begin{cases} u_\tau - \mathcal{L}u = \varphi & (11.1) \\ \varphi \geq 0, \quad u \geq h, \quad \varphi(u-h) = 0 & (11.2) \end{cases}$$

in the region $]0,T]\times]-\infty,\infty[$ with the initial condition

$$h(x) = \max(0, K - e^x).$$

Let us consider an operator splitting method which splits the equation $u_\tau - \mathcal{L}u = \varphi$

on the mth time level into two discrete equations such as:

$$\frac{\tilde{U}_i^m - U_i^{m-2}}{2k} - \left\{ \mathcal{D}_\Delta \left(\frac{\tilde{U}_i^m + U_i^{m-2}}{2} \right) + \mathcal{I}_\Delta U_i^{m-1} - (r+\lambda) U_i^{m-1} \right\} = \Psi_i^{m-1}, \tag{11.3}$$

$$\frac{U_i^m - U_i^{m-2}}{2k} - \left\{ \mathcal{D}_\Delta \left(\frac{\tilde{U}_i^m + U_i^{m-2}}{2} \right) + \mathcal{I}_\Delta U_i^{m-1} - (r+\lambda) U_i^{m-1} \right\} = \Psi_i^{m}. \tag{11.4}$$

Thus, the discrete problem for (11.1)-(11.2) is to seek a pair (U_i^m, Ψ_i^m) satisfying the discrete equations in (11.3)-(11.4) and the constraints

$$\Psi_i^m \geq 0, \quad U_i^m \geq h(x_i), \quad \Psi_i^m (U_i^m - h(x_i)) = 0, \tag{11.5}$$

where all values (U_i^j, Ψ_i^j) on the jth time level for $1 \leq j \leq m$ and $m \geq 2$ are already obtained and are reasonably good approximations to (u_i^j, φ_i^j).

The first step of the operator splitting method is to compute $\tilde{U}_i^m$ for $i = 2, \ldots, N-1$ by solving the linear system in (11.3). The goal here is to obtain an intermediate approximation $\tilde{U}_i^m$ to u_i^m by solving the discrete PIDE on the $(m-1)th$ time level with known auxiliary term Ψ_i^{m-1}. Remark that three time levels are involved in this computation, in order to force the coefficient matrices of the linear systems to be tridiagonal matrices.

The second step of the operator splitting method is to set up the relationship in (11.4) between U_i^m and Ψ_i^m. Note that the operator $u_\tau - \mathcal{L}u$ on the mth time level is approximated by the modified discrete operator on the $(m-1)th$ time level. Subtracting (11.3) to (11.4), we can rewrite the discrete equation in (11.4) with the constraints in (11.5) as a problem to find a pair (U_i^m, Ψ_i^m) such that

$$\begin{cases} \dfrac{U_i^m - \tilde{U}_i^m}{2k} = \Psi_i^m - \Psi_i^{m-1} & (11.6) \\ \Psi_i^m (U_i^m - h(x_i)) = 0 & (11.7) \end{cases}$$

with the constraints

$$U_i^m \geq h(x_i) \quad \text{and} \quad \Psi_i^m \geq 0. \tag{11.8}$$

Now, we can solve the problem (11.6)-(11.8) easily by locating the intersection point of two lines on the $U\Psi$-plane, and we have

$$(U_i^m, \Psi_i^m) = \begin{cases} \Big(h(x_i), \Psi_i^{m-1} + \dfrac{h(x_i) - \tilde{U}_i^m}{2k}\Big), & \text{if } \tilde{U}_i^m - 2k\Psi_i^{m-1} \leq h(x_i) \quad (11.9) \\ (\tilde{U}_i^m - 2k\Psi_i^{m-1}, 0), & \text{otherwise.} \quad (11.10) \end{cases}$$

In order to understand how to get the previous system, we must consider two scenarios: Either we exercise the option, being then its value equal to the payoff function (the case when $\tilde{U}_i^m = h(x_i)$) or we do not exercise the option. If the first case is verified, (11.9) is easily obtained by solving (11.6) in order to Ψ_i^m. If we do not exercise the option, then from (11.7) and (11.8), $\Psi_i^m = 0$ and we get (11.10), by solving (11.6) in order to U_i^m.

Thus, the second step can be done not by solving the discrete equation (11.4), but by updating the formula in (11.9)-(11.10) with a few counts of operations. Hence, the computational cost for the American option generally depends on solving the first step in (11.3) for $\tilde{U}_i^m$.

It is left to see how do we compute the pair (U_i^1, Ψ_i^1) on the first time level. For $m = 1$, the value U^1 is given by the payoff function and Ψ^1 is evaluated in Section 11.3. Regarding the second time level, the pair (U_i^2, Ψ_i^2) can be obtained through two steps: the first is to compute an intermediate approximation $\tilde{U}_i^1$ as below:

$$\frac{\tilde{U}_i^2 - U_i^1}{k} - \left\{\mathcal{D}_\Delta \tilde{U}_i^2 - (r+\lambda)\tilde{U}_i^2 + \mathcal{I}_\Delta U_i^1\right\} = \Psi_i^1. \quad (11.11)$$

The vector $\tilde{U}^2$ on the second time level is obtained by applying the implicit method to the differential operator and the explicit method to the integral operator. The second step is to find the pair (U_i^2, Ψ_i^2) such that

$$(U_i^2, \Psi_i^2) = \begin{cases} \Big(h(x_i), \Psi_i^1 + \dfrac{h(x_i) - \tilde{U}_i^2}{k}\Big), & \text{if } \tilde{U}_i^2 - k\Psi_i^1 \leq h(x_i), \quad (11.12) \\ \Big(\tilde{U}_i^2 - k\Psi_i^1, 0\Big), & \text{otherwise.} \quad (11.13) \end{cases}$$

11.2 Algorithm to evaluate an American put option with three time levels

To solve the LCP (10.2)-(10.4) numerically, we propose the implicit method with three time levels combined with the operator splitting method as follows:

Algorithm for pricing American options

For $m = 1$
for $i = 2, 3, ..., N-1$
$\frac{\tilde{U}_i^m - \tilde{U}_i^{m-1}}{k} = \mathcal{D}_\Delta \tilde{U}_i^m - (r+\lambda)\tilde{U}_i^m + \mathcal{I}_\Delta U_i^{m-1} + \Psi_i^{m-1}$
end for
Solve the tridiagonal linear system.

for $i = 2, 3, ..., N-1$
$U_i^m = \max\left(\tilde{U}_i^m - k\Psi_i^{m-1}, h(x_i)\right)$
$\Psi_i^m = \Psi_i^{m-1} + \frac{U_i^m - \tilde{U}_i^m}{k}$
end for
end

For $m = 2, 3, ..., M-1$
for $i = 2, 3, ..., N-1$
$\frac{\tilde{U}_i^m - U_i^{m-2}}{2k} = \mathcal{D}_\Delta\left(\frac{\tilde{U}_i^m + U_i^{m-2}}{2}\right) + \mathcal{I}_\Delta U_i^{m-1} - (r+\lambda)U_i^{m-1} + \Psi_i^{m-1}$
end for
Solve the tridiagonal linear system.

for $i = 2, 3, ..., N-1$
$U_i^m = \max\left(\tilde{U}_i^m - 2k\Psi_i^m, h(x_i)\right)$
$\Psi_i^m = \Psi_i^{m-1} + \frac{U_i^m - \tilde{U}_i^m}{2k}$
end for
end

Initial condition:
$U_i^1 = h(x_i) \quad$ for $i = 1, \ldots, N$
Boundary condition:
$U_i^m = h(x_i) \quad$ for $m = 2, 3, \ldots, M$ and $i = 1, N$

11.3 The early exercise boundary at the expiration date

We need to know the value of the auxiliary variable Ψ^1 at the expiration date in order to implement numerical methods for the American option. The early exercise boundary $S_p(\tau)$ for the American put option is given by

$$S_p(\tau) = \max\{e^x > 0 \mid u(\tau, x) \leq K - e^x, x \in \mathbb{R}\}.$$

Let us consider the early exercise boundary under the jump-diffusion model. Assume hypothetically that the asset price is in the exercise region. By definition, the price of the American put option is $u(\tau, x) = K - e^x$. Knowing u obeys to the PIDE (5.4), we compute $u_\tau - \mathcal{L}u$, using the fact that $u(\tau, x) = K - e^x$ is its solution in this case:

$$\begin{aligned}
u_\tau - \mathcal{L}u &= \frac{\partial u}{\partial \tau} - \frac{\sigma^2}{2}\frac{\partial^2 u}{\partial^2 x} - (r - q - \frac{\sigma^2}{2} - \lambda\zeta)\frac{\partial u}{\partial x} \\
&\quad + (r + \lambda)u - \lambda \int_{\mathbb{R}} u(\tau, x)p(z - x)dz \\
&= 0 + \frac{\sigma^2}{2}e^x + (r - q - \frac{\sigma^2}{2} - \lambda\zeta)e^x \\
&\quad + (r + \lambda)(K - e^x) - \lambda \int_{\mathbb{R}} u(\tau, x)p(z - x)dz \\
&= (r + \lambda)K - (q + \lambda\zeta + \lambda)e^x - \lambda \int_{\mathbb{R}} u(\tau, x)p(z - x)dz \\
&= rK - qe^x - \lambda\left[\int_{\mathbb{R}} u(\tau, x)p(z - x)dz - K + \zeta e^x + e^x\right] \\
&= rK - qe^x - \lambda\left[\int_{\mathbb{R}} u(\tau, x)p(z - x)dz - (K - e^x)\int_{\mathbb{R}} p(z - x)dz\right] \\
&= rK - qe^x - \lambda \int_{\mathbb{R}} \{u(\tau, x) - (K - e^x)\}p(z - x)dz.
\end{aligned}$$

Notice that we use the fact that p is a probability function (its integral over $\mathbb{R}$ is 1) and the definition of ζ $\left(\zeta = \int_{\mathbb{R}} e^z p(z)dz - 1\right)$.

We then got the equation:

$$u_\tau - \mathcal{L}u = rK - qe^x - \lambda \int_{\mathbb{R}} [u(\tau, z) - (K - e^z)]p(z - x)dz. \tag{11.14}$$

Let us now analyse the integrand of the integral in (11.14). At the expiration date,

$u(\tau, z) - (K - e^z) = 0$ for $z < \ln(K)$. Thus, the integral term in (11.14) is zero for $z < \ln(K)$ at the expiration date. Otherwise, $u(\tau, z) = 0$. As explained in the previous chapter, (11.1) is equal to zero, when it is not optimal to exercise the option, so, it makes sense to evaluate for which value does the right-hand side of (11.14) is null.

$$
\begin{aligned}
& rK - qe^x - \lambda \int_{\mathbb{R}} [u(\tau, z) - (K - e^z)] p(z - x) dz = 0 \\
\iff & rK - qe^x - \lambda \int_{ln(K)}^{\infty} [-(K - e^z)] p(z - x) dz = 0 \\
\iff & -qe^x - \lambda \int_{ln(K)}^{\infty} e^z p(z - x) dz = -rK - \lambda \int_{ln(K)}^{\infty} K p(z - x) dz \\
\iff & -qe^x - \lambda \int_{ln(\frac{K}{e^x})}^{\infty} e^{z+x} p(z) dz = -rK - K\lambda \int_{ln(\frac{K}{e^x})}^{\infty} p(z) dz \\
\iff & e^x \Big(-q - \lambda \int_{ln(\frac{K}{e^x})}^{\infty} e^z p(z) dz \Big) = -rK - K\lambda \int_{ln(\frac{K}{e^x})}^{\infty} p(z) dz \\
\iff & e^x = K \frac{r + \lambda \int_{ln(K/e^x)}^{\infty} p(z) dz}{q + \lambda \int_{ln(K/e^x)}^{\infty} e^z p(z) dz}.
\end{aligned}
$$

The right-hand side in (11.14) is then null when e^x is equal to:

$$
e^x = K \frac{r + \lambda \int_{ln(K/e^x)}^{\infty} p(z) dz}{q + \lambda \int_{ln(K/e^x)}^{\infty} e^z p(z) dz}.
$$

As for the case of its European counterpart, American style options should only be exercised at the expiration date if the payoff is non-zero, i.e., for the case of the put option this happens when the price of the underlying asset is lower than the strike price K. Therefore, the early exercise boundary $S_p(0)$ at the expiration date is the minimum between the strike and the expression we have just got:

$$
S_p(0) = \min \left(K, K \frac{r + \lambda \int_{ln(K/S_p(0))}^{\infty} p(z) dz}{q + \lambda \int_{ln(K/S_p(0))}^{\infty} e^z p(z) dz} \right).
$$

In conclusion, the auxiliary variable φ^1 is given by

$$
\varphi_i^1 = \begin{cases} rK - qe^{x_i} + \lambda \int_{ln(K)^{\infty}} (K - e^z) p(z - x_i) dz, & \text{if } e^{x_i} \leq S_p(0), \quad (11.15) \\ 0, & \text{if } e^{x_i} > S_p(0). \quad (11.16) \end{cases}
$$

In the case of the American call option under the jump-diffusion model, the early exercise boundary at the expiration date can be found in [13]. For our numerical simulations, we used the formula above for the auxiliary variable φ^1.

12 Numerical results for American options

In this section, we provide numerical results for the American options case, under the Merton and Kou models. We simulate results, using the implicit method with three time levels as described in the previous chapters, which leads to linear systems involving tridiagonal coefficient matrices, similarly to the European case. As previously done, we decomposed these matrices recurring to the LU decomposition with $O(N)$ operations and also applied the FFT to compute the product of a dense matrix and a column vector in the discrete integral operator. In order to treat the inequality constraints in the LCP (10.2)-(10.4) for the American option, we combined the implicit method with the operator splitting method suggested by Ikonen and Toivanen [31].

All the simulations presented here were computed in MATLAB using a Macbook Pro with processor 2,3 GHz Intel Core i5.

12.1 Results for the Merton model

We start with the Merton model, which has the density function of jump sizes given by (3.1). Parameters used in the simulation are the following:

$$\sigma = 0.15, \quad r = 0.05, \quad \sigma_J = 0.45, \quad \mu_J = -0.90,$$

$$d = 0, \quad \lambda = 0.10, \quad T = 0.25, \quad K = 100.$$

In [23], the authors develop an iterative method for the solutions of both the fully implicit and Crank–Nicolson discretizations coupled with the penalty method to solve

the LCP. The reference values are evaluated with 6 digits after the decimal point by d'Halluin, Forsyth, and Vetzal [22]. We calculated the fair price of vanilla options for three different initial values and a unique strike ($K = 100$). In order to keep the accuracy of our method larger than second-order, we build a cubic spline to evaluate the prices of the option at non-mesh points of stock prices. The results can be found in Table 3 and as the reader might notice, the pointwise errors are decreasing with the increase of the number of steps.

Table 3: Prices of standard American-style put options under the Merton jump-diffusion model

Steps	Spot	Value	Error	CPU (sec)
$M = 25$	90	9.986550	-	0.009034
$N = 128$	100	3.213501	-	0.011231
	110	1.407122	-	0.010548
$M = 50$	90	10.001556	0.015006	0.035964
$N = 256$	100	3.231315	0.017814	0.031529
	110	1.418904	0.011782	0.032150
$M = 100$	90	10.004459	0.002903	0.073928
$N = 512$	100	3.238459	0.007144	0.073035
	110	1.419276	0.000372	0.078562
$M = 200$	90	10.003350	0.001109	0.254753
$N = 1024$	100	3.240487	0.002028	0.255402
	110	1.419365	0.000089	0.228199
$M = 400$	90	10.003909	0.000559	0.834073
$N = 2048$	100	3.241060	0.000573	0.855441
	110	1.419783	0.000418	0.765838
$M = 800$	90	10.003845	0.000064	3.283936
$N = 4096$	100	3.241206	0.000146	3.236129
	110	1.419795	0.000012	3.178400

Values of American put options obtained by the implicit method with three time levels under the Merton model. The reference values are 10.003822 at $S = 90$, 3.241251 at $S = 100$, and 1.419803 at $S = 110$. The truncated domain is $[-2.5, 2.5]$, for the variable $x = \ln(S/S_0)$. M is the number of time steps and N is the number of spatial steps. The third column represents the computed values corresponding to each strike and the fourth the difference between the values computed and the ones calculated in the previous rows with less steps. The last column shows how much computation time we needed.

In Figure 12.1, we plot the prices of an American put option, for different values of time-to-maturity and spot levels. Figure 12.2 shows the early exercise boundary, for

each value of time-to-maturity, ranging from 0 to 0.25 years. As one can observe, the early exercise boundary is continuous and non-increasing. Besides, on the boundary, the Delta (first derivative of the option price with respect to the price of the underlying asset) is equal to -1, as proven in Theorem A.1. This property is known as the "smooth pasting" or "tangency" condition.

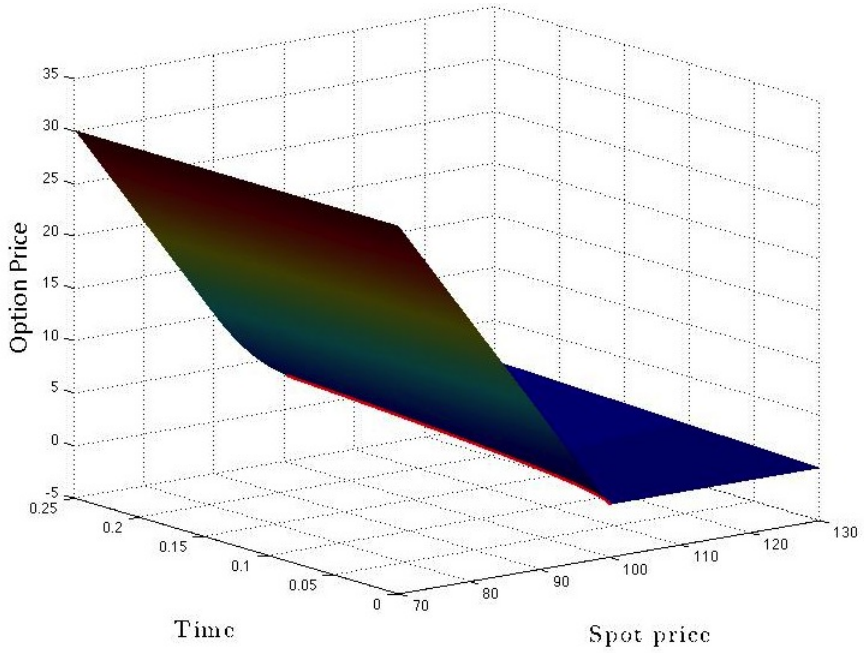

Figure 12.1: Price of an American put option under the Merton model

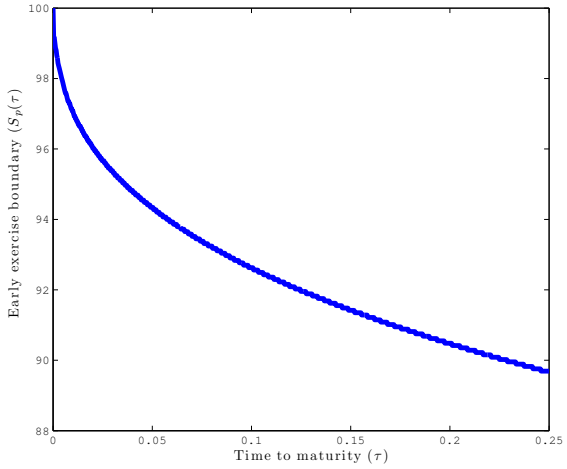

Figure 12.2: Early exercise boundary for the Merton model

12.2 Results for the Kou model

For the Kou model, whose density function is given by (3.6), the parameters used are:

$$\sigma = 0.15, \quad r = 0.05, \quad \eta_1 = 3.0465, \quad \eta_2 = 3.0775,$$
$$d = 0, \quad \lambda = 0.10, \quad T = 0.25, \quad K = 100 \quad p = 0.3445.$$

As done for the Merton model case, we compare our results with the ones presented in [22]. They are shown in Table 4.

Table 4: Prices of standard American-style put options under the Kou jump-diffusion model

Steps	Spot	Value	Error	CPU (sec)
$M = 25$	90	10.001862	-	0.007468
$N = 128$	100	2.800028	-	0.006572
	110	0.566056	-	0.008916
$M = 50$	90	10.006155	0.004293	0.017521
$N = 256$	100	2.805095	0.005067	0.017315
	110	0.562595	0.003461	0.016726
$M = 100$	90	10.004407	0.001748	0.058711
$N = 512$	100	2.807180	0.002085	0.052383
	110	0.562595	0.000735	0.057087
$M = 200$	90	10.005297	0.000890	0.188461
$N = 1024$	100	2.807682	0.000502	0.186383
	110	0.561688	0.000172	0.189463
$M = 400$	90	10.005157	0.000140	0.737610
$N = 2048$	100	2.807810	0.000128	0.670359
	110	0.561865	0.000177	0.687775
$M = 800$	90	10.005096	0.000061	3.487136
$N = 4096$	100	2.807840	0.000003	3.417621
	110	0.561865	0.000000	3.137204

Values of American put options obtained by the implicit method with three time levels under the Kou model. The reference values are 10.005071 at $S = 90$, 2.807879 at $S = 100$. and 0.561876 at $S = 110$. The truncated domain is $[-1.5, 1.5]$, for the variable $x = \ln(S/S_0)$. M is the number of time steps and N is the number of spatial steps. The third column represents the computed values corresponding to each strike and the fourth the difference between the values computed and the ones calculated in the previous rows with less steps. The last column shows how much computation time we needed.

In Figure 12.3, we plot the prices of an American put option, for different times-to-maturity and spot values. Figure 12.4 shows the early exercise boundary, for each value of time-to-maturity, ranging from 0 to 0.25 years.

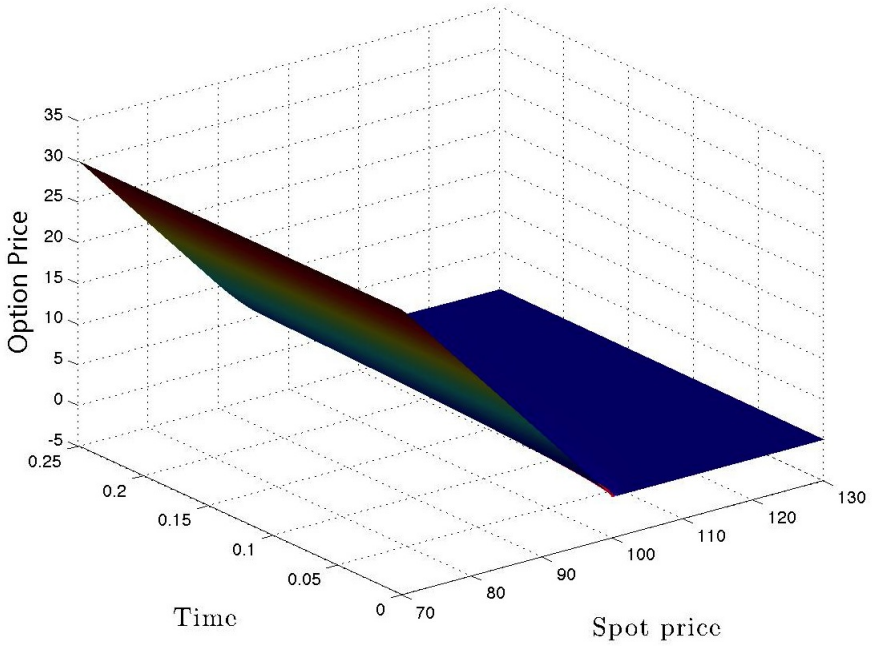

Figure 12.3: Price of an American put option under the Kou model

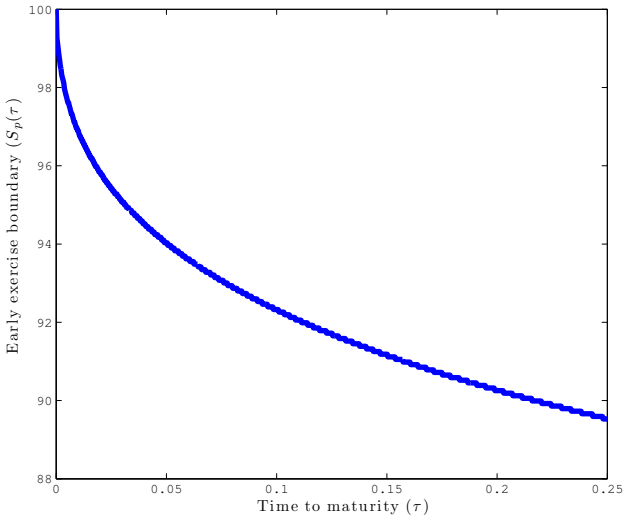

Figure 12.4: Early exercise boundary for the Kou model

13 Brief discussion on Greeks

Let us quote the following excerpt taken from the 17th Chapter of Hull [29]:

"A financial institution that sells an option to a client in the over-the-counter market is faced with the problem of managing its risk. If the option happens to be the same as one that is traded on an exchange, the financial institution can neutralize its exposure by buying on the exchange the same option as it has sold. But when the option has been tailored to the needs of a client and does not correspond to the standardized products traded by exchanges, hedging the exposure is far more difficult." One tool to manage risk is using what is called in the financial literature as "Greeks" or "sensitivities". According to [29], "each Greek letter measures a different dimension to the risk in an option position and the aim of a trader is to manage the Greeks so that all risks are acceptable. [...] Most traders use more sophisticated hedging schemes than those mentioned so far. These involved calculating measures such as delta, gamma and vega.".

In general, it is a trader's goal to immunize the portfolio value, denoted as Π, against small changes in the underlying asset price S. To achieve this, he must ensure the Delta equals zero. The Delta of an option is defined as the rate of change of the option price with respect to the price of the underlying asset. It is the slope of the curve that relates the portfolio value to the underlying price: $\frac{\partial \Pi}{\partial S}$.

In the case this is not possible, one could obviously sell the entire portfolio and invest the sum obtained in the bank, but this is neither usually feasible nor preferable. Another more interesting alternative is to add a derivative to the portfolio. Since the price of a derivative is by definition correlated with the underlying asset price, we should

be able to balance the derivative against the portfolio in such a way that the adjusted portfolio becomes Delta neutral.

On the other hand, in order not to be adjusting the Delta very frequently, it is desirable to maintain the Gamma of a portfolio the closest to zero as possible. The Gamma of a portfolio of contracts on an underlying asset is the rate of change of the portfolio's Delta with respect to the price of the underlying asset. It is the second partial derivative of the portfolio with respect to asset price: $\frac{\partial^2 \Pi}{\partial S^2}$.

If Gamma is small, Delta changes slowly and adjustments to keep a portfolio Delta neutral need to be made only relatively infrequently. However, if the absolute value of Gamma is large, Delta is highly sensitive to the price of the underlying asset. It is then quite risky to leave the portfolio highly sensitive to the price of the underlying asset.

For a more profound study on the Greeks and their intricacies, the reading of Hull [29] or Bjork [9] is suggested. We remark that given the importance of the Greeks in Risk Management, the numerical methods and models presented so far also allow for the computation of these, both for the European and American cases. As an example, we plot in Figures 13.1 and 13.2, respectively, the Delta and the Gamma of a European put option under the Merton model, while Figures 13.3 and 13.4 show the same Greeks for the American case, using the following parameters:

$$\sigma = 0.15, \quad r = 0.05, \quad \sigma_J = 0.45, \quad \mu_J = -0.90$$

$$d = 0, \quad \lambda = 0.10, \quad T = 0.25, \quad K = 100.$$

According to [38], option pricing problems frequently have nonsmooth payoff functions, which lead numerical solutions to possibly oscillate on the final time level. In order to overcome this problem, the solutions on a few initial time levels are estimated by the explicit-implicit method in the second time level and thereafter the solutions on the rest of the time levels are obtained by the implicit method with three time levels. However, with our method we can overcome this flaw, as shown in the Figures previously referred. There is no evidence of oscillations in the Delta with respect to the stock price and the Gamma too, except in the point where Delta has an angle point.

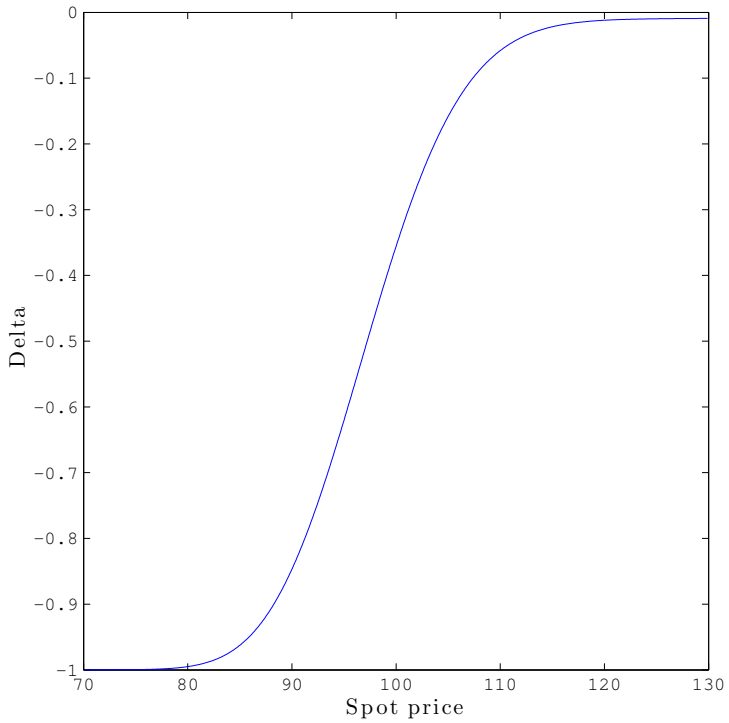

Figure 13.1: Delta of a European put option under the Merton model

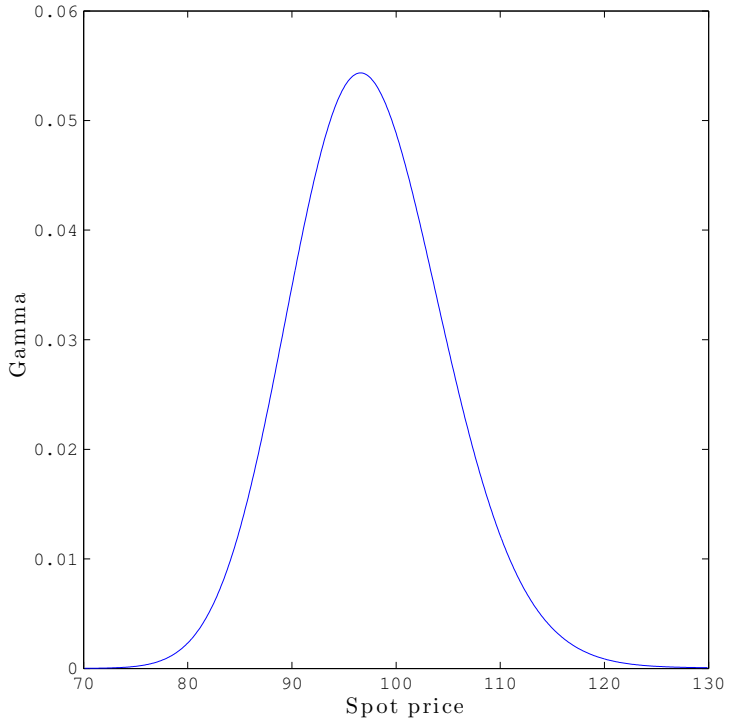

Figure 13.2: Gamma of a European put option under the Merton model

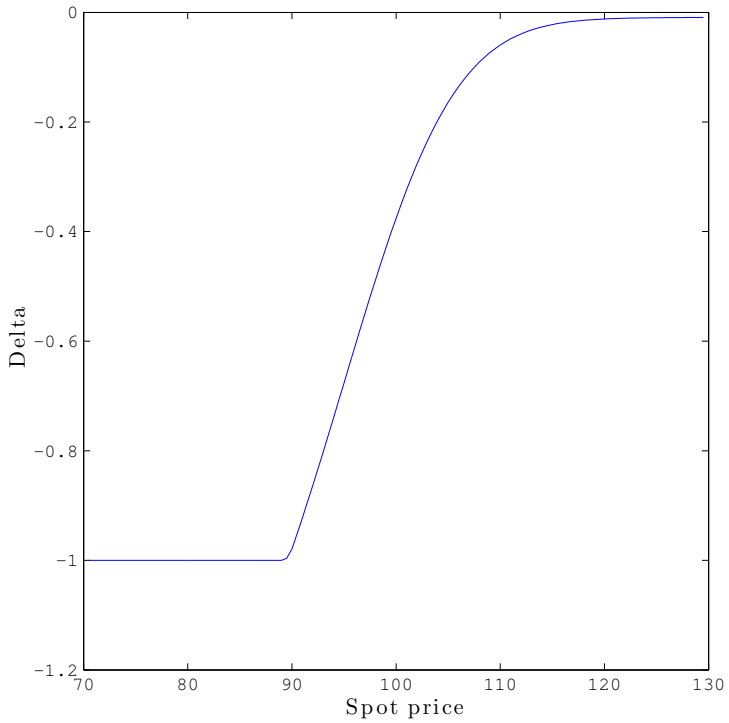

Figure 13.3: Delta of an American put option under the Merton model

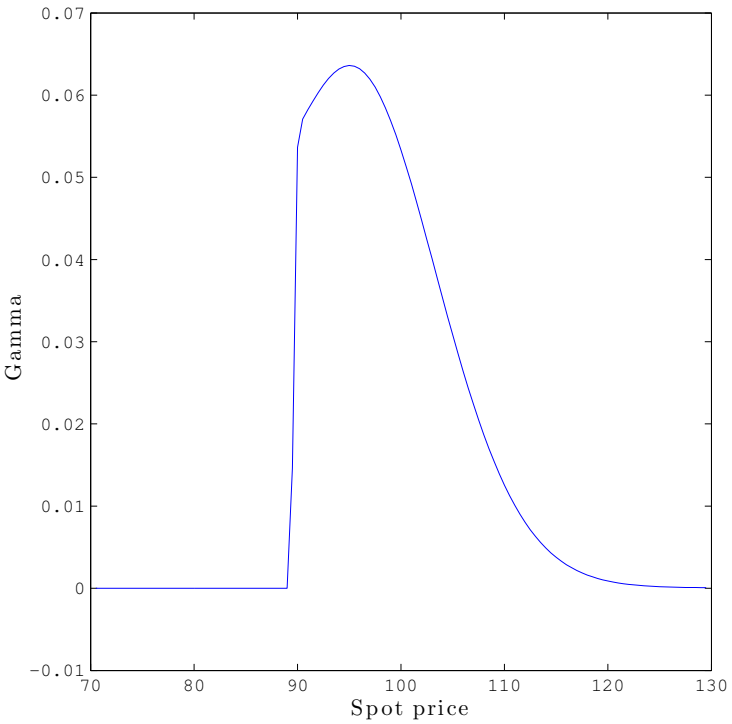

Figure 13.4: Gamma of an American put option under the Merton model

14 Conclusions & future Research

In this thesis several models and methods for pricing European and American-style options were explained. After presenting the basic concepts and results related to the Option Pricing theory, we introduced two jump-diffusion models, respectively attributing a normal and a double-exponential distribution to the jump part of equation (2.3): the Merton and the Kou models. These assume the underlying asset follows a Lévy process. The latter represents an improvement in relation to the diffusion models.

In order to apply in practice our theoretical results and accomplish our goal of obtaining prices for certain financial contracts, such as vanilla options, we established the dependency of an option's price on time and spot value through a PIDE and used the latter to build a numerical method involving three time levels. Such numerical method was first used to solve the PIDE, aiming to get prices of European-style options and, in our approach, we tried to approximate the differential operator by finite differences in such a way that we would obtain a scheme similar to the Crank-Nicolson's. The integral operator is approximated using the trapezoidal rule, thus ensuring second-order accuracy.

With this procedure, we were able to build a system of equations involving tridiagonal matrices. This way, we could easily apply LU decomposition which only requires $O(N)$ operations, where N stands for the number of spatial steps. The FFT is used to evaluate efficiently the product of a dense matrix and a column vector. These features lead our numerical experiments to be computed in just a few seconds.

We showed our method has good numerical properties too. In particular, we proved

it is consistent, stable and it has second-order convergence rate in the discrete l^2-norm.

To test the quality of our method and also the features described in the last paragraph, we then compared our results with semi-closed pricing formulae available in the literature. In particular for the Merton's case, the options' prices were given by a weighted average of Black-Scholes prices. In the Kou model, the solution was given in terms of special mathematical functions, which were evaluated using Gauss quadratures. Moreover, we also confronted our values with similar results published in the literature, corroborating the fact that our method is well conceived for, at least, the vanilla options case.

Furthermore, we were able to extend our method to American-style options, proposing a similar method to solve the LCP (10.2)-(10.4). We presented a combination of the implicit method with three time levels and the operator splitting technique.

We implemented several numerical experiments for American-style put options under the Merton and Kou models. Our results also show that our procedure leads to options prices which are consistent with the ones found in the literature, where other methods were applied.

Besides the prices of American-style put options, early exercise boundaries for both the Merton and Kou models were plotted. This way, the investor has an idea on how to make his choice, deciding for which values should he exercise the option.

With our procedure, we were also able to compute the Greeks such as the Delta and the Gamma, both for European and American-style put options, allowing a better measurement of Risk when trading options.

Despite the accomplishments attained in this thesis and the fact that vanilla options are very liquid contracts in the financial markets, it would be useful to extend this framework to more complex payoffs such as Asian and Lookback options and contracts like Correlation or Variance Swaps. For the latter, it is thoughtful to consider models that combine both jumps and stochastic volatility. As an example, Bates [5] combines the Merton jump-diffusion model and the Heston stochastic volatility model, by defining two stochastic differential equations: one for the asset's price and another for the

volatility associated with the Brownian motion.

It is also possible to consider the rate of arrival of jumps has its own distribution as it happens in [20], where its dynamics are defined through its own stochastic differential equation. However, since stochastic volatility implies correlation between the increments of the process, these models lie beyond the scope of Lévy processes.

A Appendix

A.1 Graphs

Within this part of the Appendix it is possible to visualize the graphs which were referred along the thesis.

Figure A.1: S&P500 stock prices

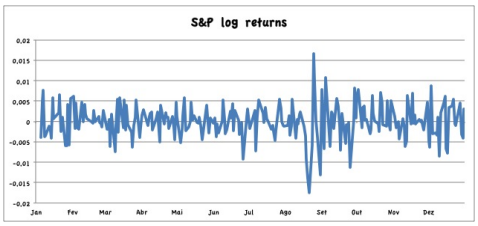

Figure A.2: S&P500 log returns

The first picture represents the closing price values of the index S&P500 from the beginning to the end of 2015. Using this data, the log-returns are computed, by simply subtracting the logarithm of exponential basis (denoted by "log" from now on) of the value in one day to the log of the following day. The chart A.2 represents these calculations.

The same was done for the EUR/CHF. Figure A.3 represents the closing price values of the EUR/CHF through the year of 2015 and Figure A.6 the log-returns.

Using the log-returns values of these data sets, an histogram of the log-returns for both asset's was conceived. They correspond to the Figures A.5 and A.6.

In Figure A.7, we plot a simulation of a Poisson process with rate $\lambda = 0.1$. The

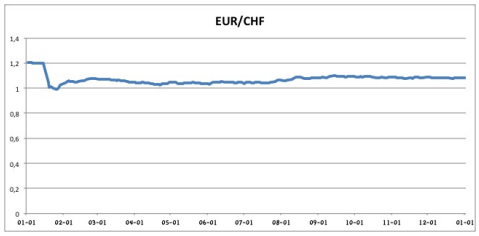

Figure A.3: EUR_CHF values

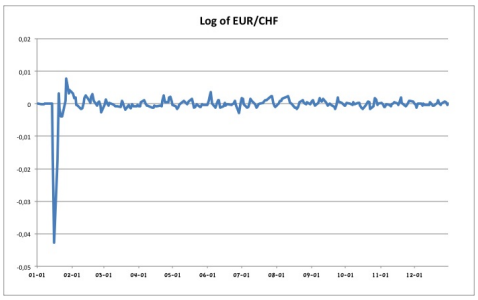

Figure A.4: EUR_CHF log returns

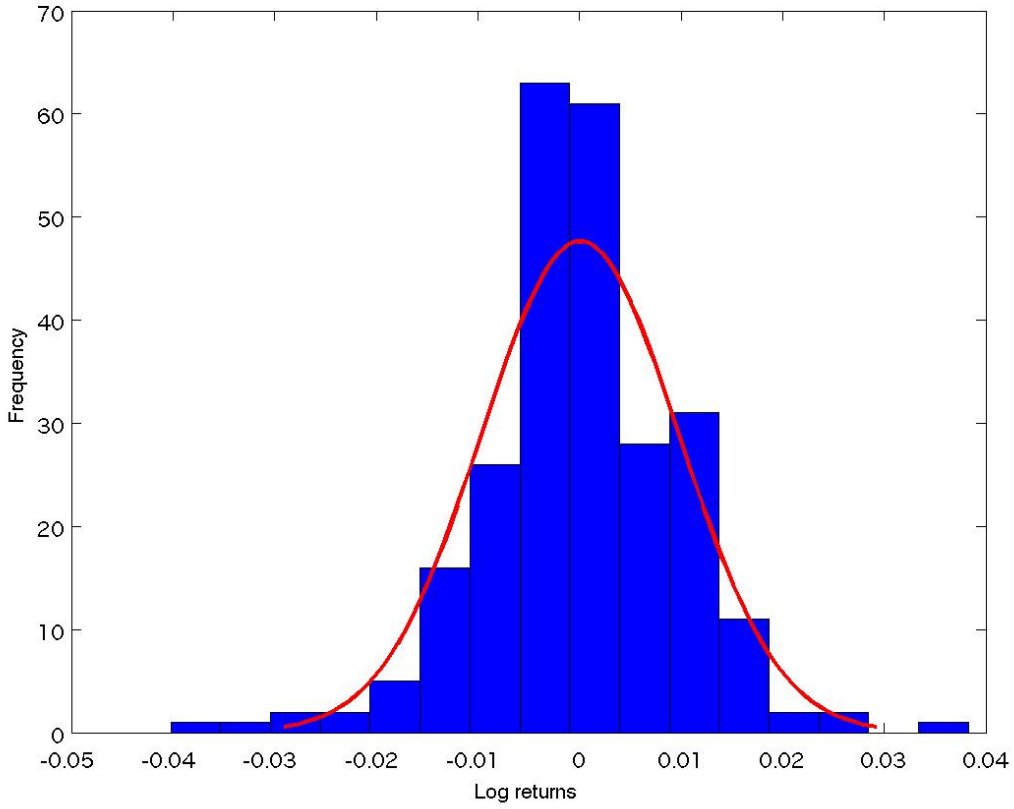

Figure A.5: S&P500 log returns histogram

horizontal axis represents the time and the vertical the value of the random variable. As one can see, the plot exhibits right-continuity and is constant between jumps.

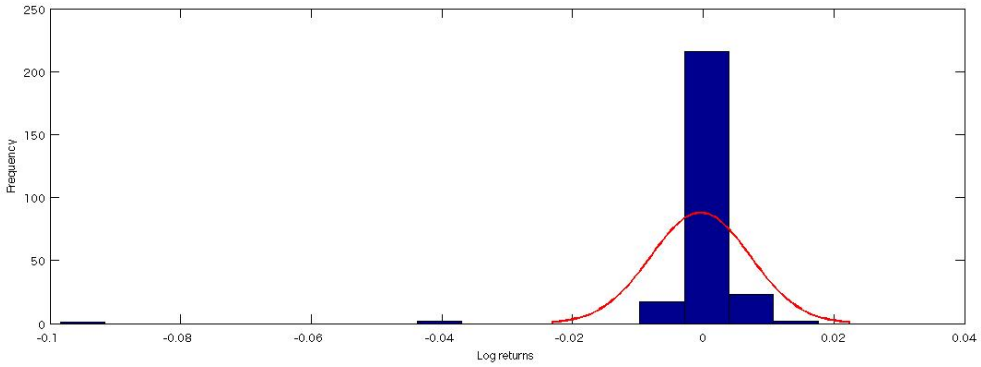

Figure A.6: EUR/CHF log returns histogram

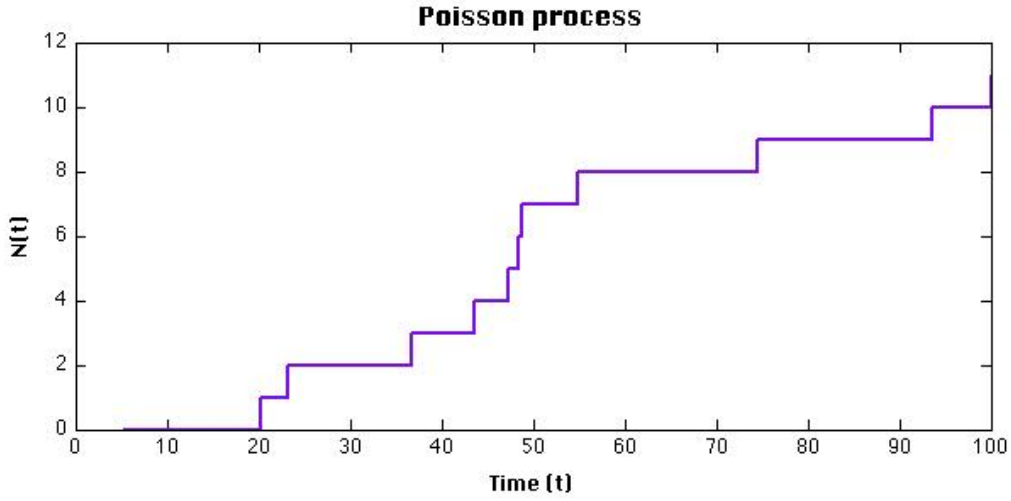

Figure A.7: Poisson process

A.2 Other results

Theorem A.1. *An American put with value p must respect the following equality, known as the "smooth pasting" or "tangency" condition:*

$$\frac{\partial p}{\partial s}(\tau, s^*(\tau)) = -1, \tag{A.1}$$

where $\tau \in [0, T]$ and $s^*(\tau)$ is the optimal exercise boundary price. For the call option, the result is similar, with the only difference that the delta of the option is 1 instead. The proof of this theorem is based on [36], but it can also be found in [41].

Proof. Take arbitrarily $\tau \in [0, T]$ and define as $F(b, s)$ the solution of (10.2)-(10.4), in the continuation region $\{(\tau, s) \in]0, T] \times \mathbb{R}^+ \mid s > b(\tau)\}$, where s stands for the underlying's price and b for the boundary, which is now assumed to be known.

$F(b, s)$ is a differentiable function and $F(b, b) = K - b$, with K being the strike price. The total derivative of F with respect to b along the boundary $s = b$ is given by:

$$\frac{dF}{db} = \frac{\partial F}{\partial s}\frac{\partial s}{\partial b}(b, s)\bigg|_{s=b} + \frac{\partial F}{\partial b}(b, s)\bigg|_{s=b}. \tag{A.2}$$

It is obvious that $\frac{\partial s}{\partial b}(b, s) = 1$, since $s = b$.

The holder of the American put chooses an early exercise policy which maximizes the value of the put for all possible continuous functions $b(\tau)$. Therefore, defining by b^* the critical value of b which maximizes F, we have $\frac{\partial F}{\partial b}(b^*, s) = 0$.

On the other hand, from the exercise payoff function of the American put option, we have

$$\frac{dF}{db}\bigg|_{b=b^*} = \frac{d}{db}(K - b)\bigg|_{b=b^*} = -1. \tag{A.3}$$

From the argumentation above, we obtain $\frac{\partial F}{\partial s}(b^*, s)\Big|_{s=b^*} = -1$. Note that the optimal choice $b^*(\tau)$ is just the optimal exercise price $s^*(\tau)$. The above condition can then be expressed in an alternative form as $\frac{\partial p}{\partial s}(\tau, s^*(\tau)) = -1$. ■

Theorem A.2. *Given two independent and continuous random variables X and Y with density functions respectively given by f_X and f_Y, the distribution of the their sum*

$Z = X + Y$ is given by the convolution of X and Y.

Proof. Given that the variables X and Y are assumed to be independent, their joint density is equal to $f_X(x)f_Y(y)$.

$$\begin{aligned} F_Z(z) &= \mathbb{P}\{X + Y \leq z\} \\ &= \int\limits_{x+y\leq z} f_X(x)f_Y(y)dxdy \\ &= \int\limits_{-\infty}^{\infty}\int\limits_{-\infty}^{z-y} f_X(x)f_Y(y)dxdy \\ &= \int\limits_{-\infty}^{\infty} F_X(z-y)f_Y(y)dy. \end{aligned}$$

Now, applying the Fundamental Theorem of Calculus, one may obtain the density of $Z = X + Y$ by differentiating the respective cumulative density function (CDF):

$$\begin{aligned} f_Z(z) &= \frac{d}{dz}F_Z(z) \\ &= \int\limits_{-\infty}^{\infty} f_X(z-y)f_Y(y)dy. \end{aligned}$$

The last expression matches the convolution of the density functions of X and Y:

$$f_Z(z) := f_{X+Y}(z) = (f_X * f_Y)(z).$$

Analogously, we can prove that

$$\begin{aligned} F_Z(z) &= \int\limits_{-\infty}^{\infty}\int\limits_{-\infty}^{z-x} f_Y(y)f_X(x)dydx \\ &= \int\limits_{-\infty}^{\infty} F_Y(z-x)f_X(x)dx. \end{aligned}$$

and therefore $f_Z(z) = \int\limits_{-\infty}^{\infty} f_Y(z-x)f_X(x)dx$.

We then conclude our claim. ■

Theorem A.3. *Let $n \in \mathbb{N}$ and $A = (a_{ij}) \in \mathcal{M}_{n\times n}, i, j \in \{1, ..., n\}$ be a strictly diagonally dominated matrix. A is invertible.*

Proof. Suppose $A = (a_{ij})$ is a strictly diagonally dominated matrix. Towards a contra-

diction, assume it is non-singular. Then, 0 is an eigenvalue of A. Let $u = (u_j) \in \mathbb{R} \setminus \{0\}$ be an eigenvector associated to the eigenvalue 0. Since u in not the null vector, there exists u_m, the greatest value of the set of u components.

Following the definition of an eigenvector and considering any row i:

$$\begin{aligned} Au &= 0 \\ \sum_j a_{ij} u_j &= 0 \\ a_{mm} u_m &= -\sum_{j \neq i} a_{ij} u_j \\ a_{mm} &= -\sum_{j \neq i} a_{ij} \frac{u_j}{u_m} \end{aligned}$$

The latter implies that

$$\begin{aligned} |a_{mm}| &\leq \sum_{j \neq i} \left| a_{ij} \frac{u_j}{u_m} \right| \\ |a_{mm}| &\leq \sum_{j \neq i} |a_{ij}|. \end{aligned}$$

Given that we reached a contradiction, we conclude A is invertible. ■

Theorem A.4. *Let $n \in \mathbb{N}$ and $A = (a_{ij}) \in \mathcal{M}_{n \times n}, i, j \in \{1, ..., n\}$ be a regular square matrix.*

$\|A\|_2 := \max_{x \neq 0} \frac{\|Ax\|_2}{\|x\|_2} = \sqrt{\lambda_{\max}(A^T A)}$, *where $\lambda_{\max}(A^T A)$ denotes the greatest eigenvalue of $A^T A$.*

Proof. First, we prove the l^2-norm does not change its value under orthogonal transformations:

Let $x \in \mathbb{R}^n$.

$$\|Qx\|_{l^2} = \sqrt{(Qx)^T (Qx)} = \sqrt{x^T Q^T Q x} = \sqrt{x^T x} = \|x\|_{l^2},$$

since $Q^T Q = I$ for an orthogonal matrix.

$\|A\|_2 = \max_{x \neq 0} \frac{\|Ax\|_{l^2}}{\|x\|_{l^2}} = \max_{x \neq 0} \frac{(x^T A^T A x)^{\frac{1}{2}}}{\|x\|_{l^2}}$.

Since $A^T A$ is symmetric, it can be decomposed as $A^T A = Q \Lambda Q^T$, where Λ is a

diagonal matrix containing the eigenvalues of A^TA, which are all real. Then

$$\|A\|_2 = \max_{x\neq 0} \frac{(x^T(Q\Lambda Q^T)x)^{\frac{1}{2}}}{\|x\|_{l^2}} = \max_{x\neq 0} \frac{((Q^Tx)^T\Lambda(Q^Tx))^{\frac{1}{2}}}{\|Q^Tx\|_{l^2}}$$

by orthogonal transformations.

On the other hand, A^TA is positive definite. In fact:

$$x^TAA^Tx = (A^Tx)^T(A^Tx) > 0, \forall x \in \mathbb{R}\setminus\{0\}.$$

Thus, all the eigenvalues of A^TA are positive.

Define $y := Q^Tx$.

$$\|A\|_2 = \max_{y\neq 0} \frac{(y^T\Lambda y)^{\frac{1}{2}}}{\|y\|_{l^2}} = \max_{y\neq 0} \sqrt{\frac{\sum_{i=1}^n \lambda_i y_i^2}{\sum_{i=1}^n y_i^2}} \leq \sqrt{\lambda_{\max}\frac{\sum_{i=1}^n y_i^2}{\sum_{i=1}^n y_i^2}} = \sqrt{\lambda_{\max}}.$$

To see we can indeed attain this value for the 2-norm, pick y as the jth column of the identity matrix, considering that λ_j is the largest eigenvalue. This leads to:
$\frac{(y^T\Lambda y)^{\frac{1}{2}}}{\|y\|_{l^2}} = \frac{\sqrt{\lambda_{\max}}}{1} = \sqrt{\lambda_{\max}}$.

We then proved our claim. ■

Theorem A.5. *Let $A = (a_{jk})$ be a regular square matrix with p rows and columns. (a_{jk}) denotes the element in the jth row and kth column of A. The following equality holds true:*

$\|A\|_\infty := \max_{x\neq 0} \frac{\|Ax\|_{l^\infty}}{\|x\|_{l^\infty}} = \max_j \sum_{k=1}^p |a_{jk}|$.

Proof. Let x be a vector of dimension p such that $\|x\|_{l^\infty} = 1$. By definition, it is obvious that $|x_k| \leq 1, \forall k \in \{1,\ldots,p\}$ and $|x_i| = 1$ for some i between 1 and p. Then:

$\|Ax\|_{l^\infty} := \max_j \left|\sum_{k=1}^p a_{jk}x_k\right| \leq \max_j \sum_{k=1}^p |a_{jk}|$.

Hence, in particular, the maximum row sum is always greater than or equal to the infinity vector norm of A.

Conversely, suppose the maximum row sum is obtained from row j of the matrix A. Then, choose the vector x defined by

$$x_k = 1, \quad \text{if } a_{jk} \geq 0,$$

$$x_k = -1, \quad \text{if } a_{jk} < 0.$$

Then $\|x\|_{l^\infty} = 1$ and

$$\|A\|_\infty \geq \|Ax\|_{l^\infty} \geq \Big|\sum_{k=1}^{p} a_{jk}x_k\Big| = \sum_{k=1}^{p} |a_{jk}|.$$

We can now conclude that $\|A\|_\infty = \max_j \sum_{k=1}^{p} |a_{jk}|$.

■

A.3 Finite differences

Here, we present some formulae which are the base for the approximations (7.10), (7.11) and (7.12).

Assume $u(\tau, x) \in C^2([0,T] \times \mathbb{R})$. According to Taylor's theorem, given $\tau_0 \in [0,T]$, a fixed $x_0 \in \mathbb{R}$ and $\Delta\tau > 0$, we have:

- $u(\tau_0 + \Delta\tau, x) \approx u(\tau_0, x) + u_\tau(\tau_0, x)\Delta\tau + \frac{u_{\tau\tau}(\tau_0,x)}{2}(\Delta\tau)^2$
- $u(\tau_0 - \Delta\tau, x) \approx u(\tau_0, x) - u_\tau(\tau_0, x)\Delta\tau + \frac{u_{\tau\tau}(\tau_0,x)}{2}(\Delta\tau)^2$

Subtracting:

$$u(\tau_0 + \Delta\tau, x) - u(\tau_0 - \Delta\tau_0, x) \approx 2u_\tau(\tau_0, x)\Delta\tau$$

$$\iff$$

$$u_\tau(\tau_0, x) \approx \frac{u(\tau_0 + \Delta\tau, x) - u(\tau - \Delta\tau, x)}{2\Delta\tau}. \tag{A.4}$$

Analogously, we have for the second variable:

$$u(\tau_0, x_0 + \Delta x) - u(\tau_0, x_0 - \Delta x) \approx 2u_\tau(\tau_0, x_0)\Delta x$$

$$\iff$$

$$u_\tau(\tau_0, x_0) \approx \frac{u(\tau_0, x_0 + \Delta x) - u(\tau, x_0 - \Delta x)}{2\Delta x}. \tag{A.5}$$

Summing up instead of subtracting, we get:

$$u(\tau_0, x_0 + \Delta x) + u(\tau_0, x_0 - \Delta x) \approx 2u(\tau_0, x_0) + u_{xx}(\tau_0, x_0)(\Delta\tau)^2$$

$$\iff$$

$$u_{xx}(\tau_0, x_0) \approx \frac{u(\tau_0, x_0) - 2u(\tau_0, x_0) + u(\tau_0, x_0 - \Delta x)}{(\Delta x)^2}. \tag{A.6}$$

A.4 Poisson processes

Let $(C_n)_{n>0}$ be a sequence of random positive variables strictly monotonically increasing ($0 = C_0 < C_1 < C_2 < \cdots$) and let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space ($\Omega, \mathcal{F}$ and $\mathbb{P}$ respectively represent a sample space, a σ-algebra and a probability measure). For each $\omega \in \Omega$, we define an indicator function as

$$\mathbb{1}_{\{C_n(\omega) \leq t\}} = \begin{cases} 1, & if \quad C_n(\omega) \leq t \\ 0, & if \quad C_n(\omega) > t \end{cases}$$

The process $C = (C_t)_{0 \leq t \leq \infty}$ defined by:

$$C_t = \sum_{n \geq 1} \mathbb{1}_{\{C_n \leq t\}}$$

is called the counting process associated to the sequence $(C_n)_{n \geq 1}$. A counting process is a stochastic process which counts the number of events that have occurred up to time t. It is non-negative and integer-valued for all $t \geq 0$. Furthermore, C_t is non-decreasing in t and $C_t - C_s$ equals the number of events in the time interval $(s, t]$, for any $s < t$. C_t could, e.g., denote the number of arrivals of customers at a railway station in $(0, t]$, or the number of accidents on a particular highway in that time interval or the number of calls to a telephone call-center during that period.

Many processes in everyday life that "count" events up to a particular point in time can be accurately described by the so-called Poisson process, named after the French scientist Siméon Poisson (1781-1840). A Poisson process is a counting process that has the desirable additional properties that the number of events in disjoint intervals are independent ("independent increments") and that the number of events in any given interval depends only on the length of that interval, and not on its particular position in time ("stationary increments"). In the case of the arrivals at the railway station, the stationary assumption is clearly not fulfilled; there will be many more arrivals between 5 P.M. and 6 P.M. than between, say, 5 A.M. and 6 A.M. Still, one might wish to study the arrival process at the railway station during the rush hour. Restricting oneself to

subsequent working days between 5 P.M. and 6 P.M. does allow one to use the stationary increments assumption.

From now on, we only consider non-explosive counting processes, that is, if $C = (C_t)_{0\leq t\leq\infty}$ is a counting process, it is defined as non-explosive when the event $\{inf\{t : C_t = \infty\} < \infty\}$ has null probability. In other words, it is explosive if the event $\{sup\{t : C_t = \infty\} = \infty\}$ has probability one almost surely (a.s.).

Definition A.6. *A stochastic process $\{N_t, t \geq 0\}$ is considered to possess independent increments if, for any $t_1, t_2, t_3, \cdots, t_n$, with $0 < t_1 < t_2 < \cdots < t_n$, we have that*

$$N_0, N_{t_1} - N_0, N_{t_2} - N_{t_1}, \cdots, N_{t_n} - N_{t_{n-1}}$$

are independent random variables. This means that the number of events occurred until the time t, N_t, is independent of the number of events that occurred between t and t+s (that is, $N_{t+s} - N_t$).

Definition A.7. *A stochastic process $\{N_t, t \geq 0\}$ has stationary increments if for any $s < t, N_t - N_s$ is equal in distribution to N_{t-s}. This implies that given $k \in \mathbb{R}$ and $s \geq t, P[N_{t+s} - N_t = k] = P[N_s = k]$, for any $t \geq 0, s > 0$.*

Theorem A.8. *Let $\{N_t, t \geq 0\}$ be a Poisson's process. Then*

$$\mathbb{P}(N_t = n) = \frac{e^{-\lambda t}(\lambda t)^n}{n!}, \quad \forall n \in \mathbb{N}$$

for some $\lambda \geq 0$. This means $N_t \sim$ Poisson(λt). Besides, $(N_t)_{t\geq 0}$ is continuous in probability $\left(\lim_{t\to u} N_t = N_u$, taking the limit in probability, i.e., $\forall \epsilon \geq 0, \lim_{t\to u} \mathbb{P}(|N_t - N_u| \geq \epsilon) = 0\right)$ and N does not have explosions.

Proof. First we prove 3 preliminary results.

(a) $\forall\, t \geq 0$, $\mathbb{P}(N_t = 0) = e^{-\lambda t}$, for some $\lambda \geq 0$.

The number of arrivals in the interval $[0, t + s]$ is zero if and only if there is no arrival in the interval $[0, t]$ or $(t, t + s]$, i.e. the event $\{N_{t+s} = 0\}$ is equal to the event $\{N_t = 0\} \cap \{N_{t+s} - N_t = 0\}$. By definition, these events are independent and

consequently,

$$\mathbb{P}[N_{s+t} = 0] = \mathbb{P}[N_t = 0, N_{t+s} - N_t = 0] = \mathbb{P}[N_t = 0]\mathbb{P}[N_{t+s} - N_t = 0]$$

On the other hand, the distribution is stationary, so

$$\mathbb{P}[N_{t+s} - N_t = 0] = \mathbb{P}[N_s = 0].$$

Then,

$$\mathbb{P}[N_{s+t} = 0] = \mathbb{P}[N_t = 0]\mathbb{P}[N_s = 0].$$

Therefore, the function $g(t) = \mathbb{P}[N_t = 0]$, fulfills the condition $g(t + s) = g(t)g(s)$ and hence $g(t) = g(t)^2$, $\forall t \geq 0$. This way, we may conclude $g(t)$ is either identically equal to one or to zero.

Assume $g(t)$ is zero for an arbitrary $t \geq 0$, towards a contradiction. Then, for any $s \geq 0$, $N_s \stackrel{d}{=} N_{t+s} - N_t \geq 1 \quad \mathbb{P}$-a.s. Now, consider $n \in \mathbb{N}$ such that $t_1 < t_2 < t_3 < \cdots < t_n = t$. We have:

$$N_t = N_{t_1} + (N_{t_2} - N_{t_1}) + \cdots + (N_{t_n} - N_{t_{n-1}}),$$

and because $N_{t_i} - N_{t_{i-1}} \geq 1, P - a.s., \quad \forall i = 1, \cdots, n$ we get $N_t(\omega) \geq n, \quad \forall n \in \mathbb{N}$, for almost every $\omega \in \Omega$. Consequently, $N_t(\omega) = \infty$ P-a.s., which leads to the conclusion that $(N_t)_{t\geq 0}$ is an explosive process, contradicting one of our hypothesis. Then $g(t)$ cannot be null. Since t was arbitrarily chosen, this means g is not zero in any point of its domain.

We now show that g is an exponential function, that is, $g(t) = e^{-\alpha t}$, for some $\alpha \geq 0$. Let $t, m \in \mathbb{N}$. $g(2) = g(1)g(1) = g(1)^2$ and $g(m) = [g(1 + 1 + ... + 1)] = g(1)g(m - 1) = g(1)g(1)g(m - 2) = g(1)^m$. Since $g(1)$ is non-negative, there exists a $\beta \in \mathbb{R}$ such that $g(1) = e^{\beta}$ and thus $g(m) = e^{m\beta}$, for $m \in \mathbb{N}$.

Now consider $n \in \mathbb{N}$. $g(1) = g\left((\frac{1}{n}) + ..., +(\frac{1}{n})\right) = g\left(\frac{1}{n}\right)^n$. Therefore, for the same β as in the last paragraph, we have $g(\frac{1}{n}) = e^{\frac{\beta}{n}}$.

Let $m, n \in \mathbb{N}. g(\frac{m}{n}) = g(m \times \frac{1}{n}) = g(\frac{1}{n})g\left((m-1)\frac{1}{n}\right) = g(\frac{1}{n})^m = e^{\beta\frac{m}{n}}$.

Now, we can show that g is an exponential function for any real positive t by taking

a sequence of rational numbers $(t_n)_{n\in\mathbb{N}}$ decreasing to t. First, notice that since N_t is a counting process, it only assumes values in $\mathbb{N}_0$ and therefore $g(t) = \mathbb{P}[N_t = 0] = \mathbb{P}[N_t \leq 0]$ represents the value, in zero, of a cumulative distribution function, which is right continuous (as mentioned in [8], for instance). Consequently, $g(t) = \lim_{n\to\infty} g(t_n) = \lim_{n\to\infty} e^{\beta t_n} = e^{\beta t}$. Since a probability is a number smaller or equal than one, β must be non-positive.

Summarily, we have shown that $g(t) = e^{-\lambda t}$, for some λ non-negative real.

(b) $\lim_{t\downarrow 0} \frac{1}{t}\mathbb{P}(N_t \geq 2) = 0.$

Define $f(t) := \mathbb{P}[N_t \geq 2]$. $N_t \geq 2$ implies $N_{t+s} \geq 2$, for any $s > 0$, which allows us to write the relation $\{N_t \geq 2\} \subset \{N_{t+s} \geq 2\}$. Since the probability is a monotonically increasing function, we have $\mathbb{P}[N_t \geq 2] \leq \mathbb{P}[N_{t+s} \geq 2]$, that is, $f(t) \leq f(t+s)$ (f is a non-decreasing function).

Let $n_t := \sup\{n \in \mathbb{N} \mid n \leq \frac{1}{t}\}$. From the definition of n_t, $t \leq \dfrac{1}{n_t}$ and $\dfrac{1}{t} \leq n_t + 1$.

Therefore, $f(t) \leq f\left(\frac{1}{n_t}\right)$ and

$$0 \leq \frac{1}{t}f(t) \leq (n_t+1)f\left(\frac{1}{n_t}\right) = \frac{n_t+1}{n_t}n_t \times f\left(\frac{1}{n_t}\right).$$

Letting $t \downarrow 0$, we have $n_t \uparrow \infty$ and $\dfrac{(n_t+1)}{n_t} \to 1$. It is sufficient to prove that $n_t f\left(\dfrac{1}{n_t}\right) \to 0$ when $n_t \to \infty$ (or equivalently, $t \to 0$) in order to demonstrate our claim.

Consider the interval $[0,1]$ and a partition of this interval in n sub-intervals of length $1/n_t$. Let $S_{n_t}(\omega)$ be the number of sub-intervals for which $N_t(\omega) \geq 2$, i.e., the number of sub-intervals where there were at least two "arrivals". S_{n_t} can be viewed as a Binomial random variable, corresponding to the sum of Bernoulli's random variables, where the number of trials is n_t. In the kth trial, with $k = 1, \cdots, n_t$, there is "success" if in the kth sub-interval, there were at least two arrivals and there is not otherwise. Given this, the probability of success for each trial is given by $p = f\left(\frac{1}{n_t}\right)$ and $E[S_n] = n_t p = n_t f\left(\dfrac{1}{n_t}\right)$.

Consider now, for each $\omega \in \Omega$, the minimal time $\delta(\omega) > 0$ between two "arrivals" (the time between the occurrence of two events) in the interval [0,1]. For n_t sufficiently large

(and equivalently, t sufficiently low), we get $\frac{1}{n_t} < \delta(\omega)$. Consequently, no sub-interval will contain two "arrivals" and we conclude that $S_{n_t}(\omega) = 0, \mathbb{P}-a.s.$. So, $\lim_{n_t\to\infty} S_{n_t}(\omega) = 0,\ \mathbb{P}-a.s.$. Besides, S_{n_t} is limited, because it cannot be greater than the number of arrivals that occur in $[0,1]$; this means, S_{n_t} is limited by N_1, whose expected value is finite: $E[S_{n_t}] \leq E[N_1] < \infty$. From this we can see that S_{n_t} is dominated and applying now the dominated convergence theorem (see [12]), we get:

$$\lim_{n_t\to\infty} n_t f\left(\frac{1}{n_t}\right) = \lim_{n_t\to\infty} E[S_{n_t}] = E[\lim_{n_t\to\infty} S_{n_t}] = E[0] = 0.$$

Finally we conclude $\lim_{t\downarrow 0} \frac{1}{t}\mathbb{P}(N_t \geq 2) = 0$.

(c) $\lim_{t\downarrow 0} \frac{1}{t}\mathbb{P}(N_t = 1) = \lambda$, where λ is the same constant mentioned in (a).

From probability theory, it is true that $\mathbb{P}[N_t = 1] = 1 - \mathbb{P}[N_t = 0] - \mathbb{P}[N_t \geq 2]$.

From (a) and (b) we have

$$\lim_{t\downarrow 0} \frac{1}{t}\mathbb{P}(N_t = 1) = \lim_{t\downarrow 0} \frac{1 - e^{-\lambda t}}{t} - \lim_{t\downarrow 0} \frac{1}{t}\mathbb{P}[N_t \geq 2] = \lambda.$$

We are now in conditions to prove the desired result:

$$\mathbb{P}(N_t = n) = \frac{e^{-\lambda t}(\lambda t)^n}{n!}, \quad \forall n \in \mathbb{N}.$$

Define $F(t) = E[\alpha^{N_t}]$, with $\alpha \in]0,1[$. Since $N_{t+s} = N_t + (N_{t+s} - N_t)$ and $N_{t+s} - N_t$ is independent of N_t, we get

$$F(t+s) = E[\alpha^{N_{t+s}}] = E[\alpha^{N_t + (N_{t+s} - N_t)}] = E[\alpha^{N_t}]E[\alpha^{N_{t+s} - N_t}] = F(t)F(s).$$

Therefore, either $F(t) = 0, \quad \mathbb{P}-a.s.$ or $F(t) = e^{t\xi}$, for some real ξ. But

$$F(t) = E[\alpha^{N_t}] = \sum_{n=0}^{\infty} \alpha^n \mathbb{P}[N_t = n] \geq P[N_t = 0] = e^{-\lambda t}, \lambda \geq 0,$$

so $F(t)$ cannot be identically null and consequently, $F(t) = e^{t\xi}$, for some real ξ. From Calculus, we know that $\lim_{t\to 0} \frac{e^{\xi t} - 1}{\xi t} = 1$. Using the fact that $F(0) = 1$, we can write: $\lim_{t\to 0} \frac{F(t) - F(0)}{t} = \xi$. This proves that ξ corresponds to the derivative of F at point 0, by noticing that the left hand side of the last equation is equal to its definition.

$F(t) = \sum_{n=0}^{\infty} \alpha^n \mathbb{P}[N_t = n]$, so its derivative is given by

$$\xi = \lim_{t\to 0} \frac{1}{t}(F(t)-1) = \lim_{t\to 0} \frac{1}{t}\left[\mathbb{P}(N_t = 0) - 1\right] + \lim_{t\to 0} \frac{1}{t}\alpha\mathbb{P}[N_t = 1] + \lim_{t\to 0} \frac{1}{t}\sum_{n=2}^{\infty} \alpha^n \mathbb{P}[N_t = n].$$

From (a), we have:

$$\lim_{t\to 0} \frac{1}{t}\left[\mathbb{P}(N_t = 0) - 1\right] = \lim_{t\to 0} \frac{1}{t}\left[e^{-\lambda t} - 1\right] = -\lambda$$

(this limit can be obtained by applying Hôpital's rule, for instance).

From (c), $\lim_{t\to 0} \frac{1}{t}\alpha\mathbb{P}[N_t = 1] = \alpha\lambda$. Since $\alpha \in]0, 1[$, we have

$$0 \leq \lim_{t\to 0} \frac{1}{t}\sum_{n=2}^{\infty} \alpha^n \mathbb{P}[N_t = n] \leq \lim_{t\to 0} \frac{1}{t}\sum_{n=2}^{\infty} \alpha^n \mathbb{P}[N_t \geq 2] = \sum_{n=2}^{\infty} \alpha^n \lim_{t\to 0} \frac{1}{t}\mathbb{P}[N_t \geq 2]$$
$$\leq \sum_{n=0}^{\infty} \alpha^n \lim_{t\to 0} \frac{1}{t}\mathbb{P}[N_t \geq 2] = \frac{1}{1-\alpha} \lim_{t\to 0} \frac{1}{t}\mathbb{P}[N_t \geq 2] = 0.$$

From this we get that $\xi = -\lambda + \lambda\alpha, \quad \forall\alpha \in]0, 1[$ and therefore $F(t) = e^{-\lambda t + \lambda t\alpha}$. Developing $e^{\lambda t\alpha}$ in Taylor series around the point 0, we may establish the equality

$$\sum_{n=0}^{\infty} \alpha^n \mathbb{P}[N_t = n] = \sum_{n=0}^{\infty} \frac{e^{-\lambda t}(\lambda t)^n}{n!}\alpha^n,$$

and, in conclusion, $\mathbb{P}[N_t = n] = \frac{e^{-\lambda t}(\lambda t)^n}{n!}$, as we wanted to prove.

It is only left to show that the Poisson's process is continuous in probability, that is, $\lim_{t\to u} \mathbb{P}(|N_t - N_u| \geq \epsilon) = 0, \quad \forall\epsilon > 0$.

Take an arbitrary $\epsilon > 0$. Since

$$\{\omega \in \Omega : |N_t(\omega) - N_u(\omega)| \geq \epsilon\} \subset \{\omega \in \Omega : |N_t(\omega) - N_u(\omega)| > 0\},$$

the inequality

$$0 \leq \mathbb{P}(|N_t - N_u| \geq \epsilon) \leq \mathbb{P}(|N_t - N_u| > 0)$$

is valid. The latter probability is equal to

$$\mathbb{P}(|N_t - N_u| > 0) = 1 - \mathbb{P}(|N_t - N_u| = 0) = 1 - e^{-\lambda|t-u|}.$$

Now, taking the limit, it follows that

$$0 \leq \lim_{t\to u} \mathbb{P}(|N_t - N_u| \geq \epsilon) \leq \lim_{t\to u} \mathbb{P}(|N_t - N_u| > 0) = \lim_{t\to u} 1 - e^{-\lambda|u-t|} = 0,$$

and we conclude that the Poisson process is continuous in probability. ■

A.5 Compound Poisson processes

In a compound Poisson process, each arrival in an ordinary Poisson process comes with an associated real-valued random variable that represents the value of the arrival in a certain sense. These variables are independent and identically distributed and also independent of the underlying Poisson process. Our interest is centered on the sum of the random variables for all the arrivals up to a fixed time t, which thus is a Poisson-distributed random sum of random variables. Distributions of this type are said to be compound Poisson distributions.

Suppose we have a Poisson process with rate $\lambda \in (0, \infty)$. We will denote the sequence of inter-arrival times by $X = (X_0, X_1, X_2, \dots)$, the sequence of arrival times by $T = (T_0, T_1, T_2, \dots)$, and the counting process by $N := \{N_t \mid t \in [0, \infty)\}$. From the section A.4, recall that X is a sequence of independent random variables, each having the exponential distribution on $[0, \infty)$ with rate λ. The sequence $(T_n)_{n \geq 0}$ is the partial sum sequence associated to X, and has stationary independent increments. For $t \in (0, \infty)$, the number of arrivals N_t in $(0, t]$ has Poisson distribution with parameter λ.

We next present some typical examples of compound Poisson processes:

The arrivals of customers at a store. Each customer spends a random amount of money. The number of visits to a website. Each visitor spends a random amount of time at the site. The number of earthquakes at a particular location. Each earthquake has a random severity, a measure of the energy released. For $n \in \mathbb{N}$, let U_n denote the value of the nth arrival. We assume that $U = (U_0, U_1, U_2, \dots)$ is a sequence of independent, identically distributed, real-valued random variables, and that U is independent of the underlying Poisson process. The common distribution may be discrete or continuous, but in either case, we let p denote the common probability density function.

Definition A.9. *The Compound Poisson process associated with a given Poisson pro-*

cess $(N_t)_{t\geq 0}$ *and a sequence* $(U_n)_{n\geq 0}$ *is the stochastic process* $V=\{V_t \mid t \in [0,\infty)\}$ *where* $V_t = \sum_{n=1}^{N_t} U_n$.

Thus, V_t is the total value for all of the arrivals in $(0,t]$. For the examples above, V_t, respectively, is the total income to the store up to time t, the total time spent at the site by the customers who arrived up to time t, and the total energy released up to time t on the earthquake.

Theorem A.10. *Consider a random variable* X*, which has a compound Poisson distribution, i.e.,* $X = \sum_{j=0}^{N_t} Y_j$*, where* N_t *has Poisson distribution with parameter* λ*, and* $(Y_j)_{j\in\mathbb{N}_0}$ *is an independent sequence of i.i.d. random variables with characteristic function* ϕ*. Then* $\Psi(s) := E[e^{isX(t)}] = e^{-\lambda t[1-\phi(s)]}$.

Proof.

$$\begin{aligned}
\Psi(s) &= E[e^{isX(t)}] \\
&= \sum_{j=0}^{\infty} E(e^{isX(t)} \mid N_t = j)P(N_t = j) \\
&= \sum_{j=0}^{\infty} E(e^{is(Y_0+\cdots+Y_{N_t})})|N_t = j)P(N_t = j) \\
&= \sum_{j=0}^{\infty} E(e^{is(Y_0+\cdots+Y_j)})P(N_t = j) \\
&= \sum_{j=0}^{\infty} \prod_{k=0}^{j} E(e^{is(Y_k)})P(N_t = j) \\
&= \sum_{j=0}^{\infty} \phi(s)^j \frac{(\lambda t)^j}{j!} \times e^{-\lambda t} \\
&= e^{\lambda t[\phi(s)-1]}.
\end{aligned}$$

Furthermore, using the moment generating function definition and assuming $(Y_j)_{j\in\mathbb{N}_0}$ has a common probability distribution p, we get:

$$\begin{aligned}
\Psi(t) &= \exp\left(\lambda t\left[\int_{\mathbb{R}} e^{isx}p(x)\,dx - 1\right]\right) \\
&= \exp\left(\lambda t\left[\int_{\mathbb{R}} e^{isx}p(x)\,dx - \int_{\mathbb{R}} p(x)dx\right]\right)
\end{aligned}$$

$$= \exp\left(\lambda t\left[\int_{\mathbb{R}}(e^{isx}-1)p(x)\,dx\right]\right).$$

■

Notice that in this proof we used the fact that $(Y_j)_{j\in\mathbb{N}_0}$ is a sequence of independent random variables and that the exponential function is equal to the following power series: $e^x = \sum_{n=0}^{\infty}\frac{x^n}{n!}$.

Theorem A.11. *Given the random variable X in the exact same conditions as in the previous theorem, its mean and variance, respectively, are: $\lambda E(Y_i)$ and $\lambda E(Y_i^2)$.*

Proof. The moment of first order of X can be obtained, by differentiating the moment generating function $\Psi(t)$ once and evaluating it at the point $t = 0$:

$$\begin{aligned}
E[X] &= (-i)\times(-1)\frac{\partial\phi(s)}{\partial t}(-\lambda)e^{-\lambda t[1-\phi(s)]}\Big|_{s=0}\\
&= \lambda t(-i)\times\frac{\partial\phi(s)}{\partial s}\Big|_{s=0}e^{-\lambda t[1-\phi(0)]}\\
&= \lambda t E(Y_i)
\end{aligned}$$

The moment of second order is obtained following a similar logic, but differentiating twice instead:

$$\begin{aligned}
E[X^2] &= (-i)^2\times\frac{\partial\phi(s)}{\partial s}\left[\frac{\partial\phi(s)}{\partial s}\lambda t e^{-\lambda t[1-\phi(s)]}\right]\Big|_{s=0}\\
&= (-i)^2\frac{\partial^2\phi(s)}{\partial s^2}\lambda t e^{-\lambda t[1-\phi(s)]}\Big|_{s=0} + (-i)^2\frac{\partial\phi(s)}{\partial s}\times\frac{\partial\phi(s)}{\partial s}(\lambda t)^2 e^{-\lambda t[1-\phi(s)]}\Big|_{s=0}\\
&= \lambda t E(Y_i^2) + \left[(-i)\frac{\partial\phi(s)}{\partial s}\Big|_{s=0}\right]^2(\lambda t)^2 e^{-\lambda t[1-\phi(s)]}\Big|_{s=0}\\
&= \lambda t E(Y_i^2) + (\lambda t)^2 E^2(Y_i)
\end{aligned}$$

The variance is then:

$Var(X) = E(X^2) - E^2(X) = \lambda t E(Y_i^2) + (\lambda t)^2 E^2(Y_i) - (\lambda t E(Y_i))^2 = \lambda t E(Y_i^2)$. ■

A.6 Brownian motion

Definition A.12. *The stochastic process* $\{W_t\}_{t\geq 0}$ *defined on a probability space* $(\Omega, \mathcal{F}, \mathbb{P})$ *is a Wiener process, or a standard Brownian motion if:*

(1) $W_0 = 0$ *almost surely;*

(2) It has independent increments;

(3) It has stationary increments;

(4) It is stochastically continuous;

(5) The increment $W_{t+s} - W_t$ *is Normally distributed with mean 0 and variance* $s > 0$*, i.e.,* $W_{t+s} - W_t \sim N(0, s)$.

This is the classical example of a diffusion process and is certainly the most studied and notorious stochastic model in quantitative finance. It was first documented by Robert Brown in 1827 (for which it was named after), then in Bachelier [2], it was used as a model for stock market prices. Five years later, Einstein consider it, as a model of particles. Only in 1923 the Brownian motion was defined and constructed rigorously by Robert Wiener, for which the process is also referred. Finally, it was thanks to Samuelson [48] that the Brownian motion was, definitely, set as the standard modelling tool in finance.

A.7 Fast Fourrier Transform

In this section, we present a few concepts and results related with the so called "Fast Fourrier Transform" (FFT). Our goal is not to provide an extensive study, but to give a brief understading on how useful FFT is to compute the product of a dense Toeplitz matrix and a column vector (as in (7.14)), because this calculation is the most time-consuming one. For more details, we suggest the reading of [27].

First, we present a few concepts and then we suggest an algorithm that allows to significantly reduce the computation time of the product refered in the last paragraph.

Definition A.13. $T \in C^{n\times n}$ *is Toeplitz if* T *is determined by the* $2n-1$ *scalars*

$t_{n-1}, \ldots, t_{n-1}$ *with* $T_{ij} = t_{j-i}$*, for all* i *and* j*.*

$$T = \begin{bmatrix} t_0 & t_{-1} & \cdots & t_{2-n} & t_{1-n} \\ t_1 & t_0 & \ddots & & t_{2-n} \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ t_{2-n} & & \ddots & \ddots & t_{-1} \\ t_{n-1} & t_{n-2} & \ldots & t_1 & t_0 \end{bmatrix}$$

For a general Toeplitz matrix, it is sufficient only to save the vector $[t_{n-1}, t_{n-2}, \cdots, t_{1-n}]^T \in C^{2n-1}$ instead of the whole matrix.

A special case of Toeplitz matrices is the so-called circulant matrix:

Definition A.14. $C \in C^{n \times n}$ *is circulant if it is a Toeplitz matrix where each column is a circular shift of its preceding column. It is of the form:*

$$C = \begin{bmatrix} c_0 & c_{n-1} & \cdots & c_2 & c_1 \\ c_1 & c_0 & \ddots & & c_2 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ c_{n-2} & & \ddots & \ddots & c_{n-1} \\ c_{n-1} & c_{n-2} & \ldots & c_1 & c_0 \end{bmatrix}$$

Circulant matrices are fully determined by their first row or column and thus we may introduce the notation $C = circ(c_0, \cdots, c_{n-1})$*. Note that once again only little storage is needed: saving the first column or row instead of the whole matrix is sufficient.*

Definition A.15. *The unitary and symmetric matrix* $F_n \in C^{n \times n}$ *is called a Fourrier matrix if it is of the form:*

$$F_n = \frac{1}{\sqrt{n}} \begin{bmatrix} 1 & 1 & \cdots & 1 & 1 \\ 1 & \omega & & \omega^{n-2} & \omega^{n-1} \\ \vdots & \vdots & & \vdots & \vdots \\ 1 & \omega^{n-2} & & & \omega^{(n-2)(n-1)} \\ 1 & \omega^{n-1} & \ldots & \omega^{(n-2)(n-1)} & \omega^{(n-1)^2} \end{bmatrix}$$

with $\omega = e^{-\frac{2\pi i}{n}}$.

Theorem A.16. *Let* $C_n \in \mathbb{R}^{n\times n}$ *be circulant. Then, it has the decomposition*

$$C_n = F_n^* \Lambda F_n, \tag{A.7}$$

where F_n^* *denotes the conjugate transpose of* F_n *and* $\Lambda = diag(\lambda_0, \cdots, \lambda_{n-1}), \lambda_j$ *being the jth eigenvalue of* $C_n, j \in \{0, \cdots, n-1\}$.

To find the eigenvalues of a circulant matrix C_n, we use the decomposition in equation (A.7). Letting c denote the first column of $C_n, e_1 := [1, 0, \ldots, 0]^T$ and $\overrightarrow{e} := [1, 1, \ldots, 1]^T$, we get:

$$\begin{aligned} & C_n = F_n^* \Lambda F_n \\ \iff & F_n C_n e_1 = \Lambda F_n e_1 \\ \iff & F_n c = \Lambda \overrightarrow{e} \frac{1}{\sqrt{n}} \\ \iff & \sqrt{n} F_n c = \Lambda \overrightarrow{e} \end{aligned} \tag{A.8}$$

Notice we used the fact that F_n is unitary.

As pointed in the fourth chapter of [27], the matrix-vector product $y = C_n x = F_n^* \Lambda F_n x$ can be computed efficiently in four steps: we start by applying an FFT to x, next we compute $\tilde{c} := [\lambda_0, \ldots, \lambda_{n-1}]^T$ by equation (A.8), multiply elementwise the so obtained vector $\tilde{c}$ containing the diagonal elements of Λ with $F_n x$, and, finally apply an inverse FFT. Formally, this looks like:

$$\begin{aligned} \tilde{x} &= F_n x \\ \tilde{c} &= \sqrt{n} F_n c, \\ z &= \tilde{c}. * \tilde{x}, \\ y &= F_n^* z. \end{aligned}$$

" .* " denotes the elements-wise multiplication.

This way of calculating the product between a circulant matrix c_n and a column vector x involves three FFTs and one vector multiplication, being then only necessary

to perform $O(n\log n)$ operations. Compared with a straightforward computation, which costs $O(n^2)$ operations, this definitely represents a substantial improvement.

Now, in the case we are dealing with a Toeplitz matrix and not a circulant one, it is still possible to apply an FFT to compute the product of a Toeplitz matrix and a vector, by "embedding" the former in a circulant matrix, i.e., the possibility to quickly compute the matrix-vector product can be extended from circulant matrices to general Toeplitz matrices. For example, one can apply the following method: for an $n \times n$ Toeplitz matrix T_n, the Toeplitz matrix-vector multiplication $T_n x$ can be computed with three FFTs by first embedding T_n into a $2n \times 2n$ circulant matrix:

$$\begin{bmatrix} T_n & B_n \\ B_n & T_n \end{bmatrix} \begin{bmatrix} x \\ 0 \end{bmatrix} = \begin{bmatrix} T_n x \\ * \end{bmatrix},$$

where

$$B_n = \begin{bmatrix} 0 & t_{n-1} & \cdots & t_2 & t_1 \\ t_{1-n} & 0 & t_{n-1} & & t_2 \\ \vdots & t_{1-n} & 0 & \ddots & \vdots \\ t_{-2} & & \ddots & \ddots & t_{n-1} \\ t_{-1} & t_{-2} & \dots & t_{1-n} & 0 \end{bmatrix}.$$

We can solve the previous system of equations using the FFT's, since the first matrix is circulant, and then "extract" the product $T_n x$ as desired.

Bibliography

[1] Almendral, A., & Oosterlee, C. (2005). Numerical Valuation of options with jumps in the underlying, *Applied Numerical Mathematics, 53*, 1-18.

[2] Bachelier, L. (1900). *Theórie de la spéculation*. Unpublished doctoral dissertation, Annales scientifiques de l'École Normale Supérieure.

[3] Bakshi, G., & Madan, D. (2000). Spanning and derivative-security valuation. *Journal of Financial Economics 55*, 205-238.

[4] Barndorff-Nielsen, O. E. (1998). Processes of normal inverse Gaussian type. *Finance Stoch.*, 2, 41–68.

[5] Bates, D. S. (1996). Jumps and stochastic volatility: Exchange rate processes implicit in Deutsche mark options. *Review of Financial Studies*, 9(1), 69-107.

[6] Bensoussan, A., & Lions. J.-L. (1982). *Contrôle Impulsionnel et Inéquations Quasi-Variotionelles*, Dunod: Paris.

[7] Bera, A. K., & Jarque, C. M. (1987). A test for normality of observations and regression residuals. *International Statistical Review*, 55, 163-172.

[8] Boes, D.C., & Mood, A.M., & Graybill, F.A. (1974). *Introduction to the Theory of Statistics*. 3rd edition McGraw-Hill, New York.

[9] Björk, T. (2009). *Arbitrage Theory in Continuous Time* (3rd ed.). Oxford University Press.

[10] Black, F., & Scholes, M. S. (1973). The Pricing of Options and Corporate Liabilities. *The Journal of Political Economy*, 81(3): 637–654.

[11] Carr, P., & Geman, H., & Madan, D. B., & Yor, M. (2002). The fine structure of asset returns: An empirical investigation, *J. Bus.*, 75, 305–332.

[12] Capinski, M., & Kopp, E. (1999) *Measure, Integral and Probability*. Springer-Verlag.

[13] Chiarella, C. & Ziogas, A. (2009). American call options under jump-diffusion processes—a Fourier transform approach, *Appl. Math. Finance*, 16, 37–79.

[14] Cont, R., & Tankov, P. (2012). *Financial modelling with jump processes*. Chapman & Hall/CRC.

[15] Cont, R., & Voltchkova, E. (2005). *A finite difference scheme for option pricing in jump diffusion and exponential Lévy models*, SIAM J. Numer. Anal., 43, 1596–1626.

[16] Cox, J. (1975). *Notes on option pricing I: Constant elasticity of diffusions* (Tech. Rep.). Stanford University.

[17] Dewynne, J., Wilmott, P., & Howison, S. (1995). *The mathematics of financial derivatives: a student introduction*, Cambridge University Press, Cambridge.

[18] Eberlein, E., & Keller, U., & Prause, P. (1998). New insights into smile, mispricing, and value at risk: The hyperbolic model, *J. Bus.*, 71, 371–405.

[19] Esche, F., & Schweizer, M. (2005). Minimal entropy preserves the Lévy property: how and why. *Stochastic Process*. Appl., 115(2): 299–327.

[20] Fang, H. (2000). *Option pricing implications of a stochastic jump rate*. Working Paper, University of Virginia.

[21] Flannery, B., & Press, H., & Teukolsky, S., & Vetterling, W., (2007). *Numerical Recipes in C* (3rd ed.). The Art of Scientific Computing. Cambridge University Press, New York, NY, USA.

[22] Forsyth P. A., & Halluin, Y., & Vetzal, K. R. (2005). Robust numerical methods for contingente claims under jump diffusion processes, Vetzal IMA J. Numer. Anal., 25, 87–112.

[23] Forsyth P. A., & Halluin, Y., & Labahn, G. (2004). A penalty method for American options with jump diffusion processes, Numer. Math., 97, 321–352.

[24] Fujiwara, T. and Miyahara, Y. (2003). The minimal entropy martingale measures for geometric Lévy processes, *Finance and Stochastics 7*, 509–531.

[25] Gerber, H.U., & Shiu, E.S.W. (1994). Option pricing by Esscher transforms. *Transactions of the Society of Actuaries XLVI*, 99–191.

[26] Glasserman, P. (2004). *Monte Carlo Methods in Financial Engineering*. NewYork: Springer-Verlag.

[27] Golub, G., & Van Loan, C. (1996). *Matrix Computations* (3rd ed.). The Johns Hopkins University Press, Baltimore.

[28] Heston, S. L. (1993). A closed-form solution for options with stochastic volatility with applications to bond and currency options. *Review of Financial Studies*, 6, 327–343.

[29] Hull, J. (2002) *Options, Futures and Other Derivatives*. (5th ed.). Upper Saddle River, New Jersey: Prentice Hall.

[30] Hudson, R., & Mandelbrot, B. (2006). *The Misbehavior of Markets: A Fractal View of Financial Turbulence*. Basic Books, annotated edition.

[31] Ikonen, S., & Toivanen, J. (2004). Operator splitting methods for American option pricing, *Appl. Math. Lett., 17*, 809–814

[32] Karatzas, I., & Shreve, S. (1991).*Brownian Motion and Stochastic Calculus*. Springer-Verlag, 2nd edition.

[33] Kou, S. G. (2002). A Jump-Diffusion Model for Option Pricing. *Manage. Sci., 48(8)*, 1086–1101.

[34] Kubrusly, C. S. (2001). *Elements of Operator Theory*. Birkhaeuser, Boston.

[35] Kulkarni, D., & Schmidt, D., & Tsui, S. (1999). Eigenvalues of tridiagonal pseudo-toeplitz matrices. *Linear Algebra Appl.*, 297: 63–80.

[36] Kwok, Y. (1999). *Mathematical Models of Financial Derivatives*. Springer-Verlag, New York.

[37] Kwon, Y., & Lee, Y. (2011). A second-order finite difference method for option pricing under jump-diffusion models, submitted.

[38] Kwon, Y., & Lee, Y. (2011). A second-order tridiagonal method for American Options under jump-diffusion models, submitted.

[39] Lamberton, D., & Lapeyre, B. (2008). *Stochastic Calculus Applied to Finance*. Chapman and Hall/CRC, second edition.

[40] Madan, D. B., & Milne, F. (1991). *Option pricing with V. G. Martingale components*, Math. Finance, 1, 139–55.

[41] Merton, R.C. (1976). Option pricing when the underlying stocks are discontinuous. *Journal of Financial Economics*, 3(1-2), 125-144.

[42] Monteiro, R. (2013). *Option pricing with Lévy processes*. Master of Science thesis - ISCTE-IUL.

[43] Øksendal, B., & Sulem, A. (2005). *Applied Stochastic Control of Jump Diffusions*. Springer-Verlag, Berlin.

[44] Oliveira, P. (2014). *The Convolution Method for Pricing American Options under Lévy Processes*. Master of Science thesis - ISCTE-IUL.

[45] Papapantoleon, A. (2008). An Introduction to Lévy processes with application in Finance. *ArXiv e-prints*.

[46] Pham, H. (1997). Optimal stopping, free boundary, and American option in a jump- diffusion model, *Applied Mathematics and Optimization* 35, 145–164.

[47] Raible, S. (2000). *Lévy Processes in Finance: Theory, Numerics, and Empirical Facts*. Ph.D. Thesis, Inst. für Mathematische Stochastik, Albert-Ludwigs-Universität Freiburg, Freiburg, Germany.

[48] Samuelson, P. A. (1965). Rational theory of warrant pricing. *Industrial Management Review*, 6, 13–31.

[49] Sato, K. (1999). *Lévy Processes and Infinitely Divisible Distributions*. Cambridge University Press: Cambridge, UK.

[50] Strikwerda, J. C. (2004). *Finite Difference Schemes and Partial Differential Equations*, SIAM, Philadelphia.

[51] Toivanen, J. (2008). Numerical valuation of European and American options under Kou's jump-diffusion model, *SIAM J. Sci. Comput., 30*, 1949–1970.

[52] *http://www.portalaction.com.br/processo-estocastico/processo-de-poisson.html*.

[53] *http://math.stackexchange.com/questions/888959/the-lévy-khintchine-formula-and-integrability-conditions-of-a-random-measure*.