CEMAPRE – ISEG

University of Lisbon

PhD Thesis

Option pricing in exponential Lévy models with transaction costs

Author:
Nicola Cantarutti

Supervisors:
João Guerra
Manuel Guerra

September 12, 2019

Contents

Introduction

The problem of pricing a European call option was first solved mathematically in the paper of [Black and Scholes, 1973]. Even if it is quite evident that this model is too simplistic to represent the real features of the market, it is still nowadays one of the most used model to price and hedge options. The reason for its success is that it gives a closed form solution for the option price, and that the hedging strategy is easily implementable. The Black-Scholes model considers a *complete* market, i.e. a market where it is possible to create a portfolio containing cash and shares of the underlying stocks, such that following a particular trading strategy it is always possible to replicate the payoff of the option. In this framework, this particular portfolio is called *replicating portfolio* and the trading strategy to hedge the option is called *delta hedge*. However, this model does not consider many features that characterize the real market.

In the Black-Scholes model the stock price follows a geometric Brownian motion. This is equivalent to assume that the log-returns are normally distributed. However, a rigorous statistical analysis of financial data reveals that the normality assumption is not a very good approximation of reality (see [Cont, 2001]). Indeed, it is easy to see that empirical log-return distributions have more mass around the origin and along the tails (*heavy tails*). This means that normal distribution underestimates the probability of large positive and negative log-returns, and considers them just as rare events. In the real market instead, log-returns manifest frequently high peaks, that come more and more evident when looking at short time scales. The log-returns peaks correspond to sudden large changes in the price, which are called *jumps*. There is a huge literature of option pricing models considering an underlying process with discontinuous paths. Most of these models assume the log-price dynamics to be a *Lévy process*.

A second issue of the Black-Scholes model is that it does not consider the presence of market frictions i.e. bid/ask spread, transaction fees and budget constraints. The securities in the market are traded with a bid-ask spread, and this means that there are two prices for the same security. But the Black-Scholes formula just gives one price. Moreover, the replicating portfolio cannot be perfectly implemented, since the delta-hedging strategy involves continuous

time trading. This is impractical because the presence of transaction costs makes it infinitely costly. Another kind of market friction that needs to be considered are the budget constraints. A limit in the budget or a restriction in the possibility of selling short, clearly restricts the set of possible trading strategies.

Many authors attempted to include the presence of proportional transaction costs in option pricing models. In [Leland, 1985], in order to avoid continuous trading, the author specifies a finite number of trading dates. He obtains a Black-Scholes-like nonlinear partial differential equation (PDE) with an adjusted volatility term, that takes into account the transaction costs. However, trading at fixed dates is not optimal, and the option price goes to infinity as the number of dates grows. Further work in this direction has been done by [Boyle and Vorst, 1992], which consider a multi-period binomial model (see [Cox et al., 1979]) with transaction costs. Here again, the cost of the replicating portfolio depends on the number of time periods. Recent developments in this direction are for instance [Mocioalca, 2007], [Florescu et al., 2014] and [Sengupta, 2014] who consider different features of the market such as jumps, stochastic volatility and stochastic interest rate respectively.

A different approach has been introduced by [Hodges and Neuberger, 1989]. The authors use an alternative definition of the option price called *indifference price*, based on the concepts of *expected utility* and *certainty equivalent*. An overview of these concepts applied to several incomplete market models can be found in [Carmona (Editor), 2009]. As long as the perfect replicating portfolio is no longer implementable in presence of transaction costs, the hedging strategy cannot be anymore riskless. The model has to take into account the risk profile of the writer/buyer to describe his trading preferences. [Hodges and Neuberger, 1989] define the option price as the value that makes an investor indifferent between holding a portfolio with an option and without, in terms of expected utility of the final wealth. They show that it is impossible to hedge perfectly the option. The optimal strategy is to keep the portfolio's values within a band called *no transaction region*. Using numerical experiments, they verify that this strategy outperforms the one proposed in [Leland, 1985]. This approach has been further developed in [Davis et al., 1993], where the problem is formulated rigorously as a singular stochastic optimal control problem. The authors prove that the value functions of the two optimization problems can be interpreted as the solutions of the associated Hamilton-Jacobi-Bellman (HJB) equation in the viscosity sense. They prove also that the numerical scheme, based on the *Markov chain approximation*, converges to the viscosity solution. Numerical methods for this model are presented in [Davis and Panas, 1994], [Clewlow and Hodges, 1997] and [Monoyios, 2003], [Monoyios, 2004]. In [Whalley and Wilmott, 1997] and [Barles and Soner, 1998] the problem is simplified by using asymptotic analysis methods for small levels of transaction costs. The authors, starting from the general HJB variational inequality, derive a simpler non-linear PDE for the option price. Further developments are presented in the thesis work of [Damgaard, 1998], where the author studies the robustness of the model from a theoretical and numerical point of view. He found that under certain conditions the model is quite robust with respect to the choice of the utility function.

In this thesis the main goal is to develop and analyze a model for pricing options when the market is incomplete due to the presence of jumps in the

stock dynamics and transaction costs. The topics of the thesis are based on the two papers [Cantarutti and Guerra, 2018], [Cantarutti et al., 2018] and the contributed chapter *Indifference pricing in a market with transaction costs and jumps* by N. Cantarutti, J. Guerra, M. Guerra and M.R. Grossinho, published in the book of [Ehrhardt et al., 2017].

If the paper is accepted I have to update the reference

Portfolio selection models with transaction costs and Lévy processes have already been introduced in the financial literature, see for instance [Framstad et al., 1999], [Benth et al., 2001] and [De Vallière et al., 2016] but these models have never been used to price options. We present and analyze the theoretical properties of the model, and prove the existence of a viscosity solution of the nonlinear Partial Integro-Differential Equation (PIDE) associated to the optimization problem. We also develop numerical methods to solve the problem under different assumptions, i.e. ignoring the default event and not. We present new numerical results and compare them with numerical values obtained from existing reference models. We also prove that the proposed numerical scheme is monotone, stable and consistent, and its solution converges to the viscosity solution of the continuous problem.

In Chapter 1, we introduce the general theory for Lévy processes, with a deeper focus on the specific Lévy processes used in this thesis i.e. the *Merton* and *Variance Gamma* processes. In Chapter 2 we make a small summary of the main assumptions and theorems of the *No arbitrage theory* for derivative pricing. After that we present the most common numerical finite differences methods used to solve the option pricing PIDEs. The Chapter 3 is a digression based on the paper [Cantarutti and Guerra, 2018] on the application of the multinomial method to solve option pricing problems under the Variance Gamma process. In Chapter 4 we present the optimal control theory for regular and singular controls, and the definition of viscosity solutions. In Chapter 5, we will develop the model for option pricing in presence of proportional transaction costs. This model is a singular stochastic control problem, which is a generalization of the model proposed in [Davis et al., 1993]. We derive the general HJB equation and prove that the value function of the optimization problem can be interpreted as the viscosity solution of the HJB equation. In Chapter 6 we present the numerical method used to solve the optimization problem and several numerical results. First, we consider the simplified problem where the number of variables is reduced by one thanks to the assumption of no default and the property of the exponential utility. The same results are also presented in [Cantarutti et al., 2018] and [Ehrhardt et al., 2017].
Then we solve the general problem with four variables, considering the possibility of default. Last, we also show that the moment matching method developed for the Variance Gamma process, can be used to solve these kind of problems with good performance.
At the end of each chapters there is a conclusive section containing a summary of the main concepts and their relevance for the thesis.

1

Modeling with Lévy Processes

In mathematical finance, Lévy processes are a powerful tool to describe the observed reality of financial markets.
Usually it is common to model the market dynamics in a continuous time setting by means of the *log-returns*

$$\log S_{t+\Delta t} - \log S_t = \log\left(\frac{S_{t+\Delta t}}{S_t}\right), \tag{1.1}$$

where S_t is the spot price of a financial asset at time t.
Log-returns are preferred to the *relative price change* $(S_{t+\Delta t} - S_t)/S_t$, because the sum of log-returns over n consecutive periods is the log-return of the period $n\Delta t$. The reason to use the log-returns rather than modeling directly the

prices $\{S_t\}_{t\geq 0}$, is that they have better statistical properties. Furthermore, log-returns can assume negative values, and thus can be modeled by distributions with "nicer" analytical properties.
Since the renowned paper of [Black and Scholes, 1973], a common assumption is that t-log-returns are normally distributed as $\mathcal{N}(\mu t, \sigma^2 t)$. This is largely due to the fact that normal distribution as well as the continuous-time process it generates (Brownian motion) has nice analytical properties. Under this assumption, the dynamics of log-returns follows the Brownian motion process

$$\log\left(\frac{S_t}{S_0}\right) = \mu t + \sigma W_t, \tag{1.2}$$

with constant drift $\mu \in \mathbb{R}$ and constant volatility $\sigma > 0$.
This model guarantees positiveness of the prices. By Itō formula, we obtain the stochastic differential equation for the price

$$\frac{dS_t}{S_t} = (\mu + \frac{1}{2}\sigma^2)dt + \sigma dW_t. \tag{1.3}$$

The process S_t is called *geometric Brownian motion*.
By the way, a thorough look at data collections from various areas of finance, reveals that the normality assumption is not a very good approximation of reality. Indeed, empirical return distributions have substantially more mass around the origin and along the tails (*heavy tails*), see [Cont, 2001]. This means that normal distribution underestimates the probability of large positive and negative returns. In the real market instead, returns manifest frequently high peaks, that come more and more evident when looking at short time scales. The return peaks correspond to sudden large changes in the price, that cannot be reproduced by the dynamics of the Brownian motion.
The Brownian motion, which is a scale invariant process with continuous paths, can be a good approximation for long time scales ($\Delta t \sim$ months to years), but is not a good model to reproduce the peaks of the log-returns at short time scales.

In the last thirty years, a lot of research has been done on processes with jumps and their applications to financial derivatives.
Lévy processes belong to the bigger family of semimartingales. If $\{X_t\}_{t\geq 0}$ is a Lévy process, there are relevant quantities such as the stochastic integral $\int_0^t \phi dX_t$ or any non-linear function $f(t, X_t)$ that are not Lévy processes anymore. For some applications it is important to consider the larger class of processes of *semimartingales*, which is closed with respect to integration and non-linear transformations. A general theory for semimartingales, is presented in the books of [Protter, 2004] and [Jacod and Shiryaev, 2002].
[Sato, 1999] is a complete reference book for the theory of Lévy processes and their analytical properties. [Applebaum, 2009] presents Lévy processes with more emphasis on stochastic calculus and stochastic differential equations (SDEs). For a comprehensive guide to applications of Lévy processes in finance some good sources are the books of [Cont and Tankov, 2003] and [Schoutens, 2003].
Among the most popular Lévy processes applied to finance, it is worth to mention:

- the Merton jump-diffusion model [Merton, 1976]
- the Kou jump-diffusion model [Kou, 2002]

- the α-stable [Mandelbrot, 1963], [Cont et al., 1997], [Kabasinskas et al., 2009]
- the Variance-Gamma (VG) [Madan and Seneta, 1990], [Madan et al., 1998]
- the Normal-Inverse-Gaussian (NIG) [Barndorff-Nielsen, 1997]
- hyperbolic Lévy processes [Eberlein and Keller, 1995]
- Carr-Geman-Madan-Yor (CGMY) model [Carr et al., 2002]

This chapter reviews the most important points concerning the theory of Lévy processes and the stochastic calculus applied to jump processes. A remarkable emphasis is given to the presentation of the exponential Lévy models used in this thesis: the *Merton model* and the *Variance Gamma model*.

1.1 Properties of Lévy processes

1.1.1 Basic definitions

Let $\{X_t\}_{t\geq 0}$ be a stochastic process defined on a probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geq 0}, \mathbb{P})$, where $\mathcal{F}_t$ is the natural filtration[1] to which the process $\{X_t\}_{t\geq 0}$ is adapted.

Definition 1.1.1. *We say that $\{X_t\}_{t\geq 0}$ is a **Lévy Process** if:*

(L1) $X_{t=0} = 0$.

(L2) *$\{X_t\}_{t\geq 0}$ has independent and stationary increments:*
For each $0 < t_1 < t_2 < ... < t < \infty$

$$X_{t_{j+1}} - X_{t_j} \text{ are independent.}$$

$$X_{t_{j+1}} - X_{t_j} \stackrel{d}{=} X_{t_{j+1}-t_j}.$$

(L3) *$\{X_t\}_{t\geq 0}$ is stochastically continuous: $\forall \epsilon > 0$ and $\forall t \geq 0$*

$$\lim_{h\to 0} \mathbb{P}(|X_{t+h} - X_t| > \epsilon) = 0.$$

It is well known (see [Protter, 2004] Chapter 1, Theorem 30) that a Lévy process has a modification with "cádlág" paths, i.e. paths which are right-continuous and have left limits.
Lévy processes are intrinsically connected with infinitely divisible distributions. In particular the Lévy-Khintchine formula for infinitely divisible random variables, is an essential tool for the classification of the Lévy processes by the form of their characteristic functions.
In the following let us present some basic definitions for random variables. In $\mathbb{R}^n$ we indicate a vector of random variables with $X = (X^1, ..., X^n)$.[2]

[1] The natural filtration is defined as $\mathcal{F}_t^X = \sigma\{X(s) : 0 \leq s \leq t\}$.

[2] Let us recall some basic definitions. We denote the inner product for $x, y \in \mathbb{R}^n$ by

$$(x, y) := \sum_{i=1}^{n} x_i y_i,$$

and the Euclidean norm by

$$|x| := \sqrt{(x, x)}.$$

Definition 1.1.2. *Let $X : (\Omega, \mathcal{F}, \mathbb{P}) \to \mathbb{R}^n$.*
*The **Characteristic function** $\phi_X : \mathbb{R}^n \to \mathbb{C}$ of X, is defined by*

$$\begin{aligned}\phi_X(u) &= \mathbb{E}[e^{i(u,X)}] \\ &= \int_\Omega e^{i(u,X)}\mathbb{P}(d\omega) \\ &= \int_{\mathbb{R}^n} e^{i(u,x)} f_X(dx). \end{aligned} \tag{1.4}$$

*for each $u \in \mathbb{R}^n$. We indicated with f_X the **probability density function (pdf)** of X.*

For each $1 \leq j \leq n$ and $p \in \mathbb{N}$, if $\mathbb{E}\big[|(X^j)^p|\big] < \infty$, then

$$\mathbb{E}\Big[(X^j)^p\Big] = i^{-p}\frac{\partial^p}{\partial u_j^p}\phi_X(u)\Big|_{u=0}. \tag{1.5}$$

With this property it is straightforward to compute the moments of each component of the random vector, as long as we know the analytic form of the characteristic function. The following properties hold for all $p > 0$:

- $\mathbb{E}\big[|X|^p\big] < \infty$ if and only if $\mathbb{E}\big[|(X^j)|^p\big] < \infty$ for each $1 \leq j \leq n$.
- If $\mathbb{E}\big[|X|^p\big] < \infty$ then $\mathbb{E}\big[|X|^q\big] < \infty$ for all $0 < q < p$.

For more information see sections 1.1.2 and 1.1.6 of [Applebaum, 2009].

Definition 1.1.3. *Let $X : (\Omega, \mathcal{F}, \mathbb{P}) \to \mathbb{R}^n$. We say that X is **infinitely divisible** if for all $m \in \mathbb{N}$ there exist i.i.d. random variables $X_1^{(m)}, ..., X_m^{(m)}$ such that*

$$X \overset{d}{=} X_1^{(m)} + ... + X_m^{(m)}. \tag{1.6}$$

The superscript (m) is to remember that the random variables depend on the initial choice of $m \in \mathbb{N}$.

Theorem 1.1.1. *A Lévy process $\{X_t\}_{t\geq 0}$ is infinitely divisible for each $t \geq 0$.*

Proof. For each $n \in \mathbb{N}$ we can write

$$X_t = \big(X_1^{(m)}\big)_t + ... + \big(X_m^{(m)}\big)_t$$

where $\big(X_k^{(m)}\big)_t = X_{\frac{kt}{m}} - X_{\frac{(k-1)t}{m}}$ are i.i.d. by definition (1.1.1, L2). □

The opposite implication also holds.

Theorem 1.1.2. *Every infinitely divisible distribution is the distribution of a Lévy process.*

A proof of the previous theorem can be found in [Applebaum, 2009] (Corollary 1.4.6).

Other properties of infinitely divisible distribution and connections with Lévy processes can be found in the chapter 2 of [Sato, 1999].

1.1.2 Lévy-Khintchine representation

We now present a beautiful formula, first established by Paul Lévy and A.Ya. Khintchine in the 1930s which gives a characterization of every infinitely divisible random variable.

Definition 1.1.4. *Let $\nu(dx)$ be a Borel measure. We say it is a **Lévy measure** if it satisfies*

$$\nu(\{0\}) = 0, \tag{1.7}$$

$$\int_{\mathbb{R}^n} (1 \wedge x^2)\nu(dx) < \infty. \tag{1.8}$$

The characteristic function of an infinitely divisible random variable has the following **Lévy Khintchine representation**:

Theorem 1.1.3. *Let X be an infinitely divisible random variable. Then there exist $b \in R^n$, a positive definite $n \times n$ matrix A and a Lévy measure ν on $\mathbb{R}^n$, such that $\forall u \in \mathbb{R}^n$:*

$$\begin{aligned}\phi_X(u) &= \mathbb{E}[e^{i(u,X)}] \\ &= e^{\eta(u)} \\ &= \exp\left[\left(i(b,u) - \frac{1}{2}(u,Au) + \int_{\mathbb{R}^n} (e^{i(u,x)} - 1 - i(u,x)\mathbb{1}_{(|x|<1)}(x))\nu(dx)\right)\right].\end{aligned} \tag{1.9}$$

A proof can be found in [Applebaum, 2009] (Theorem 1.2.14). We call the map $\eta : \mathbb{R}^n \to \mathbb{C}$, the **Lévy symbol**.

Now we can easily find the Lévy Khintchine representation for a Lévy process

Theorem 1.1.4. *If $\{X_t\}_{t\geq 0}$ is a Lévy process, then*

$$\phi_{X_t}(u) = e^{t\eta(u)}, \tag{1.10}$$

where η is the Lévy symbol of the random variable X_t at $t = 1$.

See [Applebaum, 2009] (Theorem 1.3.3). The triplet (b, A, ν) is called **Lévy triplet**, and completely characterizes the Lévy process.

1.1.3 Random measures

A convenient tool for analyzing the jumps of a Lévy process is the random measure of the jumps of the process. The jump process $\{\Delta X_t\}_{t\geq 0}$ associated to the Lévy process $\{X_t\}_{t\geq 0}$ is defined, for each $t \geq 0$, by

$$\Delta X_t = X_t - X_{t^-} \tag{1.11}$$

where $X_{t^-} = \lim_{s\uparrow t} X_s$.

Definition 1.1.5. *Consider a set $A \in \mathcal{B}(\mathbb{R}^n\backslash\{0\})$. We define the **random measure** of the jumps of the process $\{X_t\}_{t\geq 0}$ by*

$$\begin{aligned}N^X(t, A)(\omega) &= \#\{s \in [0, t] \,:\, \Delta X_s(\omega) \in A\} \\ &= \sum_{0\leq s\leq t} \mathbb{1}_A(\Delta X_s(\omega)).\end{aligned} \tag{1.12}$$

For each $\omega \in \Omega$ and for each $0 \leq t < \infty$, the map

$$A \to N^X(t, A)(w)$$

is a counting measure on $\mathcal{B}(\mathbb{R}^n \backslash \{0\})$.

We say that $A \in \mathcal{B}(\mathbb{R}^n \backslash \{0\})$ is *bounded below* if $0 \notin \bar{A}$ (zero does not belong to the closure of A).

- For each A bounded below, the process $\left\{N^X(t, A)(\omega)\right\}_{t \geq 0}$ is a Poisson process with intensity
$$\mu(A) = \mathbb{E}[N^X(1, A)] \tag{1.13}$$
- If $A_1, ..., A_m \in \mathcal{B}(\mathbb{R}^n \backslash \{0\})$ are disjoint and bounded below and $t_1, ..., t_m \in \mathbb{R}^+$ are distinct, then the random variables $N(t_1, A_1), ..., N(t_m, A_m)$ are independent.

(see [Applebaum, 2009] Theorem 2.3.5).
A random measure satisfying the properties above is called **Poisson random measure**.
If A is not bounded below, it is possible to have $\mu(A) = \infty$.

We can also define the **Compensated Poisson random measure**. For each $t \geq 0$ and A bounded below, let us define:

$$\tilde{N}(t, A) = N(t, A) - t\mu(A). \tag{1.14}$$

This is a martingale-valued measure, i.e. for each A the process $\left\{\tilde{N}(t, A)\right\}_{t \geq 0}$ is a martingale.
Now we can define the integration with respect to a random measure:

Definition 1.1.6. *Let N be the Poisson random measure associated to a Lévy process $\{X_t\}_{t \geq 0}$, and let $f : \mathbb{R}^n \to \mathbb{R}^n$ be a Borel-measurable function. For any A bounded below, we define the* ***Poisson integral*** *of f as*

$$\int_A f(x) N(t, dx)(\omega) = \sum_{x \in A} f(x) N(t, \{x\})(\omega). \tag{1.15}$$

Since $N(t, \{x\}) \neq 0 \Leftrightarrow \Delta X_s = x$ for at least one $s \in [0, t]$, we have

$$\int_A f(x) N(t, dx)(\omega) = \sum_{0 \leq s \leq t} f(\Delta X_s) \mathbb{1}_A(\Delta X_s). \tag{1.16}$$

The Poisson integral has the following important properties:

Theorem 1.1.5. *For $t \geq 0$ and for any A bounded below, the random variable $\int_A f(x) N(t, dx)$ has the characteristic function*

$$\mathbb{E}\left[\exp\left(iu \int_A f(x) N(t, dx)\right)\right] = \exp\left(t \int_{\mathbb{R}^n} [e^{iux} - 1] \mu_{A,f}(dx)\right), \tag{1.17}$$

where $\mu_{A,f}(B) = \mu(A \cap f^{-1}(B))$ for each $B \in \mathcal{B}(\mathbb{R}^n)$.

This is the Theorem 2.3.7 in [Applebaum, 2009]. By differentiation (see eq. 1.5) we can derive:

$$\mathbb{E}\left[\int_A f(x)N(t,dx)\right] = t\int_A f(x)\mu(dx) \quad \text{for} \quad f \in L^1(A,\mu_A), \tag{1.18}$$

$$Var\left[\left|\int_A f(x)N(t,dx)\right|\right] = t\int_A |f(x)|^2\mu(dx) \quad \text{for} \quad f \in L^2(A,\mu_A). \tag{1.19}$$

We can also define in the same way the **compensated Poisson integral**

$$\int_A f(x)\tilde{N}(t,dx) := \int_A f(x)N(t,dx) - t\int_A f(x)\mu(dx), \tag{1.20}$$

The random variable $\int_A f(x)\tilde{N}(t,dx)$ has characteristic function

$$\mathbb{E}\left[\exp\left(i(u,\int_A f(x)\tilde{N}(t,dx))\right)\right] = \exp\left(t\int_{\mathbb{R}^n}[e^{(i(u,x))} - 1 - i(u,x)]\mu_{A,f}(dx)\right). \tag{1.21}$$

for each $B \in \mathcal{B}(\mathbb{R}^n)$.
The process $\{\int_A f(x)\tilde{N}(t,dx)\}_{t\geq 0}$ is a martingale. For $f \in L^2(A,\mu_A)$ it holds

$$Var\left[\left|\int_A f(x)\tilde{N}(t,dx)\right|\right] = t\int_A |f(x)|^2\mu(dx). \tag{1.22}$$

Theorem 1.1.6. *The intensity measure μ is a Lévy measure.*

See Corollary 2.4.12 in [Applebaum, 2009]. From here on, we always indicate the Lévy measure with the symbol ν.

We can further define:

$$\int_{|x|<1} f(x)\tilde{N}(t,dx) := \lim_{\epsilon\to 0}\int_{\epsilon<|x|<1} f(x)\tilde{N}(t,dx), \tag{1.23}$$

that represents the compensated sum of small jumps.

1.1.4 Lévy-Itō decomposition

The following is a fundamental theorem which decomposes a general Lévy process in the superposition of independent processes: a drift term, a Brownian motion, a Poisson process with "big jumps" and a compensated Poisson process with "small jumps".

Theorem 1.1.7. *Given a Lévy process $\{X_t\}_{t\geq 0}$, there exist $b \in \mathbb{R}^n$, a Brownian motion W^A with diffusion matrix A, and an independent Poisson random measure N on $\mathbb{R}^+ \times \mathbb{R}^n$ such that*

$$X_t = bt + W_t^A + \int_{|x|<1} x\tilde{N}(t,dx) + \int_{|x|\geq 1} xN(t,dx). \tag{1.24}$$

*This is called **Lévy-Itō decomposition**.*

For a proof the reader can look at Theorem 2.4.16 in [Applebaum, 2009].
A lot of information on the features of a Lévy process can be derived from the integrability conditions of its Lévy measure. The next theorem shows that finiteness of moments depends only on the frequency of large jumps.

Theorem 1.1.8. *Let $\{X_t\}_{t\geq 0}$ be a Lévy process with Lévy measure ν. Then $\{X_t\}_{t\geq 0}$ has finite p-moment i.e. $\mathbb{E}[|X_t|^p] < \infty$ for $p > 0$ and for all $t \geq 0$, if and only if $\int_{|x|\geq 1} |x|^p \nu(dx) < \infty$.*

For a proof we refer to [Applebaum, 2009] theorem 2.5.2. In [Sato, 1999] a stronger result is proved (Theorem 25.3).

Theorem 1.1.9. *Let $\{X_t\}_{t\geq 0}$ be a Lévy process with Lévy measure ν. Let g be a non-negative measurable function on $\mathbb{R}^n$ satisfying the sub-multiplicative property[3], then $\{X_t\}_{t\geq 0}$ has finite* g-moment *i.e. $\mathbb{E}[g(X)] < \infty$ for all $t \geq 0$, if and only if $\int_{|x|>1} g(x)\nu(dx) < \infty$.*

The theorem 1.1.8 is a special case of the theorem 1.1.9 since $g(x) = \max\{|x|, 1\}$ is sub-multiplicative. If we consider the sub-multiplicative function $g(x) = e^{(p,x)}$ with $p \in \mathbb{R}^n$, it follows that $\{X_t\}_{t\geq 0}$ has finite **exponential moment** if for all $t \geq 0$:

$$\mathbb{E}\big[e^{(p,X_t)}\big] < \infty \quad \Leftrightarrow \quad \int_{|x|\geq 1} e^{(p,x)}\nu(dx) < \infty. \tag{1.25}$$

See [Sato, 1999] theorem 25.17.

The majority of Lévy processes used in finance have finite moments. For practical reasons, it makes sense to assume finite mean and variance of the price process. In this thesis we will model (see section 1.3) the 1-dimensional dynamics of the prices with the exponential of a Lévy process, i.e. $S_t = S_0 e^{X_t}$. Let us introduce the important assumption:

Assumption **EM** :
In this thesis we consider only 1-dimensional Lévy processes with finite exponential second moment.
According to 1.25 with $p = 2$ it follows that:

$$\mathbb{E}\big[S_t^2\big] < \infty \quad \Leftrightarrow \quad \int_{|x|\geq 1} e^{2x}\nu(dx) < \infty$$

The existence of the exponential 2-moment implies that $\{X_t\}_{t\geq 0}$ has finite p-moment for all $p \in \mathbb{N}$.

If we assume that $\{X_t\}_{t\geq 0}$ has finite first moment, we can simplify the Lévy-Itō formula, by adding the finite terms $\pm\int_{|x|\geq 1} xt\,\nu(dx)$ to (1.24). The last term becomes a martingale and the new drift becomes $b' = b + \int_{|x|\geq 1} x\nu(dx)$.
The new decomposition has the form:

$$X_t = b't + W_t^A + \int_{\mathbb{R}^n} x\tilde{N}(t, dx). \tag{1.26}$$

The Lévy-Khintchine formula 1.10 becomes

$$\phi_X(u) = \exp\left[t\left(i(u,b') - \frac{1}{2}(u,Au) + \int_{\mathbb{R}^n}\big(e^{i(u,x)} - 1 - i(u,x)\big)\nu(dx)\right)\right]. \tag{1.27}$$

Because in (1.26) the only non martingale term is the drift, we see that the drift term $b' = \mathbb{E}[X_{t=1}]$.

[3] A function is said to be sub-multiplicative if $\exists K > 0$ such that $g(x+y) < Kg(x)g(y)$.

1.2 Infinitesimal Generator and stochastic calculus

Lévy processes belong to the big family of Markov processes. Therefore, all the general properties of Markov processes apply to Lévy processes as well. In this section, we do not go too deep in the big field of semigroups theory, but we just present the fundamental definitions and theorems necessary to define the infinitesimal generator of a Lévy process.

We also present the Itō formula for any Lévy stochastic integral, and the sufficient conditions for the existence of solutions of the SDEs considered in this chapter.

1.2.1 Infinitesimal generator of a Markov process

Let us denote with $B_b(\mathbb{R}^n)$ the linear space of all bounded Borel measurable function $f : \mathbb{R}^n \to \mathbb{R}$. Let $C_0(\mathbb{R}^n)$ be the subspace of $B_b(\mathbb{R}^n)$ containing continuous functions such that $\lim_{|x|\to\infty} f(x) = 0$. They are both Banach spaces under the norm $||f|| = \sup\{|f(x)| : x \in \mathbb{R}^n\}$. We also denote with $C_0^n(\mathbb{R}^n)$ the set of $f \in C_0(\mathbb{R}^n)$ such that f is n time differentiable and the partial derivatives of f with order $\leq n$ belong to $C_0(\mathbb{R}^n)$.
Let us also recall some definitions of operators in a Banach space B. A linear operator $\mathcal{L}$ is a mapping from a linear subspace $D_\mathcal{L}$ of B into B such that

$$\mathcal{L}(af + bg) = a\mathcal{L}f + b\mathcal{L}g \quad f, g \in D_\mathcal{L} \quad a, b \in \mathbb{R}.$$

The set $D_\mathcal{L}$ is the domain of $\mathcal{L}$. A linear operator is called bounded if $D_\mathcal{L} = B$ and $||\mathcal{L}|| := \sup_{||f||\leq 1} ||\mathcal{L}f||$ is finite. If the previous conditions are not satisfied, the operator is called unbounded.

Definition 1.2.1. *Let $\{X_t\}_{t\geq 0}$ be an adapted process on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ equipped with a filtration $\mathcal{F}_t$. We say that $\{X_t\}_{t\geq 0}$ is a* ***Markov process*** *if for all $f \in B_b(\mathbb{R}^n)$ and $0 \leq s \leq t < \infty$*

$$\mathbb{E}\big[f(X_t)\big|\mathcal{F}_s\big] = \mathbb{E}\big[f(X_t)\big|X_s\big]. \tag{1.28}$$

The property (1.28) is called **Markov property**. We can define the **stochastic evolution** of a Markov process as

$$(T_{s,t}f)(x) = \mathbb{E}[f(X_t)|X_s = x], \tag{1.29}$$

for every $f \in B_b(\mathbb{R}^n)$ and $0 \leq s \leq t < \infty$. The family of operators $T_{s,t} : B_b(\mathbb{R}^n) \to B_b(\mathbb{R}^n)$ satisfies

1. $T_{s,t}$ is linear for each $0 \leq s \leq t < \infty$.
2. $T_{s,s} = I$ for each $s \geq 0$.
3. $T_{r,s}T_{s,t} = T_{r,t}$ for each $0 \leq r \leq s \leq t < \infty$
4. $f > 0 \Rightarrow T_{s,t}f > 0$.
5. $T_{s,t}$ is a contraction, i.e. $||T_{s,t}|| < 1$ for each $0 \leq s \leq t < \infty$.
6. $T_{s,t}(1) = 1$.

See Theorem 3.1.2 in [Applebaum, 2009]. An immediate consequence of condition 5 is that $T_{s,t}$ is a bounded operator.

Definition 1.2.2. *We can define the* ***transition probability*** *as the mapping* $p_{s,t}(x, B) : \mathbb{R}^n \times \mathcal{B}(\mathbb{R}^n) \to [0, 1]$*, with* $0 \leq s \leq t < \infty$ *as:*

$$p_{s,t}(x, B) := T_{s,t}\mathbb{1}_B(x) = P(X_t \in B | X_s = x). \tag{1.30}$$

It is related with the semigroup operator as follows:

$$(T_{s,t}f)(x) = \int_{\mathbb{R}^n} f(y)p_{s,t}(x, dy). \tag{1.31}$$

The transition probabilities satisfy the properties

1. The maps $x \to p_{s,t}(x, A)$ are measurable for each $A \in \mathcal{B}(\mathbb{R}^n)$.
2. $p_{s,t}(x, \cdot)$ is a probability measure on $\mathcal{B}(\mathbb{R}^n)$ for each $x \in \mathbb{R}^n$.
3. $p_{s,s}(x, B) = \delta_x(B)$ for $s \geq 0$.
4. For $0 \leq r \leq s \leq t$ it satisfies the **Chapman-Kolmogorov** equation

$$p_{r,t}(x, B) = \int_{\mathbb{R}^n} p_{r,s}(x, dy)p_{s,t}(y, B). \tag{1.32}$$

Definition 1.2.3. *We say that the transition probability is* ***time homogeneous*** *if*

$$p_{s,t}(x, B) = p_{0,t-s}(x, B). \tag{1.33}$$

Definition 1.2.4. *We say that the transition probability is* ***translation invariant*** *if*

$$p_{s,t}(x, B) = p_{s,t}(0, B - x), \tag{1.34}$$

where $B - x = \{y - x : y \in B\}$.

Theorem 1.2.1. *Every Lévy process is a time homogeneous and translation invariant Markov process.*

See Theorems (10.4) and (10.5) in [Sato, 1999].

A Markov process is said to be time homogeneous if the associated stochastic evolution operator satisfies $T_{s,t} = T_{0,t-s}$ for all $0 \leq s \leq t < \infty$. This is analogous to require that the transition probabilities are time homogeneous (can be verified by using (1.2.3) and (1.31)). We indicate the operator $T_{0,t}$ as T_t. For a time homogeneous Markov process, the condition 3 for the stochastic evolution operators (1.29) becomes:

$$T_{s+t} = T_s T_t \quad \forall s, t \geq 0. \tag{1.35}$$

Any family of linear operators on a Banach space that satisfies (1.35) is called **semigroup**.

Definition 1.2.5. *The operator* T_t *associated with a time homogeneous Markov process, is called* ***Feller semigroup*** *if:*

1. $T_t : C_0(\mathbb{R}^n) \to C_0(\mathbb{R}^n)$ *for all* $t \geq 0$,
2. *The map* $t \to T_t$ *with* $t \geq 0$ *is strongly continuous at 0, i.e.*

$$\lim_{t \downarrow 0} ||T_t f - f|| = 0.$$

The homogeneous Markov process associated with the Feller semigroup is called **Feller process**.

In the definition of Feller semigroup, we used the space $C_0(\mathbb{R}^n)$ in place of $C_b(\mathbb{R}^n)$. The space $C_0(\mathbb{R}^n)$ has nicer analytical properties than $C_b(\mathbb{R}^n)$ and allows the proof of important probabilistic theorems. In particular, when replacing $C_0(\mathbb{R}^n)$ by $C_b(\mathbb{R}^n)$, the condition 2 above may fail.

Definition 1.2.6. *The* ***infinitesimal generator*** $\mathcal{L}$ *of the Feller semigroup* T_t *is defined by*

$$\mathcal{L}f = \lim_{t \downarrow 0} \frac{T_t f - f}{t}. \tag{1.36}$$

The domain $D_{\mathcal{L}}$ *of* $\mathcal{L}$ *is the subspace of* $C_0(\mathbb{R}^n)$ *such that the limit above exists.*

In general $\mathcal{L}$ is a linear (possibly unbounded) operator. Among the many properties of the infinitesimal generator, it is possible to prove (see Hille-Yosida, Theorem 31.3 in [Sato, 1999]) that $\mathcal{L}$ is closed and that $D_{\mathcal{L}}$ is dense in $C_0(\mathbb{R}^n)$. For more information we refer to Chapter 1 of [Ethier and Kurtz, 2005].

The following theorem gives an explicit form for the infinitesimal generator of a general Lévy process.

Theorem 1.2.2. *Let* $\{X_t\}_{t \geq 0}$ *be a Lévy process with Lévy triplet* (b, A, ν). *Let* T_t *be the associated Feller-semigroup with generator* $\mathcal{L}$. *For each* $f \in C_0^2(\mathbb{R}^n)$, $t \geq 0$, $x \in \mathbb{R}^n$, $\mathcal{L}$ *has the form:*

$$\begin{aligned}(\mathcal{L}f)(x) = &\sum_{j=1}^{n} b_j \frac{\partial f}{\partial x_j}(x) + \frac{1}{2} \sum_{i,j=1}^{n} A_{i,j} \frac{\partial^2 f}{\partial x_i \partial x_j}(x) \\ &+ \int_{\mathbb{R}^n} \left(f(x+y) - f(x) - \sum_{j=1}^{n} y_j \frac{\partial f}{\partial x_j}(x) \mathbb{1}_{\{|y|<1\}}(y) \right) \nu(dy).\end{aligned} \tag{1.37}$$

A proof of this theorem can be found in [Sato, 1999] (Theorem 31.5).

In the more general framework of section 4.1 we will show that the integral term of $\mathcal{L}$ is well defined also for continuous functions that do not vanish at infinity, but have polynomial growth (see definition (2.1.7)).

1.2.2 The Itō formula

Following the section 4.2.2 of [Applebaum, 2009], we call $\mathcal{P}_2(T, E)$ with $E \in \mathcal{B}(\mathbb{R}^n)$, the set of all functions $f : [0, T] \times E \times \Omega \to \mathbb{R}$ satisfying the two conditions:

1. Predictable[4].

[4] Given the probability space $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbb{P})$ and $E \in \mathcal{B}(\mathbb{R}^n)$. A function $f : [0, T] \times E \times \Omega \to \mathbb{R}$ is said to be **predictable**, if for each $0 \leq t \leq T$ the mapping $(x, \omega) \to F(t, x, \omega)$ is $\mathcal{B}(E) \otimes \mathcal{F}_t$-measurable, and for each $x \in E$ and $\omega \in \Omega$ the mapping $t \to f(t, x, \omega)$ is left-continuous.
If f is predictable, the process $t \to f(t, x, \cdot)$ for each $x \in E$, is adapted.

2. $\mathbb{P}\left(\int_0^T \int_E |f(t,x)|^2 \nu(dx)dt < \infty\right) = 1.$

An $\mathbb{R}^n$ valued stochastic process $\{Y_t\}_{t\geq 0}$ is a **Lévy stochastic integral** if can be written as a superposition of an ordinary integral, an Itō integral and a Poisson and compensated Poisson integrals.

$$\begin{aligned} Y_t^i = Y_0^i + \int_0^t G_s^i ds + \int_0^t F_s^i dW(s) \\ + \int_0^t \int_{|x|<1} H^i(s,x)\tilde{N}(ds,dx) \\ + \int_0^t \int_{|x|\geq 1} K^i(s,x)N(ds,dx), \end{aligned} \tag{1.38}$$

for $1 \leq i \leq n$ and $t > 0$. Y_0 is a $\mathcal{F}_0$-measurable random variable. The expression (1.38) has the differential form

$$\begin{aligned} dY_t = G_t dt + F_t dW_t \\ + \int_{|x|<1} H(t,x)\tilde{N}(dt,dx) + \int_{|x|\geq 1} K(t,x)N(dt,dx). \end{aligned} \tag{1.39}$$

where we dropped the indexes and used the convention that integrals are taken on each components. In order for (1.38) to be well defined, for each $1 \leq i \leq n$, the processes $|G_s^i|^{1/2}$ and F_s^i are in $\mathcal{P}_2\big(T, \{0\}\big)$. The function $H^i(s,x) \in \mathcal{P}_2\big(T, \{|x| < 1\}\big)$ and $K^i(s,x)$ have to be predictable.

Now let us introduce the most important formula in stochastic calculus: the **Itō's formula**.

Theorem 1.2.3. *If Y_t is the Lévy stochastic integral (1.38), for each $f \in C^2(\mathbb{R}^n)$ we have*

$$\begin{aligned} df(Y_t) = \sum_{j=1}^n \frac{\partial f}{\partial y_j}(Y_{t^-})G_t^j dt + \sum_{j=1}^n \frac{\partial f}{\partial y_j}(Y_{t^-})F_t^j dW_t \\ + \frac{1}{2}\sum_{j=1}^n \frac{\partial^2 f}{\partial y_j^2}(Y_{t^-})(F_t^j)^2 dt \\ + \int_{|x|\geq 1} \big[f\big(Y_{t^-} + K(t,x)\big) - f(Y_{t^-})\big]N(dt,dx) \\ + \int_{|x|<1} \big[f\big(Y_{t^-} + H(t,x)\big) - f(Y_{t^-})\big]\tilde{N}(dt,dx) \\ + \int_{|x|<1} \big[f\big(Y_{t^-} + H(t,x)\big) - f(Y_{t^-}) - \sum_{j=1}^n \frac{\partial f}{\partial y_j}(Y_{t^-})H^j(t,x)\big]\nu(dx)dt. \end{aligned} \tag{1.40}$$

For a complete proof see [Applebaum, 2009] Theorem 4.4.7. The terms in the first two lines are the same as in the diffusion case. The other terms are due to the discontinuous part of the process.

Let us introduce also the **Itō product rule**.

Theorem 1.2.4. *If Y_t^1 and Y_t^2 are $\mathbb{R}$-valued stochastic integrals of the form (1.38), then for all $t \geq 0$, with probability 1 we have*

$$d\big(Y_t^1 \cdot Y_t^2\big) = Y_{t^-}^1\, dY_t^2 + Y_{t^-}^2\, dY_t^1 + d\big[Y_t^1, Y_t^2\big], \tag{1.41}$$

where the quadratic variation term is

$$\begin{aligned} d\big[Y_t^1, Y_t^2\big] = F_t^1 F_t^2 dt \\ + \int_{|x|<1} H^1(t,x)H^2(t,x)N(dt,dx) \\ + \int_{|x|\geq 1} K^1(t,x)K^2(t,x)N(dt,dx). \end{aligned}$$

The proof of this theorem can be obtained by a direct application of the Itō's formula (1.40) to the product of Y_t^1 and Y_t^2. See Theorem 4.4.13 of [Applebaum, 2009].

The next theorem establishes the useful **Dynkin formula** for Lévy processes:

Theorem 1.2.5. *Let $\{X_t\}_{t\geq 0}$ be a Lévy process, and let $f \in C_0^2(\mathbb{R}^n)$. Let τ be a stopping time*[5] *such that $\mathbb{E}_x[\tau] < \infty$, then*

$$\mathbb{E}_x[f(X_\tau)] = f(x) + \mathbb{E}_x\left[\int_0^\tau \mathcal{L}f(X_s)ds\right], \tag{1.42}$$

where $\mathcal{L}$ is the infinitesimal generator as in eq. (1.37).

Proof. This result comes by applying Itō's formula (1.40) to $f(X_s)$, integrating in $[0,\tau]$ and taking expectation conditioned by $X_0 = x$. □

1.2.3 Existence and uniqueness

Let us consider, for simplicity, a time-homogeneous SDE like:

$$\begin{aligned} dY_t = b(Y_{t^-})dt + \sigma(Y_{t^-})dW_t \\ + \int_{|x|<c} F(Y_{t^-},x)\tilde{N}(dt,dx) + \int_{|x|\geq c} G(Y_{t^-},x)N(dt,dx), \end{aligned} \tag{1.43}$$

with $\{W_t\}_{t\geq 0}$ a d-dimensional Brownian motion. The functions $b : \mathbb{R}^n \to \mathbb{R}^n$, $\sigma : \mathbb{R}^n \to \mathbb{R}^{n\times d}$, $F : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n$ and $G : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n$ are measurable. The constant $c \in [0,\infty]$ give us the freedom to specify the size of the big and small jumps. Usually, a common choice is $c = 1$, as we did for the differential form of a Lévy-type stochastic integral (1.39), and the Lévy-Itō decomposition (1.24). If we want to put both large and small jumps in the same integral, we choose $c = \infty$ or $c = 0$.

Let us choose $c = \infty$ and write the SDE in the following form:

$$\begin{aligned} dY_t = b(Y_{t^-})dt + \sigma(Y_{t^-})dW_t \\ + \int_{\mathbb{R}^n} F(Y_{t^-},x)\tilde{N}(dt,dx). \end{aligned} \tag{1.44}$$

Let us introduce two conditions:

[5] A stopping time is a random variable $\tau : \Omega \to \mathbb{R}_+$ such that $\{w \in \Omega : \tau(\omega) \leq t\} \in \mathcal{F}_t$.

(C1) **Lipschitz condition** There exist $K_1 > 0$, such that $\forall y_1, y_2 \in \mathbb{R}^n$,

$$\begin{aligned} &|b(y_1) - b(y_2)| + ||\sigma(y_1) - \sigma(y_2)|| \\ &+ \int_{\mathbb{R}^n} |F(y_1, x) - F(y_2, x)|\nu(dx) \ \leq \ K_1|y_1 - y_2|. \end{aligned} \tag{1.45}$$

(C2) **Linear growth condition**. There exist $K_2 > 0$, such that $\forall y \in \mathbb{R}^n$,

$$|b(y)|^2 + ||\sigma(y)||^2 + \int_{\mathbb{R}^n} |F(y, x)|^2\nu(dx) \ \leq \ K_2(1 + |y|^2). \tag{1.46}$$

Where for every $(d \times n)$ matrix, the norm is defined as

$$||\sigma||^2 = \sum_{i=1}^{d}\sum_{j=1}^{n}[\sigma_{i,j}]^2. \tag{1.47}$$

If ν is finite, then condition **C2** is a consequence of **C1**. If it is possible to write $F(y, x) = H(y)\rho(x)$, with H Lipschitz and ρ satisfying:

$$\int_{\mathbb{R}^n} |\rho(x)|^2\nu(dx) < \infty, \tag{1.48}$$

then again the growth condition is a consequence of the Lipschitz condition.

Theorem 1.2.6. *Given the conditions **C1**, **C2**, there exist a unique strong solution Y_t of (1.44) with initial condition $Y_{t=0} = Y_0$. The process $\{Y_t\}_{t\geq 0}$ is cádlág and $\mathcal{F}_t$-adapted.*

A proof of existence and uniqueness based on Picard iteration, can be found in [Applebaum, 2009] (see Theorem 6.2.3). For more information we refer to Chapter 3.2 of [Gihman and Skorohod, 1979] where the time inhomogeneous case is considered.

Theorem 1.2.7. *Let us consider the Feller semigroup associated with the Feller process described by the SDE (1.43). For each $f \in C_0^2(\mathbb{R}^n)$, $t \geq 0$, $x \in \mathbb{R}^n$, the infinitesimal generator $\mathcal{L}$ has the form:*

$$\begin{aligned} (\mathcal{L}f)(x) = &\sum_{j=1}^{n} b_j(x)\frac{\partial f}{\partial x_j}(x) + \frac{1}{2}\sum_{i,j=1}^{n} A_{i,j}(x)\frac{\partial^2 f}{\partial x_i \partial x_j}(x) \\ &+ \int_{|y|<c}\left(f\big(x + F(x, y)\big) - f(x) - \sum_{j=1}^{n} F^j(x, y)\frac{\partial f}{\partial x_j}(x)\right)\nu(dy) \\ &+ \int_{|y|\geq c}\Big(f\big(x + G(x, y)\big) - f(x)\Big)\nu(dy). \end{aligned} \tag{1.49}$$

with $A(x) = \sigma(x)\sigma^T(x)$.

The proof can be obtained by the application of the Itō lemma, see Theorem 6.7.4 of [Applebaum, 2009], where the author also show that $C_0^2(\mathbb{R}^n) \subseteq D_{\mathcal{L}}$.

1.3 Exponential Lévy models

Finally we are able to generalize the equation (1.2) for the process of the log-returns[6]. We write:

$$\log\left(\frac{S_t}{S_0}\right) = X_t, \tag{1.50}$$

where X_t is a one dimensional Lévy process with triplet (b, σ^2, ν).

The name **exponential Lévy model** comes from the expression written as:

$$S_t = S_0 e^{X_t}, \tag{1.51}$$

1.3.1 Exponential Lévy SDE

In order to obtain an SDE for the process (1.51), we apply Itō formula (1.40), and we consider the Lévy-Itō decomposition (1.24) for X_t, written in the differential form like (1.39).

$$\begin{aligned} dS_t &= S_0 e^{X_{t^-}} b dt + S_0 e^{X_{t^-}} \sigma dW_t + \frac{1}{2} S_0 e^{X_{t^-}} \sigma^2 dt \\ &+ \int_{|x|\geq 1} (S_0 e^{X_{t^-}+x} - S_0 e^{X_{t^-}}) N(dt, dx) \\ &+ \int_{|x|<1} (S_0 e^{X_{t^-}+x} - S_0 e^{X_{t^-}}) \tilde{N}(dt, dx) \\ &+ \int_{|x|<1} (S_0 e^{X_{t^-}+x} - S_0 e^{X_{t^-}} - x S_0 e^{X_{t^-}}) \nu(dx) dt. \end{aligned}$$

After some substitutions we can see that the resulting equation, as expected, is a generalization of the equation (1.3).

$$\begin{aligned} \frac{dS_t}{S_{t^-}} &= (b + \frac{1}{2}\sigma^2) dt + \sigma dW_t \\ &+ \int_{|x|<1} (e^x - x - 1) \nu(dx) dt \\ &+ \int_{|x|\geq 1} (e^x - 1) N(dt, dx) + \int_{|x|<1} (e^x - 1) \tilde{N}(dt, dx). \end{aligned} \tag{1.52}$$

Thanks to the assumption **EM** (section 1.1.4) we can simplify this equation. First we look at the integrability conditions:

- $\int_{|x|\geq 1} e^x \nu(dx) < \infty$ by **EM**.
- $\int_{|x|\geq 1} 1 \, \nu(dx) < \infty$ by definition (1.1.4) of ν.

We can add and subtract $\pm \int_{|x|\geq 1} (e^x - 1) \nu(dx) dt$ and obtain the final form

$$\begin{aligned} \frac{dS_t}{S_{t^-}} &= \left(b + \frac{1}{2}\sigma^2 + \int_{\mathbb{R}} (e^x - 1 - x \mathbb{1}_{|x|<1}) \nu(dx) \right) dt \\ &+ \sigma dW_t + \int_{\mathbb{R}} (e^x - 1) \tilde{N}(dt, dx). \end{aligned} \tag{1.53}$$

[6] It is also possible to generalize the differential equation (1.3). The solution of the modified SDE is the **Doleans-Dade exponential** of a Lévy process. The two approaches are equivalent (see propositions 8.22 in [Cont and Tankov, 2003]). In this thesis we choose to use exponential Lévy models.

If we set

$$\mu := b + \frac{1}{2}\sigma^2 + \int_{\mathbb{R}} (e^x - 1 - x\mathbb{1}_{|x|<1})\nu(dx) \tag{1.54}$$

we have an SDE of type (1.44).

$$dS_t = \mu S_{t^-}\, dt + \sigma S_{t^-}\, dW_t \; + \int_{\mathbb{R}} S_{t^-}(e^x - 1)\tilde{N}(dt, dx). \tag{1.55}$$

The same equation can be derived quickly by considering the Lévy-Itō form (1.26) for X_t:

$$\begin{aligned} dS_t \; &= S_0 e^{X_{t^-}} \left(b + \int_{|x|\geq 1} x\nu(dx) \right) dt \; + \; S_0 e^{X_{t^-}} \sigma dW_t \; + \; \frac{1}{2} S_0 e^{X_{t^-}} \sigma^2 dt \\ &\quad + \int_{\mathbb{R}} (S_0 e^{X_{t^-}+x} - S_0 e^{X_{t^-}})\tilde{N}(dt, dx) + \int_{\mathbb{R}} (S_0 e^{X_{t^-}+x} - S_0 e^{X_{t^-}} - x S_0 e^{X_{t^-}})\nu(dx)dt \\ &= S_{t^-} \left[\mu dt + \sigma dW_t \; + \int_{\mathbb{R}} (e^x - 1)\tilde{N}(dt, dx) \right]. \end{aligned}$$

It is easy to check that the coefficients of this equation satisfy the conditions (1.45) (1.46) For this purpose let us define $\rho(x) = e^x - 1$ and let us verify that it satisfies the integrability condition (1.48): We write

$$\int_{\mathbb{R}} (e^x - 1)^2 \nu(dx) = \int_{|x|\geq 1} (e^x - 1)^2 \nu(dx) + \int_{|x|<1} (e^x - 1)^2 \nu(dx),$$

- For $|x| \geq 1$ the three integrals $\int_{|x|\geq 1} e^{2x}\nu(dx)$, $\int_{|x|\geq 1}(-2e^x)\nu(dx)$, $\int_{|x|\geq 1} \nu(dx)$ are finite by assumption **EM** and by (1.8).
- For $|x| < 1$:

  $$(e^x - 1)^2 \; = \; x^2 \left(\frac{e^x - 1}{x} \right)^2 \; < \; x^2 (e - 1)^2.$$

  The definition (1.8) of Lévy measure says $\int_{|x|<1} |x|^2 \nu(dx) < \infty$, so

  $$\int_{|x|<1} (e^x - 1)^2 \nu(dx) \; < \; (e - 1)^2 \int_{|x|<1} |x|^2 \nu(dx) \; < \; \infty.$$

So we have checked that the equation (1.55) admits a unique solution which is given by the exponential Lévy process (1.51).

1.3.2 The Merton Model

The first jump-diffusion model for the log-prices is the *Merton model*, presented in [Merton, 1976]. In the same paper the author also obtains a closed form solution for the price of an European vanilla option. The Merton model describes the log-prices evolution as a Lévy process with a nonzero diffusion component and a finite activity jump process with normal distributed jumps.

$$X_t = \bar{b}t + \sigma W_t + \sum_{i=1}^{N_t} Y_i, \tag{1.56}$$

where N_t is a Poisson random variable counting the jumps of X_t in $[0,t]$, and $Y_i \sim \mathcal{N}(\alpha, \xi^2)$ represents the size of the jumps. Using the Poisson integral notation (Def. 1.1.6), the process

$$X_t = \bar{b}t + \sigma W_t + \int_{\mathbb{R}} xN(t,dx)$$

corresponds to the Lévy-Itō decomposition (1.24), where we defined the drift $\bar{b} := b - \int_{|x|<1} x\nu(dx)$.
The Lévy measure of a finite activity Lévy process, can be factorized in the activity λ of the Poisson process and the pdf of the jump size:

$$\begin{aligned}\nu(dx) &= \lambda f_Y(dx),\\ &= \frac{\lambda}{\xi\sqrt{2\pi}} e^{-\frac{(x-\alpha)^2}{2\xi^2}} dx.\end{aligned}$$

such that $\int_{\mathbb{R}} \nu(dx) = \lambda$.
Since the term $\int_{|x|<1} x\nu(dx)$ is finite, the jump process has **finite variation**. However, the Merton model has infinite variation due to the presence of the diffusion component.
The Lévy exponent has the following form:

$$\eta(u) = i\bar{b}u - \frac{1}{2}\sigma^2u^2 + \lambda\left(e^{i\alpha u - \frac{\xi^2u^2}{2}} - 1\right). \tag{1.57}$$

Using the formula for the moments (1.5) we obtain:

$$\begin{aligned}\mathbb{E}[X_t] &= t(\bar{b} + \lambda\alpha). \\ \mathrm{Var}[X_t] &= t(\sigma^2 + \lambda\xi^2 + \lambda\alpha^2).\\ \mathrm{Skew}[X_t] &= \frac{t\lambda(3\xi^2\alpha + \alpha^3)}{\big(\mathrm{Var}[X_t]\big)^{3/2}}.\\ \mathrm{Kurt}[X_t] &= \frac{t\lambda(3\xi^3 + 6\alpha^2\xi^2 + \alpha^4)}{\big(\mathrm{Var}[X_t]\big)^{2}}.\end{aligned} \tag{1.58}$$

The stock price SDE (1.55) has the following form:

$$\frac{dS_t}{S_{t-}} = \bar{\mu}dt + \sigma dW_t + \int_{\mathbb{R}} (e^x - 1)\tilde{N}(dt,dx).^7 \tag{1.59}$$

with

$$\bar{\mu} := \bar{b} + \frac{1}{2}\sigma^2 + \int_{\mathbb{R}} (e^x - 1)\nu(dx).$$

[7] In the literature, the jump part is often indicated with the not rigorous notation $(J-1)dN_t$, where J is lognormal distributed and dN_t is the infinitesimal variation of the Poisson process.

1.3.3 The Variance Gamma process

The *variance gamma* process is a pure jump Lévy process with infinite activity. The first presentation with applications in finance is due to [Madan and Seneta, 1990]. The model presented in their paper is however a symmetric VG model, where there is only an additional parameter which controls the kurtosis, while the skewness is still not considered.
The non-symmetric VG process is described in [Madan et al., 1998] where a closed form solution for European vanilla options is also presented.

The VG process is obtained by time changing a Brownian motion with drift. The new time variable is a random variable T_t whose increments are Gamma distributed with density $T_t \sim \Gamma(\mu t, \kappa t)$ [8].

$$f_{T_t}(x) = \frac{(\frac{\mu}{\kappa})^{\frac{\mu^2 t}{\kappa}}}{\Gamma(\frac{\mu^2 t}{\kappa})} x^{\frac{\mu^2 t}{\kappa}-1} e^{-\frac{\mu x}{\kappa}}. \tag{1.60}$$

The process $\{T_t\}_{t\geq 0}$ is called **subordinator**. In general a subordinator is a one dimensional Lévy process that is non-decreasing almost surely. Therefore it is consistent to be a time variable.
The characteristic function of T_t is:

$$\phi_{T_t}(u) = \left(\frac{1}{1 - iu\frac{\kappa}{\mu}}\right)^{\frac{\mu^2 t}{\kappa}}. \tag{1.61}$$

The Lévy measure is:

$$\nu^{T_t}(dx) = \begin{cases} \frac{\mu^2 e^{-\frac{\mu}{\kappa}x}}{\kappa x} dx, & \text{for } x > 0, \\ 0 & \text{otherwise.} \end{cases} \tag{1.62}$$

If we consider a Brownian motion with drift $X_t = \theta t + \bar{\sigma} W_t$ and substitute the time variable with the gamma subordinator $T_t \sim \Gamma(t, \kappa t)$ ($\mu = 1$), we obtain the **variance gamma** process:

$$X_{T_t} = \theta T_t + \bar{\sigma} W_{T_t}. \tag{1.63}$$

It depends on three parameters:

- $\bar{\sigma}$, the volatility of the Brownian motion
- κ, the variance of the Gamma process
- θ, the drift of the Brownian motion

The VG is a process with **finite variation**. Every process with finite variation can be written as the difference of two increasing processes. In this case the two increasing processes are Gamma processes:

$$X_t = Y_t^p - Y_t^n, \tag{1.64}$$

[8] Usually the Gamma distribution is parametrized by a shape and scale positive parameters $X \sim \Gamma(\rho, \theta)$. The random variable $X_t \sim \Gamma(\rho t, \theta)$ has pdf $f_{X_t}(x) = \frac{\theta^{-\rho t}}{\Gamma(\rho t)} x^{\rho t - 1} e^{-\frac{x}{\theta}}$ and has $\mathbb{E}[X_t] = \rho\theta t$ and $\mathrm{Var}[X_t] = \rho\theta^2 t$. Here we use a parametrization as in [Madan et al., 1998] such that $\mathbb{E}[X_t] = \mu t$ and $\mathrm{Var}[X_t] = \kappa t$, so $\theta = \frac{\kappa}{\mu}$, $\rho = \frac{\mu^2}{\kappa}$.

with $Y_t^p \sim \Gamma(\mu_p t, \kappa_p t)$ and $Y_t^n \sim \Gamma(\mu_n t, \kappa_n t)$. For the specific relation between the parameters $\mu_p, \kappa_p, \mu_n, \kappa_n$ and $\bar{\sigma}, \kappa, \theta$ we refer to [Madan et al., 1998].

The pdf of X_t can be computed conditioning on the realization of T_t:

$$\begin{aligned} f_{X_t}(x) &= \int_y f_{X_t,T_t}(x,y)dy = \int_y f_{X_t|T_t}(x|y) f_{T_t}(y)dy \qquad (1.65)\\ &= \int_0^\infty \frac{1}{\bar{\sigma}\sqrt{2\pi y}} e^{-\frac{(x-\theta y)^2}{2\bar{\sigma}^2 y}} \frac{y^{\frac{t}{\kappa}-1}}{\kappa^{\frac{t}{\kappa}}\Gamma(\frac{t}{\kappa})} e^{-\frac{y}{\kappa}}\, dy \\ &= \frac{2\exp(\frac{\theta x}{\bar{\sigma}^2})}{\kappa^{\frac{t}{\kappa}}\sqrt{2\pi}\bar{\sigma}\Gamma(\frac{t}{\kappa})} \left(\frac{x^2}{2\frac{\bar{\sigma}^2}{\kappa}+\theta^2}\right)^{\frac{t}{2\kappa}-\frac{1}{4}} K_{\frac{t}{\kappa}-\frac{1}{2}}\left(\frac{1}{\bar{\sigma}^2}\sqrt{x^2\big(\frac{2\bar{\sigma}^2}{\kappa}+\theta^2\big)}\right), \end{aligned}$$

where the function K is a modified Bessel function of the second kind (see [Madan et al., 1998] for the computations).
The characteristic function can be obtained by conditioning too:

$$\begin{aligned} \phi_{X_t}(u) &= \left(1 - i\kappa\big(u\theta + \frac{i}{2}\bar{\sigma}^2 u^2\big)\right)^{-\frac{t}{\kappa}} \\ &= \left(1 - i\theta\kappa u + \frac{1}{2}\bar{\sigma}^2\kappa u^2\right)^{-\frac{t}{\kappa}}. \end{aligned}$$

The VG Lévy measure is

$$\nu^{X_t}(dx) = \frac{e^{\frac{\theta x}{\bar{\sigma}^2}}}{\kappa|x|} \exp\left(-\frac{\sqrt{\frac{2}{\kappa}+\frac{\theta^2}{\bar{\sigma}^2}}}{\bar{\sigma}}|x|\right) dx, \qquad (1.66)$$

and the Lévy exponent is

$$\eta(u) = -\frac{1}{\kappa}\log(1 - i\theta\kappa u + \frac{1}{2}\bar{\sigma}^2\kappa u^2). \qquad (1.67)$$

Using the formula for the moments (1.5) we obtain:

$$\begin{aligned} \mathbb{E}[X_t] &= t\theta. \qquad (1.68)\\ \mathrm{Var}[X_t] &= t(\bar{\sigma}^2 + \theta^2\kappa). \\ \mathrm{Skew}[X_t] &= \frac{t(2\theta^3\kappa^2 + 3\bar{\sigma}^2\theta\kappa)}{\big(\mathrm{Var}[X_t]\big)^{3/2}}. \\ \mathrm{Kurt}[X_t] &= \frac{t(3\bar{\sigma}^4\kappa + 12\bar{\sigma}^2\theta^2\kappa^2 + 6\theta^4\kappa^3)}{\big(\mathrm{Var}[X_t]\big)^2}. \end{aligned}$$

The Lévy-Itō decomposition (1.24) for any pure jump finite variation process i.e $\int_{|x|<1} x\nu(dx) < \infty$, can be written as

$$X_t = \tilde{b}t + \int_{\mathbb{R}} xN(t,dx) \qquad (1.69)$$

with $\tilde{b} = b - \int_{|x|<1} x\nu(dx)$. We can apply the Itō formula to (1.51) to obtain the stock price SDE for a finite variation process:

$$\begin{aligned}\frac{dS_t}{S_{t-}} &= \tilde{b}dt + \int_{\mathbb{R}}(e^x - 1)N(dt, dx). \\ &= \left(b + \int_{\mathbb{R}}\big(e^x - 1 - x\mathbb{1}_{|x|<1}(x)\big)\nu(dx)\right)dt + \int_{\mathbb{R}}(e^x - 1)\tilde{N}(dt, dx).\end{aligned} \tag{1.70}$$

Consider the process (1.63). We can take its expectation

$$\mathbb{E}[X_{T_t}] = \theta\mathbb{E}[T_t] + \bar{\sigma}\mathbb{E}[W_{T_t}] = \theta t,$$

which is equal to the expectation of (1.69). Using (1.18) we obtain

$$\begin{aligned}\mathbb{E}[X_t] &= \tilde{b}t + \mathbb{E}\left[\int_{\mathbb{R}} xN(t, dx)\right] \\ &= t\left(\tilde{b} + \int_{\mathbb{R}} x\,\nu(dx)\right),\end{aligned} \tag{1.71}$$

and therefore $\tilde{b} = \theta - \int_{\mathbb{R}} x\nu(dx)$.
We can compute the integral using the explicit formula (1.66) for the Lévy measure. Let us call

$$A = \frac{\theta}{\bar{\sigma}^2} \qquad \text{and} \qquad B = \frac{|\theta|}{\bar{\sigma}^2}\sqrt{1 + \frac{2\bar{\sigma}^2}{\kappa\theta^2}}$$

with $A < B$, and solve the integral:

$$\begin{aligned}\int_{\mathbb{R}} \frac{x}{\kappa|x|}e^{Ax - B|x|} &= \int_0^\infty \frac{1}{\kappa}e^{(A-B)x} - \int_{-\infty}^0 \frac{1}{\kappa}e^{(A+B)x} \\ &= \frac{1}{\kappa}\frac{2A}{B^2 - A^2} \\ &= \theta.\end{aligned}$$

As expected, $\tilde{b} = 0$.
The Lévy-Itō decomposition for the VG process in (1.63) is simply

$$X_t = \int_{\mathbb{R}} xN(t, dx). \tag{1.72}$$

All the information is contained in the Lévy measure (1.66), which completely describes the process. Even if the process has been created by Brownian subordination, it has no diffusion component. The Lévy triplet is

$$\left(\int_{|x|<1} x\nu(dx), 0, \nu\right). \tag{1.73}$$

The SDE for the stock price following an **exponential VG** is

$$\frac{dS_t}{S_{t-}} = \int_{\mathbb{R}}(e^x - 1)N(dt, dx). \tag{1.74}$$

1.3.4 Infinitesimal Generator for exponential Lévy processes

This section derives the infinitesimal generator for the stock price process under the *Merton* and the *Variance Gamma* models.

The stock price SDE (1.55), has the form (1.43) with $c = \infty$. Using (1.49), the corresponding infinitesimal generator is:

$$\begin{aligned}\mathcal{L}^S f(s) = \mu s \frac{\partial f(s)}{\partial s} + \frac{1}{2}\sigma^2 s^2 \frac{\partial^2 f(s)}{\partial s^2} \\ + \int_{\mathbb{R}} \left[f(se^x) - f(s) - s(e^x - 1)\frac{\partial f(s)}{\partial s} \right] \nu(dx).\end{aligned} \tag{1.75}$$

Let us derive the infinitesimal generators for the Merton and VG processes described by the SDEs (1.59) and (1.74):

- **Merton model generator**:
  Under the representation with $c = \infty$, the 1-dim generator $\mathcal{L}^M$ has the form:

$$\begin{aligned}(\mathcal{L}^M f)(s) = \ \bar{\mu} s \frac{\partial f(s)}{\partial s} + \frac{1}{2}\sigma^2 s^2 \frac{\partial^2 f(s)}{\partial s^2} \\ + \ \int_{\mathbb{R}} \big(f(se^x) - f(s) - s(e^x - 1)\frac{\partial f(s)}{\partial s} \big) \nu(dx),\end{aligned} \tag{1.76}$$

  with

$$\bar{\mu} = \bar{b} + \frac{1}{2}\sigma^2 + m$$

  and we defined the parameter

$$m := \int_{\mathbb{R}} (e^x - 1)\nu(dx) = \lambda \left(e^{\alpha + \frac{1}{2}\xi^2} - 1 \right). \tag{1.77}$$

  Under the equivalent representation with $c = 0$, the generator is:

$$\begin{aligned}(\mathcal{L}^M f)(s) = \ \left(\bar{b} + \frac{1}{2}\sigma^2 \right) s \frac{\partial f(s)}{\partial s} + \frac{1}{2}\sigma^2 s^2 \frac{\partial^2 f(s)}{\partial s^2} \\ + \ \int_{\mathbb{R}} \big(f(se^x) - f(s) \big) \nu(dx).\end{aligned} \tag{1.78}$$

- **Variance Gamma generator**:
  Using the representation with $c = 0$, the generator $\mathcal{L}^{VG}$ is:

$$(\mathcal{L}^{VG} f)(s) = \int_{\mathbb{R}} \big(f(se^x) - f(s) \big) \nu(dx). \tag{1.79}$$

  Under $c = \infty$ we obtain the equivalent form:

$$(\mathcal{L}^{VG} f)(s) = ws \frac{\partial f(s)}{\partial s} + \int_{\mathbb{R}} \left(f(se^x) - f(s) - s(e^x - 1)\frac{\partial f(s)}{\partial s} \right) \nu(dx), \tag{1.80}$$

where we introduced the new parameter

$$w := \int_{\mathbb{R}} (e^x - 1)\nu(dx) = -\frac{1}{\kappa}\log\left(1 - \theta\kappa - \frac{1}{2}\bar{\sigma}^2\kappa\right). \tag{1.81}$$

In order to calculate the integral, we use the following relation between the Lévy measure and the transition probability (1.2.2) of the process:

$$\nu(dx) = \lim_{t\to 0}\frac{1}{t}p_{0,t}(0, dx). \tag{1.82}$$

This relation is presented by [Cont and Tankov, 2003] in Chapter 3.6, and a proof can be found in Corollary 8.9 of [Sato, 1999].
Let us compute first the expected value of the exponential VG process

$$\begin{aligned}\mathbb{E}[e^{X_t}] = \phi_{X_t}(-i) &= \exp\left(-\frac{t}{\kappa}\log(1 - \theta\kappa - \frac{1}{2}\bar{\sigma}^2\kappa)\right)\\ &= e^{wt}\end{aligned}$$

where we called $w = -\frac{1}{\kappa}\log(1 - \theta\kappa - \frac{1}{2}\bar{\sigma}^2\kappa)$. The integral becomes

$$\begin{aligned}\int_{\mathbb{R}}(e^x - 1)\nu(dx) &= \int_{\mathbb{R}}(e^x - 1)\lim_{t\to 0}\frac{1}{t}p_{0,t}(0, dx)\\ &= \lim_{t\to 0}\frac{1}{t}\mathbb{E}[e^{X_t} - 1]\\ &= w.\end{aligned}$$

Remember that since the VG has finite variation, the integral is finite because the integrand is $e^x - 1 = x + \mathcal{O}(x^2)$, so we can always take the limit outside the integral.

1.4 Cumulants

The cumulant generating function $H_{X_t}(u)$ of X_t is defined as the natural logarithm of its characteristic function (see [Cont and Tankov, 2003]). Using the Lévy-Khintchine representation for the characteristic function (1.10), it is easy to find its relation with the Lévy symbol:

$$\begin{aligned}H_{X_t}(u) &= \log(\phi_{X_t}(u))\\ &= t\eta(u)\\ &= \sum_{n=1}^{\infty} c_n\frac{(iu)^n}{n!}\end{aligned} \tag{1.83}$$

The **cumulants** of a Lévy process are thus defined by

$$c_n = \frac{t}{i^n}\frac{\partial^n\eta(u)}{\partial u^n}\bigg|_{u=0}. \tag{1.84}$$

The cumulants are closely related to the central moments μ_n:

$$\mu_0 = 1, \quad \mu_1 = 0, \quad \mu_n = \sum_{k=1}^{n}\binom{n-1}{k-1}c_k\mu_{n-k} \quad \text{for } n > 1. \tag{1.85}$$

For a Poisson process with finite first n moments, all the information about the cumulants is contained inside the Lévy measure. Let us expand in Taylor series the exponential

$$e^{iux} \approx 1 + iux - \frac{u^2x^2}{2} - \frac{iu^3x^3}{3!} + \frac{u^4x^4}{4!} + \dots$$

The Lévy symbol corresponding to the representation (1.69), for a process with finite variation with $\tilde{b} = b - \int_{|x|<1} x\nu(dx)$ becomes

$$\begin{aligned} t\eta(u) &= i\tilde{b}ut + t\int_{\mathbb{R}}(e^{iux} - 1)\nu(dx) \qquad (1.86)\\ &= i\Big(b - \int_{|x|<1} x\nu(dx)\Big)ut + iut\int_{\mathbb{R}} x\nu(dx) - \frac{u^2}{2}t\int_{\mathbb{R}} x^2\nu(dx)\\ &\quad - \frac{iu^3}{3!}t\int_{\mathbb{R}} x^3\nu(dx) + \frac{u^4}{4!}t\int_{\mathbb{R}} x^4\nu(dx) + \dots\\ &= ic_1u - \frac{c_2u^2}{2} - \frac{ic_3u^3}{3!} + \frac{c_4u^4}{4!} + \dots \end{aligned}$$

with $c_1 = t\big(b + \int_{|x|\geq 1} x\nu(dx)\big)$.

1.5 Chapter conclusions

In this chapter we present the main theoretical features of Lévy processes in a quite general setting. In this thesis we model the log-price dynamics by using stochastic differential equations involving Lévy processes. For this purpose it is important to present a complete framework containing all the tools needed in the next chapters. Starting with the basic definition of a Lévy process and the form of its characteristic function, we then introduce the concept of *random measures* which is a method to describe the jumps of the process, and the important *Lévy-Itō decomposition* that permits to decompose any Lévy processes into the superposition of a Brownian motion with drift and two jump process. We then introduce the assumption of finite exponential second moment, called assumption **EM**, which is recurrent throughout the thesis.
The two concepts of *infinitesimal generator* and *Itō formula* are presented in a general form in section 1.2, where we also discussed the hypothesis for the existence of a unique solution of the Lévy-type SDE. All these concepts are applied in section (1.3) to the specific case of exponential Lévy processes. We derive the exponential Lévy SDE and the form of the infinitesimal generator associated to the process. These formulas are fundamental in the thesis and will be used several times in chapter 2 for the derivation of the pricing PIDE, and in chapter 5 for the derivation of the dynamic programming equation associated with the main problem of the thesis.
In the end of the chapter we introduce two of the most popular Lévy processes used in finance, i.e. the *Merton model* and the *Variance Gamma model*. They are used throughout the thesis as practical examples in the numerical calculations. In chapter 2 we present the numerical solution of the Merton and VG pricing PIDEs and in chapter 6 we solve numerically the pricing problem under transaction costs considering these two processes. Furthermore, the whole 3 is

dedicated to the application of the multinomial method to the pricing problem under VG process.

2

The martingale approach to option pricing

In mathematical finance, one of the key issues is to find the fair price of a financial derivative. A financial derivative is a security whose value depends on the price of one or more basic securities such as stocks or bonds (the so called **underlying assets**). A **European call** option on a security with price process $\{S_t\}_{t\geq 0}$, is the right to buy the security at the predetermined exercise price K (also called **strike**). This right may be exercised at the expiration date T (also called **maturity**) of the option. The call option can be purchased at the price C_{t_0} at time $t_0 = 0 < T$. A **European put** option is similar, but gives the owner the right to sell the underlying asset at the strike price at maturity. In contrast to European options, **American options** can be exercised at any time between the writing and the expiration of the contract. The value of the option at the exercise time τ such that $0 \leq \tau \leq T$, is called **payoff**, and has the form

$$\text{Call:} \quad C_\tau = \max\{S_\tau - K, 0\}$$

$$\text{Put:} \quad P_\tau = \max\{K - S_\tau, 0\}$$

Because of the max operator in the payoff, the options are nonlinear instruments.

Determining the correct price of an option is not a simple task. It requires a stochastic model for the dynamics of the underlying price and several assumptions on the market. The solution of this problem has been presented for the first time in the celebrated paper of [Black and Scholes, 1973]. The Black-Scholes (BS) model is built on the concept of an **ideal market**, where all the following conditions are fulfilled:

1. *There are no arbitrage possibilities (see [2.1.3] below).*
2. *Exists a risk free asset B with dynamics*

$$dB_t = r_t B_t dt \tag{2.1}$$

   where r_t is the deterministic risk free interest rate.
3. *It is possible to borrow and lend any amount of cash, even fractional, at the risk free rate.*
4. *It is possible to buy and sell any amount, even fractional, of the stock. This includes short selling.*
5. *The market is frictionless.*
6. *For $\mu \in \mathbb{R}$ and $\sigma > 0$, the underlying stock process $\{S_t\}_{t\geq 0}$ follows the geometric Brownian motion*

$$\frac{dS_t}{S_t} = \mu dt + \sigma dW_t. \tag{2.2}$$

7. *The stock does not pay dividends.*

By relaxing one or more of the previous assumptions, it is possible to develop new models that usually are generalizations of the BS model.

The hypothesis of *frictionless market* implies that the cost of trading is zero. This means that: all investors are price takers (no market impact), all parties have the same access to the relevant information, there are no transaction costs or commissions, and all assets are assumed to be perfectly divisible and liquid. Transaction costs can be divided in two categories: costs that are proportional to the price of the traded security, and fixed costs.

In this thesis we relax the hypothesis (5) by introducing proportional transaction costs, and the hypothesis (6) by replacing the geometric Brownian motion with an exponential Lévy process.

In this chapter, we present the basic concepts of the "No-arbitrage" pricing theory and the numerical algorithms based on the finite difference method. We will see that the pricing model can always be represented by a partial integro-differential equation (PIDE) when the stock dynamics follows an exponential Lévy process.

2.1 No arbitrage theory

All the mathematical framework for derivative pricing is based on the concept of **No-Arbitrage**.
It is convenient to introduce some useful definitions.

Definition 2.1.1. *Let us consider a financial market containing a stock with price process* $\{S_t\}_{t\geq 0}$ *defined on the probability space* $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geq 0}, \mathbb{P})$.
A ***financial derivative*** *(or contingent claim)*[1] *with maturity* T *is any* $\mathcal{F}_T$*-measurable random variable* $\mathcal{X}$.

A financial derivative is called "simple" if it is of the form $\mathcal{X} = \Phi(S_T)$. *The function* Φ *is called the* ***payoff****.*

We see that European call and put options are simple financial derivatives. All European path dependent options are examples of non-simple derivatives. In thesis we will consider only simple derivatives. Also American-style contracts are included into this definition if we replace the maturity date T with the exercise date τ such that $\tau \leq T$. However, in this chapter we do not consider such instruments because their pricing formula requires the introduction of a more complex framework i.e. optimal stopping theory.

Definition 2.1.2. *Let us consider a stock with (cádlág) price process* $\{S_t\}_{t\geq 0}$ *and a risk free asset with process* $\{B_t\}_{t\geq 0}$.

1. *A* ***portfolio*** *is a pair of predictable processes* $(\alpha_t, \beta_t)_{t\geq 0}$ *describing the amount of each asset held by the investor.*
2. *The value of such a portfolio at time* t *is*

$$\Theta_t = \alpha_t S_t + \beta_t B_t.$$

   The process $\{\Theta_t\}_{t\geq 0}$ *is called* ***portfolio value process****.*
3. *A portfolio* $(\alpha_t, \beta_t)_{t\geq 0}$ *is said to be* ***self-financing*** *if*

$$\Theta_t = \Theta_0 + \int_{0^+}^{t} \alpha_u dS_u + \int_{0^+}^{t} \beta_u dB_u. \tag{2.3}$$

The meaning of equation (2.3) is that the value of the portfolio at time t is equal to the initial value Θ_0 plus the capital gain between 0 and t.

Definition 2.1.3. *An* ***arbitrage*** *is a self-financing portfolio value process* $\{\Theta_t\}_{t\geq 0}$ *satisfying* $\Theta_0 = 0$ *and also for some* $T > 0$

$$\mathbb{P}\big(\Theta_T \geq 0\big) = 1 \quad \textit{and} \quad \mathbb{P}\big(\Theta_T > 0\big) > 0.$$

An arbitrage is a trading strategy such that an investor starts with zero capital and at some later time T he is sure to have lost no money and furthermore has a positive probability of having made a profit.

We can define the discount factor for $0 \leq s \leq t \leq T$ as

$$D(s,t) = e^{-\int_s^t r_u du}. \tag{2.4}$$

In the following of this thesis we assume a constant interest rate $r_u = r$ for all $u \in [0, T]$.
It is common to indicate with $\mathbb{P}$ the physical probability measure and with $\mathbb{Q}$ a risk neutral measure, also called equivalent martingale measure (EMM).

[1] With the term "contingent claim" some authors indicate only contracts with an optionality. In this thesis we follow the nomenclature of [Björk, 2009] and [Musiela and Rutkowski, 2005].

Definition 2.1.4. *Given the asset price process $\{S_t\}_{t\geq 0}$ defined on the probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geq 0}, \mathbb{P})$, we say that the probability measure $\mathbb{Q}$ is an **EMM** if it verifies the following two properties:*

$$\mathbb{Q} \sim \mathbb{P}: \quad \forall A \in \mathcal{F} \qquad \mathbb{Q}(A) = 0 \Leftrightarrow \mathbb{P}(A) = 0, \tag{2.5}$$

$$D(0,t)S_t = \mathbb{E}^{\mathbb{Q}}\big[D(0,T)S_T \big| \mathcal{F}_t\big] \quad \textit{for} \quad 0 \leq t \leq T. \tag{2.6}$$

Our main problem is to determine the fair price $\Pi_t(\mathcal{X})$ of the derivative $\mathcal{X}$ at time t (with $0 \leq t \leq T$). There are some minimal requirements that $\Pi_t(\mathcal{X})$ should verify to qualify as a pricing rule:

1. It should be possible to price a derivative using only the information given at time t.
2. A derivative with a positive payoff should have a positive value.
3. Linearity.

In [Cont and Tankov, 2003] (see Proposition 9.1), the authors prove that any arbitrage-free linear pricing rule $\Pi_t(\mathcal{X})$ satisfying the properties above, is represented by the *risk neutral* pricing rule

$$D(0,t)\Pi_t(\mathcal{X}) = \mathbb{E}^{\mathbb{Q}}\big[D(0,T)\mathcal{X} \big| \mathcal{F}_t\big] \quad \text{for} \quad 0 \leq t \leq T, \tag{2.7}$$

where $\mathbb{Q}$ is an EEM.

The concept of arbitrage is related with the existence of an equivalent martingale measures through the **first fundamental theorem of asset pricing**.

Theorem 2.1.1. *A market model does not admit arbitrage if and only if there exists a risk-neutral probability measure.*

For a detailed proof we refer to the academic literature on this topic i.e. [Harrison and Kreps, 1979], [Harrison and Pliska, 1981], [Schachermayer, 2002], [Delbaen and Schachermayer, 1998]. The general No-arbitrage theory of asset pricing is a fundamental and sophisticated theory of mathematical finance with several important developments in the last forty years, and we do not discuss it in details in this thesis. For a general presentation of this theory we refer to [Björk, 2009] (Chapter 10.2). For a comprehensive introduction, considering discontinuous processes, we refer to [Cont and Tankov, 2003] (Chapter 9.1).

Another important concept originating in the Black-Scholes model is the concept of perfect hedge.

Definition 2.1.5. *A self-financing portfolio (α_t, β_t) is said to be a **perfect hedge** (or **replicating portfolio**) for a derivative $\mathcal{X}$ if the associated portfolio value process $\{\Theta_t\}_{t\geq 0}$ satisfies*

$$\mathcal{X} = \Theta_T \qquad \mathbb{P} \text{ - almost surely.} \tag{2.8}$$

If the market is arbitrage-free and (2.8) holds under $\mathbb{P}$, it must hold also under $\mathbb{Q}$, since they are equivalent.

The initial value Θ_0 is the price of the derivative at time 0. Using (2.8) and (2.3), under $\mathbb{Q}$ we can write $D(0,T)\mathcal{X} = \Theta_0 + \int_{0+}^{T} \alpha_t \, d\big(D(0,t)S_t\big)$. Since

the discounted price $D(0,t)S_t$ under $\mathbb{Q}$ is a martingale, the stochastic integral $\int_{0+}^{T} \alpha_t \, d\big(D(0,t)S_t\big)$ is a martingale as well. By taking the expectation we obtain:

$$\Pi_0(\mathcal{X}) = \mathbb{E}^{\mathbb{Q}}\big[D(0,T)\mathcal{X}\big|\mathcal{F}_t\big] = \Theta_0. \tag{2.9}$$

Moreover, the initial value Θ_0 is unique, since two replicating strategies with different initial values lead to an arbitrage.

Definition 2.1.6. *A market model is said to be* ***complete*** *if the payoff of every derivative security can be perfectly hedged.*

In a complete market, the unique price of a financial derivative corresponds to the initial capital Θ_0 needed to set up a perfect hedge.

The completeness of a market is connected with the uniqueness of the EMM through the **second fundamental theorem of asset pricing**.

Theorem 2.1.2. *Consider a market model that has a risk-neutral probability measure. The model is complete if and only if the risk-neutral probability measure is unique.*

The theorem as stated above holds in discrete time models. In continuous time this formulation is not rigorous. It is necessary to carefully define the set of admissible trading strategies, contingent claims and the notion of martingale measure. In the case where the stock process has unbounded jumps, which is the case of most exponential Lévy models, a rigorous formulation is quite difficult and we refer to [Cherny and Shiryaev, 2002] and [Kabanov, 2001] for more information.

The ideal market assumed by Black-Scholes is complete (see Theorem 8.3 of [Björk, 2009]). However, the majority of the models used in finance are not.

In this thesis we analyze a market model that is not complete. The introduction of proportional transaction costs prevents to trade continuously in time as suggested by the replicating portfolio strategy, because the cost of trading would be infinity. On the other hand, if we consider exponential Lévy models, we again obtain an incomplete market. The reason is that the unpredictable jumps of a Lévy process are a source of risk that cannot be hedged by the replicating strategy.

In a complete market there is only one arbitrage-free way to price a financial derivative, and the price is defined as the cost to replicate the derivative's payoff. In an incomplete market, instead, the notion of perfect replication does not exist. Therefore the hedging strategy will only approximate the payoff of the derivative. Different ways to measure the risk of trading lead to different ways to hedge the derivative. In this thesis we focus on the approach based on the concept of *utility maximization* (see chapter 5).

In the following of this chapter we consider a frictionless market where the stock price follows an exponential Lévy process. In in such a market, the class of EMM is infinite, i.e. there are infinite EMMs such that the discounted stock prices are martingales. This means that for every financial derivative there are infinite prices satisfying the condition of no-arbitrage.

In order to overcome this problem, there are several methods to select the best EMM to use in the pricing formula (2.7) (see [Cont and Tankov, 2003], chapter 10). However, the best approach is to derive the model parameters

directly from the prices of derivatives (usually European call and put options with different strikes and maturities) already quoted in the market. The process of choosing the risk neutral parameters for a model that reproduces the prices in the market is called *model calibration*. In this chapter all the parameters of the Lévy processes used for the pricing purpose, are intended to be risk neutral parameters obtained by a calibration process.

2.1.1 Derivation of the price PIDE

In an arbitrage-free market, if the price process $\{S_t\}_{t\geq 0}$ follows an exponential Lévy process, we can express the price of any simple financial derivative $\mathcal{X}$ as a function of the current time $t \in [0,T]$ and current stock price $s = S_t$, i.e. $\Pi_t(\mathcal{X}) = V(t,s)$. This is a direct consequence of the Markov property (1.28) applied to the formula (2.7). In this section we show that $V(t,s)$ can be obtained by solving a partial integro-differential equation (PIDE).

First of all, let us define the space of continuous functions that have polynomial growth of order p at infinity.

Definition 2.1.7. *For $E \subseteq \mathbb{R}^n$ and for $p \geq 0$, let us define the space:*

$$\mathcal{C}_p\big(E\big) = \left\{ \phi \in C^0\big(E\big) : \sup_E \frac{|\phi(x)|}{1+|x|^p} < \infty \right\}. \tag{2.10}$$

For functions with time dependence, e.g. $\phi : [0,T] \times \mathbb{R}^n \to \mathbb{R}$, we maintain the same notation but specify the domain:

$$\mathcal{C}_p\big([0,T]\times\mathbb{R}^n\big) = \left\{ \phi \in C^0\big([0,T]\times\mathbb{R}^n\big) : \sup_{[0,T]\times\mathbb{R}^n} \frac{|\phi(t,x)|}{1+|x|^p} < \infty \right\}. \tag{2.11}$$

We consider only the case $p = 2$, because we are working with underlying processes with finite second moment. (see assumption **EM** in section 1.1.4).

The following theorem will be useful.

Theorem 2.1.3. *Let $\{X_t\}_{t\geq 0}$ be a Lévy process with Lévy triplet (b,σ,ν), satisfying the assumption EM. The process $\{S_t\}_{t\geq 0}$ defined by $S_t = e^{X_t}$ is a martingale if and only if*

$$b + \frac{1}{2}\sigma^2 + \int_{\mathbb{R}} \big(e^z - 1 - z\mathbb{1}_{\{|z|<1\}}\big)\nu(dz) = 0. \tag{2.12}$$

Proof. The SDE for the process $\{e^{X_t}\}_{t\geq 0}$ has been derived in Eq. (1.53). The exponential Lévy process is a martingale if and only if the drift is zero. □

Let us consider a stock price process described by the *exponential Lévy model*

$$S_t = S_0 e^{L_t} = S_0 e^{rt + X_t} \tag{2.13}$$

where $\{X_t\}_{t\geq 0}$ is a Lévy process with Lévy triplet (b,σ,ν). Under a risk neutral measure $\mathbb{Q}$, the process $\{L_t\}_{t\geq 0}$ is a Lévy process with triplet $(r+b,\sigma,\nu)$ satisfying (2.12). The discounted price is a $\mathbb{Q}$-martingale:

$$\mathbb{E}^{\mathbb{Q}}\big[e^{-rt}S_t\big|S_0\big] = \mathbb{E}^{\mathbb{Q}}\big[S_0 e^{X_t}\big|S_0\big] = S_0, \tag{2.14}$$

such that $\mathbb{E}^{\mathbb{Q}}[e^{X_t}|X_0 = 0] = 1$.

In Chapter 1 we derived the infinitesimal generator for an exponential Lévy process in Eq. (1.75) and the parameter μ in (1.54). We can repeat the same computation that led to Eq. (1.53) for the process $L_t = X_t + rt$ and define the new parameter

$$\mu := r + b + \frac{1}{2}\sigma^2 + \int_{\mathbb{R}} (e^z - 1 - z1_{|z|<1})\nu(dz) \tag{2.15}$$

Using the condition (2.12) we obtain the fundamental relation

$$\mu = r. \tag{2.16}$$

The risk neutral dynamics of (2.13) is described by the SDE:

$$dS_t = rS_{t^-}dt + \sigma S_{t^-}dW_t + \int_{\mathbb{R}} S_{t^-}(e^z - 1)\tilde{N}(dt, dz). \tag{2.17}$$

For $f \in C^2(\mathbb{R}^+) \bigcap \mathcal{C}_2(\mathbb{R}^+)$, the associated infinitesimal generator is:

$$\begin{aligned}\mathcal{L}^S f(s) = rs\frac{\partial f(s)}{\partial s} + \frac{1}{2}\sigma^2 s^2 \frac{\partial^2 f(s)}{\partial s^2} \\ + \int_{\mathbb{R}} \left[f(se^z) - f(s) - s(e^z - 1)\frac{\partial f(s)}{\partial s} \right] \nu(dz).\end{aligned} \tag{2.18}$$

The derivative pricing function $V(t, s)$ with $t \in [0, T]$ and $s \in \mathbb{R}^+$, can be obtain by solving a pricing PIDE according to the following theorem.

Theorem 2.1.4. *Let us consider an arbitrage free market, where the underlying stock price follows the exponential Lévy process (2.13). Let also assume that* $V(t, s) \in C^{1,2}([t_0, T] \times \mathbb{R}^+) \bigcap \mathcal{C}_2([t_0, T] \times \mathbb{R}^+)$, $\frac{\partial V}{\partial t}$ *and* $\frac{\partial V}{\partial s}$ *are in* $\mathcal{C}_1([t_0, T] \times \mathbb{R}^+)$ *and* $\frac{\partial^2 V}{\partial s^2}$ *is bounded. The payoff* $\Phi(\cdot)$ *is such that* $\mathbb{E}[\Phi(S_T)] < \infty$. *Therefore* $V(t, s)$ *satisfies the PIDE*

$$\frac{\partial V(t, s)}{\partial t} + \mathcal{L}^S V(t, s) - rV(t, s) = 0 \tag{2.19}$$

$$V(T, s) = \Phi(s), \tag{2.20}$$

where $\mathcal{L}^S$ *is the infinitesimal generator in (2.18).*

Proof. Let us consider the formula (2.7) together with the Markov property (1.28). For any stopping time τ such that $0 \le t \le \tau \le T$, we can use the law of iterated expectations:

$$D(0, t)V(t, s) = \mathbb{E}^{\mathbb{Q}}\Big[\mathbb{E}^{\mathbb{Q}}\big[D(0, T)V(T, S_T)\big|S_\tau\big]\Big|S_t\Big] = \mathbb{E}^{\mathbb{Q}}\big[D(0, \tau)V(\tau, S_\tau)\big|S_t\big].$$

We can write $D(0, \tau)V(\tau, S_\tau) = D(0, t)V(t, s) + \int_t^\tau d\big(D(t, u)V(u, S_u)\big)du$. Using the Itō product rule (1.41) we get:

$$\mathbb{E}^{\mathbb{Q}}\left[\int_t^\tau e^{-r(u-t)}\left(\frac{\partial V(u, S_u)}{\partial u} + \mathcal{L}^S V(u, S_u) - rV(u, S_u)\right)du\middle|S_t = s\right] = 0,$$

where all the martingales terms have zero expectation. In the last step we also introduced the explicit expression of the discount factor (2.4) with constant r.

By definition, the terms inside the integral are all continuous and are all dominated by some quadratic functions that do not depend on τ. We can divide both sides by $(\tau - t)$ and take the limit for $\tau \to t$. Using the mean value theorem, for some $u \in [t, \tau]$ we get

$$\lim_{u \to t} \mathbb{E}^{\mathbb{Q}} \left[e^{-r(u-t)} \left(\frac{\partial V(u, S_u)}{\partial u} + \mathcal{L}^S V(u, S_u) - rV(u, S_u) \right) \middle| S_t = s \right] = 0.$$

When $\tau \to t$ also $u \to t$. Thanks to the dominated convergence theorem we can take the limit inside the expectation and conclude the proof. □

In practice, the hypothesis of the theorem above are rarely satisfied. The payoff Φ is usually not in the domain of $\mathcal{L}$ and sometimes is not even differentiable, e.g. call/put options. For this reasons, the option price should be considered a solution of (2.19) in a weaker sense. The notion of *viscosity solution* allows to cover this case. We will introduce it in section (4.3). For a complete exposition on this topic, we refer to [Cont and Voltchkova, 2005b]. The authors prove that, in a general setting, option prices in exponential Lévy models correspond to viscosity solutions of the pricing PIDE.

Putting together the Eq. (2.19) and (2.18) we obtain the PIDE for the option price with the associated boundary conditions.

$$\begin{aligned} &\frac{\partial V(t,s)}{\partial t} - rV(t,s) + rs\frac{\partial V(t,s)}{\partial s} + \frac{1}{2}\sigma^2 s^2 \frac{\partial^2 V(t,s)}{\partial s^2} \\ &+ \int_{\mathbb{R}} \left[V(t, se^z) - V(t,s) - s(e^z - 1)\frac{\partial V(t,s)}{\partial s} \right] \nu(dz) = 0. \end{aligned} \tag{2.21}$$

CALL:

- Terminal:
$$V(T,s) = \max(s - K, 0),$$
- Lateral:
$$V(t,0) = 0 \quad \text{and} \quad V(t,s) \underset{s \to \infty}{\sim} s - Ke^{-r(T-t)}.$$

PUT:

- Terminal:
$$V(T,s) = \max(K - s, 0),$$
- Lateral:
$$V(t,0) = Ke^{-r(T-t)} \quad \text{and} \quad V(t,s) \underset{s \to \infty}{=} 0.$$

While the terminal conditions are simply the payoffs of the respective options, in order to obtain the lateral condition we need to consider the asymptotic behavior of the option for $s = 0$ and for $s \to \infty$. When the stock price is much higher than the strike, a call option is almost sure to be exercised. The current value of the call option will thus be approximately equal to the stock price minus the (discounted) strike price. The opposite applies to a put option, which is almost sure to not be exercised. Its value will be approximately zero.

On the other hand, if the stock price is much less than the strike, a call option is likely to not be exercised, and has approximately zero value. A put option instead, will be exercised with high probability and its value corresponds to the (discounted) exercise price.

2.1.2 PIDE in log-variable

In order to have a simpler PIDE expression, it turns out that it is better to work with a Lévy process instead of its exponential. So let us invert Eq. (1.51) and consider $X_t = \log\left(\frac{S_t}{S_0}\right)$ with dynamics described by the SDE

$$dX_t = \left(b + \int_{|x|\geq 1} x\nu(dx)\right)dt \ + \sigma dW_t + \int_{\mathbb{R}} z\tilde{N}(dt,dz), \tag{2.22}$$

of the form in Eq. (1.26). It is better to reformulate the equation with the parameter μ, since it will be easier to make the substitution (2.16). Using the Itō formula, and considering Eq. (1.55) representing the dynamics of S_t, we obtain

$$\begin{aligned}
dX_t = d\left(\log\frac{S_t}{S_0}\right) =& \frac{1}{S_t}S_t\mu dt + \frac{1}{S_t}S_t\sigma dW - \frac{1}{2}\frac{1}{S_t^2}S_t^2\sigma^2 dt \\
&+ \int_{\mathbb{R}}\big(\log(S_t + S_t(e^z-1)) - \log(S_t)\big)\tilde{N}(dt,dz) \\
&+ \int_{\mathbb{R}}\big(\log(S_t + S_t(e^z-1)) - \log(S_t) - \frac{1}{S_t}S_t(e^z-1)\big)\nu(dz)dt \\
=&(\mu - \frac{1}{2}\sigma^2)dt + \sigma dW + \int_{\mathbb{R}} z\tilde{N}(dt,dz) \\
&+ \int_{\mathbb{R}}\big(z - (e^z-1)\big)\nu(dz)dt \\
=&\left(\mu - \frac{1}{2}\sigma^2 - \int_{\mathbb{R}}\big(e^z - 1 - z\big)\nu(dz)\right)dt + \sigma dW + + \int_{\mathbb{R}} z\tilde{N}(dt,dz).
\end{aligned}$$

For $f \in C^2(\mathbb{R})\bigcap \mathcal{C}_2(\mathbb{R})$, the corresponding infinitesimal generator is:

$$\begin{aligned}
\mathcal{L}^X f(x) =& \left(\mu - \frac{1}{2}\sigma^2 - \int_{\mathbb{R}}\big(e^z - 1 - z\big)\nu(dz)\right)\frac{\partial f(x)}{\partial x} \\
&+ \frac{1}{2}\sigma^2\frac{\partial^2 f(x)}{\partial x^2} + \int_{\mathbb{R}}\big(f(x+z) - f(x) - z\frac{\partial f(x)}{\partial x}\big)\nu(dz).
\end{aligned} \tag{2.23}$$

Alternatively, this can be obtained by a change of variables in Eq (2.21), $s = e^x$ and $\tilde{V}(t,x) := V(t,s)$. The differential operators change according to

$$s\frac{\partial}{\partial s} = \frac{\partial}{\partial x}, \qquad s^2\frac{\partial^2}{\partial s^2} = \frac{\partial^2}{\partial x^2} - \frac{\partial}{\partial x}. \tag{2.24}$$

After the change of variables, the option PIDE, using (2.16), becomes

$$\begin{aligned}
&\frac{\partial\tilde{V}(t,x)}{\partial t} - r\tilde{V}(t,x) + \left(r - \frac{1}{2}\sigma^2 - \int_{\mathbb{R}}\big(e^z - 1 - z\big)\nu(dz)\right)\frac{\partial\tilde{V}(t,x)}{\partial x} \\
&+ \frac{1}{2}\sigma^2\frac{\partial^2\tilde{V}(t,x)}{\partial x^2} + \int_{\mathbb{R}}\big(\tilde{V}(t,x+z) - \tilde{V}(t,x) - z\frac{\partial\tilde{V}(t,x)}{\partial x}\big)\nu(dz) = 0.
\end{aligned} \tag{2.25}$$

With boundary conditions:
CALL:

- Terminal:
$$\tilde{V}(T,x) = \max(e^x - K, 0),$$
- Lateral:
$$\tilde{V}(t,x) \underset{x\to-\infty}{=} 0 \quad \text{and} \quad \tilde{V}(t,x) \underset{x\to\infty}{\sim} e^x - Ke^{-r(T-t)}.$$

PUT:

- Terminal:
$$\tilde{V}(T,x) = \max(K - e^x, 0),$$
- Lateral:
$$\tilde{V}(t,x) \underset{x\to-\infty}{\sim} = Ke^{-r(T-t)} \quad \text{and} \quad \tilde{V}(t,x) \underset{x\to\infty}{=} 0.$$

2.2 Finite difference methods

Finite difference methods are a technique for obtaining numerical solutions of PDEs and PIDEs. The idea underlying finite-difference methods is to replace the partial derivatives occurring in the PDE by the finite differences approximations defined below. Let $V(t,x) \in C^{1,2}$, for a small but not infinitesimal $\Delta x > 0$, we can define the **forward difference** as

$$\frac{\partial V(t,x)}{\partial x} \approx \frac{V(t,x+\Delta x) - V(t,x)}{\Delta x}. \tag{2.26}$$

The **backward difference** is defined as

$$\frac{\partial V(t,x)}{\partial x} \approx \frac{V(t,x) - V(t,x-\Delta x)}{\Delta x} \tag{2.27}$$

and the **central difference** is defined as

$$\frac{\partial V(t,x)}{\partial x} \approx \frac{V(t,x+\Delta x) - V(t,x-\Delta x)}{2\Delta x}. \tag{2.28}$$

Analogous approximations can be introduced for the time derivative $\frac{\partial V(t,x)}{\partial t}$ for a small $\Delta t > 0$. The use of the forward or backward difference approximation for $\frac{\partial V(t,x)}{\partial t}$ leads to the **explicit** or **implicit** finite difference schemes respectively. Usually the central difference is not used for the time derivative because it may lead to bad numerical schemes (specifically, unstable schemes). For the space variable instead, it is quite common to approximate the first order derivative by central differences.

For second order derivatives in space, we can define the **symmetric central difference** approximation as

$$\begin{aligned}\frac{\partial^2 V(t,x)}{\partial x^2} &\approx \frac{\frac{V(t,x+\Delta x)-V(t,x)}{\Delta x} - \frac{V(t,x)-V(t,x-\Delta x)}{\Delta x}}{\Delta x} \\ &= \frac{V(t,x+\Delta x) + V(t,x-\Delta x) - 2V(t,x)}{\Delta x^2}.\end{aligned}$$

When considering PIDEs there is an additional integral term to discretize, and can be replaced by Riemann sums. But first it is necessary to localize the problem and truncate the integral.

Since numerical computations can only be performed on a finite domain, the first step is to reduce the PIDE to a bounded domain. The initial problem

$$\frac{\partial V(t,x)}{\partial t} + \mathcal{L}^X V(t,x) - rV(t,x) = 0 \quad \text{for} \quad t,x \in [t_0,T] \times \mathbb{R}$$

is restricted to the finite domain $[t_0,T] \times [A_1,A_2]$, with $A_1, A_2 \in \mathbb{R}$. In order to have a well posed problem, we need to impose the boundary conditions for $V(t,x)$ everywhere outside $[t_0,T] \times [A_1,A_2]$, and not only at the lateral boundaries A_1 and A_2.

The next step is to replace the domain $[t_0,T] \times [A_1,A_2]$ by a discrete grid: For $n = 0,1,...N \in \mathbb{N}$, define the discrete time step $\Delta t = \frac{T-t_0}{N}$ such that $t_n = t_0 + n\Delta t$. For $i = 0,1,...M \in \mathbb{N}$, define the discrete space step $\Delta x = \frac{A_2-A_1}{M}$ such that $x_i = A_1 + i\Delta x$. The grid is divided into equally spaced **nodes** of distance Δx in the x-axis, and of distance Δt in the t-axis. The mesh points have the form $(t_0 + n\Delta t, A_1 + i\Delta x)$. At this point we concern ourselves only with the values of $V(t,x)$ on the mesh nodes. We write

$$V(t_0 + n\Delta t, A_1 + i\Delta x) = V_i^n.$$

The integral terms in (2.25) are restricted to the bounded domain $[-B_1, B_2]$, with $B_1, B_2 \in \mathbb{R}$.

$$\begin{aligned} &\frac{\partial V(t,x)}{\partial t} - rV(t,x) + \left(r - \frac{1}{2}\sigma^2 - \int_{-B_1}^{B_2} \big(e^z - 1 - z\big)\nu(dz)\right)\frac{\partial V(t,x)}{\partial x} \qquad (2.29)\\ &+ \frac{1}{2}\sigma^2 \frac{\partial^2 V(t,x)}{\partial x^2} + \int_{-B_1}^{B_2} \big(V(t,x+z) - V(t,x) - z\frac{\partial V(t,x)}{\partial x}\big)\nu(dz) = 0\\ &\text{for} \quad t,x \in [t_0,T] \times]-A_1,A_2[. \end{aligned}$$

The computational domain of interest becomes $[t_0,T] \times [A_1 - B_1, A_2 + B_2]$, where in the regions $[t_0,T] \times [A_1 - B_1, A_1]$ and $[t_0,T] \times [A_2, A_2 + B_2]$ we need to define the boundary conditions.

In order to solve the problem (2.29) we consider the **IMEX** (Implicit-Explicit) method proposed in [Cont and Voltchkova, 2005a]. The integro-differential operator

$$\frac{\partial V(t,x)}{\partial t} + \mathcal{L}V(t,x) - rV(t,x) = 0$$

is split in two parts:

$$\frac{\partial V(t,x)}{\partial t} + DV(t,x) + JV(t,x) - rV(t,x) = 0$$

where D and J stand for the differential and integral parts of $\mathcal{L}$, respectively. We can replace $DV(t,x)$ by a finite difference approximated operator $D_\Delta V(t,x)$ and $JV(t,x)$ by the operator $J_\Delta V(t,x)$ obtained through **Riemann sums**[2] approximation, and use the following IMEX time-stepping scheme:

$$\frac{V_i^{n+1} - V_i^n}{\Delta t} + D_\Delta V^n + J_\Delta V^{n+1} - rV^n = 0. \qquad (2.30)$$

[2]The discretization of the integral depends on the activity of the Lévy process under consideration. We will see in the next sections the two cases of finite and infinite activity.

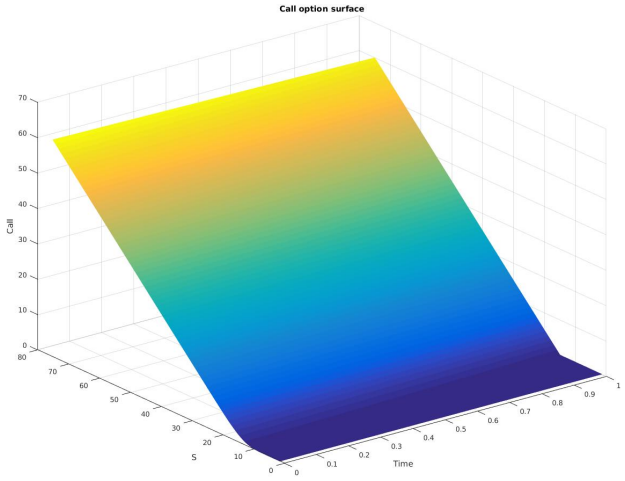

Figure 2.1: Call option surface obtained by solving the Black-Scholes PDE. It is computed with the parameters in table 2.1

We treat the integral part with an explicit time stepping in order to avoid the inversion of the dense matrix J_Δ. In the next sections let us discuss about the application of the IMEX scheme to the numerical solution of the PIDE (2.25) for the BS, Merton and VG processes.

2.2.1 Black and Scholes PDE

The [Black and Scholes, 1973] model assumes a geometric Brownian motion for the dynamics of the underlying, as we saw in (2.2). It corresponds to the exponential of a Lévy process $\{X_t\}_{t\geq 0}$ with triplet $(b, \sigma, 0)$, with Lévy measure $\nu = 0$. The BS PDE (2.25) in log-variables turns out to be

$$\frac{\partial V(t,x)}{\partial t} + \left(r - \frac{1}{2}\sigma^2\right)\frac{\partial V(t,x)}{\partial x} + \frac{1}{2}\sigma^2\frac{\partial^2 V(t,x)}{\partial x^2} - rV(t,x) = 0. \tag{2.31}$$

The Lévy measure is identically null and therefore there is no integral term. The domain is restricted to $[t_0, T] \times [A_1, A_2]$. We apply the IMEX scheme, that in this case is a fully implicit scheme. The discretized equation becomes

$$\begin{aligned} &\frac{V_i^{n+1} - V_i^n}{\Delta t} + (r - \frac{1}{2}\sigma^2)\frac{V_{i+1}^n - V_{i-1}^n}{2\Delta x} \\ &+ \frac{1}{2}\sigma^2\frac{V_{i+1}^n + V_{i-1}^n - 2V_i^n}{\Delta x^2} - rV_i^n = 0. \end{aligned} \tag{2.32}$$

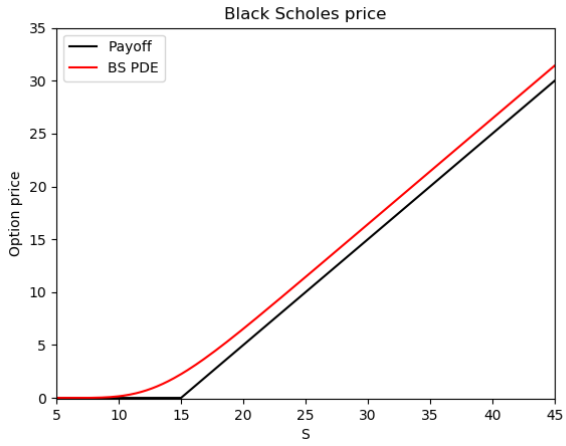

Figure 2.2: Price of a call option at time $t = 0$ with parameters in table 2.1.

Rearranging the terms:

$$\begin{aligned} V_i^{n+1} = V_i^n &\left(1 + r\Delta t + \sigma^2 \frac{\Delta t}{h_x^2}\right) \\ &+ V_{i+1}^n \left(-(r - \frac{1}{2}\sigma^2)\frac{\Delta t}{2\Delta x} + \frac{1}{2}\sigma^2 \frac{\Delta t}{\Delta x^2}\right) \\ &+ V_{i-1}^n \left((r - \frac{1}{2}\sigma^2)\frac{\Delta t}{2\Delta x} + \frac{1}{2}\sigma^2 \frac{\Delta t}{\Delta x^2}\right). \end{aligned}$$

We can rename the coefficients such that:

$$V_i^{n+1} = aV_{i-1}^n + bV_i^n + cV_{i+1}^n,$$

and write it in matrix form:

$$\begin{pmatrix} V_1^{n+1} \\ V_2^{n+1} \\ \vdots \\ V_{M-2}^{n+1} \\ V_{M-1}^{n+1} \end{pmatrix} = \underbrace{\begin{pmatrix} b & c & 0 & \cdots & 0 \\ a & b & c & 0 & 0 \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & 0 & a & b & c \\ 0 & 0 & 0 & a & b \end{pmatrix}}_{\mathcal{D}} \cdot \begin{pmatrix} V_1^n \\ V_2^n \\ \vdots \\ V_{M-2}^n \\ V_{M-1}^n \end{pmatrix} + \underbrace{\begin{pmatrix} aV_0^n \\ 0 \\ \vdots \\ 0 \\ cV_M^n \end{pmatrix}}_{\text{B (boundary terms)}}$$

The system

$$V_i^{n+1} = \mathcal{D}V_i^n + B \quad \text{for} \quad 1 \leq i \leq M-1$$

can be solved easily for V_i^n by inverting[3] the matrix $\mathcal{D}$.

Using the parameters in table (2.1) we can solve the linear system and plot the solution in Figures 2.1 and 2.2.

[3] Matrix inversion is a slow operation and there are plenty of efficient algorithms that permit to solve a linear system with no need of matrix inversion. In this thesis we solved all the linear systems occurring after the implicit discretization of a PDE (or PIDE) with the LU decomposition method.

BS Parameters			
K	T	r	σ
15	1	0.1	0.25

Table 2.1: Option parameters and diffusion process parameters.

2.2.2 Merton PIDE

We presented the Merton model in section (1.3.2). Let us recall that the jump component of the Merton process has finite activity, $\nu(\mathbb{R}) = \lambda < \infty$. Following Eq. (2.25), the Merton PIDE in log-variables has the following form:

$$\begin{aligned} &\frac{\partial V(t,x)}{\partial t} - rV(t,x) + \left(r - \frac{1}{2}\sigma^2 - m\right)\frac{\partial V(t,x)}{\partial x} \\ &+ \frac{1}{2}\sigma^2\frac{\partial^2 V(t,x)}{\partial x^2} + \int_{\mathbb{R}} V(t,x+z)\nu(dz) - \lambda V(t,x) = 0. \end{aligned} \tag{2.33}$$

with m defined in (1.77). Since we have to restrict the problem to a bounded region, we consider the equation (2.29):

$$\begin{aligned} &\frac{\partial V(t,x)}{\partial t} - rV(t,x) + \left(r - \frac{1}{2}\sigma^2 - \hat{m}\right)\frac{\partial V(t,x)}{\partial x} \\ &+ \frac{1}{2}\sigma^2\frac{\partial^2 V(t,x)}{\partial x^2} + \int_{-B_1}^{B_2} V(t,x+z)\nu(dz) - \hat{\lambda}V(t,x) = 0. \end{aligned}$$

with $\hat{m} = \int_{-B_1}^{B_2}\big(e^z - 1\big)\nu(dz)$ and $\hat{\lambda} = \int_{-B_1}^{B_2}\nu(dz)$.

For $0 < K_1 < K_2$ we choose B_1, B_2 such that $\big[-B_1, B_2\big] = \big[(-K_1 - 1/2)\Delta x, (K_2 + 1/2)\Delta x\big]$. Let us discretize the integral as follows:

$$\int_{-B_1}^{B_2} V(t_n, x_i + z)\nu(dz) \approx \sum_{k=-K_1}^{K_2} \nu_k V_{i+k}^n \tag{2.34}$$

where

$$\nu_k = \int_{(k-\frac{1}{2})\Delta x}^{(k+\frac{1}{2})\Delta x} \nu(z)dz, \quad \text{for} \quad -K_1 \le k \le K_2. \tag{2.35}$$

We have that $\hat{\lambda} = \sum_{k=-K_1}^{K_2}\nu_k$. For large values of B_1 and B_2, the parameter $\hat{\lambda}$ is a good approximation for λ, since

$$\lambda = \lim_{B_1,B_2\to\infty}\hat{\lambda} = \lim_{B_1,B_2\to\infty}\int_{-B_1}^{B_2}\nu(dz).$$

The discretized equation using the IMEX scheme becomes:

$$\begin{aligned} &\frac{V_i^{n+1} - V_i^n}{\Delta t} + (r - \frac{1}{2}\sigma^2 - \hat{m})\frac{V_{i+1}^n - V_{i-1}^n}{2\Delta x} \\ &+ \frac{1}{2}\sigma^2\frac{V_{i+1}^n + V_{i-1}^n - 2V_i^n}{\Delta x^2} - (r + \hat{\lambda})V_i^n + \sum_{k=-K_1}^{K_2}\nu_k V_{i+k}^{n+1} = 0. \end{aligned} \tag{2.36}$$

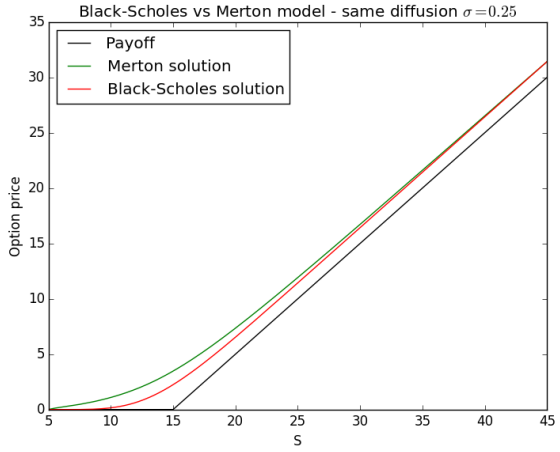

Figure 2.3: Comparison of prices of a call option at time $t = 0$ for BS and Merton models. The parameters are in Tables 2.1 and 2.2 (first line).

Rearranging the terms:

$$\underbrace{V_i^{n+1} + \Delta t \sum_{k=-K_1}^{K_2} \nu_k V_{i+k}^{n+1}}_{\tilde{V}_i^{n+1}} = V_i^n \left(1 + (r + \hat{\lambda})\Delta t + \sigma^2 \frac{\Delta t}{h_x^2}\right)$$
$$+ V_{i+1}^n \left(-(r - \frac{1}{2}\sigma^2 - \hat{m})\frac{\Delta t}{2\Delta x} + \frac{1}{2}\sigma^2 \frac{\Delta t}{\Delta x^2}\right)$$
$$+ V_{i-1}^n \left((r - \frac{1}{2}\sigma^2 - \hat{m})\frac{\Delta t}{2\Delta x} + \frac{1}{2}\sigma^2 \frac{\Delta t}{\Delta x^2}\right).$$

We can rename the coefficients:

$$\tilde{V}_i^{n+1} = aV_{i-1}^n + bV_i^n + cV_{i+1}^n,$$

and solve the system for V_i^n:

$$\begin{cases} \tilde{V}_i^{n+1} = V_i^{n+1} + \Delta t \sum_{k=-K_1}^{K_2} V_{i+k}^{n+1} \nu_k \\ V_i^n = \mathcal{D}^{-1}\left(\tilde{V}_i^{n+1} - B\right) \quad \text{for} \quad 1 \le i \le M-1 \end{cases}$$

where $\mathcal{D}$ is the tridiagonal matrix formed by the coefficients a, b, c, and with boundary terms $B = (aV_0^n, 0, ..., 0, cV_M^n)$.

In Figure 2.3 we computed the BS and Merton curves using the parameters in tables 2.1 and 2.2. The Merton curve is everywhere higher than the BS curve. This is a consequence of the additional jump component in the Merton process that increases the total variance. The two processes have same diffusion component σ, but the Merton process has total variance $\sigma^2 + \lambda\xi^2 + \lambda\alpha^2$, while the diffusion process has only variance σ^2.

In order to investigate the effect of the heavy tails on the shape of the option curve, let us consider the second set of parameters in table 2.2. Under these

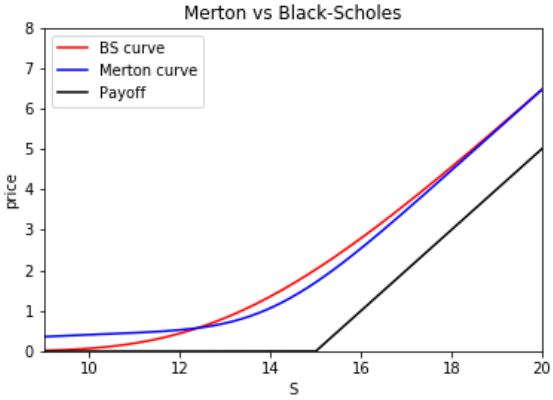

Figure 2.4: Comparison of prices of a call option at time $t = 0$ for Merton and BS models with same standard deviation. The Merton parameters are those in the second line of the table 2.2, while the diffusion parameter σ for the BS prices is chosen such that the variances of the two processes are the same.

Merton parameters							
Figure	K	T	r	σ	α	ξ	λ
Fig. 2.3	15	1	0.1	0.25	0	0.5	0.8
Fig. 2.4	15	1	0.1	0.1	0	1.8	0.01

Table 2.2: Option's parameters and Merton process parameters.

parameters, the Merton distribution has a very high kurtosis ($\kappa = 97.32$). In figure 2.4 we compare the Merton curve with a BS curve computed using the volatility parameter $\sigma_{BS} = \sqrt{\sigma^2 + \lambda\xi^2 + \lambda\alpha^2}$. In this way, the differences in the shape are only due to the kurtosis, since both processes have zero skewness and equal mean and variance.

As expected, the BS curve is smaller than the Merton curve in the *deep out of the money* region (in the picture 2.4 this region corresponds to $S < 12$). This is a consequence of the heavy tails distribution of the Merton process, that assigns more probability to large movements of the underlying, i.e. a deep out of the money option has more probability to return *in the money* (where $S > K$).

2.2.3 Variance Gamma PIDE

We introduced the Variance Gamma process in Section 1.3.3. The VG process has infinite activity i.e. $\nu(\mathbb{R}) = \infty$ and has the triplet presented in (1.73).

From the general PIDE pricing formula (2.25), we obtain the VG PIDE for a function $V \in C^{1,1}([0,T] \times \mathbb{R}) \bigcap \mathcal{C}_2([0,T] \times \mathbb{R})$:

$$\frac{\partial V(t,x)}{\partial t} + (r-w)\frac{\partial V(t,x)}{\partial x} + \int_{\mathbb{R}} \big[V(t,x+z) - V(t,x)\big]\nu(dz) = rV(t,x). \quad (2.37)$$

with w defined in (1.81).

Unfortunately, it is not possible to apply the IMEX discretization directly to this equation. The Lévy measure has a singularity in the origin, that should be removed before the discretization.

An idea to overcome this problem, that can be applied to any Lévy processes with infinite activity, is presented in [Cont and Voltchkova, 2005a]. The authors propose to approximate the process $\{X_t\}_{t\geq 0}$ by an appropriate finite activity process with a modified diffusion component. The "small jumps" martingale component is approximated by a Brownian motion with same variance. After fixing a truncation parameter $\epsilon > 0$, the integrals in the SDE are split in two domains: $\{|z| < \epsilon\}$ and $\{|z| \geq \epsilon\}$. The integrand on the domain $\{|z| < \epsilon\}$, is approximated by the Taylor expansion $e^z - 1 - z = \frac{z^2}{2} + \mathcal{O}(z^3)$ such that

$$\begin{aligned} dX_t =& \Big(\mu - \frac{1}{2}\sigma^2 - \int_{\mathbb{R}} \big(e^z - 1 - z\big)\nu(dz)\Big)dt + \sigma dW + + \int_{\mathbb{R}} z\tilde{N}(dt, dz) \\ =& \Big(\mu - \frac{1}{2}\sigma^2 - \int_{|z|<\epsilon} (e^z - 1 - z)\nu(dz) - \int_{|z|\geq\epsilon} (e^z - 1 - z)\nu(dz)\Big)dt \\ &+ \sigma dW_t + \underbrace{\int_{|z|<\epsilon} z\tilde{N}(dt, dz)}_{\sigma_\epsilon dW_t} + \int_{|z|\geq\epsilon} z\tilde{N}(dt, dz) \\ =& \Big(\mu - \frac{1}{2}(\sigma^2 + \sigma_\epsilon^2) - w_\epsilon + \lambda_\epsilon\theta_\epsilon\Big)dt + \big(\sigma + \sigma_\epsilon\big)dW_t + \int_{|z|\geq\epsilon} z\tilde{N}(dt, dz), \end{aligned} \tag{2.38}$$

where we defined the new parameters

$$\begin{aligned} \sigma_\epsilon^2 &:= \int_{|z|<\epsilon} z^2\nu(dz), \qquad w_\epsilon := \int_{|z|\geq\epsilon} (e^z - 1)\nu(dz), \\ \lambda_\epsilon &:= \int_{|z|\geq\epsilon} \nu(dz), \qquad \theta_\epsilon := \frac{1}{\lambda_\epsilon}\int_{|z|\geq\epsilon} z\nu(dz). \end{aligned} \tag{2.39}$$

The process $\int_{|z|\geq\epsilon} z\tilde{N}(dt, dz)$ is a compensated Poisson process with finite activity λ_ϵ and variance $\sigma_J^2 = \int_{|z|\geq\epsilon} z^2\nu(dz)$.

For the VG process, where $\sigma = 0$ the approximated dynamics is thus

$$dX_t = \Big(\mu - \frac{1}{2}\sigma_\epsilon^2 - w_\epsilon + \lambda_\epsilon\theta_\epsilon\Big)dt + \sigma_\epsilon dW_t + \int_{|z|\geq\epsilon} z\tilde{N}(dt, dz), \tag{2.40}$$

where the parameters are obtained from the Lévy measure (1.66).

For any $V \in C^2(\mathbb{R})\bigcap C_2(\mathbb{R})$, the infinitesimal generator associated with (2.40) has a "jump-diffusion" form

$$\begin{aligned} \mathcal{L}^{VG}V(x) \;=& \big(\mu - \frac{1}{2}\sigma_\epsilon^2 - w_\epsilon\big)\frac{\partial V}{\partial x} + \frac{1}{2}\sigma_\epsilon^2\frac{\partial^2 V}{\partial x^2} \\ &+ \int_{|z|\geq\epsilon} V(x + z)\nu(dz) - \lambda_\epsilon V(x). \end{aligned} \tag{2.41}$$

The same result can be obtained directly from the infinitesimal generator (2.23)

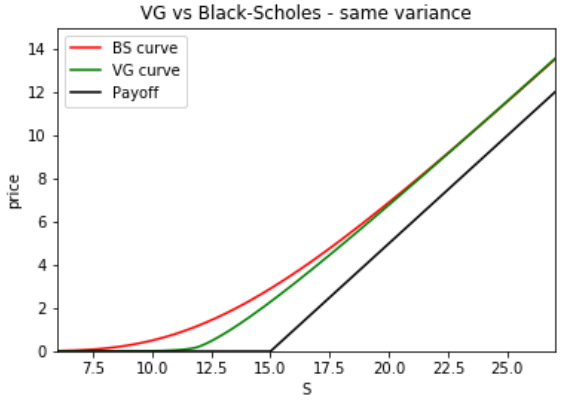

Figure 2.5: **Negative skewness:** Comparison of call option curves at $t = 0$ for VG and BS models. The VG parameters are in table (2.3). The BS curve is computed with a $\sigma_{BS} = \sqrt{\sigma^2 + \theta^2\kappa}$, such that the two processes have same variance.

with $\sigma = 0$:

$$\mathcal{L}^{VG}V(x) = \mu\frac{\partial V}{\partial x} + \int_{|z|<\epsilon}\left[V(x+z) - V(x) - (e^z - 1)\frac{\partial V}{\partial x}\right]\nu(dz) + \int_{|z|\geq\epsilon}\left[V(x+z) - V(x) - (e^z - 1)\frac{\partial V}{\partial x}\right]\nu(dz),$$

In the integral term on the domain $\{|z| < \epsilon\}$, we have to assume a smooth enough V such that we can use the following Taylor approximation

- $V(x+z) = V(x) + \frac{\partial V}{\partial x}z + \frac{1}{2}\frac{\partial^2 V}{\partial x^2}z^2 + \mathcal{O}(z^3)$.
- $e^z - 1 = z + \frac{z^2}{2} + \mathcal{O}(z^3)$.

Considering only the terms up to the second order, the integral for $\{|z| < \epsilon\}$ is

$$\int_{|z|<\epsilon}\frac{z^2}{2}\left[\frac{\partial^2 V}{\partial x^2} - \frac{\partial V}{\partial x}\right]\nu(dz) = \frac{\sigma_\epsilon^2}{2}\left[\frac{\partial^2 V}{\partial x^2} - \frac{\partial V}{\partial x}\right],$$

and we get again the infinitesimal generator (2.41). Using this generator and equations (2.19) and (2.16), the final PIDE is thus

$$\frac{\partial V(t,x)}{\partial t} + \left(r - \frac{1}{2}\sigma_\epsilon^2 - w_\epsilon\right)\frac{\partial V(t,x)}{\partial x} + \frac{1}{2}\sigma_\epsilon^2\frac{\partial^2 V(t,x)}{\partial x^2} + \int_{|z|\geq\epsilon} V(t,x+z)\nu(dz) = (\lambda_\epsilon + r)V(t,x). \tag{2.42}$$

This equation is almost identical to equation (2.33), except for the truncation in the integral. At this point we can restrict the computational domain on $[A_1, A_2]$ and the integral region on $[-B_1, B_2]$ and using the same discretization used for the Merton PIDE, we can solve the problem with the IMEX scheme.

VG parameters						
Figure	K	T	r	θ	$\bar{\sigma}$	κ
Fig. 2.5	15	1	0.1	-0.2	0.2	2.5
Fig. 2.6	15	1	0.1	0.2	0.2	2.5

Table 2.3: Option's parameters and VG process parameters.

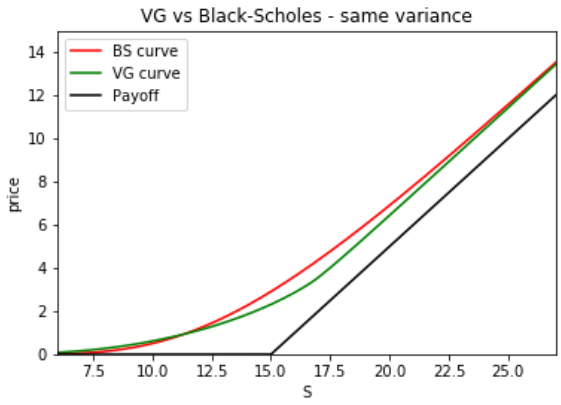

Figure 2.6: **Positive skewness:** Comparison of call option curves at $t = 0$ for VG and BS models. The VG parameters are in table (2.3). The BS curve is computed with a $\sigma_{BS} = \sqrt{\bar{\sigma}^2 + \theta^2\kappa}$, such that the two processes have same variance.

Using the parameters in Table 2.3 we compute VG call option prices and compare them with BS prices with volatility $\sigma_{BS} = \sqrt{\bar{\sigma}^2 + \theta^2\kappa}$, such that the two processes have same variance. In this way, only the features coming from skewness and kurtosis should pop up. Under this choice of parameters the VG process has standard deviation equal to 0.37, skewness (in absolute value) 3.05 and kurtosis 14.38. The sign of the skewness is determined by the sign of the parameter θ.

In figures 2.5 and 2.6 we compare the BS and VG curves. This example wants to show how different the shape of the VG curves can be by modifying the skewness parameter.

The reader can consult [Schoutens, 2003] for further information on Merton and VG processes and on general numerical tests and calibration of exponential Lévy processes.

2.2.4 Numerical convergence analysis

In table 2.4 we compare the ATM (*at the money*) prices of a European call option computed with the algorithms presented in the previous sections, with the prices obtained by closed formulas. For the BS model, we used the well known closed formula presented in the paper [Black and Scholes, 1973]. In order to compute the Merton price we use the semi-closed formula derived in [Merton, 1976], and for the VG price we used the semi-closed formula of [Madan et al., 1998]. The

	Closed formula	**PDE/PIDE**
BS	2.246368	2.246352
Merton	3.477645	3.477468
VG	1.987006	1.987089

Table 2.4: Option prices with $S_0 = K = 15$ and $T = 1$. Black Scholes parameters are in tables 2.1. Merton parameters are in the first line of table 2.2. The Variance Gamma parameters are $\theta = -0.1$, $\bar{\sigma} = 0.2$, $\kappa = 0.1$.

PIDE prices are obtained by solving the equations (2.31), (2.33) and (2.42). In this analysis we consider the same sets of parameters that we will use in chapter 6 (see table 6.2).

It is a common practice to select the upper-bound and lower-bound of the price as a multiple of the strike price. In our case we set $S_{max} = 6K$ and $S_{min} = K/6$. According to this choice, we set $A_1 = \log(K/6)$ and $A_2 = \log(6K)$ for all the three considered problems.

For the Merton PIDE we set $B_1 = B_2 = 5\sigma_M$, with $\sigma_M = \sqrt{\lambda\xi^2 + \lambda\alpha^2}$. For the VG PIDE we set $B_1 = B_2 = 5\sigma_{VG}$, with $\sigma_{VG} = \sqrt{\bar{\sigma}^2 + \theta^2\kappa}$. In both cases, the parameters B_1 and B_2 are multiple of the standard deviation of the respective jump processes. The parameters A_1, A_2, B_1 and B_2 are very important and should be set as large as possible in order to have an accurate result.

space steps	**BS**	**Merton**	**VG**
50	2.268400	3.495694	2.032549
100	2.251734	3.481931	2.014171
200	2.247682	3.478573	2.004308
400	2.246682	3.477744	1.998356
800	2.246433	3.477537	1.994912
1600	2.246371	3.477486	1.992661
3200	2.246356	3.477473	1.990863
6400	2.246352	3.477469	1.989213

Table 2.5: Convergence table. Fixed 12000 time steps.

time steps	**BS**	**Merton**	**VG**
50	2.242021	3.440676	0.998308
100	2.244216	3.459215	1.295962
200	2.245297	3.468434	1.554740
400	2.245834	3.473032	1.741888
800	2.246102	3.475328	1.858731
1600	2.246235	3.476475	1.924853
3200	2.246302	3.477048	1.960160
6400	2.246335	3.477335	1.978422

Table 2.6: Convergence table. Fixed 16000 space steps.

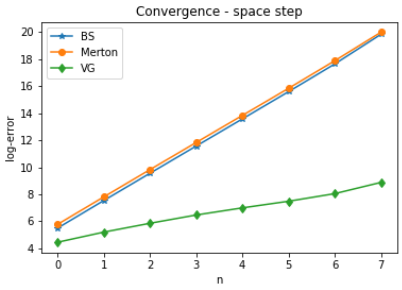

Figure 2.7: Log-error of the prices in table 2.5.

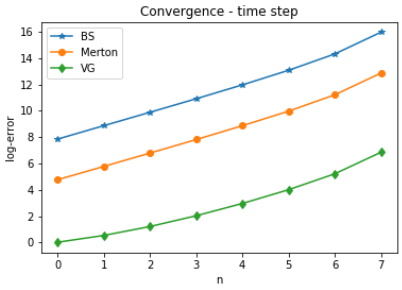

Figure 2.8: Log-error of the prices in table 2.6.

The PDE/PIDE prices in table 2.4 are obtained by dividing the space interval $[A_1, A_2]$ in $M = 16000$ steps and the time interval $[t_0, T]$ in $N = 12000$ steps. We can see that BS and VG prices are accurate up to the fourth decimal digit and Merton is accurate up to the third decimal digit. In the following analysis we assume they are the correct limiting value of the numerical algorithm, and indicate them by $V^* = \lim_{M,N\to\infty} V^{M,N}$.

In tables 2.5 we computed option prices V^M, by varying the number of space steps M and keeping the time steps fixed $N = 12000$. In table 2.6 we did the opposite i.e. we computed V^N for different values of N and constant $M = 16000$. From these values it is possible to investigate how the error changes with respect to the discretization step.

Let us consider for instance the space discretization. If we define the error $\varepsilon_M = |V^M - V^*|$, we can assume that it has a polynomial relation with the discretization step $\Delta x = \frac{A_2 - A_1}{M}$ i.e. $\varepsilon_M \propto (\frac{1}{M})^p$, where p is the **order of convergence**. In the table 2.5 we chose M in a smart way such that $M = 50 \cdot 2^n$, for $0 \leq n \leq 7$. We can take the $\log_2$ on both sides such that

$$-\log_2 \varepsilon_n = C + np, \tag{2.43}$$

where $C \in \mathbb{R}$ is a constant of proportionality. The same analysis can be done for the time discretization.

The theoretical order of convergence of the implicit scheme applied to the BS PDE is $p = 2$ for the space discretization and $p = 1$ for the time discretization. For a discussion on the convergence rate for IMEX schemes applied to PIDEs we refer to [Cont and Voltchkova, 2005a] (section 6.4). Since the theoretical analysis of convergence for IMEX schemes is a wide topic, in this thesis we just perform a numerical convergence analysis.

In the two pictures 2.7 and 2.8 we plotted the log-error $-\log_2 \varepsilon_n$ for the values in the tables 2.5 and 2.6 respectively. It turns out that both the BS and Merton prices have a quadratic error ($p = 2$) in space, while VG prices have a convergence order much smaller ($p \sim 0.8$). It is worth to mention that in our algorithm we choose the truncation parameter $\epsilon = 1.5\Delta x$. This choice not only influences the convergence of the numerical scheme, but also the convergence of the Brownian approximation, since the parameters in (2.39) depend on ϵ.
According to the results of figure 2.8, the time error is linear ($p = 1$) for all the three models.

2.3 Chapter conclusions

This chapter exposes in brief the main concepts of the "No-arbitrage", or "martingale", pricing theory.

This is the most common approach used to price financial derivatives. The well known Black-Scholes model is included in this framework as a special case when the stock dynamics follows a geometric Brownian motion. The martingale pricing theory applies to all exponential Lévy processes that satisfy the assumption **EM**. For the purpose of this thesis, the contents of this chapter are quite important because we will use them to make comparisons with the more complex models introduced in the next chapters.

In section 2.1 we present the fundamental concepts necessary to define the pricing function of a derivative contract. Then we prove that this pricing function is the solution of a partial integro-differential equation when the stock dynamics follows an exponential Lévy process.

In the second part of the chapter, section 2.2, we describe the algorithm used for the numerical solution of the Black-Scholes PDE and the Merton and VG PIDEs. We explain how to discretize the PIDE using the IMEX finite difference method proposed in [Cont and Voltchkova, 2005a]. We also provide a numerical convergence analysis of this method. The numerical results of this chapter will be used for comparison in chapters 3 and 6. In specific, in chapter 3 we compare the prices obtained by the multinomial method with those obtained by solving the VG PIDE. In chapter 6, instead, we will see that when the transaction costs go to zero, the prices obtained from the transaction costs model are equal to the prices obtained by the martingale pricing theory.

3

Multinomial methods for Variance Gamma

In this thesis we decided to work with the Variance Gamma (VG) process because it has nice analytical properties and it reproduces quite well the statistical features of the stock dynamics (see for instance [Cont and Tankov, 2003] and [Ait-Sahalia and Jacod, 2012]).

To support this statement, we present in Figure 3.1 some examples of histograms of daily log-returns of the four indices: the **S&P 500** Stock Index, the **KOSPI** (Korea Composite Stock Price Index), **XAO** (All Ordinaries Australian Index) and **TAIEX** (Taiwan Capitalization weighted Stock Index). In the pictures we show the fit of the Normal and Variance Gamma (VG) densities, using the market data. It is clear that the VG density reproduces much better the high peaks near the origin and the heavy tails of the empirical distribution.

The VG process was first presented in the context of option pricing in [Madan and Milne, 1991], where it has been used for pricing European options. European vanilla options can be easily priced by the analytical formula presented in [Madan et al., 1998] and exotics can be priced numerically by several common techniques. Monte Carlo methods for VG are presented in [Fu, 2000]. A finite difference scheme for the VG Partial Integro-Differential Equation (PIDE) is described in [Cont and Voltchkova, 2005a]. In [Carr and Madan, 1998], the authors show how to price options by a Fourier transform approach. The problem

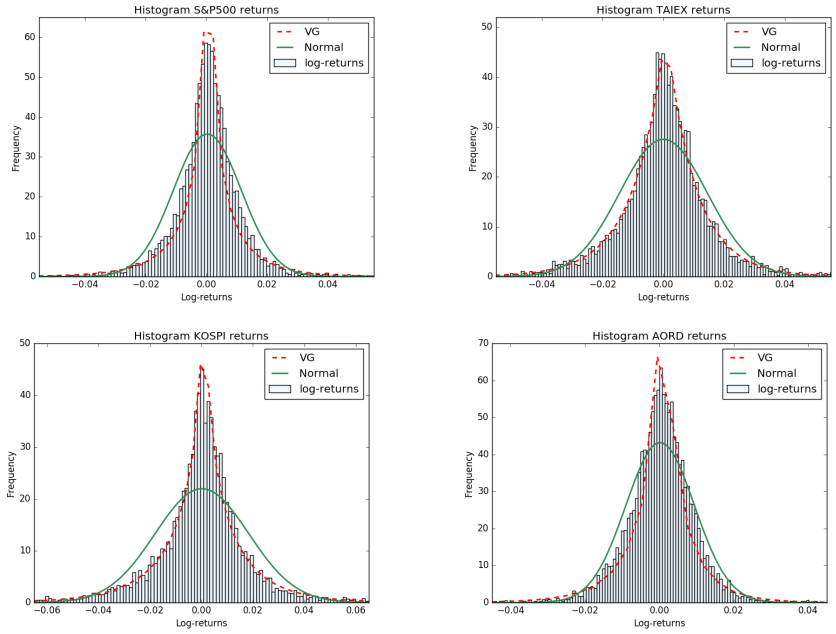

Figure 3.1: Histograms of daily log-returns for S&P500, KOSPI, XAO and TAIEX, from 1 January 1988 to 9 December 2016. Source: Yahoo Finance. The dashed line corresponds to the VG density (1.65). The continuous line is the normal density. The parameters are obtained by the *method of moments*. For more details on the parameter estimation for the VG density we refer to [Seneta, 2004].

for American options is considered in [Almendral, 2005], [Almendral and Oosterlee, 2007] and [Hirsa and Madan, 2001], where the authors present different finite difference schemes to solve the American VG PIDE.

The tree method was first introduced by [Cox et al., 1979] for a market where the log-price can change only in two different ways: an upward jump, or a downward jump. For this reason the model is called *binomial model*. The authors prove that when the number of time steps goes to infinity, the discrete random walk of the log-price converges to the Brownian motion and the option price converges to the Black-Scholes price. The *multinomial model* is a generalization of the binomial model, and at each time step it considers more than just two possible future states. A general multinomial method for pricing European and American options under exponential Lévy processes is described in [Maller et al., 2006]. In [Kellezi and Webber, 2006] the authors consider a multinomial method for general exponential Lévy processes based on the moment matching condition. Other methods based on the moment matching condition are for instance [Hainaut and MacGilchrist, 2010], with applications to the Normal Inverse Gaussian process, and [Ssebugenyi and Konlack, 2013] with applications to the VG process. In the present work we consider a multinomial discretization based on the cumulant matching condition as explained in [Yamada and Primbs, 2001], [Yamada and Primbs, 2003] and [Yamada and Primbs, 2004].

In Section 3.1 we review the construction of the multinomial tree, following the method of moment matching proposed in [Yamada and Primbs, 2001]. We prove that the multinomial tree converges to the continuous time jump process that we have introduced to approximate the VG process. In Section 3.2, we describe the algorithm for pricing options with the multinomial method and show the numerical results for European and American options.

3.1 The multinomial method

In this section we introduce the multinomial method proposed in [Yamada and Primbs, 2004]. The stock price is represented by a Markov chain with L possible future states at each time. In this setting, the time $t \in [t_0, T]$ is discretized as $t_n = t_0 + n\Delta t$ for $n = 0, ..., N$ and $\Delta t = (T - t_0)/N$. We denote the stock price at time t_n as $S_{t_n} = S_n$.

Let us consider the up/down factors $u > d > 0$, and write the discrete evolution of the stock price S_n as:

$$S_{n+1} = u^{L-l} d^{l-1} S_n \qquad l = 1, ..., L \tag{3.1}$$

where each future state has transition probability p_l, satisfying $\sum_{l=1}^{L} p_l = 1$. The value of the stock at time t_n can assume $j \in [1, ..., n(L-1)+1]$ possible values:

$$S_n^{(j)} = u^{n(L-l)+1-j} d^{j-1} S_0. \tag{3.2}$$

The multinomial tree is recombining if $u/d = c$, for $c > 1$. In the present work we only consider five branches, $L = 5$. As we will see in the next sections this number of branches is enough to model the features of a stochastic process up to its fourth moment.

3.1.1 Moment matching

To determine the parameters of the Markov chain we require that its local moments are equal to that of the continuous process. Let us consider a VG process with drift $(r - w)$, where w is the martingale correction defined in (1.81):

$$\begin{aligned} Y_{t+\Delta t} - Y_t &= (r - w)\Delta t + \int_{\mathbb{R}} x N(\Delta t, dx) \\ &= (r - w + \theta)\Delta t + \int_{\mathbb{R}} x \tilde{N}(\Delta t, dx) \end{aligned} \tag{3.3}$$

The parameter $\theta = \int_{\mathbb{R}} x\nu(dx) = \mathbb{E}\big[\int_{\mathbb{R}} xN(1, dx)\big]$ is the expected value of the VG process in (1.63), when $\Delta t = 1$. The integral with respect to the compensated Poisson measure $\tilde{N}(\Delta t, dx)$ is a martingale as discussed in Section 1.1.3.

We can pass to log-prices $Y_n = \log(S_n)$ in the discrete Eq. (3.1), and write it as the sum of a drift component and a random variable with L possible outcomes:

$$\begin{aligned} \Delta Y = Y_{n+1} - Y_n &= (L - l)\log(u) + (l - 1)\log(d) \\ &= \bar{b}\,\Delta t + (L - 2l + 1)\alpha(\Delta t). \end{aligned} \tag{3.4}$$

The term $\bar{b}\,\Delta t$ is the drift term, while the term l is a random variable that takes values in $\{1, 2, ..., L\}$ with probability p_l. It has to satisfy the martingale condition:

$$\mathbb{E}\big[(L-2l+1)\alpha(\Delta t)\big] = \alpha(\Delta t)\sum_{l=1}^{L} p_l(L-2l+1) = 0,$$

and $\alpha(\Delta t)$ is a function of Δt.

The corresponding up/down factors have the following representation:

$$u = \exp\bigg(\frac{b}{L-1} + \alpha(\Delta t)\bigg) \qquad d = \exp\bigg(\frac{b}{L-1} - \alpha(\Delta t)\bigg), \tag{3.5}$$

and we can readily see that if u/d is a constant, then the tree recombines.

Given the mean $c_1 = \mathbb{E}[\Delta Y] = \bar{b}\Delta t$, the k-central moment is:

$$\mathbb{E}\big[(\Delta Y - c_1)^k\big] = \alpha(\Delta t)^k\,\mathbb{E}\big[(L-2l+1)^k\big]. \tag{3.6}$$

The moment matching condition requires that the central moments of the discrete process (3.4) are equal to the central moments of the continuous process (3.3):

$$\alpha(\Delta t)^k\,\mathbb{E}\big[(L-2l+1)^k\big] = \mu_k. \tag{3.7}$$

We fix $L = 5$, and using the relation between central moments and cumulants (Eq. (1.85)) we solve the linear system of equations for the transition probabilities:

$$\begin{aligned}
p_1 &= \frac{1}{196\alpha(t)^4}\left[\frac{3}{2}c_2^2 - 2c_2\alpha(t)^2 + 2c_3\alpha(t) + \frac{1}{2}c_4\right] \\
p_2 &= \frac{1}{196\alpha(t)^4}\left[-6c_2 + 32c_2\alpha(t)^2 - 4c_3\alpha(t) - 2c_4\right] \\
p_3 &= 1 + \frac{1}{196\alpha(t)^4}\left[3c_4 + 9c_2^2 - 60c_2\alpha(t)^2\right] \\
p_4 &= \frac{1}{196\alpha(t)^4}\left[-6c_2 + 32c_2\alpha(t)^2 + 4c_3\alpha(t) - 2c_4\right] \\
p_5 &= \frac{1}{196\alpha(t)^4}\left[\frac{3}{2}c_2^2 - 2c_2\alpha(t)^2 - 2c_3\alpha(t) + \frac{1}{2}c_4\right].
\end{aligned} \tag{3.8}$$

The drift parameter corresponds to $\bar{b} = r - w + \theta$. The only missing term to find is $\alpha(\Delta t)$. This is a function of the time increment Δt and can be determined using the higher order terms in the moment matching condition together with the condition of positive probabilities.

Recall that the well known binomial model [Cox et al., 1979] assumes the value $\alpha(\Delta t) = \sigma\sqrt{\Delta t}$, that represents the volatility of the increments in the time interval Δt. In the trinomial model, the parameter $\alpha(\Delta t)$ assumes value $\alpha(\Delta t) = \frac{1}{2}\sigma\sqrt{3\Delta t}$, see for instance [Yamada and Primbs, 2001]. For the multinomial method a good representation for $\alpha(\Delta t)$ is:

$$\alpha(\Delta t) = \sqrt{c_2}\sqrt{\frac{3+\bar{\kappa}}{12}}, \tag{3.9}$$

where $\bar{\kappa} = c_4/c_2^2$ is the excess of kurtosis[1]. We refer to [Yamada and Primbs, 2004] for the derivation. This choice guarantees that the probabilities p_i for $i = 1...5$ are always positive and sum to one. We can replace the expression (3.9) inside (3.8), to obtain the simpler form:

$$[p_1, p_2, p_3, p_4, p_5] \approx \left[\frac{3+\bar{\kappa}+s\sqrt{9+3\bar{\kappa}}}{4(3+\bar{\kappa})^2}, \frac{3+\bar{\kappa}-s\sqrt{9+3\bar{\kappa}}}{2(3+\bar{\kappa})^2}, \right. \tag{3.10}$$
$$\left. \frac{3+2\bar{\kappa}}{2(3+\bar{\kappa})}, \frac{3+\bar{\kappa}+s\sqrt{9+3\bar{\kappa}}}{2(3+\bar{\kappa})^2}, \frac{3+\bar{\kappa}-s\sqrt{9+3\bar{\kappa}}}{4(3+\bar{\kappa})^2}\right],$$

where $s = c_3/\sqrt{c_2^3}$ is the skewness.

Remark 1. *The standard deviation of every Lévy process with finite second moment follows the square root propagation rule. This means that the term $\alpha(\Delta t)$ has to be proportional to the square root of Δt. In the binomial and trinomial models, the proportionality constant is explicit, while for the pentanomial method it is implicit in the formula (3.9). Expanding the formula using the expression (1.68) for the cumulants, it is possible to check that the square root rule is satisfied at first order in $\sqrt{\Delta t}$.*

3.1.2 Convergence

We call a generic jump process (1.69) with first four cumulants c_1,c_2,c_3,c_4 as in (1.68), the *approximated process* X^A. The cumulant generating function of the increment ΔX^A has the following series representation (see Section 1.4):

$$H_{\Delta X^A}(u) = ic_1u - \frac{c_2u^2}{2} - \frac{ic_3u^3}{3!} + \frac{c_4u^4}{4!} + \mathcal{O}(u^5). \tag{3.11}$$

We are interested in the approximation of a VG process with drift (3.3), therefore we require that $c_1 = \bar{b}\Delta t = (r - w + \theta)\Delta t$.

Theorem 3.1.1. *The increments of the discrete Markov chain (3.4) and the increments of the approximated process X^A have the same distribution by construction.*

Proof. The idea of the proof is to show that the cumulant generating function of the discrete process (3.4) coincides with that of the approximated process (3.11). We prove it by using the moment matching condition (3.7).

$$\begin{aligned} H_{\Delta Y}(u) &= \log\big(\phi_{\Delta Y}(u)\big) = \log\Big(\mathbb{E}\big[e^{iu\Delta Y}\big]\Big) \\ &= \log\left(\mathbb{E}\left[e^{iu\left(\bar{b}\Delta t+(L-2l+1)\alpha(\Delta t)\right)}\right]\right) \\ &= iu\bar{b}\Delta t + \log\left(\mathbb{E}\left[e^{iu\left((L-2l+1)\alpha(\Delta t)\right)}\right]\right). \end{aligned} \tag{3.12}$$

[1] We use the bar over κ, to distinguish the kurtosis from the variance of the gamma process κ.

We can expand the exponential function in Taylor series and use the moment matching condition (3.7) to obtain:

$$
\begin{aligned}
H_{\Delta Y}(u) &= iu\bar{b}\Delta t + \log\left(\sum_{k=0}^{\infty}\frac{(iu)^k}{k!}\big(\alpha(\Delta t)\big)^k\mathbb{E}\Big[\big(L-2l+1\big)^k\Big]\right) \qquad (3.13)\\
&= iu\bar{b}\Delta t + \log\left(\sum_{k=0}^{\infty}\frac{(iu)^k}{k!}\mu_k\right)\\
&= iuc_1 + \sum_{k=0}^{\infty}\frac{(iu)^k}{k!}c_k\\
&= H_{\Delta X^A}(u),
\end{aligned}
$$

□

Remark 2. *All the cumulants of ΔX^A are equal to the cumulants of the Markov chain (3.4) by construction, but only the first four are equal to the VG cumulants. When all the cumulants c_i, for $0 \leq i \leq n$, are equal to the VG cumulants, the approximated process X^A converges to the original VG process for $n \to \infty$. In order to describe n cumulants, we need $n+1$ branches. Therefore, when the number of cumulants of ΔX^A that are equal to those of the VG goes to infinity, the number of branches has to go to infinity as well. We assume that five branches ($L = 5$) are enough to describe the features of the underlying process and, at the same time, keep the numerical problem simple.*

Theorem 3.1.2. *The distribution of the pentanomial tree at time N converges to the distribution of a compound Poisson process at time N with $L = 5$ possible jump sizes and activity $\lambda = \frac{3}{2\bar{\kappa}N}$, when $\Delta t \to 0$.*

For the proof of this theorem we refer to Section 4.2 of [Yamada and Primbs, 2004]. The authors first define the jump sizes and their respective probabilities, and then prove that when $\Delta t \to 0$ the characteristic function of the pentanomial tree converges to the characteristic function of the compound Poisson process.

3.2 Numerical results

In this section we present the steps to implement the algorithm for pricing European and American options with the multinomial method. Then we compare the results with those obtained by the PIDE method and Black-Scholes model.

3.2.1 Algorithm

We suggest the following algorithm for pricing with the multinomial method:

1. Compute the transition probabilities vector (3.10).
2. Compute the up/down factors u and d (3.5) and the vector of prices S_N at terminal time N as in Eq. (3.2).
3. Evaluate the payoff of the option $V^N(S_N)$ at terminal time N.

4. Given the option values at time t_{n+1} compute the values at time t_n. The value is the conditional expectation:

$$V^n(s_n^{(k)}) = e^{-r\Delta t}\mathbb{E}^{\mathbb{Q}}\left[V^{n+1}(S_{n+1})\middle|S_n^{(k)} = s_n^{(k)}\right]. \tag{3.14}$$

5. If computing the price of an American option, the value at the previous time level is the maximum between the conditional expectation and the intrinsic value of the option. For an American put we have:

$$V^n(s_n^{(k)}) = \max\left\{e^{-r\Delta t}\mathbb{E}^{\mathbb{Q}}\left[V^{n+1}(S_{n+1})\middle|S_n^{(k)} = s_n^{(k)}\right], K - s_n^{(k)}\right\}. \tag{3.15}$$

6. Iterate the algorithm until the initial time t_0.

In all the numerical computations of this chapter we consider the risk neutral VG parameters in Table 3.1. These parameters correspond to the parameters used in [Cantarutti and Guerra, 2018].

Parameters			
r	θ	σ	κ
0.06	-0.1	0.2	0.2

Table 3.1: r is the risk free interest rate and θ, σ, κ are the risk neutral VG parameters.

3.2.2 European options

We compare the numerical results for European call and put options obtained with the multinomial and the PIDE approaches. We solve the VG PIDE following the IMEX method introduced in section 2.2.3. The details of the implementation are presented in section 2.2.4.

Following the algorithm proposed in the previous section, we solve the option pricing problem using the multinomial method. The number of time steps for all the computations presented in the following pictures is $N = 2000$. The table 3.2 shows a convergence analysis for the prices of European calls, puts and American puts. We can see that the convergence is quite fast. Figures (3.2) and (3.3) show the single prices obtained by the multinomial method compared with the curve obtained by solving the PIDE. In table 3.3 we compare directly the call/put numerical values obtained with the two methods.

Remark 3. *Pricing vanilla call and put European options is quite simple and the best approach is to use the closed formula derived in [Madan et al., 1998]. The big advantage of the multinomial method is in the computation of American options prices, where there is no closed formula and all the other approaches, such as PIDEs and Least Squares Monte Carlo, are difficult to implement and much slower.*

N	Eu. Call	Eu. Put	Time	Am. Put	Time
50	4.41873125	2.08928091	0.001	2.36765911	0.007
100	4.41960265	2.09015381	0.002	2.37255454	0.02
200	4.41997010	2.09052201	0.004	2.37480218	0.07
400	4.42013640	2.09068869	0.01	2.37587117	0.29
800	4.42021515	2.09076762	0.03	2.37639131	1.09
1000	4.42023054	2.09078306	0.04	2.37649417	1.67
1500	4.42025089	2.09080345	0.06	2.37663070	3.79
2000	4.42026098	2.09081357	0.10	2.37669869	6.80
2500	4.42026701	2.09081962	0.16	2.37673941	10.65
3000	4.42027102	2.09082364	0.2	2.37676652	14.78

Table 3.2: Convergence table for ATM European and American options with strike $K = 40$ and $T = 1$. The time unit is in seconds.

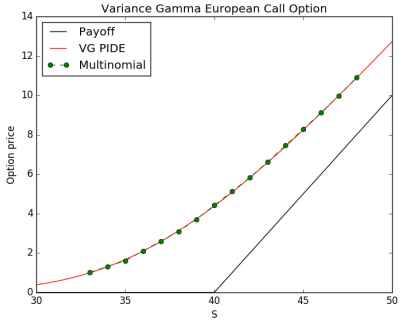

Figure 3.2: European call option with strike $K = 40$ and time to maturity 1 year.

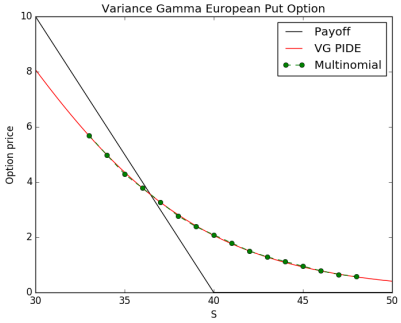

Figure 3.3: European put option with strike $K = 40$ and time to maturity 1 year.

3.2.3 American options

In this section we present the numerical results obtained with the multinomial method algorithm for American put options, and compare them with the PIDE method (see fig. 3.4). The PIDE (2.42) is modified in order to take in account the early exercise feature:

$$\min\left\{-\frac{\partial V(t,x)}{\partial t} - \left(r - \frac{1}{2}\sigma_\epsilon^2 - w_\epsilon\right)\frac{\partial V(t,x)}{\partial x} - \frac{1}{2}\sigma_\epsilon^2\frac{\partial^2 V(t,x)}{\partial x^2} + (\lambda_\epsilon + r)V(t,x)\right. \tag{3.16}$$

$$\left. - \int_{|z|\geq\epsilon} V(t,x+z)\nu(dz)\,,\ \Big(V(t,x) - (K - e^x)^+\Big)\right\} = 0.$$

To solve this equation we use the same discretization and same settings used in sections 2.2.3 and 2.2.4 for the European option problem.

The numerical values of the prices obtained with the multinomial and PIDE methods are collected in Tab. 3.2.3. The run times for the multinomial algorithm are shown in the convergence table 3.2.

Different methods				
S_0	PIDE Call	Multi Call	PIDE Put	Multi Put
36	2.1036	2.1131	3.7842	3.7837
38	3.1163	3.1051	2.7893	2.7756
40	4.4162	4.4203	2.0852	2.0908
42	5.8335	5.8309	1.5050	1.5014
44	7.4417	7.4524	1.1132	1.1229

Table 3.3: European Options, with strike $K = 40$ and $T = 1$.

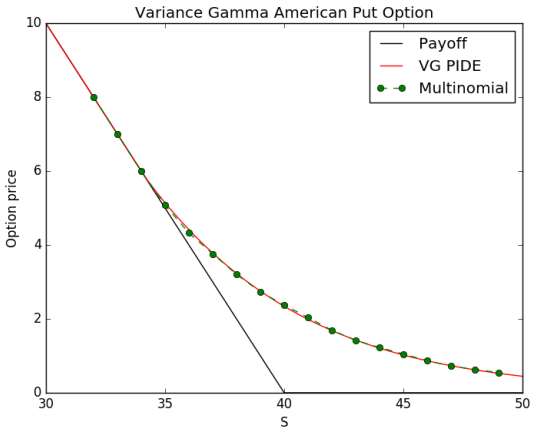

Figure 3.4: American put option with strike $K = 40$ and time to maturity 1 year. Comparison of PIDE prices and multinomial prices.

In Table 3.2.3, we consider also European and American put option prices calculated with the Black-Scholes (BS) models. The BS volatility is chosen equal to the VG volatility $\sigma^{BS} = (\sigma^2 + \theta^2\kappa)$. As expected, the deep OTM (out of the money) prices computed under VG are higher than the corresponding prices computed under BS. This is a consequence of the shape of the VG density function (1.65), which has *heavier tails* than the normal distribution. This means that the probability of a deep OTM option to return in the money, is higher if calculated with the VG model than BS, and therefore we get higher option prices.

The Black-Scholes prices are computed using a binomial algorithm. The same values can be obtained using the multinomial algorithm for the VG process and setting $\theta = \kappa = 0$ and $\sigma = \sigma^{BS}$. Recall that under the Black-Scholes model, the log-returns follow a Brownian motion. Looking at the definition of the VG process (1.63), it is evident that when θ and κ are zero, the process becomes a Brownian motion:

$$X_t^{VG} \underset{\theta,\kappa\to 0}{\to} \sigma W_t.$$

It follows that the price process converges to the Geometric Brownian Motion:

$$S_t = S_0 e^{(r-w)t+X_t} \underset{\theta,\kappa\to 0}{\to} S_0 e^{(r-\frac{1}{2}\sigma^2)t+\sigma W_t}$$

Prices comparison					
S_0	BS Eu. Put	VG Eu. Put	BS Am. Put	VG Am. Put	VG Am. PIDE Put
30	8.1316	8.0809	10	10	10
32	6.5292	6.4055	8	8	8
34	5.1169	4.9851	6.0894	6	6
36	3.9150	3.7837	4.5415	4.3173	4.3982
38	2.9263	2.7756	3.3264	3.2034	3.2195
40	2.1388	2.0908	2.3924	2.3767	2.3531
42	1.5322	1.5014	1.6911	1.6947	1.6849
44	1.0766	1.1229	1.1755	1.2267	1.2118
46	0.7433	0.7858	0.8043	0.8699	0.8650
48	0.5049	0.5787	0.5425	0.6310	0.6221
50	0.3384	0.4259	0.3612	0.4661	0.4480
52	0.2238	0.3015	0.2376	0.3242	0.3222
55	0.1178	0.1909	0.1243	0.2051	0.1990
60	0.0386	0.0880	0.0404	0.0942	0.0913

Table 3.4: Values for European and American put options using Black-Scholes and Variance Gamma model. Strike $K = 40$ and $T = 1$. The BS volatility have same value of the VG volatility: $\sigma^{BS} = (\sigma^2 + \theta^2\kappa) = 0.2049$.

where:

$$\begin{aligned}\lim_{\theta,\kappa\to 0} w &= \lim_{\theta,\kappa\to 0} -\frac{1}{\kappa}\log(1-\theta\kappa-\frac{1}{2}\sigma^2\kappa)\\ &= -\frac{1}{2}\sigma^2.\end{aligned}$$

3.3 Chapter conclusions

In this chapter we make a small digression from the main theme of the thesis by presenting the application of a multinomial approximation for option pricing using the Variance Gamma model. The main results of this chapter are published in the paper [Cantarutti and Guerra, 2018].

The VG process is approximated by a general jump process that has the same first four cumulants of the original VG process. We show that by construction the multinomial method has the same distribution of the approximated process. We obtained numerical results for European and American options and compared them with results from the PIDE method and with Black–Scholes prices. It turns out that the multinomial method is easier to implement than the finite differences method. The algorithm does not involve any matrix multiplication, matrix inversion or decomposition, as in the case of the IMEX method for PIDEs. This means that the computational time is much smaller.

In order to show the ease of use of this algorithm, in Chapter 6 we applied it to solve the option pricing problem with transaction costs, with good performance.

4

HJB equation and viscosity solution

In this chapter we present a brief introduction to the use of the **dynamic programming principle** for solving stochastic control problems. The fundamental equation of dynamic programming is a nonlinear evolution equation for the **value function**. The value function is the optimum value of the payoff considered as a function of the initial data. This principle was introduced in the 1950s by Bellman, see [Bellman, 1957]. For controlled Lévy processes, the approach yields a certain nonlinear PIDE, usually of first or second order, called **Hamilton-Jacobi-Bellman** (HJB) equation.

The theory of viscosity solutions, provides a convenient framework in which to study HJB equations. Typically, the value function is not smooth enough to satisfy the HJB in the classical sense. However, under certain assumptions, it is the unique viscosity solution of the HJB equation with appropriate boundary conditions. A general review of the theory of viscosity solution is presented in [Crandall et al., 1992].

The first notion of viscosity solution has been introduced by [Crandall and Lions, 1983] for first order partial differential equations. The theory has been immediately extended by [Lions, 1983] for second order PDEs. Uniqueness results for general second order equations are presented in [Jensen, 1988], [Ishii, 1989]

and [Ishii and P.L., 1990]. The main development of these works is the *Ishii's lemma*, which plays a key role in most of the uniqueness proofs. The viscosity solution theory has been further extended by [Soner, 1986a], [Soner, 1986b] and [Sayah, 1991] for piecewise deterministic processes with random jumps. An important paper for viscosity solution of Lévy-type PIDEs (PIDEs involving the generator of a Lévy process) is [Barles and Imbert, 2008], where the Ishii's lemma is extended. Further important contributions that deserve to be mentioned are [Pham, 1998], that analyzes HJB equations for optimal stopping problems under Lévy processes and [Cont and Voltchkova, 2005b] that analyze the solution of linear PIDEs derived from option prices problems (plain vanilla and barrier options) for finite and infinite activity Lévy processes.

4.1 Optimal control framework

We consider a framework where the state of the system X_t, is governed by the following controlled SDE with values in $\mathbb{R}^n$:

$$\begin{aligned} dX_t =& b(t, X_{t-}, \alpha_t)dt + \sigma(t, X_{t-}, \alpha_t)dW_t \\ &+ \int_{\mathbb{R}} \gamma(t, X_{t-}, \alpha_t, z)\tilde{N}(dt, dz). \end{aligned} \tag{4.1}$$

where we consider a d-dimensional Brownian motion and a l-dimensional compensated Poisson random measure. The coefficients $b : [0,T] \times \mathbb{R}^n \times A \to \mathbb{R}^n$, $\sigma : [0,T] \times \mathbb{R}^n \times A \to \mathbb{R}^{n\times d}$, $\gamma : [0,T] \times \mathbb{R}^n \times A \times \mathbb{R}^n \to \mathbb{R}^{n\times l}$ are continuous functions with respect to (t,x,a) and $\gamma(t,x,a,\cdot)$ is also bounded uniformly in $a \in A$ in any neighborhood of $z = 0$.

In this chapter we use the same compact notation introduced for (1.39), where all the indexes are suppressed. The expression (4.1) corresponds to the component-wise equation

$$\begin{aligned} dX_t^i =& b^i(t, X_{t-}, \alpha_t)dt + \sum_{j=1}^{d} \sigma^{i,j}(t, X_{t-}, \alpha_t)dW_t^j \\ &+ \sum_{k=1}^{l} \int_{\mathbb{R}} \gamma^{i,k}(t, X_{t-}, \alpha_t, z)\tilde{N}^k(dt, dz). \end{aligned} \tag{4.2}$$

for $1 \le i \le n$.

The set $\mathcal{A}$ is the set of all processes $\alpha : [0,T] \times \Omega \to A$, predictable with respect to the filtration generated by $\{X_t\}_{t\le 0}$ solution of (4.1). The set A is compact. In this thesis we only consider Markovian controls such that $\alpha_t := \alpha(t, X_{t^-})$, with $t \in [0,T]$.

In section 1.2.3, we presented existence and uniqueness conditions for a time-homogeneous SDE (we considered only a one-dimensional compensated Poisson random measure). Following chapter 3.3 of [Gihman and Skorohod, 1979], we can extend those conditions for the more general SDE in (4.1). Let us consider a function $\rho : \mathbb{R}^n \to \mathbb{R}$ satisfying (1.48). We have the following:

(C1) **Lipschitz condition** There exist $K > 0$, such that $\forall x, y \in \mathbb{R}^n$, $\forall t \in [0,T]$

and $\forall a \in A$

$$|b(t,x,a) - b(t,y,a)| + ||\sigma(t,x,a) - \sigma(t,y,a)|| \leq K|x-y|, \tag{4.3}$$
$$||\gamma(t,x,a,z) - \gamma(t,y,a,z)|| \leq |\rho(z)|\,|x-y|. \tag{4.4}$$

Theorem 4.1.1. *The assumptions (4.3) and (4.4) ensure that for each control $\alpha \in \mathcal{A}$, there exists a unique strong solution of (4.1) with given initial conditions.*

See Chapter 3.3 of [Gihman and Skorohod, 1979] for a complete proof.

Note that the Lipschitz conditions (4.3) and (4.4) together with the continuity of b, σ, γ, imply the growth conditions:

$$|b(t,x,a)| + ||\sigma(t,x,a)|| \leq K|1+|x||$$
$$||\gamma(t,x,a,z)|| \leq |\rho(z)|(1+|x|)$$

Definition 4.1.1. *Let us indicate with $\mathcal{T}_{t_1,t_2}$ the* ***set of all stopping times*** *in $[t_1,t_2]$ adapted to $\{\mathcal{F}_s\}_{s\in[t_1,t_2]}$, where $\mathcal{F}_{s\in[t_1,t_2]}$ is the natural filtration associated with the solution of (4.1).*

Theorem 4.1.2. *Let us consider the process described by (4.1) and the hypothesis (4.3) and (4.4) satisfied. For any $k \in [0,2]$, there exist $C > 0$ such that $\forall h, t \in [0,T]$, $x, y \in \mathbb{R}^n$, $\alpha \in \mathcal{A}$ and $\tau \in \mathcal{T}_{0,h}$*

$$\mathbb{E}_{t,x}\big[|X_\tau|^k\big] \leq C(1+|x|^k)$$
$$\mathbb{E}_{t,x}\big[|X_\tau - x|^k\big] \leq C(1+|x|^k)h^{k/2}$$
$$\mathbb{E}_{t,x}\left[\left(\sup_{0\leq s\leq h}|X_s - x|\right)^k\right] \leq C(1+|x|^k)h^{k/2}$$
$$\mathbb{E}_{t,x,y}\big[|X_\tau - Y_\tau|^k\big] \leq C|x-y|^2$$

where we used the simple notation $\mathbb{E}_{t,x}[\cdot]$ to indicate the expectation conditioned on the initial value $X_t = x$. A detailed proof of the previous theorem can be found in the appendix of [Pham, 1998].

Lévy processes and, more in general, all the controlled processes with dynamics (4.1), are **stochastically continuous**, and satisfy the following theorem.

Theorem 4.1.3. *A process $\{X_t\}_{t\geq 0}$ starting at $x = X_0$ and solution of (4.1), for each $\alpha \in \mathcal{A}$ taking values in the compact set A, and for each $\rho > 0$, satisfies*

$$\mathbb{P}\Big(|X_t - x| \geq \rho\Big) \underset{t\to 0}{\longrightarrow} 0. \tag{4.11}$$

Proof. By the Markov inequality

$$\mathbb{P}\Big(|X_t - x| \geq \rho\Big) \leq \frac{\mathbb{E}_{\bar{x}}\Big[|X_t - \bar{x}|\Big]}{\rho} \leq \frac{C}{\rho}(1+|x|)\sqrt{t}$$

where we used (4.8), with $C > 0$. Taking the limit $t \to 0$ proves the theorem. □

For $\delta > 0$, let us define the following two useful integral operators. For $\phi \in C^{1,2}([0,T] \times \mathbb{R}^n)$ consider the operator:[1]

$$\mathcal{I}^{1,\delta,a}(t,x,\phi) = \int_{|z|\leq\delta} \Big[\phi(t,x+\gamma(t,x,a,z)) - \phi(t,x) - D_x\phi(t,x)\cdot\gamma(t,x,a,z)\Big]\nu(dz) \tag{4.12}$$

where $D_x\phi$ corresponds to the gradient of ϕ.
For $\phi \in \mathcal{C}_2([0,T]\times\mathbb{R}^n)$ (see definition (2.1.7)) and $p \in \mathbb{R}^n$, consider the operator:

$$\mathcal{I}^{2,\delta,a}(t,x,p,\phi) = \int_{|z|\geq\delta} \Big[\phi(t,x+\gamma(t,x,a,z)) - \phi(t,x) - p\cdot\gamma(t,x,a,z)\Big]\nu(dz) \tag{4.13}$$

We can check that the integral operators $\mathcal{I}^{1,\delta,a}$ and $\mathcal{I}^{2,\delta,a}$ are well defined.

In the integral (4.12), for $|z| \leq \delta$, we know that $|\gamma(t,x,a,z)|$ is bounded by $\bar{\gamma} = \sup_{|z|\leq\delta}\gamma(t,x,a,z)$, for fixed t, x, uniformly in a. We can use a first order Taylor approximation on $\phi(t,x+\gamma(t,x,a,z))$ and consider the Lagrange remainder. For $y \in \big(x,\gamma(t,x,a,z)\big)$, we can write:

$$\begin{aligned}
&\Big|\phi(t,x+\gamma(t,x,a,z)) - \phi(t,x) - D_x\phi(t,x)\cdot\gamma(t,x,a,z)\Big| \\
&= \left|\frac{1}{2}\gamma(t,x,a,z)^T \cdot D_{xx}\phi(t,y)\cdot\gamma(t,x,a,z)\right| \\
&\leq \frac{1}{2}\left|\left|D_{xx}\phi(t,y)\right|\right| \left|\gamma(t,x,a,z)\right|^2 \\
&\leq \frac{1}{2}\left|\left|D_{xx}\phi(t,y)\right|\right| \left|\rho(z)(1+|x|)\right|^2 \\
&\leq \frac{1}{2}M\left|\rho(z)(1+|x|)\right|^2.
\end{aligned}$$

where we used the Schwarz inequality and $M = \sup_{y\in[x,\bar{\gamma}]}||D_{xx}\phi(t,y)||$. Thanks to (1.48) we can see that the integral is well defined.

Let us consider the integral (4.13) on $|z| \geq \delta$. By definition $\phi \in \mathcal{C}_2([0,T]\times\mathbb{R}^n)$ has quadratic growth. Thanks to Theorem (1.1.8) and (4.6) each term inside the integral is well defined.

Finally, for $\phi \in C^{1,2}([0,T]\times\mathbb{R}^n)\bigcap\mathcal{C}_2([0,T]\times\mathbb{R}^n)$ we can define the **Integral operator**:

$$\begin{aligned}
\mathcal{I}^a(t,x,\phi) &= \mathcal{I}^{1,\delta,a}(t,x,\phi) + \mathcal{I}^{2,\delta,a}(t,x,D_x\phi,\phi) \\
&= \int_{\mathbb{R}^n}\Big[\phi(t,x+\gamma(t,x,a,z)) - \phi(t,x) - D_x\phi(t,x)\cdot\gamma(t,x,a,z)\Big]\nu(dz).
\end{aligned} \tag{4.14}$$

The 1 dimensional **Lévy operator** can be obtained with $\gamma(t,x,a,z) = z$, $\delta = 1$ and $p = 0$ in this way:

$$\begin{aligned}
\mathcal{I}_L(t,x,\phi) &= \mathcal{I}^{1,1,0}(t,x,\phi) + \mathcal{I}^{2,1,0}(t,x,0,\phi) \\
&= \int_{\mathbb{R}^n}\Big[\phi(t,x+z) - \phi(t,x) - D_x\phi(t,x)\, z\mathbb{1}_{|z|<1}(z)\Big]\nu(dz)
\end{aligned}$$

[1] Inside the operators $\mathcal{I}^{1,\delta,a}$ and $\mathcal{I}^{1,\delta,a}$, we intend only the k-column (for $1 \leq k \leq l$) of the $n \times l$ matrix $\gamma(t,x,a,z)$. As explained at the beginning of the section, we suppressed the indexes to simplify the notation.

corresponding to the integral term in (1.37). The **controlled integro-differential operator** is defined as:

$$\mathcal{L}^a(t,x,\phi) = A^a\phi(t,x) + \mathcal{I}^a(t,x,\phi) \tag{4.15}$$

with

$$A^a\phi(t,x) = D_x\phi(t,x)\cdot b(t,x,a) + \frac{1}{2}\mathrm{Tr}\Big[\sigma(t,x,a)^T\, D_x^2\phi(t,x)\,\sigma(t,x,a)\Big],$$

where **Tr** is the trace of the matrix.

4.1.1 Dynamic Programming Principle

Define the cylindrical region $Q = [t_0, T)\times\mathcal{O}$, with $\mathcal{O}\subseteq\mathbb{R}^n$ an open set. With abuse of notation we define the first exit time from Q as

$$\tau = \inf\{s : (s, X_s)\notin Q\}. \tag{4.16}$$

Note that τ is the exit time from O if X_s exits before time T. If $X_s\in\mathcal{O}$ $\forall s\in[t_0,T)$ than $\tau = T$. If $\mathcal{O} = \mathbb{R}^n$, than the set Q does not have lateral boundaries but only terminal boundary at $\{T\}\times R^n$.
Let $f : [0,T]\times\mathbb{R}^n\times A\to\mathbb{R}$ and $g : \big([0,T)\times\mathbb{R}^n\backslash\mathcal{O}\big)\bigcup\big(\mathbb{R}^n\times\{T\}\big)\to\mathbb{R}$ two continuous functions. We assume that there exists $C<\infty$, they satisfy the condition:

$$|f(t,x,a)| + |g(t,x)| \le C(1+|x|^2) \qquad \forall a\in A. \tag{4.17}$$

Denote by $\mathcal{A}_{t,x}$ the subset of $\mathcal{A}$ dependent on the current state (t,x) such that $\mathbb{E}_{t,x}[\int_t^\tau |f(s,X_s,\alpha_s)|ds] < \infty$. We define the **objective function**

$$J(t,x;\alpha) = \mathbb{E}_{t,x}\left[\int_t^\tau f(s,X_s,\alpha_s)ds + g(\tau,X_\tau)\right]. \tag{4.18}$$

Thanks to the growth condition 4.17 and to equation (4.9), we can verify easily that the objective function is well defined. We want to maximize the objective function over the set of control processes $\mathcal{A}_{t,x}$. We introduce the **value function**

$$V(t,x) = \sup_{\alpha\in\mathcal{A}_{t,x}} J(t,x,\alpha). \tag{4.19}$$

If the optimal control α^* exists, we have $V(t,x) = J(t,x,\alpha^*)$. We also always assume that the value function is measurable.

The dynamic programming principle (DPP) is a fundamental principle in the theory of stochastic control. It is formulated as follows:

Theorem 4.1.4. ***Dynamic Programming Principle**:*

- *For all $\alpha\in\mathcal{A}_{t,x}$ and all stopping time $\theta\in\mathcal{T}_{t,\tau}$:*

$$V(t,x) \ge \mathbb{E}_{t,x}\left[\int_t^\theta f(s,X_s,\alpha_s)ds + V(\theta,X_\theta)\right]. \tag{4.20}$$

- *For all $\epsilon > 0$, there exists $\alpha \in \mathcal{A}_{t,x}$ such that $\forall \theta \in \mathcal{T}_{t,\tau}$:*

$$V(t,x) \leq \mathbb{E}_{t,x}\left[\int_t^\theta f(s, X_s, \alpha_s)ds + V(\theta, X_\theta)\right] + \epsilon. \tag{4.21}$$

Since ϵ is arbitrary, we can write the DPP in the following form:

$$V(t,x) = \sup_{\alpha \in \mathcal{A}_{t,x}} \mathbb{E}_{t,x}\left[\int_t^\theta f(s, X_s, \alpha_s)ds + V(\theta, X_\theta)\right]. \tag{4.22}$$

The proof of the DPP is quite technical, and several proofs can be found in many textbooks on stochastic control theory. See for instance Chapter 4.7 of [Fleming and Soner, 2005], where the authors only considered diffusion processes. A proof of the DPP that considers a controlled process as in (4.1), and bounded function f and g, can be found in [Goldys and Wu, 2016]. Note that the two cited proofs assume both a compact control set A.

An admissible control $\alpha \in \mathcal{A}_{t,x}$ is ϵ-optimal conditioned on (t,x) if and only if it is ϵ-optimal on every $\mathcal{A}_{\theta,X_\theta}$ with $\theta \in \mathcal{T}_{t,\tau}$. In order to determine the ϵ-optimal control α_t, it suffices to consider the DPP with a stopping time θ arbitrarily close to t.

4.1.2 HJB equation (formal derivation)

In this section we assume that the value function is continuously differentiable and obtain a formal nonlinear PIDE satisfied by the value function. In general however, the value function is not necessarily differentiable, and the notion of *viscosity solution* should be considered. The Hamilton-Jacobi-Bellman equation (HJB) is the infinitesimal version of the DPP. It describes the local behavior of the value function when the stopping time θ converges to t. The HJB equation is also called **dynamic programming equation**.

Let us consider the stopping time $\theta = t + h$, with $h > 0$, in the DPP (4.20) and a constant control $\alpha_s = a$ for $s \in [t, \theta]$. Subtract $V(t,x)$ from both sides and then divide by h. This yields

$$\begin{aligned} 0 &\geq \mathbb{E}_{t,x}\left[\int_t^{t+h} f(s, X_s, a)ds + V(t+h, X_{t+h}) - V(t,x)\right]. \\ &\geq \mathbb{E}_{t,x}\left[\frac{1}{h}\int_t^{t+h} f(s, X_s, a) + \left(\frac{\partial V}{\partial s} + \mathcal{L}^a V\right)(s, X_s)\, ds\right]. \end{aligned}$$

where we used the Dynkin formula (Eq. 1.2.5) assuming $V \in C^{1,2}$, with $\mathcal{L}^\alpha$ the integro-differential operator 4.15. By sending $h \to 0$ and using the mean value theorem we obtain the equation:

$$\frac{\partial V(t,x)}{\partial t} + \left(f(t,x,a) + \mathcal{L}^a V(t,x)\right) \leq 0. \tag{4.23}$$

Since this hold true for every $a \in A$, we can write

$$-\frac{\partial V(t,x)}{\partial t} - \sup_{a \in A}\left(f(t,x,a) + \mathcal{L}^a V(t,x)\right) \geq 0. \tag{4.24}$$

On the other hand, if we suppose that the optimal control α^* exists, then:

$$0 = \mathbb{E}_{t,x}\bigg[\int_t^{t+h} f(s, X_s^*, \alpha_s^*)ds + V(t+h, X_{t+h}^*) - V(t,x)\bigg], \tag{4.25}$$

where X^* is the optimal state, solution of (4.1), under the control α^*. Using similar arguments as above we obtain:

$$-\frac{\partial V(t,x)}{\partial t} - \Big(f(t,x,\alpha_t^*) + \mathcal{L}^{\alpha_t^*}V(t,x)\Big) = 0. \tag{4.26}$$

We can combine the two results (4.24), (4.26) in a single compact equation i.e. the **dynamic programming equation**:

$$-\frac{\partial V(t,x)}{\partial t} - \sup_{a\in A}\Big(f(t,x,a) + \mathcal{L}^a V(t,x)\Big) = 0. \tag{4.27}$$

Remark 4. *In the previous sections we worked under the assumption that A is compact. But in general, the HJB (4.27) is well defined also for non-compact sets A, as long as the supremum in $a \in A$ is finite. When A is compact and b, σ, γ and f are continuous with respect to a, the supremum is always attained. In section 4.2 we will see that when A is unbounded, and the supremum in $a \in A$ is not finite, the HJB equation becomes a variational inequality.*

Let us define the Hamiltonian function

$$0 = -\frac{\partial V(t,x)}{\partial t} - \mathcal{H}\big(t, x, D_x V, D_{xx}V, \mathcal{I}^a(t,x,V)\big)$$

where for $(t,x,p,M,I^a) \in Q \times \mathbb{R}^n \times S^n \times \mathbb{R}$:[2]

$$\begin{aligned}\mathcal{H}(t,x,p,M,I^a) = \sup_{a\in A}\Big(& f(t,x,a) + b(t,x,a)p \\ &+ \frac{1}{2}\mathrm{Tr}\big(\sigma(t,x,a)\sigma^T(t,x,a)\,M\big) + I^a\Big)\end{aligned} \tag{4.28}$$

The regular terminal and lateral boundary conditions of this nonlinear PIDE are:

$$V(T,x) = g(T,x). \tag{4.29}$$

$$V(t,x) = g(t,x) \quad \text{for} \quad (t,x) \in [0,T) \times (\mathbb{R}^n \backslash \mathcal{O}). \tag{4.30}$$

4.2 Singular control

In the previous section, the theory of stochastic control for generalized jump-diffusion processes is presented under the assumption that the control process takes values in a compact space. In this section this assumption is relaxed, and the control space A is assumed to be unbounded. When the problem coefficients are linear functions of the control, the form (4.27) of the HJB equation is no more valid. In section 4.2.1 we present a formal derivation of the HJB equation, that under these assumptions has the form of a **variational inequality**. A

[2] S^n is the set of symmetric matrices of dimension n.

rigorous basis to this derivation can be given a posteriori by means of a verification theorem, or by considering the viscosity solution framework. For more details we refer to Chapter 8.4 of [Fleming and Soner, 2005] or to Theorem 5.2 of [Øksendal and Sulem, 2007].

When considering A unbounded and a linear dependence of the coefficients on α, in general there are no optimal controls, and ϵ-optimal controls take arbitrarily large values. For this reason it is convenient to reformulate the problem by introducing a new control variable ξ defined as

$$\xi_t = \int_0^t \hat{\alpha}_s du_s, \tag{4.31}$$

with

$$\hat{\alpha}_s := \begin{cases} \frac{\alpha_s}{|\alpha_s|} & \text{if } \alpha_s \neq 0 \\ 0 & \text{if } \alpha_s = 0. \end{cases} \quad \text{and} \quad u_t = \int_0^t |\alpha_s| ds. \tag{4.32}$$

In order to obtain optimal controls we enlarge the class of control to admit ξ_t which may not be an absolutely continuous function of t. But we assume that ξ_t is a cádlág, predictable, non-decreasing function with bounded variation on every time interval $[0, T]$. In section (4.2.2) we present the general formulation for **singular control** problems. The name *singular* comes from the fact that the control process ξ can be singular with respect to the Lebesgue measure dt.

4.2.1 Derivation of the variational inequality

When the control space A is unbounded, the Hamiltonian (4.28) may take infinite values. Let us assume that A is a closed cone in $\mathbb{R}^m$ i.e.

$$a \in A, \quad \lambda > 0 \Longrightarrow \lambda a \in A. \tag{4.33}$$

Let us assume also that the control influences linearly the dynamics of the system and the running function.

$$\begin{aligned} &b(t,x,a) = \hat{b}(t,x) + \kappa(t,x)a, \quad \sigma(t,x,a) = \hat{\sigma}(t,x), \\ &\gamma(t,x,a,z) = \hat{\gamma}(t,x,z), \quad f(t,x,a) = \hat{f}(t,x) + h(t,x)a, \end{aligned} \tag{4.34}$$

where $\kappa : [0,T] \times \mathbb{R}^n \to \mathbb{R}^{n \times m}$ and $h : [0,T] \times \mathbb{R}^n \to \mathbb{R}^m$ are continuous. Let us indicate with $\hat{I}$ the integral operator (4.14) containing $\hat{\gamma}(t,x,z)$. The Hamiltonian becomes

$$\begin{aligned} \mathcal{H}(t,x,p,M,\hat{I}) = &\Big(\hat{f}(t,x) + p\,\hat{b}(t,x) \\ &+ \frac{1}{2}\mathrm{Tr}\big(\sigma(t,x)\sigma^T(t,x)\,M\big) + \hat{I} \Big) + \underbrace{\sup_{a \in A} \Big(\big(p\,\kappa(t,x) + h(t,x)\big)a \Big)}_{\hat{H}(t,x,p)} \end{aligned}$$

If for some $a \in A$ and for fixed (t,x) we have $\big(p\,\kappa(t,x) + h(t,x)\big)a > 0$, then $\hat{H}(t,x,p) = \infty$. We can define

$$H(t,x,p) = \sup_{a \in \hat{K}} \Big(\big(p\,\kappa(t,x) + h(t,x)\big)a \Big) \quad \text{with} \quad \hat{K} = \{a \in A : |a| = 1\} \tag{4.35}$$

such that

$$\hat{H}(t,x,p) = \begin{cases} \infty, & \text{if } H(t,x,p) > 0 \\ 0, & \text{if } H(t,x,p) \leq 0. \end{cases} \tag{4.36}$$

Using this last equation together with the HJB (4.27) with $a = 0$ we have the following two equations:

$$\begin{cases} \frac{\partial V(t,x)}{\partial t} + \hat{f}(t,x) + \mathcal{L}^{a=0}V(t,x) \leq 0. \\ H\big(t,x,D_xV(t,x)\big) \leq 0. \end{cases} \tag{4.37}$$

Now, suppose $H(t,x,D_xV(t,x)) < 0$. Then $\hat{H}(t,x,p) = 0$ and the optimal control is indeed $a = 0$, with the uncontrolled HJB equal to zero:

$$H\big(t,x,D_xV(t,x)\big) < 0 \quad \Longrightarrow \quad \frac{\partial V(t,x)}{\partial t} + \hat{f}(t,x) + \mathcal{L}^{a=0}V(t,x) = 0. \tag{4.38}$$

The last equation can be written in a more compact form. We have the following **variational inequality**:

$$\max\left\{ \frac{\partial V(t,x)}{\partial t} + \hat{f}(t,x) + \mathcal{L}^{a=0}V(t,x)\,,\, H\big(t,x,D_xV(t,x)\big) \right\} = 0 \quad \text{for} \quad (t,x) \in Q. \tag{4.39}$$

4.2.2 Singular control formulation

Let us consider a state system governed by the following SDE:

$$\begin{aligned} dX_t =& \hat{b}(t,X_{t-})dt + \hat{\sigma}(t,X_{t-})dW_t \\ &+ \int_{\mathbb{R}} \hat{\gamma}(t,X_{t-},z)\tilde{N}(dt,dz) + \kappa(t,X_{t-})d\xi_t. \end{aligned} \tag{4.40}$$

The m-dimensional process $\xi : [0,T]\times\Omega \to \mathbb{R}^m$ is a predictable, cádlág, bounded variation, non-decreasing process, with $\xi_{0^-} = 0$. Let us denote with Π the space of all controls ξ and with $\Pi_{t,x}$ the subset dependent on the current state (t,x).

For τ defined in (4.16), the objective function is:

$$J^{\xi}(t,x) = \mathbb{E}_{t,x}\left[\int_t^{\tau} f(s,X_s)ds + \int_t^{\tau} h(s,X_{s^-})d\xi_s + g(\tau,X_\tau) \right]. \tag{4.41}$$

The value function is

$$V(t,x) = \sup_{\Pi_{t,x}} J^{\xi}(t,x). \tag{4.42}$$

Under certain assumptions, it can be proved that the value function (4.42) is a viscosity solution of a variational inequality with form (4.39).

Let us assume that the value function satisfies the DPP. A straightforward modification of (4.20) and (4.21) yields the expressions:

Theorem 4.2.1. ***DPP for singular control:***

- *For all $\xi \in \Pi_{t,x}$ and all stopping time $\theta \in \mathcal{T}_{t,\tau}$:*

$$V(t,x) \geq \mathbb{E}_{t,x}\left[\int_t^{\theta} f(s,X_s)ds + \int_t^{\theta} h(s,X_{s^-})d\xi_s + V(\theta,X_\theta) \right]. \tag{4.43}$$

- *For all $\epsilon > 0$, there exists $\xi \in \Pi_{t,x}$ such that $\forall \theta \in \mathcal{T}_{t,\tau}$:*

$$V(t,x) \leq \mathbb{E}_{t,x}\left[\int_t^\theta f(s,X_s)ds + \int_t^\tau h(s,X_{s^-})d\xi_s + V(\theta,X_\theta)\right] + \epsilon. \quad (4.44)$$

The singular control problems we will introduce in chapter 5 are formulated according to the general framework introduced in [De Vallière et al., 2016]. We refer to this article for the proof of the DPP for singular control problems under the hypothesis introduced in chapter 5.

4.3 Viscosity solutions theory

This section is dedicated to the definition of viscosity solutions. In the literature there are different definitions of viscosity solution depending on the context. For instance, the theory presented in [Pham, 2009] considers only the PDE case, while in [Cont and Tankov, 2003] only linear PIDEs are considered. Other important references are [Fleming and Soner, 2005], [Pham, 1998] and [Barles and Imbert, 2008] among others. In the general discontinuous viscosity solutions approach, there is no need to prove a priori the continuity of the value function V. The continuity will actually follow from a strong comparison principle.
Here we present the definition for continuous viscosity solutions introduced in [De Vallière et al., 2016], which is suitable for the problem proposed in this thesis (Chapter 5).

4.3.1 Definition viscosity solution

Let us consider a general parabolic problem:

$$\begin{cases} F(t,x,u,D_tu,D_xu,D_{xx}u,\mathcal{I}(t,x,u)) = 0 \quad \text{for} \quad (t,x) \in Q \\ u(t,x) = g(t,x) \quad \text{for} \quad (t,x) \notin Q \end{cases} \quad (4.45)$$

where $g \in C^0 \bigcap \mathcal{C}_2([0,T] \times (\mathbb{R}^n \backslash \mathcal{O}))$ is a given function and F is a continuous function that satisfies the following elliptic/parabolic local and non local conditions. For all $t \in [0,T)$; $x \in \mathcal{O}$; $r, \hat{r} \in \mathbb{R}$; $q, \hat{q} \in \mathbb{R}$; $p \in \mathbb{R}^n$; $M, \hat{M} \in \mathcal{S}^n$; $\mathcal{I}, \hat{\mathcal{I}} \in \mathbb{R}$:

- $r \leq \hat{r} \implies F(t,x,r,q,p,M,\mathcal{I}) \leq F(t,x,\hat{r},q,p,M,\mathcal{I})$
- $q \leq \hat{q} \implies F(t,x,r,q,p,M,\mathcal{I}) \geq F(t,x,r,\hat{q},p,M,\mathcal{I})$
- $M \leq \hat{M} \implies F(t,x,r,q,p,M,\mathcal{I}) \geq F(t,x,r,q,p,\hat{M},\mathcal{I})$
- $\mathcal{I} \leq \hat{\mathcal{I}} \implies F(t,x,r,q,p,M,\mathcal{I}) \geq F(t,x,r,q,p,M,\hat{\mathcal{I}})$.

where the matrix ordering is intended with this meaning:

$$\hat{M} \geq M \Leftrightarrow \hat{M} - M \text{ is positive semi-definite.}$$

Having in mind our specific problem, we will assume that the viscosity subsolution and supersolution are continuous on $[0,T] \times \mathbb{R}^n$. We can now introduce the definitions:

Definition 4.3.1. *A continuous function u is a* ***viscosity subsolution*** *of (4.45) if for any $(\bar{t},\bar{x}) \in [0,T]\times\mathbb{R}^n$ and any test function $\phi \in C^{1,2}([0,T]\times\mathbb{R}^n)\bigcap \mathcal{C}_2([0,T]\times\mathbb{R}^n)$ such that $u-\phi$ has a global maximum at $(\bar{t},\bar{x})$ the following is satisfied:*

$$F\Big(\bar{t},\bar{x},u(\bar{t},\bar{x}),D_t\phi(\bar{t},\bar{x}),D_x\phi(\bar{t},\bar{x}),D_{xx}\phi(\bar{t},\bar{x}),\mathcal{I}(\bar{t},\bar{x},\phi(\bar{t},\bar{x}))\Big)\leq 0 \qquad (4.46)$$
$$\text{for} \quad (\bar{t},\bar{x})\in Q.$$

In the same way we define:

Definition 4.3.2. *A continuous function u is a* ***viscosity supersolution*** *of (4.45) if for any $(\bar{t},\bar{x}) \in [0,T]\times\mathbb{R}^n$ and any test functions $\phi \in C^{1,2}([0,T]\times\mathbb{R}^n)\bigcap \mathcal{C}_2([0,T]\times\mathbb{R}^n)$ such that $u-\phi$ has a global minimum at $(\bar{t},\bar{x})$ the following is satisfied:*

$$F\Big(\bar{t},\bar{x},u(\bar{t},\bar{x}),D_t\phi(\bar{t},\bar{x}),D_x\phi(\bar{t},\bar{x}),D_{xx}\phi(\bar{t},\bar{x}),\mathcal{I}(\bar{t},\bar{x},\phi(\bar{t},\bar{x}))\Big)\geq 0 \qquad (4.47)$$
$$\text{for} \quad (\bar{t},\bar{x})\in Q.$$

When a function is both a viscosity subsolution and supersolution, it is called a **viscosity solution**.

4.4 Chapter conclusions

This chapter serves as a brief introduction to the theory of stochastic control for processes with a jump and a diffusion components.

The topics in the initial section 4.1 follow closely the presentations given in [Gihman and Skorohod, 1979] and [Pham, 1998]. In the same section, we introduce the important *dynamic programming principle*, and derive the HJB equation for "regular" controls. In this presentation we assume a compact control set.

In section 4.2 we extend the theory to *singular controls*. In this framework the control set is assumed to be unbounded and the class of controls is enlarged in order to contain also non absolute continuous processes.

For completeness, we started the presentation of this chapter with the theory for regular controls. However, in this thesis we are more interested in the theory for singular controls! The reason is that it includes the portfolio optimization problems with transaction costs that we will formulate in chapter 5.

In the end of the chapter we present the definition of viscosity solution, that will be used in the proofs of chapter 5.

5

Option pricing with transaction costs

5.1 Transaction costs models for the portfolio selection problem

In the mathematical finance literature, the first application of stochastic optimal control theory to finance appears in the seminal paper [Merton, 1969]. This is the classical *Merton optimal portfolio problem*, and is formulated as follows. An investor has a portfolio consisting of two assets, one "risk free" B, and the other "risky" S. The assets in the portfolio have dynamics

$$\begin{cases} dB_t &= rB_t dt \\ dS_t &= S_t \left(\mu dt + \sigma dW_t\right), \end{cases} \tag{5.1}$$

with $\mu > r$ and $\sigma > 0$. If we denote with $\mathcal{W}_t = B_t + S_t$ the wealth at time t, and with π_t the fraction of wealth invested in the risky asset, such that $S_t = \pi_t \mathcal{W}_t$ and $B_t = (1 - \pi_t)\mathcal{W}_t$. The model also considers the investor consumption $\{c_t\}_{0 \leq t \leq T}$. The optimization problem can be formulated considering

the utility function of the consumption and the terminal condition. It consists in maximizing the objective function

$$J(x;\pi,c)=\mathbb{E}_x\left[\int_0^T e^{-\beta t}\mathcal{U}(c_t)dt+g(T,\mathcal{W}_T)\right], \tag{5.2}$$

where $\beta>0$ and $\mathcal{U}:[0,\infty)\to\mathbb{R}$ is a concave and increasing utility function, such that $\mathcal{U}(0)=0$. In the original article, the function g is called *bequest valuation function* and is assumed to be concave. We indicate the initial portfolio state with $x=(B_0,S_0)$. The wealth $\mathcal{W}$ takes value in $\mathcal{O}=(0,\infty)$ and has state dynamics

$$d\mathcal{W}_t=(1-\pi_t)\mathcal{W}_t r dt+\pi_t\mathcal{W}_t(\mu dt+\sigma dW_t)-c(t)dt. \tag{5.3}$$

The control α_t is a two dimensional vector $(\alpha_t^1,\alpha_t^2)=(\pi_t,c_t)$, with values in $A=\mathbb{R}\times[0,\infty)$. This is a finite horizon problem with an objective function of type (4.18) and HJB equation as in (4.27). Even if in this case the control space A is not compact, the supremum in (4.27) is attained, thanks to the concave structure of the problem.

Merton, using the utility $\mathcal{U}(c)=\frac{1}{\gamma}c^\gamma$ with $0<\gamma<1$ (of HARA type) obtained an explicit solution for the problem. In particular he found that the optimal control

$$\pi_t^*=\frac{\mu-r}{(1-\gamma)\sigma^2} \tag{5.4}$$

is a constant and does not depend on the state variables. This means that the optimal fraction of the risky and risk free assets in the portfolio is constant. The line containing all the optimal points in the (B,S)-plane is the so called **Merton line**.

The Merton portfolio selection problem is the first optimal portfolio problem solved by stochastic control theory methods. In [Merton, 1971] the author extends the results to different utility functions. All the successive models are generalizations of the Merton model, where by the word "generalizations", we mean that under some particular choices of the parameters, the Merton problem is included in those models as a special case. For what concerns this thesis, we review the main portfolio selection problems that have appeared in the literature, considering the presence of proportional transaction costs. The first portfolio selection model with transaction costs and consumption was introduced by [Constantinides, 1986] and then extended by [Davis and Norman, 1990]. We refer to [Shreve and Soner, 1994] for a viscosity solution approach to the same problem. Other important contributions are [Dai and Yi, 2009], [Dai et al., 2010], [Liu and M., 2002], [Liu and M., 2007], [Benth et al., 2001], [Framstad et al., 1999], [De Vallière et al., 2016] where the last four articles use Lévy processes to model the stock dynamics. All the cited works are portfolio optimization models formulated as infinite horizon problems, and do not involve the pricing of the options.

Further assumptions are needed when the portfolio process is introduced for the purpose of pricing a derivative contract. The main contributions on option pricing with transaction costs, using the *indifference pricing* approach are [Hodges and Neuberger, 1989] and [Davis et al., 1993] (DPZ). The problem is formalized as a finite horizon singular stochastic control problem. In this

section we present the DPZ model, which is the building block of this thesis, and then extend it following the framework introduced by [De Vallière et al., 2016], that consider the stock dynamics described by Lévy processes.

5.1.1 Davis - Panas - Zariphopoulou (DPZ)

This is the market model with proportional transaction costs presented in [Davis et al., 1993]. The authors consider a portfolio composed by one risk-free asset B (bank account) paying a fixed interest rate $r > 0$ and a stock S. The symbol Y denotes the number of shares of the stock S that the investor holds. The state of the portfolio at time $t \in [t_0, T]$ is (B_t^π, Y_t^π, S_t), and is the solution of the following the SDE:

$$\begin{cases} dB_t^\pi &= rB_t^\pi dt - (1+\theta_b)S_t dL_t + (1-\theta_s)S_t dM_t \\ dY_t^\pi &= dL_t - dM_t \\ dS_t &= S_t\left(\mu dt + \sigma dW_t\right). \end{cases} \tag{5.5}$$

for a particular strategy $\{\pi_t\}_{t\in[t_0,T]} := \{(L_t, M_t)\}_{t\in[t_0,T]}$. The parameters θ_b, $\theta_s \geq 0$ are the proportional transaction costs when buying and selling respectively. This portfolio equation is a generalization of the portfolio in the Merton problem (5.1), where there is a new state variable Y and the action of the controls is represented with a different approach. The control process $\{\pi_t\}_{t\in[t_0,T]} := \{(L_t, M_t)\}_{t\in[t_0,T]}$ represents the trading strategy and indicates the cumulative number of shares bought and sold respectively in $[t_0, T]$. The strategy π_t is a cádlág, $\mathcal{F}_t$-progressively measurable, nondecreasing process with bounded variation in every finite time interval and such that $\pi_{t_0^-} = (L_{t_0^-}, M_{t_0^-}) = (0, 0)$, allowing for an initial transaction at t_0.
In the discontinuity points of L_t and M_t we indicate the process variation with $\Delta L_t = L_t - L_{t^-}$ and $\Delta M_t = M_t - M_{t^-}$.

Definition 5.1.1. *We define the **cash value** function $c(y, s)$ as the value in cash when the shares in the portfolio are liquidated, i.e. long positions are sold and short positions are covered.*

$$c(y,s) = \begin{cases} (1+\theta_b)ys, & \text{if } y \leq 0 \\ (1-\theta_s)ys, & \text{if } y > 0. \end{cases} \tag{5.6}$$

Definition 5.1.2. *For $t \in [t_0, T]$, we define the **total wealth** process:*

$$\mathcal{W}_t^\pi = B_t^\pi + c(Y_t^\pi, S_t). \tag{5.7}$$

We say that a portfolio is solvent if the portfolio's wealth $\mathcal{W}_t$ is greater than a fixed constant $-C$, with $C \geq 0$ that may depend on the initial wealth and on the parameters in (5.5). This constant may be interpreted as the *credit availability* of the investor.

Definition 5.1.3. *We define the **solvency region**:*

$$\mathcal{S} := \Big\{(b, y, s) \in \mathbb{R} \times \mathbb{R} \times \mathbb{R}^+ : b + c(y, s) > -C\Big\}. \tag{5.8}$$

Definition 5.1.4. *The set of* ***admissible trading strategies*** $\Pi(t_0, b, y, s)$*, is defined as the set of all right-continuous, nondecreasing, $\mathcal{F}_t$-progressively measurable processes $\{\pi_t\}_{t\in[t_0,T]}$, such that (B_t^π, Y_t^π, S_t) is a solution of (5.5) with initial values $(B_{t_0}^\pi = b, Y_{t_0}^\pi = y, S_{t_0} = s)$, and*

$$\mathcal{W}_t^\pi \in \mathcal{S} \tag{5.9}$$

for $t \in [t_0, T]$ almost surely.

Remark 5. *In the first problem formulation of [Hodges and Neuberger, 1989], the authors consider a portfolio starting with zero total wealth. The writer (or buyer) of the option creates a portfolio at time t_0 with the purpose of hedging the option. It is reasonable to assume that no shares of the underlying stock are contained in the portfolio at time t_0. In the numerical computations of Chapter 6, we will consider portfolios with zero initial shares and initial cash amount B_0. The model is able to compute option prices for an investor with any initial number Y_0 of shares if needed.*

Utility maximization

Let us assume an investor builds a portfolio with cash, shares of a stock and in addition he sells or purchases a European call option written on the same stock, with strike price K and expiration date T.
From now on, we introduce the superscripts w, b and 0 to indicate the writer, buyer and zero-option portfolios respectively. In the zero-option portfolio, the wealth process $\{\mathcal{W}_t^{0;\pi}\}_{t\in[t_0,T]}$ corresponds to (5.7).

Definition 5.1.5. *For $t \in [t_0, T]$, we define the wealth processes for the writer:*

$$\begin{aligned}\mathcal{W}_t^{w;\pi} := B_t^\pi + c(Y_t^\pi, S_t)\mathbb{1}_{\{t<T\}} \\ + c(Y_t^\pi, S_t)\mathbb{1}_{\{t=T,\, S_t(1+\theta_b)\le K\}} + \Big(c\big(Y_t^\pi - 1, S_t\big) + K\Big)\mathbb{1}_{\{t=T,\, S_t(1+\theta_b)>K\}}\end{aligned} \tag{5.10}$$

and the buyer:

$$\begin{aligned}\mathcal{W}_t^{b;\pi} := B_t^\pi + c(Y_t^\pi, S_t)\mathbb{1}_{\{t<T\}} \\ + c(Y_t^\pi, S_t)\mathbb{1}_{\{t=T,\, S_t(1+\theta_b)\le K\}} + \Big(c\big(Y_t^\pi + 1, S_t\big) - K\Big)\mathbb{1}_{\{t=T,\, S_t(1+\theta_b)>K\}}.\end{aligned} \tag{5.11}$$

In the case the option is exercised, $S_T(1+\theta_b) > K$, the buyer pays to the writer the strike value K in cash, and the writer delivers one share to the buyer. In a market with transaction costs the real value (in cash) of a share incorporates the transaction costs. Therefore the buyer does not exercise when $S_T > K$, but when $S_T(1+\theta_b) > K$. In figure 5.1 there is a comparison between the payoff in a market with zero costs (such as in the Black-Scholes model) and a market with transaction costs.

The objective of the investor is to maximize the expected utility of the wealth at terminal time T over all the admissible strategies. This expectation is conditioned on the current value of cash, number of shares and value of the stock. The sets of trading strategies corresponding to each of the three portfolios, is obtained by using the respective wealth process inside Definition [5.1.4].

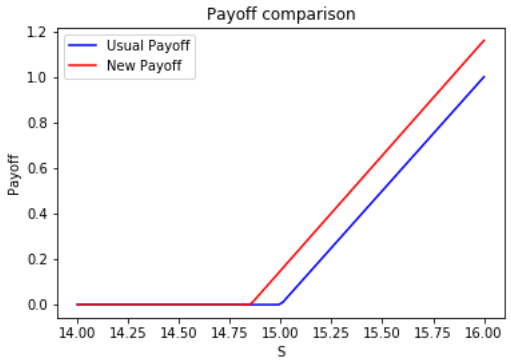

Figure 5.1: Comparison of payoffs of an European Call option with strike $K = 15$, in a market with zero costs and in a market with $\theta_b = 0.01$.

Definition 5.1.6. *The **value function** of the maximization problem for $j = w, b, 0$ is defined as:*

$$V^j(t,b,y,s) := \sup_{\pi \in \Pi^j(t,b,y,s)} \mathbb{E}_{t,b,y,s}\left[\mathcal{U}(\mathcal{W}_T^{j;\pi})\right], \tag{5.12}$$

where $\mathcal{U} : \mathbb{R} \to \mathbb{R}$ is a concave increasing utility function such that $\mathcal{U}(0) = 0$.

The writer (buyer) option price is defined as the amount of cash to add (subtract) to the bank account, such that the maximal expected utility of wealth of the writer (buyer) is the same he could get with the zero-option portfolio.

Definition 5.1.7. *The writer price is the value $p^w > 0$ such that*

$$V^0(t,b,y,s) = V^w(t,b+p^w,y,s), \tag{5.13}$$

and the buyer price is the value $p^b > 0$ such that

$$V^0(t,b,y,s) = V^b(t,b-p^b,y,s). \tag{5.14}$$

HJB Equation

According to the theory of stochastic optimal control presented in Chapter 4, we can identify the DPZ model within the class of singular control problems described in Section (4.2). Let us re-write the state equation (5.5) for $X_t^\pi = (B_t^\pi, Y_t^\pi, S_t)$ in the matrix form

$$\begin{aligned} dX_t^\pi =& d\begin{pmatrix} B_t^\pi \\ Y_t^\pi \\ S_t \end{pmatrix} \\ =& \begin{pmatrix} rB_t^\pi \\ 0 \\ \mu S_t \end{pmatrix} dt + \begin{pmatrix} 0 \\ 0 \\ \sigma S_t \end{pmatrix} dW_t \\ &+ \underbrace{\begin{pmatrix} -S_t(1+\sigma_b) & S_t(1-\sigma_s) \\ 1 & -1 \\ 0 & 0 \end{pmatrix}}_{\kappa(t,X_t)} \underbrace{\begin{pmatrix} dL_t \\ dM_t \end{pmatrix}}_{d\xi_t}, \end{aligned} \tag{5.15}$$

which can be easily identified with the general equation (4.40). The state X_t^π takes values in $\mathcal{O}$ corresponding to the solvency region $\mathcal{S}$. The value function (5.12) has a cost functional as in (4.41) with f and h equal to zero.

According to singular control theory, the HJB equation of a singular control problem is a variational inequality. Here we can obtain the HJB equation of this problem from the general equation (4.39). We omit the subscript of V^j when the discussion refers to all the three problems. Let us indicate $x = (b, y, s)$, then

- $D_x V(t,x) = \left(\frac{\partial V(t,b,y,s)}{\partial b}, \frac{\partial V(t,b,y,s)}{\partial y}, \frac{\partial V(t,b,y,s)}{\partial s} \right)$
- $H(t,x,p) = \sup_{a \in \hat{K}} \big(p \cdot \kappa(t,x) \cdot a \big)$ and define $\hat{K} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \bigcup \begin{pmatrix} 0 \\ 1 \end{pmatrix}$
- $H\big(t,x,D_xV(t,x)\big) = \max\left\{ \frac{\partial V}{\partial y} - (1+\theta_b) s \frac{\partial V}{\partial b}, -\big(\frac{\partial V}{\partial y} - (1-\theta_s) s \frac{\partial V}{\partial b} \big) \right\}$

The resulting HJB equation is

$$\begin{aligned} \max \Bigg\{ & \frac{\partial V}{\partial t} + rb \frac{\partial V}{\partial b} + \mu s \frac{\partial V}{\partial s} + \frac{1}{2} \sigma^2 s^2 \frac{\partial^2 V}{\partial s^2}, \\ & \frac{\partial V}{\partial y} - (1+\theta_b) s \frac{\partial V}{\partial b} \ , \ -\left(\frac{\partial V}{\partial y} - (1-\theta_s) s \frac{\partial V}{\partial b} \right) \Bigg\} = 0, \end{aligned} \tag{5.16}$$

for $(t,b,y,s) \in [t_0, T] \times \mathcal{S}$. The terminal conditions for the three portfolio problems $j = 0, w, b$ are:

$$V^j(T,b,y,s) = \mathcal{U}\big(w^j(b,y,s) \big) \quad \text{for} \quad (b,y,s) \in \mathcal{S}^j \tag{5.17}$$

with

$$w^0(b,y,s) = b + c(y,s).$$

$$w^w(b,y,s) = b + c(y,s) \mathbb{1}_{\{s(1+\theta_b) \le K\}} + \Big(c\big(y-1,s\big) + K \Big) \mathbb{1}_{\{s(1+\theta_b) > K\}}.$$

$$w^b(b,y,s) = b + c(y,s) \mathbb{1}_{\{s(1+\theta_b) \le K\}} + \Big(c\big(y+1,s\big) - K \Big) \mathbb{1}_{\{s(1+\theta_b) > K\}}.$$

The variational inequality (5.16) says that the maximum of three operators is equal to zero. This feature can be interpreted better if we consider the state space divided into three different regions: the **Buy**, the **Sell** and the **No Transaction** (NT) regions.

- **Buy**

$$\begin{cases} \frac{\partial V}{\partial t} + rb \frac{\partial V}{\partial b} + \mu s \frac{\partial V}{\partial s} + \frac{1}{2} \sigma^2 s^2 \frac{\partial^2 V}{\partial s^2} \le 0 \\ \frac{\partial V}{\partial y} - (1+\theta_b) s \frac{\partial V}{\partial b} = 0 \\ -\left(\frac{\partial V}{\partial y} - (1-\theta_s) s \frac{\partial V}{\partial b} \right) \le 0. \end{cases}$$

- **Sell**

$$\begin{cases} \frac{\partial V}{\partial t} + rb \frac{\partial V}{\partial b} + \mu s \frac{\partial V}{\partial s} + \frac{1}{2} \sigma^2 s^2 \frac{\partial^2 V}{\partial s^2} \le 0 \\ \frac{\partial V}{\partial y} - (1+\theta_b) s \frac{\partial V}{\partial b} \le 0 \\ -\left(\frac{\partial V}{\partial y} - (1-\theta_s) s \frac{\partial V}{\partial b} \right) = 0. \end{cases}$$

- **No Transaction**

$$\begin{cases} \frac{\partial V}{\partial t} + rb\frac{\partial V}{\partial b} + \mu s\frac{\partial V}{\partial s} + \frac{1}{2}\sigma^2 s^2 \frac{\partial^2 V}{\partial s^2} = 0 \\ \frac{\partial V}{\partial y} - (1+\theta_b)s\frac{\partial V}{\partial b} \leq 0 \\ -\left(\frac{\partial V}{\partial y} - (1-\theta_s)s\frac{\partial V}{\partial b}\right) \leq 0. \end{cases}$$

The optimization problem is a free boundary problem, and its solution consists of finding the value function V and the optimal boundaries that divide the three regions. We assume that the Buy and Sell regions are separated by the NT region. This assumption is reasonable since buying and selling a share at the same time just decreases the total wealth (because of the transaction costs) and therefore makes no sense.

The free boundaries of the NT region completely characterize the investor's trading strategy. The optimal strategy consists in keeping the portfolio process inside the NT region. If the portfolio exits the NT region, the optimal strategy is to trade in order to bring it back to the NT region.

In the Buy and Sell regions the value functions is constant along the directions of the trades. We get respectively:

$$\text{Buy} : \qquad V(t,b,y,s) = V(t, b - s(1+\theta_b)\Delta L_t, y + \Delta L_t, s).$$

$$\text{Sell} : \qquad V(t,b,y,s) = V(t, b + s(1-\theta_s)\Delta M_t, y - \Delta M_t, s).$$

where ΔL_t and ΔM_t are the number of shares respectively bought or sold in the trade.

5.1.2 DPZ with jumps

The purpose of this thesis is to extend the DPZ model in order to include jumps in the stock dynamics. Following the framework of [De Vallière et al., 2016] let us introduce a market model with proportional transaction costs that consider an exponential Lévy process for the stock dynamics, as in 1.55. The state of the portfolio at time $t \in [t_0, T]$ is (B_t^π, Y_t^π, S_t) and evolves following the SDE:

$$\begin{cases} dB_t^\pi &= rB_t^\pi dt - (1+\theta_b)S_{t^-}dL_t + (1-\theta_s)S_{t^-}dM_t \\ dY_t^\pi &= dL_t - dM_t \\ dS_t &= S_{t^-}\left(\mu dt + \sigma dW_t + \int_{\mathbb{R}}(e^z - 1)\tilde{N}(dt,dz)\right). \end{cases} \tag{5.18}$$

The strategy $\{\pi_t\}_{t\in[t_0,T]}$ is a cádlág, **predictable**, nondecreasing process with bounded variation, such that $\pi_{t_0^-} = \left(L_{t_0^-}, M_{t_0^-}\right) = (0,0)$. Under these assumptions the portfolio process $\left\{(B_t^\pi, Y_t^\pi, S_t)\right\}_{t\in[t_0,T]}$ is cádlág.

If at time t there is an unpredictable jump in the stock price $\Delta S_t = S_t - S_{t^-}$, a possible transaction should happen immediately after the jump. The control process $\{\pi_t\}_{t\in[t_0,T]}$ is assumed to be predictable, i.e. measurable with respect to the left-continuous filtration generated by $\{S_{t^-}\}_{t\in[t_0,T]}$. Therefore, a jump in the price and a jump in the control cannot occur simultaneously, almost surely. A deeper digression on this topic can be found in Section 2 of [De Vallière et al., 2016]. If the investor at time t observes a jump in the price and decides to rebalance his portfolio, he will trade at some time $u > t$ at the

price S_{u^-}. Under this framework, as explained in [De Vallière et al., 2016], the optimal strategy cannot exist.

The definitions of cash value (5.6), of total wealth (5.7), (5.11), (5.10) and the solvency regions (5.8) are kept the same. The set of admissible trading strategy will be changed.

Since the underlying stock follows a process with jumps, it is not guaranteed that the portfolio stays solvent for all $t \in [t_0, T]$. When holding short positions, it is possible that a sudden increase in the stock price cause the total wealth to jump out of the solvency region. The same can happen with a downward jump when the investor is long in stocks and negative in cash. The immediate decrease of the stock's price makes him unable to pay the debts. If the investor goes bankrupt, there are no trading strategies to save him.

Definition 5.1.8. *The first **exit time** from the solvency region is defined as:*

$$\tau := \inf\big\{t \in [t_0, T] : \mathcal{W}_t^\pi \notin \mathcal{S}\big\}. \tag{5.19}$$

Definition 5.1.9. *The set of **admissible trading strategies** $\Pi(t_0, b, y, s)$ is the set of all cádlág, nondecreasing, predictable, bounded variation processes $\{\pi_t\}_{t\in[t_0,T]}$ such that (B_t^π, Y_t^π, S_t) is a solution of (5.18) with initial values $(B_{t_0}^\pi = b, Y_{t_0}^\pi = y, S_{t_0} = s)$ and such that $\mathcal{W}_t^\pi \in \mathcal{S}$ almost surely for all $t \in [t_0, \tau)$, and $\pi_t = \pi_\tau$ for all $t \geq \tau$.*

The objective of the investor is to maximize the expected utility of the wealth at $\tau^j \wedge T$ over all the admissible strategies. This expectation is conditioned on the current value of cash, number of shares and value of the stock.

Definition 5.1.10. *The **value function** of the maximization problem for $j = w, b, 0$ is defined as:*

$$\begin{aligned} V^j(t,b,y,s) := \sup_{\pi\in\Pi^j(t,b,y,s)} \mathbb{E}_{t,b,y,s}\Big[&\mathcal{U}(\mathcal{W}_T^{j;\pi})\ \mathbb{1}_{\{\tau^j>T\}} \\ &+ e^{\beta(T-\tau^j)}\mathcal{U}(\mathcal{W}_{\tau^j}^{j;\pi})\ \mathbb{1}_{\{\tau^j\leq T\}}\Big], \end{aligned} \tag{5.20}$$

where $\mathcal{U} : \mathbb{R} \to \mathbb{R}$ is a concave increasing utility function such that $\mathcal{U}(0) = 0$, and $\beta \geq 0$.

The HJB equation associated to this stochastic control problem is obtained following the same steps we used to derive (5.16). The infinitesimal generator of the price dynamics has the form (1.75) for exponential Lévy processes. The HJB variational inequality is

$$\begin{aligned} \max\Bigg\{ &\frac{\partial V^j}{\partial t} + rb\frac{\partial V^j}{\partial b} + \mu s\frac{\partial V^j}{\partial s} + \frac{1}{2}\sigma^2 s^2\frac{\partial^2 V^j}{\partial s^2} \\ &+ \int_{\mathbb{R}}\left[V^j(t,b,y,se^z) - V^j(t,b,y,s) - s(e^z-1)\frac{\partial V^j}{\partial s}\right]\nu(dz)\ , \\ &\frac{\partial V^j}{\partial y} - (1+\theta_b)s\frac{\partial V^j}{\partial b}\ ,\ -\left(\frac{\partial V^j}{\partial y} - (1-\theta_s)s\frac{\partial V^j}{\partial b}\right)\Bigg\} = 0, \end{aligned} \tag{5.21}$$

for $(t, b, y, s) \in [t_0, T)\times\mathcal{S}^j$ and $j = 0, w, b$. The terminal boundary conditions are given by Eq. (5.17). Since this HJB equation is a PIDE, the non-local integral

operator implies to define the lateral conditions not only on the boundary of the solvency region, but also beyond:

$$V^j(t,b,y,s) = e^{\beta(T-t)}\mathcal{U}\big(b+c(y,s)\big) \quad \text{for} \quad t\in[t_0,T),\ (b,y,s)\notin\mathcal{S},\quad j=0,w,b. \tag{5.22}$$

5.1.3 Variable reduction

In the DPZ model introduced in Section 5.1.1, the portfolio is solvent for every $t\in[t_0,T]$ and it is always possible to calculate the utility of the wealth at the terminal time $\mathcal{U}(\mathcal{W}_T^\pi)$. In the extended model of Section 5.1.2 the stock process can jump, and in presence of short positions the portfolio can go bankrupt at any time before the maturity T.

With the intention of simplifying the maximization problem (5.20) and reducing the number of variables, we restrict our attention to the case of no bankruptcy. A possible idea is to consider a positive initial wealth, and define the restricted set of admissible strategies as the set of $\{\pi_t\}_{t\in[t_0,T]}$ such that $B_t^\pi \geq 0$ and $Y_t^\pi \geq 0$ for all $t\in[t_0,T]$ (see [Benth et al., 2002]). However, in order to implement a hedging strategy, we are interested in portfolios containing short positions as well. So, we can assume that the investor has a very large credit availability C in the sense that

$$\mathbb{P}(\tau > T) \underset{C\to\infty}{\approx} 1. \tag{5.23}$$

In practical terms, we ignore the possibility of default. The solvency region becomes $\mathcal{S}=\mathbb{R}^2\times\mathbb{R}^+$ and no lateral boundary conditions are imposed.

As in [Davis et al., 1993], for $\gamma>0$, we consider the exponential utility function

$$\mathcal{U}(x) := 1-e^{-\gamma x}. \tag{5.24}$$

Thanks to (5.23) and (5.24) we can remove $\{B_t^\pi\}_{t\in[t_0,T]}$ from the state dynamics. By solving (5.18) we get

$$B_T^\pi = \frac{B_t^\pi}{\delta(t,T)} - \int_t^T (1+\theta_b)\frac{S_u}{\delta(u,T)}dL_u + \int_t^T (1-\theta_s)\frac{S_u}{\delta(u,T)}dM_u \tag{5.25}$$

where $\delta(u,T)=e^{-r(T-u)}$. Using together (5.23), (5.24) and (5.25), and the wealth processes (5.7),(5.10),(5.11), we obtain for $B_t^\pi=b$, $Y_t^\pi=y$, $S_t=s$ and $j=0,w,b$:

$$V^j(t,b,y,s) = \sup_\pi\ \mathbb{E}_{t,b,y,s}\left[1-e^{-\gamma\mathcal{W}^j(T)}\right] = 1-e^{-\gamma\frac{b}{\delta(t,T)}}Q^j(t,y,s), \tag{5.26}$$

where

$$Q^j(t,y,s) = \inf_\pi\ \mathbb{E}_{t,y,s}\Bigg[\, e^{-\gamma\left[-\int_t^T(1+\theta_b)\frac{S_u}{\delta(u,T)}dL_u+\int_t^T(1-\theta_s)\frac{S_u}{\delta(u,T)}dM_u\right]} \times H^j(Y_T^\pi,S_T)\Bigg] \tag{5.27}$$

is our new minimization problem. The exponential term inside the expectation can be considered as a discount factor, and the second term $H^j(y,s) = Q^j(T,y,s)$ is the terminal payoff:

- No option:
$$H^0(y,s) = e^{-\gamma\, c(y,s)}. \tag{5.28}$$

- Writer:
$$H^w(y,s) = e^{-\gamma\left[c(y,s)\mathbb{1}_{\{s(1+\theta_b)\le K\}} + \left(c(y-1,s)+K\right)\mathbb{1}_{\{s(1+\theta_b)>K\}}\right]}. \tag{5.29}$$

- Buyer:
$$H^b(y,s) = e^{-\gamma\left[c(y,s)\mathbb{1}_{\{s(1+\theta_b)\le K\}} + \left(c(y+1,s)-K\right)\mathbb{1}_{\{s(1+\theta_b)>K\}}\right]}. \tag{5.30}$$

Using conditions (5.13), (5.14) together with (5.26), we obtain the explicit formulas for the option prices:

$$p^w(t_0,y,s) = \frac{\delta(t_0,T)}{\gamma}\log\left(\frac{Q^w(t_0,y,s)}{Q^0(t_0,y,s)}\right), \tag{5.31}$$

$$p^b(t_0,y,s) = \frac{\delta(t_0,T)}{\gamma}\log\left(\frac{Q^0(t_0,y,s)}{Q^b(t_0,y,s)}\right). \tag{5.32}$$

Since $Q^j(t,y,s)$ is independent on b, let us write $Q^j(t,y,s) := 1 - V^j(t,0,y,s)$. It is convenient to pass to the log-variable $x = \log(s)$, such that

$$s\frac{\partial}{\partial s} = \frac{\partial}{\partial x}, \qquad s^2\frac{\partial^2}{\partial s^2} = \frac{\partial^2}{\partial x^2} - \frac{\partial}{\partial x}. \tag{5.33}$$

For $j = 0, w, b$, the HJB Eq. (5.21) becomes:

$$\begin{aligned} \min\Bigg\{ & \frac{\partial Q^j}{\partial t} + (\mu - \frac{1}{2}\sigma^2)\frac{\partial Q^j}{\partial x} + \frac{1}{2}\sigma^2\frac{\partial^2 Q^j}{\partial x^2} \\ & + \int_{\mathbb{R}}\left[Q^j(t,y,x+z) - Q^j(t,y,x) - (e^z-1)\frac{\partial Q^j}{\partial x}\right]\nu(dz)\ , \\ & \frac{\partial Q^j}{\partial y} + (1+\theta_b)e^x\frac{\gamma}{\delta(t,T)}Q^j\ ,\ -\left(\frac{\partial Q^j}{\partial y} + (1-\theta_s)e^x\frac{\gamma}{\delta(t,T)}Q^j\right)\Bigg\} = 0. \end{aligned} \tag{5.34}$$

5.2 Existence of viscosity solution

The general fact that value functions of control problems can be characterized as viscosity solutions of certain partial differential equations is a direct consequence of the dynamic programming principle. For singular control problems, however, the classical approach of [Lions, 1983] fails because the state process may jump due to the singular control and it needs thus not stay in a small ball[1] for a small Δt. This problem can be circumvented by relying on the existence of the optimal control, as done in [Davis et al., 1993]. However, in our proof we will not assume the existence of the optimal control, which in this framework does not exist. We will assume that the **DPP** for the singular control problem (5.18), (5.20) holds, and that the value function is continuous. We refer to sections 4 and 9 of [De Vallière et al., 2016] for the proofs of these statements.

[1] For processes of type (4.40) it is not possible to apply the theorem (4.1.3).

In this section we prove that the value function (5.20), can be interpreted as the viscosity solution of the HJB equation (5.21). The proof follows the approach in [Framstad et al., 1999].

Let us call $\tau_{\mathcal{S}}$ the first exit time from $\mathcal{S}$. Since the stock dynamics is stochastically continuous and (by definition) the control process π cannot be the cause of bankruptcy, we can exclude an immediate jump out of the solvency region, i.e. $t_0 \neq \tau_{\mathcal{S}}$.

We have to interpret the HJB equation (5.21) with boundary conditions (5.22) and (5.17) as a parabolic problem of the type (4.45). We use the dummy variable $x = (b, y, s)$ to indicate a point in $\mathbb{R}^2 \times \mathbb{R}^+$. In order to satisfy the parabolic conditions we reformulate the HJB as

$$\begin{aligned} \min \Bigg\{ -\Bigg(& \frac{\partial V}{\partial t} + rb\frac{\partial V}{\partial b} + \mu s\frac{\partial V}{\partial s} + \frac{1}{2}\sigma^2 s^2 \frac{\partial^2 V}{\partial s^2} \\ & + \int_{\mathbb{R}} \left[V(t,b,y,se^z) - V(t,b,y,s) - s(e^z-1)\frac{\partial V}{\partial s} \right] \nu(dz) \Bigg), \\ & - \left(\frac{\partial V}{\partial y} - (1+\theta_b)s\frac{\partial V}{\partial b} \right), + \left(\frac{\partial V}{\partial y} - (1-\theta_s)s\frac{\partial V}{\partial b} \right) \Bigg\} = 0. \end{aligned} \tag{5.35}$$

We indicate this function with $F(t, x, V, D_t V, D_x V, D_{xx} V, \mathcal{I}(t, x, V)) = 0$ in the domain $(t, x) \in [t_0, T) \times \mathcal{S}$. To simplify the notation, let us introduce the integro-differential infinitesimal generator:

$$\begin{aligned} \mathcal{L}V(t,x) := & -\left(\frac{\partial V}{\partial t}(t,x) + rb\frac{\partial V}{\partial b}(t,x) + \mu s\frac{\partial V}{\partial s}(t,x) + \frac{1}{2}\sigma^2 s^2 \frac{\partial^2 V}{\partial s^2}(t,x) \right) \\ & - \int_{\mathbb{R}} \left[V(t,b,y,se^z) - V(t,b,y,s) - s(e^z-1)\frac{\partial V}{\partial s} \right] \nu(dz), \end{aligned}$$

such that the equation (5.35) has the form

$$\min \left\{ \mathcal{L}V(t,x),\ -(V_y - (1+\theta_b)sV_b),\ V_y - (1-\theta_s)sV_b \right\} = 0. \tag{5.36}$$

In order to prove that $V(t, x)$ is a viscosity solution we need to verify both the subsolution and supersolution properties.

5.2.1 Subsolution

In this section we prove that the value function can be interpreted as a viscosity subsolution of the HJB equation associated to the maximization problem. We enunciate the following theorem:

Theorem 5.2.1. *The value function of the maximization problem (5.20), is a viscosity subsolution of the Eq. (5.36).*

Proof. Let us consider a test function $\phi \in C^2([t_0, T] \times \mathbb{R}^2 \times \mathbb{R}^+) \bigcap \mathcal{C}_2([t_0, T] \times \mathbb{R}^2 \times \mathbb{R}^+)$ such that $(\bar{t}, \bar{x}) \in [t_0, T] \times \mathcal{S}$ is a maximum point for $V - \phi$:

$$V(\bar{t}, \bar{x}) - \phi(\bar{t}, \bar{x}) \geq V(t, x) - \phi(t, x) \qquad \forall (t, x) \in [t_0, T] \times \mathcal{S} \tag{5.37}$$

and we assume without loss of generality that:

$$V(\bar{t}, \bar{x}) = \phi(\bar{t}, \bar{x}) \tag{5.38}$$

so we can write:

$$V(t,x) \leq \phi(t,x) \tag{5.39}$$

We want to prove that

$$F(\bar{t},\bar{x},V(\bar{t},\bar{x}),D_t\phi(\bar{t},\bar{x}),D_x\phi(\bar{t},\bar{x}),D_{xx}\phi(\bar{t},\bar{x}),\mathcal{I}(\bar{t},\bar{x},\phi(\bar{t},\bar{x}))) \leq 0$$

Reductio ad absurdum:
Let us assume that

$$F(\bar{t},\bar{x},V(\bar{t},\bar{x}),D_t\phi(\bar{t},\bar{x}),D_x\phi(\bar{t},\bar{x}),D_{xx}\phi(\bar{t},\bar{x}),\mathcal{I}(\bar{t},\bar{x},\phi(\bar{t},\bar{x}))) > 0.$$

This means that all the terms inside the minimum in Eq. (5.36) are positive, i.e. exist $\kappa > 0$ such that:

1. $\mathcal{L}\phi(\bar{t},\bar{x}) > \kappa$,
2. $-\left(\frac{\partial\phi}{\partial y}(\bar{t},\bar{x}) - (1+\theta_b)\bar{s}\frac{\partial\phi}{\partial b}(\bar{t},\bar{x})\right) > \kappa$,
3. $\frac{\partial\phi}{\partial y}(\bar{t},\bar{x}) - (1-\theta_s)\bar{s}\frac{\partial\phi}{\partial b}(\bar{t},\bar{x}) > \kappa$.

Since the test function is smooth by definition, there exist a ball $\mathcal{B}(\bar{x},\rho) \subset \mathcal{S}$ where the three previous conditions are satisfied $\forall x \in \mathcal{B}(\bar{x},\rho)$. We can define the exit time of the process $\{X_t\}_{t\in[\bar{t},T\wedge\tau_{\mathcal{S}}]}$ from the ball $\mathcal{B}(\bar{x},\rho)$:

$$\tau_\rho = \inf\{t \in [t_0,T] : X_t \notin \mathcal{B}(\bar{x},\rho)\} \wedge N, \tag{5.40}$$

with $N > 0$ fixed. Since $\mathcal{B}(\bar{x},\rho) \subset \mathcal{S}$ it follows that $\tau_\rho \leq \tau_{\mathcal{S}}$.

In this proof, we will take advantage from the fact that transaction costs are linear. This means that any transaction can be split into two or more simultaneous smaller transactions. Therefore, the jump $\Delta\pi_t > 0$ of the control can be divided into smaller jumps i.e. $\Delta\pi_t = \sum_i c_i\Delta\pi_t$, with $c_i \geq 0$ and $\sum_i c_i = 1$. The coefficients c_i can be chosen such that after the action of each $c_i\Delta\pi_t$ the state process is contained in an arbitrarily small closed ball.

Let us consider an admissible ϵ-optimal control $\{\pi_t^\epsilon\}_{t_0\leq t\leq T}$. To lighten the notation, in the following we suppress the apex ϵ. At the point $(\bar{t},\bar{x})$, for a $\rho > 0$ there are two alternative cases:

A: $\Delta\pi_t = 0$ for $\bar{t} \leq t \leq \tau_\rho$.

B: $\Delta\pi_{t_k} > 0$ for some $\bar{t} \leq t_k \leq \tau_\rho$.
In this case we select a small enough $\rho > 0$ such that $\tau_\rho = \bar{t}$.

The action of the control can be split in the following way. For a control jump at time $\bar{t}$, let us introduce

$$x_0 = (b_0,y_0,s_0) = (\bar{b},\bar{y},\bar{s}) = \bar{x}$$
$$x_{i+1} = x_i + \left(-(1+\theta_b)\bar{s}\,c_i\Delta L_{\bar{t}} + (1-\theta_s)\bar{s}\,c_i\Delta M_{\bar{t}}, c_i\Delta L_{\bar{t}} - c_i\Delta M_{\bar{t}}, 0\right).$$

For each x_i, we choose a ray $\rho_i > 0$. We can now define the coefficients c_i as:

$$c_i = \inf\Bigg\{c \in [0,1] : \tag{5.41}$$
$$x_i + \left(-(1+\theta_b)\bar{s}\,c\Delta L_{\bar{t}} + (1-\theta_s)\bar{s}\,c\Delta M_{\bar{t}}, c\Delta L_{\bar{t}} - c\Delta M_{\bar{t}}, 0\right) \notin \mathcal{B}(x_i,\rho_i)\Bigg\}. \tag{5.42}$$

By the DPP (4.44), for every $\epsilon > 0$, there exists an an ϵ-optimal control $\pi \in \Pi(\bar{t}, \bar{x})$ such that

$$V(\bar{t}, \bar{x}) \leq \mathbb{E}_{\bar{x}}\Big[V(\tau_\rho, X^\pi_{\tau_\rho})\Big] + \epsilon. \tag{5.43}$$

Using (5.38) and (5.39) we can write:

$$\phi(\bar{t}, \bar{x}) = V(\bar{t}, \bar{x}) \ \leq \mathbb{E}_{\bar{x}}[V(\tau_\rho, X^\pi_{\tau_\rho})] + \epsilon \ \leq \mathbb{E}_{\bar{x}}[\phi(\tau_\rho, X^\pi_{\tau_\rho})] + \epsilon.$$

In the following, in order to get a more readable notation, when necessary we indicate X_t with $X(t)$. The test function ϕ is smooth enough to use the generalized Itō formula.

$$\begin{aligned}
0 \ &\leq \mathbb{E}_{\bar{x}}\Big[\phi(\tau_\rho, X^\pi_{\tau_\rho}) - \phi(\bar{t}, \bar{x})\Big] + \epsilon \qquad (5.44)\\
&= \mathbb{E}_{\bar{x}}\Bigg[\int_{\bar{t}}^{\tau_\rho} -\mathcal{L}\phi(t, X^\pi_t)dt \\
&\quad + \int_{\bar{t}}^{\tau_\rho} \left(\frac{\partial \phi}{\partial y}(t, X^\pi_t) - (1+\theta_b)S_{t^-}\frac{\partial \phi}{\partial b}(t, X^\pi_t)\right) dL^c_t \\
&\quad - \int_{\bar{t}}^{\tau_\rho} \left(\frac{\partial \phi}{\partial y}(t, X^\pi_t) - (1-\theta_s)S_{t^-}\frac{\partial \phi}{\partial b}(t, X^\pi_t)\right) dM^c_t \\
&\quad + \sum_{\bar{t} \leq t_k \leq \tau_\rho} \Big(\phi\big(t_k, B^\pi(t_k^-) - (1+\theta_b)S(t_k^-)\,\Delta L_{t_k}, Y^\pi(t_k^-) + \Delta L_{t_k}, S(t_k)\big) \\
&\qquad\qquad - \phi\big(t_k, B^\pi(t_k^-), Y^\pi(t_k^-), S(t_k)\big)\Big) \\
&\quad + \sum_{\bar{t} \leq t_k \leq \tau_\rho} \Big(\phi\big(t_k, B^\pi(t_k^-) + (1-\theta_s)S(t_k^-)\,\Delta M_{t_k}, Y^\pi(t_k^-) - \Delta M_{t_k}, S(t_k)\big) \\
&\qquad\qquad - \phi\big(t_k, B^\pi(t_k^-), Y^\pi(t_k^-), S(t_k)\big)\Big)\Bigg] + \epsilon
\end{aligned}$$

where again t_k denotes the jump times of $\{\pi_t\}_{t \in [\bar{t}, \tau_\rho]}$ and we indicate the continuous parts of π as

$$L^c_t = L_t - \sum_{\bar{t} \leq t_k \leq \tau_\rho} \Delta L_{t_k} \quad \text{and} \quad M^c_t = M_t - \sum_{\bar{t} \leq t_k \leq \tau_\rho} \Delta M_{t_k}. \tag{5.45}$$

Case A: $\Delta \pi_t = 0$ for each $\bar{t} \leq t \leq \tau_\rho$. Equation (5.44) becomes:

$$\begin{aligned}
0 \ &\leq E_{\bar{x}}\Bigg[\int_{\bar{t}}^{\tau_\rho} -\mathcal{L}\phi(t, X^\pi_t)dt \qquad (5.46)\\
&\quad + \int_{\bar{t}}^{\tau_\rho} \left(\frac{\partial \phi}{\partial y}(t, X^\pi_t) - (1+\theta_b)S_{t^-}\frac{\partial \phi}{\partial b}(t, X^\pi_t)\right) dL^c_t \\
&\quad - \int_{\bar{t}}^{\tau_\rho} \left(\frac{\partial \phi}{\partial y}(t, X^\pi_t) - (1-\theta_s)S_{t^-}\frac{\partial \phi}{\partial b}(t, X^\pi_t)\right) dM^c_t\Bigg] + \epsilon \\
&< \underbrace{\left(-\kappa\,\mathbb{E}_{\bar{x}}\big[\tau_\rho - \bar{t}\big] - \kappa\,\mathbb{E}_{\bar{x}}\big[L^c_{\tau_\rho} - L^c_{\bar{t}}\big] - \kappa\,\mathbb{E}_{\bar{x}}\big[M^c_{\tau_\rho} - M^c_{\bar{t}}\big]\right)}_{<0} + \epsilon.
\end{aligned}$$

Since this must hold true for all ϵ, when taking the limit $\epsilon \to 0$ we get a contradiction.

Case B: In this case we have $\tau_\rho = \bar{t}$.
Let us consider for simplicity $\Delta\pi_{\bar{t}} = (\Delta L_{\bar{t}}, 0)$. The case with $\Delta\pi_{\bar{t}} = (0, \Delta M_{\bar{t}})$ is analogous.

We can split the trade and consider only the first action $c_0\Delta L_{\bar{t}}$ of the control, where c_0 is defined in (5.41) with $\rho_0 = \rho$. After the control jump the state moves from x_0 to $x_1 \in \partial\mathcal{B}(x_0, \rho)$.
Now we want to investigate the relation between ρ and ϵ. Remember that ρ should be small enough in order to guarantee the immediate exit from the ball i.e. $\bar{t} = \tau_\rho$. The size of the initial jump $\Delta L_{\bar{t}}$ however, depends on ϵ. In general there are two possibilities:

1. $\Delta L_{\bar{t}} \geq \epsilon$.
   If at the point $(\bar{t}, \bar{x})$ the size of the control jump $\Delta L_{\bar{t}}$ is bigger or equal than ϵ, we can choose c_0 such that
   $$c_0\Delta L_{\bar{t}} = \epsilon. \tag{5.47}$$

2. $\Delta L_{\bar{t}} < \epsilon$.
   If $0 < \Delta L_{\bar{t}} < \epsilon$ holds true for every $\epsilon > 0$, then for $\epsilon \to 0$ it follows that $\Delta L_{\bar{t}} \to 0$, leading to the previous **case A**.

Let us consider the case of a control with a jump bigger than ϵ i.e. $\Delta L_{\bar{t}} \geq \epsilon$. Thanks to the relation (5.47), when we take the limit $\epsilon \to 0$ then $c_0 \to 0$ as well, and consequently also the distance between x_0 and x_1 i.e. $\rho \to 0$.
It is convenient to rescale the parameter ϵ:
$$c_0\Delta L_{\bar{t}} = \epsilon = 2\tilde{\epsilon}/\kappa. \tag{5.48}$$

Using the DPP between the two states x_0 and x_1, the equation (5.44) becomes:

$$\begin{aligned}
0 \ &\leq \phi(\bar{t}, x_1) - \phi(\bar{t}, x_0) + \tilde{\epsilon} \\
&= \frac{\phi\big(\bar{t},\ \bar{b} - (1+\theta_b)\bar{s}\, c_0\Delta L_{\bar{t}},\ \bar{y} + c_0\Delta L_{\bar{t}},\ \bar{s}\big) - \phi(\bar{t}, \bar{x})}{c_0\Delta L_{\bar{t}}} + \frac{\tilde{\epsilon}}{c_0\Delta L_{\bar{t}}} \\
&\underset{\epsilon\to 0}{=} \frac{\phi\big(\bar{t},\ \bar{b} - (1+\theta_b)\bar{s}\,\epsilon,\ \bar{y} + \epsilon,\ \bar{s}\big) - \phi(\bar{t}, \bar{x})}{\epsilon} + \frac{\kappa}{2} \\
&= \left(\frac{\partial\phi}{\partial y}(\bar{t}, \bar{x}) - (1+\theta_b)\bar{s}\frac{\partial\phi}{\partial b}(\bar{t}, \bar{x})\right) + \frac{\kappa}{2} \\
&< -\kappa + \frac{\kappa}{2}
\end{aligned}$$

In the second line we expressed x_1 in explicit form, and divided both sides by $c_0\Delta L_{\bar{t}}$. In the third line we used (5.48) and took the limit for $\epsilon \to 0$.
We obtained a contradiction. □

5.2.2 Supersolution

Let us prove that the value function is a viscosity supersolution of the HJB equation (5.36).

Theorem 5.2.2. *The value function of the maximization problem (5.20), is a viscosity supersolution of the Eq. (5.36).*

Proof. Let us consider a test function $\phi \in C^2([t_0,T]\times\mathbb{R}^2\times\mathbb{R}^+)\bigcap \mathcal{C}_2([t_0,T]\times\mathbb{R}^2\times\mathbb{R}^+)$ such that $(\bar{t},\bar{x})\in[t_0,T]\times\mathcal{S}$ is a minimum point for $V-\phi$.

$$V(\bar{t},\bar{x})-\phi(\bar{t},\bar{x})\leq V(t,x)-\phi(t,x)\qquad \forall(t,x)\in[t_0,T]\times\mathcal{S} \tag{5.49}$$

and we assume without loss of generality that:

$$V(\bar{t},\bar{x})=\phi(\bar{t},\bar{x}) \tag{5.50}$$

so we can write:

$$V(t,x)\geq\phi(t,x) \tag{5.51}$$

We want to prove that

$$F(\bar{t},\bar{x},V(\bar{t},\bar{x}),D_t\phi(\bar{t},\bar{x}),D_x\phi(\bar{t},\bar{x}),D_{xx}\phi(\bar{t},\bar{x}),\mathcal{I}(\bar{x},\phi(\bar{t},\bar{x})))\geq 0$$

that is analogous to prove that all the following terms are non-negative:

1. $\mathcal{L}\phi(\bar{t},\bar{x})\geq 0$,
2. $-\left(\frac{\partial\phi}{\partial y}(\bar{t},\bar{x})-(1+\theta_b)\bar{s}\frac{\partial\phi}{\partial b}(\bar{t},\bar{x})\right)\geq 0$,
3. $\frac{\partial\phi}{\partial y}(\bar{t},\bar{x})-(1-\theta_s)\bar{s}\frac{\partial\phi}{\partial b}(\bar{t},\bar{x})\geq 0$.

By the DPP (4.43), for all $\tau\in\mathcal{T}_{\bar{t},\tau_{\mathcal{S}}}$ and **for all** $\pi\in\Pi(\bar{t},\bar{x})$, we can write:

$$V(\bar{t},\bar{x})\geq\mathbb{E}_{\bar{x}}\Big[V(\tau,X_\tau^\pi)\Big]. \tag{5.52}$$

Using (5.50) and (5.51) we can write:

$$\phi(\bar{t},\bar{x})=V(\bar{t},\bar{x})\ \geq\mathbb{E}_{\bar{x}}[V(\tau,X_\tau^\pi)]\ \geq\mathbb{E}_{\bar{x}}[\phi(\tau,X_\tau^\pi)]$$

The test function ϕ is smooth enough to use the generalized Itō formula.

$$\begin{aligned}
0\ &\geq\mathbb{E}_{\bar{x}}\left[\phi(\tau,X_\tau^\pi)-\phi(\bar{t},\bar{x})\right] &(5.53)\\
&=\mathbb{E}_{\bar{x}}\Bigg[\int_{\bar{t}}^{\tau}-\mathcal{L}\phi(t,X_t^\pi)dt\\
&\quad+\int_{\bar{t}}^{\tau}\left(\frac{\partial\phi}{\partial y}(t,X_t^\pi)-(1+\theta_b)S_{t^-}\frac{\partial\phi}{\partial b}(t,X_t^\pi)\right)dL_t^c\\
&\quad-\int_{\bar{t}}^{\tau}\left(\frac{\partial\phi}{\partial y}(t,X_t^\pi)-(1-\theta_s)S_{t^-}\frac{\partial\phi}{\partial b}(t,X_t^\pi)\right)dM_t^c\\
&\quad+\sum_{\bar{t}\leq t_k\leq\tau}\Big(\phi\big(t_k,B^\pi(t_k^-)-(1+\theta_b)S(t_k^-)\,\Delta L_{t_k},Y^\pi(t_k^-)+\Delta L_{t_k},S(t_k)\big)\\
&\qquad\qquad-\phi\big(t_k,B^\pi(t_k^-),Y^\pi(t_k^-),S(t_k)\big)\Big)\\
&\quad+\sum_{\bar{t}\leq t_k\leq\tau}\Big(\phi\big(t_k,B^\pi(t_k^-)+(1-\theta_s)S(t_k^-)\,\Delta M_{t_k},Y^\pi(t_k^-)-\Delta M_{t_k},S(t_k)\big)\\
&\qquad\qquad-\phi\big(t_k,B^\pi(t_k^-),Y^\pi(t_k^-),S(t_k)\big)\Big)\Bigg] &(5.54)
\end{aligned}$$

where, as in (5.44) the t_k denote the jump times of $\{\pi_t\}_{t\in[\bar{t},\tau]}$ and L^c, M^c indicate the continuous parts of π.

Now we can consider the two cases:

A There is a jump $\Delta\pi_{\bar{t}} > 0$ at the initial time $\bar{t}$, and then the process is continuous i.e. $\Delta\pi_t = 0$ for $t \in (\bar{t}, \tau]$. We consider the two cases $\Delta L_{\bar{t}} = l > 0$ and $\Delta M_{\bar{t}} = 0$ or $\Delta L_{\bar{t}} = 0$ and $\Delta M_{\bar{t}} = m > 0$.

B The control π_t is constant for all $t \in [\bar{t}, \tau]$.

Case A: taking the limit $\tau \to \bar{t}$ and writing $\bar{x} = (\bar{b}, \bar{y}, \bar{s})$ we obtain the following expression

$$0 \geq \phi\big(\bar{t}, \bar{b} - (1+\theta_b)\bar{s}\, l, \bar{y} + l, \bar{s}\big) - \phi(\bar{t}, \bar{b}, \bar{y}, \bar{s}),$$

and all the other terms go to zero.
The previous equation holds true for all $l > 0$, so we can take the limit $l \to 0$ and get

$$\frac{\partial\phi}{\partial y}(\bar{t}, \bar{x}) - (1+\theta_b)\bar{s}\frac{\partial\phi}{\partial b}(\bar{t}, \bar{x}) \leq 0,$$

and after multiplying by -1, we verify the second inequality.

$$-\left(\frac{\partial\phi}{\partial y}(\bar{t}, \bar{x}) - (1+\theta_b)\bar{s}\frac{\partial\phi}{\partial b}(\bar{t}, \bar{x})\right) \geq 0.$$

The third inequality can be obtained following the same steps but choosing $\Delta M_{\bar{t}} = m > 0$ and $\Delta L_{\bar{t}} = 0$.

$$\frac{\partial\phi}{\partial y}(\bar{t}, \bar{x}) - (1-\theta_s)\bar{s}\frac{\partial\phi}{\partial b}(\bar{t}, \bar{x}) \geq 0.$$

Case B: for a constant control the only surviving term in (5.44) is

$$\int_{\bar{t}}^{\tau} -\mathcal{L}\phi(t, X_t^\pi)dt \leq 0.$$

Let us divide by $\tau - \bar{t}$, send $\tau \to \bar{t}$ and use the mean value theorem to get:

$$\mathcal{L}\phi(\bar{t}, \bar{x}) \geq 0.$$

This proves the first inequality. □

We conclude this chapter with the following main theorem:

Theorem 5.2.3. *The value function of the maximization problem (5.20), is a viscosity solution of the Eq. (5.36).*

The proof is a direct consequence of theorems 5.2.1 and 5.2.2. Since the value function is both a viscosity subsolution and a supersolution, it is a viscosity solution.

5.3 Chapter conclusions

This chapter contains the main topic of the thesis i.e. a model for pricing options under proportional transaction costs and when the stock price follows an exponential Lévy process.

This model is an extension of the celebrated model of [Davis et al., 1993], that we recall in section 5.1.1. This model can be identified in the category of singular stochastic problems presented in chapter 4.

Following the framework of [De Vallière et al., 2016], we extend the DPZ model by replacing the geometric Brownian motion with an exponential Lévy process. Following the general singular control theory, we derive the associated HJB variational inequality. The complete equation (5.21) and the reduced equation (5.34) are the main equations of this thesis and are also presented in the paper [Cantarutti et al., 2018].

The end of the chapter is dedicated to the proof of existence of a viscosity solution of the HJB (5.21). Specifically, we prove that the value function of the problem satisfies the HJB in the viscosity sense.

6

Numerical Methods

This chapter presents the numerical methods for solving the option pricing problems with transaction costs presented in section (5.1.2). The optimization problems (5.20) and (5.27) are solved by using the Markov chain approximation method. The same approach has been used frequently in the literature in the case of diffusion processes, (see for instance [Hodges and Neuberger, 1989], [Davis et al., 1993], [Damgaard, 1998], [Monoyios, 2004] and [Palczewsky et al., 2015]). We present results for the particular cases of diffusion process, Merton jump-diffusion and Variance Gamma process, although our scheme works for any Lévy process with finite variance. We show that our numerical scheme, applied to the problem (5.27), is monotone, consistent and stable, and that its solution

converges to the viscosity solution of the HJB eq. (5.34). Further analysis, such as numerical convergence rate and time complexity of the algorithm are presented. In the section 6.4 are presented numerical results obtained for the general problem (5.20) where also the default feature is considered. In the final section 6.5 we apply the multinomial method introduced in chapter 3 to the problem (5.27).

6.1 Markov chain approximation

To solve the problems (5.20) and (5.27) we use the Markov chain approximation method developed by [Kushner and Dupuis, 2001]. The numerical technique for singular controls has been originally developed in the article of [Kushner and Martins, 1991]. The portfolio dynamics (5.18) is approximated by a discrete state controlled Markov chain in discrete time. The method consists in creating a backward recursive dynamic programming algorithm, in order to compute the value function at time t, given its value at time $t + \Delta t$. [Kushner and Dupuis, 2001] prove that the value function obtained through the discrete dynamic programming algorithm converges to the value function of the original continuous time problem as $\Delta t \to 0$. Their proof uses a weak convergence in probability argument. Another approach to prove convergence has been introduced by [Barles and Souganidis, 1991]. It consider the convergence of the discrete value function to the viscosity solution of the original HJB equation. In the work of [Davis et al., 1993] the authors prove existence and uniqueness of the viscosity solution of the HJB Eq. (5.16) (diffusion case), and using the method developed by [Barles and Souganidis, 1991] prove that the value function obtained through the Markov chain approximation converges to it.

In this chapter we propose a discretization scheme and prove that it is monotone, consistent, stable, and its solution converges to the continuous viscosity solution of (5.34).

In this work we model the stock dynamics with a general exponential Lévy process. For practical computations we need to specify which Lévy process we are using, and this is equivalent to choose a Lévy triplet. Since every Lévy process satisfies the Markov property, we are allowed to use the Markov chain approximation approach. A possible way to construct the Markov chain is to discretize the infinitesimal generator by using an explicit finite difference method (see for instance [Kushner and Dupuis, 2001] or [Fleming and Soner, 2005]). This is straightforward for Lévy processes of jump-diffusion type with finite jump activity. But for Lévy processes with infinite jump activity, it is not straightforward to obtain the transition probabilities from the discretization of the generator. A common procedure is to approximate the small jumps with a Brownian motion, as explained in Section 2.2.3, in order to remove the singularity of the Lévy measure near the origin.
In the next section we focus our attention on the simplified problem (5.27).

6.1.1 The discrete model

Thanks to the variable reduction introduced in the previous section, the optimization problem (5.27) only depends on two state variables. The portfolio

dynamics (5.18) has the simpler form (using $X_t = \log S_t$):

$$\begin{cases} dY_t^\pi &= dL_t - dM_t \\ dX_t &= \Big(\mu - \frac{1}{2}\sigma^2 - \int_{\mathbb{R}}(e^z - 1 - z)\nu(dz)\Big)dt + \sigma dW_t + \int_{\mathbb{R}} z\tilde{N}(dt, dz). \end{cases} \tag{6.1}$$

where the SDE for the log-variable corresponds to (2.22) with μ defined in (1.54). If the process has finite activity $\lambda := \int_{\mathbb{R}} \nu(dz)$, thanks to assumption **EM** (in Section 1.1.4), we can define with an abuse of notation[1] $m := \int_{\mathbb{R}}\big(e^z - 1\big)\nu(dz)$ and $\lambda\alpha := \int_{\mathbb{R}} z\nu(dz)$ such that the SDE of $\{X_t\}_{t\in[t_0,T]}$ can be written as

$$dX_t = \left(\mu - \frac{1}{2}\sigma^2 - m + \lambda\alpha\right)dt + \sigma dW_t + \int_{\mathbb{R}} z\tilde{N}(dt, dz). \tag{6.2}$$

If the process has infinite activity $\lambda = \int_{\mathbb{R}} \nu(dz) = \infty$, we can approximate the "small jumps" martingale component by a Brownian motion, following the arguments in section (2.2.3), and get the equation

$$dX_t = \left(\mu - \frac{1}{2}(\sigma^2 + \sigma_\epsilon^2) - \omega_\epsilon + \lambda_\epsilon\theta_\epsilon\right)dt + \big(\sigma + \sigma_\epsilon\big)dW_t + \int_{|z|\geq\epsilon} z\tilde{N}(dt, dz), \tag{6.3}$$

with parameters defined in (2.39).

Now we can discretize the time and space to create a Markov chain approximation of the portfolio process (6.1). For $n = 0, 1, ...N \in \mathbb{N}$, we define the discrete time step $\Delta t := \frac{T-t_0}{N}$ such that $t_n = t_0 + n\Delta t$. We assume that the controls $\big(L_u, M_u\big)$ are constant for $u \in [t_n, t_{n+1})$, and allow for a possible variation at t_n for each n.

From now on, we indicate with X_n the value of X_t at t_n and with Y_n the value of Y_t at the time t_n^- immediately before the possible transaction.

Let us define the set $\Sigma_x := \{-K_1 h_x, ..., -h_x, 0, h_x, ..., +K_2 h_x\}$, where $h_x > 0$ is the discrete log-return step. The values $K_1, K_2 \in \mathbb{N}$ can be different to capture the possible asymmetry in the jump sizes. Its dimension is $\bar{L} = \#(\Sigma_x) = K_1 + K_2 + 1$. Let us define also the set $\Sigma_y := \{-K_3 h_y, ..., -h_y, 0, h_y, ..., +K_4 h_y\}$, where $h_y > 0$ is the discrete shares step and $K_3, K_4 \in \mathbb{N}$. Its dimension is $\bar{M} = \#(\Sigma_y) = K_3 + K_4 + 1$.

The discretized version of the SDE (6.1) is:

$$\begin{cases} \Delta Y_n &= \Delta L_n - \Delta M_n \\ \Delta X_n &= \hat{\mu}\Delta t + \hat{\sigma}\Delta W_n + \Delta\tilde{J}_n \;=\; \Delta\Xi_n + \Delta\tilde{J}_n, \end{cases} \tag{6.4}$$

where $\Delta X_n := X_{n+1} - X_n \in \Sigma_x$, $\hat{\mu} \in \mathbb{R}$ and $\hat{\sigma} > 0$. The term $\Delta\Xi_n := \hat{\mu}\Delta t + \hat{\sigma}\Delta W_n$ takes values in $\{-h_x, 0, h_x\}$[2] and satisfies $\mathbb{E}\big[\Delta\Xi_n\big] = \hat{\mu}\Delta t$ and $\mathbb{E}\big[(\Delta\Xi_n)^2\big] = \hat{\sigma}\Delta t$, at first order in Δt. The term $\Delta\tilde{J}_n$ is the discrete version of the compensated Poisson jump term, and satisfies $\mathbb{E}\big[\Delta\tilde{J}_n\big] = 0$ and $\mathbb{E}\big[(\Delta\tilde{J}_n)^2\big] = \tilde{\sigma}\Delta t$, at first order in Δt, with $\tilde{\sigma} > 0$. When the continuous time jump term is $\int_{\mathbb{R}} z\tilde{N}(dt, dz)$, the corresponding discrete version $\Delta\tilde{J}_n$ can assume all the values

[1] We have already defined m in (1.77) and α in (1.56) for the Merton model. We extend this notation for any Lévy measure with finite activity.

[2] A common alternative is to consider a binomial discretization with $\Delta\Xi \in \{-h_x, h_x\}$, as in [Davis et al., 1993].

in Σ_x. If instead the integral has a truncation term ϵ, i.e. $\int_{|z|\geq\epsilon} z\tilde{N}(dt,dz)$, we can define the subset $\Sigma_x^\epsilon := \Sigma_x \setminus \{-h_x, 0, h_x\}$, such that $\Delta\tilde{J}_n \in \Sigma_x^\epsilon$.

The Markov chain $\{X_n\}_{n\in\mathbb{N}}$ has the shape of a recombining multinomial tree, where each node has $\bar{L}$ branches. The number of nodes at time n is $n(\bar{L}-1)+1$. We derive the transition probabilities by an explicit discretization of the infinitesimal generator (see Section 6.2.2). Following [Kushner and Dupuis, 2001], the process $\{X_n\}_{n\in\mathbb{N}}$ has to satisfy the following two conditions in order to be admissible:

1. the transition probabilities have the representation:

$$p^X\big(X_n, X_{n+1}\big) = \big(1-\lambda\Delta t\big)p^D\big(X_n, X_{n+1}\big) + \big(\lambda\Delta t\big)p^J\big(X_n, X_{n+1}\big) \quad (6.5)$$

   where $\lambda > 0$, and p^D, p^J are respectively the diffusion and jump transition probabilities (see Section 6.2.1).

2. (*local consistency*) The moments of the discrete increments match those of the continuous increments, at first order in Δt:

$$\mathbb{E}_n\big[\Delta X_n\big] = \mathbb{E}_t\big[\Delta X_t\big], \quad \mathbb{E}_n\big[(\Delta X_n)^2\big] = \mathbb{E}_t\big[(\Delta X_t)^2\big]. \quad (6.6)$$

The process $\{Y_n\}_{n\in\mathbb{N}}$ assumes values in Σ_y and ΔL_n, ΔM_n are non-negative multiples of h_y. The two increments $\Delta L_n := L(t_n) - L(t_n^-)$ and $\Delta M_n := M(t_n) - M(t_n^-)$ can occur instantaneously at time t_n. They cannot assume values different from 0 at the same time, and must satisfy the condition $Y_{n+1} = Y_n + \Delta Y_n \in \Sigma_y$[3] for all n.

6.1.2 Discrete dynamic programming algorithm

We can formulate a discrete backward algorithm by applying the dynamic programming principle to (5.27) on the discrete nodes of the chain $\{(Y_n, X_n)\}_n$:

$$\begin{aligned} Q^j(t_n, Y_n, X_n) = \min \Bigg\{ & \mathbb{E}_n\Big[Q\big(t_{n+1}, Y_n, X_n + \Delta X_n\big)\Big], \quad (6.7)\\ & \min_{\Delta L_n} \exp\!\Big(\frac{\gamma}{\delta(t_n,T)}(1+\theta_b)e^{X_n}\Delta L_n\Big)\mathbb{E}_n\Big[Q^j\big(t_{n+1}, Y_n + \Delta L_n, X_n + \Delta X_n\big)\Big],\\ & \min_{\Delta M_n} \exp\!\Big(\frac{-\gamma}{\delta(t_n,T)}(1-\theta_s)e^{X_n}\Delta M_n\Big)\mathbb{E}_n\Big[Q^j\big(t_{n+1}, Y_n - \Delta M_n, X_n + \Delta X_n\big)\Big]\Bigg\}. \end{aligned}$$

The variations of $\{Y_t\}_{t\in[t_0,T]}$ are instantaneous at t_n for each n, while the process $\{X_t\}_{t\in[t_0,T]}$ changes in the interval $[t_n, t_{n+1}]$ according to its Lévy dynamics. This feature suggests to introduce a numerical scheme based on two steps: an evolution step and a control step.

From now on we drop the superscript j from Q^j. We introduce the discretization parameter $\rho = (\Delta t, h_x, h_y)$ and indicate the discretized value function with Q^ρ. For a fixed ρ, we adopt the common short notation $Q^n_{j,i} := Q^\rho(t_n, y_j, x_i)$. We set the initial value $x_i = X_{n=0}$ for $i = 0$. At time n, the index i assumes values in $\{-nK_1, -nK_1+1, ..., nK_2-1, nK_2\}$ and j assumes values in $\{-K_3, -K_3+1, ..., K_4-1, K_4\}$.

[3] The values attainable by ΔL_n and ΔM_n depend on the current value of $Y_n \in \Sigma_y$. For instance, if $Y_n = -K_3 h_y$, then $\Delta L_n \in \{0, h_y, ..., (\bar{M}-1)h_y\}$ and $\Delta M_n \in \{0\}$.

We can define the auxiliary functions:

$$F(x_i, l, t_n) := e^{\left(\frac{\gamma}{\delta(t_n,T)}(1+\theta_b)e^{x_i} l h_y\right)} \qquad (6.8)$$
$$G(x_i, m, t_n) := e^{\left(-\frac{\gamma}{\delta(t_n,T)}(1-\theta_s)e^{x_i} m h_y\right)},$$

such that $l \in \{0, ..., K_4 - j\}$ and $m \in \{0, ..., K_3 + j\}$ for each fixed j.

Algorithm 1 Backward algorithm

Input: $r, (b, \sigma, \nu), X_0, K, T, \theta_b, \theta_s, \gamma, N, \bar{L}, \bar{M}$,
Output: $Q^j(t_0, y, X_0)$ for $j = 0, w, b$

1: Create the lattice for (6.4) with appropriate discrete steps $\Delta t, h_y, h_x$.
2: Create the vector of probabilities p_k as defined in 6.19.
3: Use (5.28) or (5.29) or (5.30) to initialize a $\bar{M} \times \left(N(\bar{L}-1)+1\right)$ grid for $Q^N_{j,i}$.
4: **for** n = N-1 to 0 **do**
5: $W_{j,i} = \sum_{k=-K_1}^{K_2} p_k \, Q^{n+1}_{j,i+k}$
6: $Q^n_{j,i} = \min\left\{W_{j,i}, \min_l F(x_i, l, t_n) W_{j+l,i}, \min_m G(x_i, m, t_n) W_{j-m,i}\right\}$
7: **end for**

The computational complexity of the algorithm (1) is

$$\mathcal{O}\left((N+1)\left[\frac{N(\bar{L}-1)}{2}+1\right] \times \bar{M} \times \bar{M}\right).$$

The first factor comes from the loop over all the nodes of the tree i.e. $\sum_{n=0}^{N} n(\bar{L}-1)+1$. The second factor, $\bar{M}$, comes from the loop over all the values y_j, and the third factor, $\bar{M}$, comes from the minimum search.

For a simple diffusion process the number of branches is fixed to $\bar{L} = 3$, but for processes with jumps it is proportional to $\sqrt{N}$. The standard deviation of every Lévy process satisfying the finite second moment assumption grows as the square root of time. Therefore the size of a space step $h_x \propto \sqrt{\mathbb{E}[\Delta X^2]} \propto \sqrt{\Delta t} \propto \frac{1}{\sqrt{N}}$. Let us consider for instance the integral term in Eq. (6.2) or (6.3). For computational reasons we have to reduce the region of integration to the bounded domain $[-B_1, B_2]$, with $B_1, B_2 > 0$ (see Section 6.2.2). The number of branches to cover this region is $\bar{L} = \frac{B_1+B_2}{h_x} \propto \sqrt{N}$.

In order to have a more accurate result, it is better to choose $h_y \propto h_x$ and consequently $\bar{M} \propto N$. In this way, the number h_y of shares to buy or sell is more sensitive to the resolution h_x in the log-price tree. If we set $\bar{L} = \sqrt{N}$ and $\bar{M} = N$ we have total computational complexity $\mathcal{O}(N^{4.5})$. For a fixed $\bar{L}$, the total complexity is reduced to $\mathcal{O}(N^4)$.

6.2 Properties of the Markov chain

We explained in the Section 6.1.1 that the Markov chain approximation of a continuous time jump-diffusion process has to satisfy two properties. This section makes a summary of the key concepts and refers to [Kushner and Dupuis, 2001] for detailed definitions and proofs of convergence.

6.2.1 Transition probabilities

Let us indicate the transition probabilities of $\{X_n\}_{n\in\mathbb{N}}$ as:

$$p(x_i, x_j) := \mathbb{P}(X_{n+1} = x_j | X_n = x_i). \tag{6.9}$$

The number of jumps of a jump-diffusion process is Poisson distributed $N_t \sim \text{Po}(\lambda t)$, with $\lambda > 0$, i.e.

$$\mathbb{P}(N_t = n) = e^{-\lambda t} \frac{(\lambda t)^n}{n!}. \tag{6.10}$$

For a small Δt, we can compute the first order approximated probabilities:

- $\mathbb{P}(N_{t+\Delta t} - N_t = 0) \overset{d}{=} \mathbb{P}(N_{\Delta t} = 0) = e^{-\lambda \Delta t} \approx 1 - \lambda \Delta t$,
- $\mathbb{P}(N_{\Delta t} = 1) = e^{-\lambda \Delta t}(\lambda \Delta t) \approx \lambda \Delta t$,
- $\mathbb{P}(N_{\Delta t} > 0) = 1 - \mathbb{P}(N_{\Delta t} = 0) \approx \lambda \Delta t$.

Let us consider the discrete dynamics of $\{X_n\}_{n\in\mathbb{N}}$ in Eq. (6.4). We assume that in a small time step Δt the process jumps exactly once ($N_{\Delta t} = 1$), or does not jump at all ($N_{\Delta t} = 0$). The two possible mutually exclusive events are:

- **Diffusion**. The transition probability is $p^D(x_i, x_i + \Delta\Xi)$ and $\Delta\Xi \in \{-h_x, 0, h_x\}$. $p^D(x_i, x_{i+k}) = 0$ for $k \notin \{-1, 0, +1\}$.
- **Jumps**. The transition probability is $p^J(x_i, x_i + \Delta\tilde{J})$. The random variable $\Delta\tilde{J}$ takes values in Σ_x (or Σ_x^ϵ).

By conditioning on the values of $N_{\Delta t}$, the total transition probability is

$$\begin{aligned} p(x_i, x_j) &= p^D(x_i, x_j)\, \mathbb{P}(N_{\Delta t} = 0) + p^J(x_i, x_j)\, \mathbb{P}(N_{\Delta t} = 1) \\ &= (1 - \lambda \Delta t)\, p^D(x_i, x_j) + (\lambda \Delta t)\, p^J(x_i, x_j). \end{aligned} \tag{6.11}$$

The request of a positive probability impose a restriction on the time step size $\Delta t \leq \frac{1}{\lambda}$. In this section we showed that the transition probability of $\{X_n\}_{n\in\mathbb{N}}$ is a convex combination of p^D and p^J, as required by the property (6.5). We refer to Chapter 5.6 of [Kushner and Dupuis, 2001] for more details.

6.2.2 Infinitesimal generator discretization and local consistency

In this section we provide an explicit form for the transition probabilities. This can be achieved by discretizing the infinitesimal generator of the process $\{X_t\}_{t\in[t_0,T]}$ in (6.1), which corresponds to the first term inside the "min" in the HJB equation (5.34). In the following steps we consider only the finite activity case, but the same idea works for an infinite activity process approximated by a jump-diffusion. In fact, the only difference between (6.2) and (6.3) is the truncation in the integral.

In this section we drop the variable y_j from $Q(t_n, y_j, x_i)$, because we are interested only in the uncontrolled log-price dynamics. Let us assume for convenience that Q is smooth enough, the derivatives are discretized by the finite differences:

- Backward approximation in time: $\frac{\partial Q}{\partial t} \approx \frac{Q_i^{n+1} - Q_i^n}{\Delta t}$.
- Central approximation in space: $\frac{\partial Q}{\partial x} \approx \frac{Q_{i+1}^{n+1} - Q_{i-1}^{n+1}}{2h_x}$.
- Second order in space: $\frac{\partial^2 Q}{\partial x^2} \approx \frac{Q_{i+1}^{n+1} + Q_{i-1}^{n+1} - 2Q_i^{n+1}}{h_x^2}$.

The integral terms in (5.34) are truncated and restricted to the domain $\big[-B_1, B_2\big] = \big[(-K_1 - 1/2)h_x, (K_2 + 1/2)h_x\big]$[4]. The discretization is obtained by approximating with Riemann sums (see [Cont and Voltchkova, 2005a]):

$$\int_{-B_1}^{B_2} Q(t_{n+1}, y_j, x_i + z)\nu(dz) \approx \sum_{k=-K_1}^{K_2} \nu_k Q_{i+k}^{n+1}, \tag{6.12}$$

where

$$\nu_k = \int_{(k-\frac{1}{2})h_x}^{(k+\frac{1}{2})h_x} \nu(z)dz, \quad \text{for} \quad -K_1 \le k \le K_2. \tag{6.13}$$

We define the discrete version of $\hat{m} := \int_{-B_1}^{B_2}(e^z - 1)\nu(dz)$, $\hat{\lambda} := \int_{-B_1}^{B_2}\nu(dz)$ and $\hat{\alpha} := \frac{1}{\hat{\lambda}}\int_{-B_1}^{B_2} z\nu(dz)$:

$$\hat{m} \approx \sum_{k=-K_1}^{K_2}(e^{kh_x} - 1)\nu_k, \quad \hat{\lambda} \approx \sum_{k=-K_1}^{K_2}\nu_k, \quad \hat{\alpha} \approx \frac{h_x}{\hat{\lambda}}\sum_{k=-K_1}^{K_2} k\nu_k. \tag{6.14}$$

The jump transition probabilities can be defined as:

$$p_k^J := \frac{\nu_k}{\hat{\lambda}}. \tag{6.15}$$

The discretized equation becomes:

$$\begin{aligned} &\frac{Q_i^{n+1} - Q_i^n}{\Delta t} + (\mu - \frac{1}{2}\sigma^2 - \hat{m})\frac{Q_{i+1}^{n+1} - Q_{i-1}^{n+1}}{2h_x} \\ &+ \frac{1}{2}\sigma^2\frac{Q_{i+1}^{n+1} + Q_{i-1}^{n+1} - 2Q_i^{n+1}}{h_x^2} + \sum_{k=-K_1}^{K_2}\nu_k Q_{i+k}^{n+1} - \hat{\lambda}Q_i^n = 0. \end{aligned} \tag{6.16}$$

Rearranging the terms we get:

$$\begin{aligned} \Big(1 + \hat{\lambda}\Delta t\Big)Q_i^n = p_{-1}^D Q_{i-1}^{n+1} + p_0^D Q_i^{n+1} + p_{+1}^D Q_{i+1}^{n+1} \\ + (\hat{\lambda}\Delta t)\sum_{k=-K_1}^{K_2} p_k^J Q_{i+k}^{n+1}. \end{aligned}$$

where we defined:

$$\begin{aligned} p_{-1}^D &:= \big(-(\mu - \frac{1}{2}\sigma^2 - \hat{m})\frac{\Delta t}{2h_x} + \frac{1}{2}\sigma^2\frac{\Delta t}{h_x^2}\big) \ge 0 \\ p_0^D &:= \big(1 - \sigma^2\frac{\Delta t}{h_x^2}\big) \ge 0 \\ p_{+1}^D &:= \big((\mu - \frac{1}{2}\sigma^2 - \hat{m})\frac{\Delta t}{2h_x} + \frac{1}{2}\sigma^2\frac{\Delta t}{h_x^2}\big) \ge 0 \end{aligned} \tag{6.17}$$

[4] If the integral has a truncation parameter as in (2.41), we choose $\epsilon = 1.5h_x$ and the restricted domain becomes $\big[-B_1, -\epsilon\big]\bigcup\big[\epsilon, B_2\big] = \big[(-K_1 - 1/2)h_x, -3/2h_x\big]\bigcup\big[3/2h_x, (K_2 + 1/2)h_x\big]$.

and $p_k^D := 0$ for $k \notin \{-1, 0, +1\}$. From p_0^D we obtain an important restriction on the time step size: $\Delta t \leq \frac{h_x^2}{\sigma^2}$, while the condition obtained from p_{-1}^D and p_{+1}^D i.e. $h_x \leq \frac{\sigma^2}{|\mu - \frac{1}{2}\sigma^2 - \hat{m}|}$ is easily satisfied. If we bring the term $\left(1 + \hat{\lambda}\Delta t\right)$ on the right hand side and use the first order Taylor approximation $\left(1 + \hat{\lambda}\Delta t\right)^{-1} \approx 1 - \hat{\lambda}\Delta t$, we obtain:

$$\begin{aligned} Q_i^n &\approx \left(1 - \hat{\lambda}\Delta t\right) \sum_{k=-1}^{1} p_k^D \, Q_{i+k}^{n+1} + \left(\hat{\lambda}\Delta t\right) \sum_{k=-K_1}^{K_2} p_k^J \, Q_{i+k}^{n+1} \qquad (6.18) \\ &= \sum_{k=-K_1}^{K_2} p_k \, Q_{i+k}^{n+1} \end{aligned}$$

where

$$p_k = (1 - \hat{\lambda}\Delta t) p_k^D + (\hat{\lambda}\Delta t) p_k^J \qquad (6.19)$$

is the total transition probability, written in the form (6.5). It is straightforward to check that $\sum_k p_k = 1$. Let us check that also the *local consistency* conditions (6.6) are satisfied at first order in Δt:

$$\begin{aligned} \mathbb{E}\left[\Delta X_n\right] =& \left(1 - \hat{\lambda}\Delta t\right) \sum_{k=-1}^{1} p_k^D \, k h_x + \left(\hat{\lambda}\Delta t\right) \sum_{k=-K_1}^{K_2} p_k^J \, k h_x \\ =& \left(1 - \hat{\lambda}\Delta t\right)\left(\mu - \frac{1}{2}\sigma^2 - \hat{m}\right)\Delta t + \left(\hat{\lambda}\Delta t\right)\hat{\alpha} \\ \approx& \left(\mu - \frac{1}{2}\sigma^2 - \hat{m} + \hat{\lambda}\hat{\alpha}\right)\Delta t, \end{aligned}$$

$$\begin{aligned} \mathbb{E}\left[\left[\Delta X_n\right]^2\right] =& \left(1 - \hat{\lambda}\Delta t\right) \sum_{k=-1}^{1} p_k^D \, (k h_x)^2 + \left(\hat{\lambda}\Delta t\right) \sum_{k=-K_1}^{K_2} p_k^J \, (k h_x)^2 \\ =& \left(1 - \hat{\lambda}\Delta t\right)\sigma^2 \Delta t + \left(\hat{\lambda}\Delta t\right)\hat{\eta}^2 \\ \approx& \left(\sigma^2 + \hat{\lambda}\hat{\eta}^2\right)\Delta t. \end{aligned}$$

We introduced $\hat{\eta}^2 := \frac{1}{\hat{\lambda}} \int_{-B_1}^{B_2} z^2 \nu(dz) \approx \frac{h_x^2}{\hat{\lambda}} \sum_{k=-K_1}^{K_2} k^2 \nu_k$, and indicate $\eta^2 := \frac{1}{\lambda} \int_{\mathbb{R}} z^2 \nu(dz) = \frac{1}{\lambda}\mathbb{E}\left[\left|\int_{\mathbb{R}} z \tilde{N}(dt, dz)\right|^2\right]$. The discrete moments match the continuous moments when $h_x \to 0$ and $K_1, K_2 \to \infty$ such that $\hat{\lambda} \to \lambda$, $\hat{\alpha} \to \alpha$ and $\hat{\eta} \to \eta$.

6.2.3 Convergence of the numerical scheme

In Section 6.1.2 we introduced the discretization parameter $\rho = (\Delta t, h_x, h_y)$ and the discretized value function Q^ρ. For a fixed ρ, we indicate $Q_{j,i}^n := Q^\rho(t_n, y_j, x_i)$.
Using the functions in (6.8), let us define the numerical scheme corresponding to the algorithm 1:

Scheme 1. *Let us define a two steps numerical scheme* $\mathbb{S}$ *such that*

$$\mathbb{S}\left(\rho, (t_n, y_j, x_i), Q^\rho(t_n, y_j, x_i), [Q^\rho]_{t_n, y_j, x_i}\right) = 0, \qquad (6.20)$$

where $[Q^\rho]_{t_n,y_j,x_i}$ *indicates all the values of* Q^ρ *not in* (t_n, y_j, x_i).

step 1: $\quad Q^n_{j,i} = \sum_{k=-K_1}^{K_2} p_k\, Q^{n+1}_{j,i+k} \quad$ *for all* $\quad j, i$

step 2: $\quad \mathbb{S} = Q^n_{j,i} - \min\Big\{ Q^n_{j,i},\, \min_l F(x_i, l, t_n) Q^n_{j+l,i},\, \min_m G(x_i, m, t_n) Q^n_{j-m,i} \Big\}$

Let us indicate the Eq. (5.34) by:

$$F\big(\mathrm{x}, Q(\mathrm{x}), DQ(\mathrm{x}), D^2Q(\mathrm{x}), \mathcal{I}(\mathrm{x}, Q)\big) = 0, \tag{6.21}$$

where $\mathrm{x} := (t, y, x)$.

Theorem 6.2.1. *The scheme [1] (with transition probabilities* p_k *defined in 6.19) is monotone, stable and consistent.*

Let us prove the three properties separately.
With the square brackets around $[\varepsilon^n_{j,i}]$, we indicate all the possible values $\varepsilon^{n'}_{j',i'}$ such that $(n', j', i') \neq (n, j, i)$. The scheme is **monotone** i.e. for all $[\varepsilon^n_{j,i}] \geq 0$, then $\mathbb{S}\big(\rho, (n, j, i), Q^n_{j,i}, [Q^n_{j,i}] + [\varepsilon^n_{j,i}]\big) \leq \mathbb{S}\big(\rho, (n, j, i), Q^n_{j,i}, [Q^n_{j,i}]\big)$.

Proof. Let us write the scheme [1] as

$$\mathbb{S} = Q^n_{j,i} - \min\Big\{ \sum_{k=-K_1}^{K_2} p_k\, Q^{n+1}_{j,i+k},\, \min_l F(x_i, l, t_n) Q^n_{j+l,i},\, \min_m G(x_i, m, t_n) Q^n_{j-m,i} \Big\}$$

Since $F(x_i, l, t_n) > 0$, $G(x_i, m, t_n) > 0$ for all x_i, l, m, n, and $p_k > 0$ for all k, the scheme $\mathbb{S}$ is a decreasing function of $[Q^n_{j,i}]$. □

The scheme is **stable** i.e. for any $\rho > 0$ there exists a bounded solution Q^ρ, with bound independent of ρ. This is equivalent to prove that $||Q^n||_\infty \leq C$, for any $0 \leq n \leq N$ and for C independent on ρ.

Proof. The terminal conditions (5.28), (5.29), (5.30) are positive bounded functions in a bounded domain. Therefore we can write $0 \leq Q^N_{j,i} \leq C$ for all i, j, and C does not depend on ρ. Since all the coefficients are positive, it follows that the scheme is *sign preserving* i.e. $Q^n_{j,i} \geq 0$ for all n. We can write:

$$\begin{aligned} Q^n_{j,i} &= \min\Big\{ \sum_{k=-K_1}^{K_2} p_k\, Q^{n+1}_{j,i+k},\, \min_l F(x_i, l, t_n) Q^n_{j+l,i},\, \min_m G(x_i, m, t_n) Q^n_{j-m,i} \Big\} \\ &\leq \sum_{k=-K_1}^{K_2} p_k\, Q^{n+1}_{j,i+k} \;\leq\; ||Q^{n+1}||_\infty. \end{aligned}$$

This holds for all i, j, then $||Q^n||_\infty \leq ||Q^{n+1}||_\infty$. Iterating we obtain:

$$||Q^n||_\infty \leq ||Q^N||_\infty \leq C.$$

□

The scheme is **consistent** i.e. for any smooth function ϕ

$$\mathbb{S}\big(\rho, \mathrm{x}_\rho, \phi^\rho(\mathrm{x}_\rho), [\phi^\rho]_{\mathrm{x}_\rho}\big) \underset{\substack{\rho\to 0\\ \mathrm{x}_\rho\to \mathrm{x}}}{\longrightarrow} F\big(\mathrm{x}, \phi(\mathrm{x}), D\phi(\mathrm{x}), D^2\phi(\mathrm{x}), I(\mathrm{x},\phi)\big).$$

with $\mathrm{x}_\rho := (t_n, y_j, x_i)$.

Proof. Now we look at the following cases, corresponding to each minimum value in the scheme [1]:
1) For some $l > 0$, it holds

$$\begin{aligned}
0 &= e^{\left(\frac{\gamma}{\delta(t_n,T)}(1+\theta_b)e^{x_i} l h_y\right)}\phi\big(t_n, y_j + l h_y, x_i\big) - \phi\big(t_n, y_j, x_i\big)\\
&= \left(1 + \frac{\gamma(1+\theta_b)e^{x_i}}{\delta(t_n,T)} l h_y + \mathcal{O}(h_y)\right)\left(\phi(t_n, y_j, x_i) + \frac{\partial\phi}{\partial y}\bigg|_{y_j} l h_y + \mathcal{O}(h_y)\right) - \phi(t_n, y_j, x_i)\\
&= \frac{\partial\phi}{\partial y}\big(t_n, y_j, x_i\big) + \frac{\gamma}{\delta(t_n,T)}(1+\theta_b)e^{x_i}\,\phi\big(t_n, y_j, x_i\big) + \mathcal{O}(h_y).
\end{aligned}$$

2) For some $m > 0$, an analogous computation leads to

$$-\frac{\partial\phi}{\partial y}\big(t_n, y_j, x_i\big) - \frac{\gamma}{\delta(t_n,T)}(1-\theta_s)e^{x_i}\,\phi\big(t_n, y_j, x_i\big) + \mathcal{O}(h_y) = 0.$$

3) When $\sum_{k=-K_1}^{K_2} p_k\ \phi(t_{n+1}, y_j, x_{i+k}) - \phi(t_n, y_j, x_i) = 0$ let us consider the expression (6.19), and expand p^D (6.17) and p^J (6.15):

$$\begin{aligned}
&(1-\sigma^2\frac{\Delta t}{h_x^2})\left(\phi + \frac{\partial\phi}{\partial t}\bigg|_{t_n}\Delta t + \mathcal{O}(\Delta t^2)\right) - \hat\lambda\Delta t\phi(t_{n+1}, y_j, x_i) - \phi(t_n, y_j, x_i)\\
&+ \left((\mu - \frac{1}{2}\sigma^2 - \hat m)\frac{\Delta t}{2h_x} + \frac{1}{2}\sigma^2\frac{\Delta t}{h_x^2} + \mathcal{O}(\Delta t^2)\right)\left(\phi + \frac{\partial\phi}{\partial x}\bigg|_{x_i} h_x + \frac{1}{2}\frac{\partial^2\phi}{\partial x^2}h_x^2 + \mathcal{O}(h_x^3)\right)\\
&+ \left(-(\mu - \frac{1}{2}\sigma^2 - \hat m)\frac{\Delta t}{2h_x} + \frac{1}{2}\sigma^2\frac{\Delta t}{h_x^2} + \mathcal{O}(\Delta t^2)\right)\left(\phi - \frac{\partial\phi}{\partial x}\bigg|_{x_i} h_x + \frac{1}{2}\frac{\partial^2\phi}{\partial x^2}h_x^2 + \mathcal{O}(h_x^3)\right)\\
&+ \hat\lambda\Delta t\sum_{k=-K_1}^{K_2}\frac{\nu_k}{\hat\lambda}\,\phi(t_{n+1}, y_j, x_{i+k}) = 0.
\end{aligned}$$

Let us replace all the terms at t_{n+1} by using the first order Taylor approximation $\phi(t_{n+1},\cdot,\cdot) = \phi(t_n,\cdot,\cdot) + \frac{\partial\phi}{\partial t}\Big|_{t_n}\Delta t + \mathcal{O}(\Delta t^2)$. The two terms in $\hat\lambda$ can be rewritten as $\Delta t\sum_{k=-K_1}^{K_2}\nu_k\,\big(\phi(t_n, y_j, x_{i+k}) - \phi(t_n, y_j, x_i)\big)$. Using the approximation (6.12) and (6.13) we obtain

$$\begin{aligned}
&\frac{\partial\phi}{\partial t} + (\mu - \frac{1}{2}\sigma^2 - \hat m)\frac{\partial\phi}{\partial x} + \frac{1}{2}\sigma^2\frac{\partial^2\phi}{\partial x^2}\\
&+ \int_{-B_1}^{B_2}\big(\phi(t_n, y_j, x_i + z) - \phi(t_n, y_j, x_i)\big)\nu(z)dz + \mathcal{O}(\Delta t) + \mathcal{O}(h_x) = 0.
\end{aligned}$$

When sending $\rho \to 0$, $\mathrm{x}_\rho \to \mathrm{x}$ and $B_1, B_2 \to \infty$ we obtain the desired result for all 1) 2) and 3). □

Theorem 6.2.2. *The solution Q^ρ of (6.20) converges uniformly to the unique viscosity solution of (5.34).*

The proof follows closely [Barles and Souganidis, 1991].

Proof. We only prove the subsolution case, since the arguments for the supersolution are identical. Let $\bar{\mathrm{x}}$ the strict global maximum of $Q - \phi$ for some $\phi \in C^{1,1,2} \bigcap C_2$, and such that $Q(\bar{\mathrm{x}}) = \phi(\bar{\mathrm{x}})$. Then there exist sequences ρ_n and x_n such that for $n \to \infty$:
$\rho_n \to 0$, $\mathrm{x}_n \to \bar{\mathrm{x}}$, $Q^{\rho_n}(\mathrm{x}_n) \to Q(\bar{\mathrm{x}})$ and x_n is a global maximum of $Q^{\rho_n}(\cdot) - \phi(\cdot)$. Let us define $\xi_n := Q^{\rho_n}(\mathrm{x}_n) - \phi(\mathrm{x}_n)$, such that $\xi_n \to 0$ when $n \to \infty$. For any x it holds $Q^{\rho_n}(\mathrm{x}) \leq \phi(\mathrm{x}) + \xi_n$. Let us consider the scheme [1]:

$$\begin{aligned} 0 &= \mathbb{S}\big(\rho_n, \mathrm{x}_n, Q^{\rho_n}(\mathrm{x}_n), [Q^{\rho_n}]_{\mathrm{x}_n}\big) \\ &\geq \mathbb{S}\big(\rho_n, \mathrm{x}_n, \phi(\mathrm{x}_n) + \xi_n, [\phi + \xi_n]_{\mathrm{x}_n}\big), \end{aligned}$$

where we used the monotonicity property. By sending $n \to \infty$ and thanks to the consistency property, we obtain:

$$F\big(\bar{\mathrm{x}}, \phi(\bar{\mathrm{x}}), D\phi(\bar{\mathrm{x}}), D^2\phi(\bar{\mathrm{x}}), I(\bar{\mathrm{x}}, \phi)\big) \leq 0.$$

□

6.3 Numerical results

In this section we implement the algorithm [1] described in Section 6.1.2 and calculate the prices of European call options for the writer and the buyer. The prices are computed under the assumption that the stock log-price follows three different Lévy processes: a Brownian motion, a Merton jump-diffusion and a Variance Gamma, with parameters in table 6.1.

Details			**Diffusion parameters**					
K	T	r	μ	σ	γ			
15	1	0.1	0.1	0.25	0.001			
			Merton parameters					
			μ	σ	α	ξ	λ	γ
			0.1	0.25	0	0.5	0.8	0.04
			VG parameters					
			μ	θ	$\bar{\sigma}$	κ	γ	
			0.1	-0.1	0.2	0.1	0.05	

Table 6.1: Option details and parameters for diffusion, Merton and VG processes.

We compute the option prices using the standard *martingale pricing theory* presented in Chapter 2. In the table 6.2 we show the *at the money* values obtained with the closed formula and by solving the respective PIDE (see table 2.4). The PIDE prices are obtained by solving the equations (2.31), (2.33) and (2.37). Of course, the parameter μ has not been used to compute the prices in

Diffusion price		Merton price		VG price	
Closed formula	PDE	Closed formula	PIDE	Closed formula	PIDE
2.2463	2.2463	3.4776	3.4775	1.9870	1.9871

Table 6.2: At the money prices with $S_0 = K = 15$ with parameters in tables 6.1.

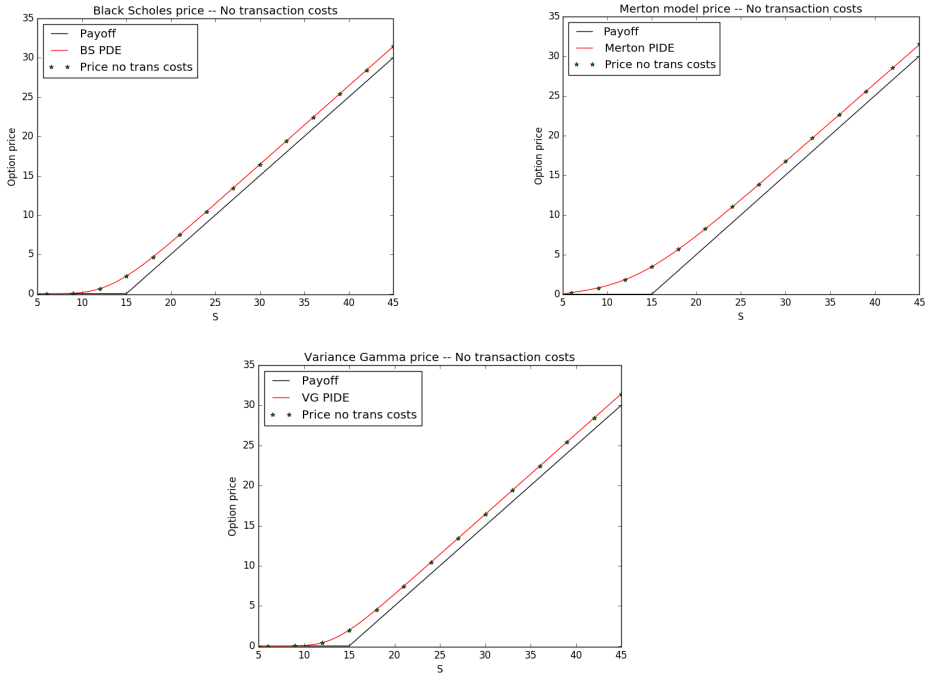

Figure 6.1: Writer prices with zero transaction costs for diffusion (top-left), Merton (top-right) and VG (bottom) process. Parameters are in the table 6.1.

Tab. 6.2. In Section (6.3.5), we prove with a numerical experiment that even in the model with transaction costs, the drift μ does not play an important role for the option price. In the following analysis, we consider the PIDE prices as our benchmarks for comparisons. In all the computations we use equal transaction costs for buying and selling, $\theta_b = \theta_s$.

In Fig. 6.1 we show that model prices replicate the PIDE prices for zero transaction costs and small values of γ. The values of γ in the table 6.1, are chosen very small[5] for this purpose. An intuitive argument to justify this choice is that for $\gamma \to 0$, the utility function can be approximated by a linear utility $\mathcal{U}(w) = 1 - e^{-\gamma w} \approx \gamma w$ and the investor can be considered risk neutral. A rigorous argument can be found in [Barles and Soner, 1998], where the authors use asymptotic analysis for small values of θ_b, θ_s and γ to derive a nonlinear

[5] In chapter 5 of [Grinold and Kahn, 1999] are presented some common values for the risk aversion coefficient: $\gamma = 0.3$, $\gamma = 0.2$ and $\gamma = 0.1$ for high, medium and low level of risk aversion respectively.

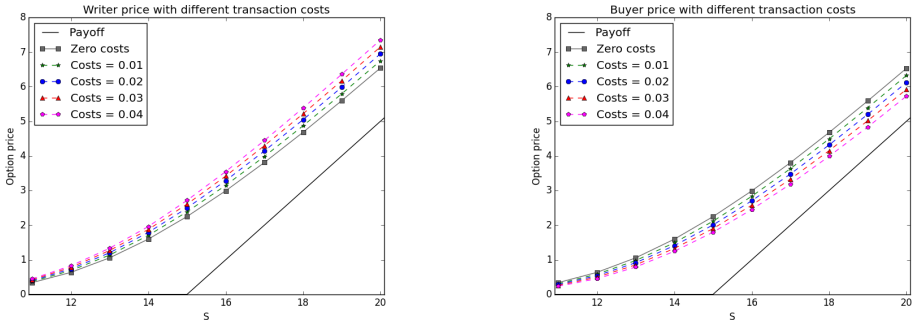

Figure 6.2: Writer and buyer prices for different levels of transaction costs. The continuous line is the solution of the Black-Scholes PDE.

PDE for the option price. For zero transaction costs this equation corresponds to the Black-Scholes PDE. Their argument can be extended also to PIDEs.

Convergence table				
$N = \bar{M}$	$\gamma = 0.0001$	$\gamma = 0.001$	$\gamma = 0.01$	Execution time
50	2.241214	2.241764	2.247311	0.01 ± 0.004
100	2.249142	2.249506	2.253159	0.02 ± 0.005
200	2.245422	2.245676	2.248216	0.11 ± 0.02
400	2.246784	2.246959	2.248717	0.85 ± 0.04
800	2.246288	2.246271	2.247635	8.63 ± 0.1
1600	2.246576	2.246662	2.247515	82.44 ± 2.71
3200	2.246412	2.246471	2.247068	910.8 ± 10.5
3500	2.246366	2.246423	2.246993	1291.3 ± 13

Table 6.3: Convergence table for ATM diffusion prices with zero transaction costs.

6.3.1 Diffusion results

In the figure 6.2 we show the diffusion writer and buyer prices with different transaction costs. We can see that a higher transaction cost corresponds to a higher writer price, while a lower transaction cost corresponds to a lower buyer price. In fact, the writer and buyer prices are respectively increasing and decreasing functions of the transaction cost, as already verified in [Clewlow and Hodges, 1997]. The prices figures 6.2, are calculated with $N = 1500$ time steps and $\bar{M} = N$. In the Table 6.3 we show ATM option prices for different values of N, with $\theta_s = \theta_b = 0$ and different risk aversion coefficients. For $\gamma = 0.0001$ and $N = 3500$ the price is identical, up to the fourth decimal digit, to the original Black-Scholes price in table 6.2. Using the values in the table 6.3 it is possible to perform a *numerical convergence analysis* (see section 6.3.4). We also present the execution times, from which we can estimate the asymptotic *time complexity* of the algorithm. In Section (6.1.2) we

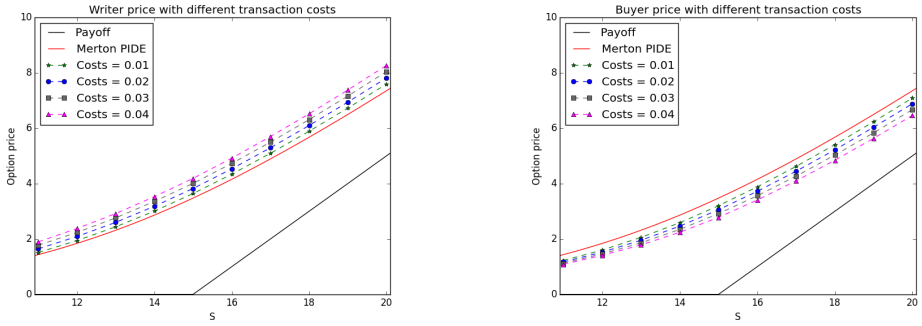

Figure 6.3: Writer and buyer prices for different transaction costs. The continuous line is the solution of the Merton PIDE.

estimated a complexity of $\mathcal{O}(N^4)$. From the table 6.3, we obtain the exponent $\frac{\log(1291.3/910.8)}{\log(3500/3200)} = 3.9$, which is very close to the predicted value. The algorithm is written in Matlab using vectorized operations, and runs on an Intel i7 (7th Gen) with Linux.

6.3.2 Merton results

In the Figure 6.3 we show the writer and buyer prices for the Merton process, with parameters in Tab. 6.1. An interesting feature of the multinomial tree construction for jump-diffusion processes is that $\bar{L} \propto \sqrt{N}$. The integral domain is restricted to the bounded domain $[-B_1, B_2]$ with length $B_1 + B_2 = \bar{L}\, h_x$. We choose the size of a space step $h_x = \sqrt{\mathbb{E}[\Delta X^2]} = \sigma_X \sqrt{\Delta t}$ and $\sigma_X^2 = \sigma^2 + \tilde{\sigma}_J^2$ with $\tilde{\sigma}_J^2 = \int_{-B_1}^{B_2} z^2 \nu(dz)$. However, the size of the Poisson jumps does not scale with Δt. So the number $\bar{L}$ has to be chosen big enough in order to have $L\, h_x \geq B_1 + B_2$.

In general, for a fixed h_x, the interval $[-B_1, B_2]$ should be chosen as big as possible. In practice, the choice of the truncation depends on the shape of the Lévy measure. The figures 6.4, 6.5 show two examples with $[-B_1, B_2] = [\sqrt{\lambda}\xi, \sqrt{\lambda}\xi]$ and $[-B_1, B_2] = [-3\sqrt{\lambda}\xi, 3\sqrt{\lambda}\xi]$ respectively. These choices correspond to $\bar{L} = 17$ and $\bar{L} = 52$. For the Merton Lévy measure (a scaled Normal distribution), a good choice is $[-B_1, B_2] = [-3\sigma_J, 3\sigma_J]$, where $\sigma_J^2 = \int_{\mathbb{R}} z^2 \nu(dz) = \lambda(\alpha^2 + \xi^2)$ is the variance of the jump component of the Merton process. The length of the interval is $\bar{L}\, h_x = 6\sigma_J$. It is well known that the integral over this region is about the 99.74% of the total area. Using this interval and the parameters in Tab. 6.1 we obtain the relation $\bar{L} \geq 5.86\sqrt{N}$. In the calculation of the Merton prices in Fig. 6.3, we used a discretization with $N = \bar{M} = 100$, and $\bar{L} = 81$, with a good balance between small computational time and small price error.

The convergence Tab. 6.4 shows different Merton prices for different values of N and $\bar{L}$. Looking at the table from left to right, for each fixed N it is possible to note how the truncation error decreases when $\bar{L}$ increases.

In Table 6.5 we show several prices with increasing values of N and $\bar{L}$. We choose $\bar{L}$ big enough, such that the truncation error can be ignored. Given the

Truncation error table					
N	$\bar{L}=51$	$\bar{L}=71$	$\bar{L}=91$	$\bar{L}=101$	$\bar{L}=111$
50	3.481318	3.481616	3.481617	3.481617	3.481617
100	3.468774	3.478806	3.479141	3.479146	3.479146
150	3.439090	3.474403	3.477574	3.477714	3.477742
200	3.399442	3.466338	3.476439	3.477234	3.477457

Table 6.4: Truncation error for ATM Merton prices with zero transaction costs.

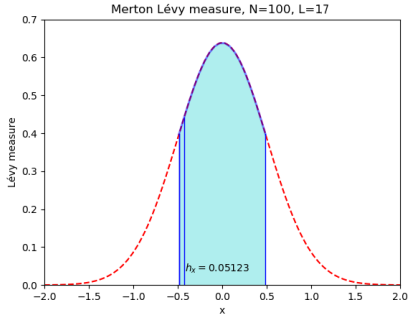

Figure 6.4: Merton Lévy measure computed using parameters in Tab. 6.1, $N = 100$ and $\bar{L} = 17$. The domain $[-B_1, B_2] = [-\sqrt{\lambda}\xi, \sqrt{\lambda}\xi]$ has length $\bar{L}h_x \approx 2\sqrt{\lambda}\xi$.

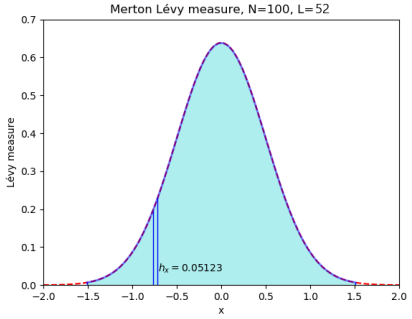

Figure 6.5: Merton Lévy measure computed using parameters in Tab. 6.1, $N = 100$ and $\bar{L} = 52$. The domain $[-B_1, B_2] = [-3\sqrt{\lambda}\xi, 3\sqrt{\lambda}\xi]$ has length $\bar{L}h_x \approx 6\sqrt{\lambda}\xi$.

high computational complexity of the algorithm, it is difficult to present prices with bigger values of N,$\bar{L}$. For larger values of N and smaller γ, we expect a convergence to the Merton price in Tab. 6.2. The computational complexity in this case is expected to be $\mathcal{O}(N^{4.5})$. From the table 6.5 we get the exponent equal to $\frac{\log(1106.0/585.7)}{\log(200/175)} = 4.76$. Considering the errors, the results is not so different from the predicted value.

6.3.3 VG results

In the Figure 6.6 we show how the writer and buyer prices for the VG process change for several level of transaction costs (parameters in Tab. 6.1). In these computations we used $N = \bar{M} = 150$ and $\bar{L} = 43$, such that the program can run in a reasonable computational time. The integration region in 2.41 is restricted to $[-B_1, -\epsilon] \bigcup [\epsilon, B_2]$ with $\epsilon = 1.5h_x$. The choice of B_1 and B_2 depends on the shape of the Lévy measure. In Fig. 6.7 and 6.8 we show two examples for the VG Lévy measure (using parameters in table 6.1) with $N = 150$, $h_x = 0.0165$ and $N = 1000$, $h_x = 0.0064$. The two Lévy measures are normalized, such that the integral on the region $[-\infty, -\epsilon] \bigcup [\epsilon, +\infty]$ is equal to one. The area underlying the functions on $[-B_1, -\epsilon] \bigcup [\epsilon, B_2]$ is highlighted for clarity. For $\bar{L} = 43$, we can see that in both cases it is possible to cover a very

Convergence table			
$N = \bar{M}$	$\bar{L}$	Price	Execution time
50	61	3.481600	2.20 ± 0.08
75	75	3.479980	15.15 ± 0.07
100	91	3.479141	63.04 ± 0.49
125	97	3.478254	148.4 ± 1.16
150	105	3.477731	315.3 ± 5.58
175	113	3.477610	585.7 ± 10.57
200	121	3.477513	1106.0 ± 12.2

Table 6.5: Convergence table for ATM Merton prices with zero transaction costs.

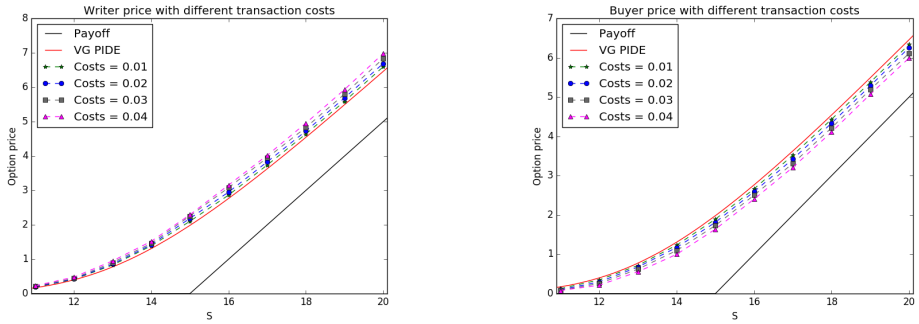

Figure 6.6: Writer and buyer prices for different transaction costs. The continuous line is the solution of the VG PIDE.

high percentage of the initial unrestricted region. We can conclude that, unlike the Merton measure, we do not need a big truncation interval. Given the space step $h_X = \sigma_X\sqrt{\Delta t}$, with $\sigma_X^2 = \hat{\sigma}_J^2 + \sigma_\epsilon^2$ and $\hat{\sigma}_J^2 = \int_{[-B_1,-\epsilon]\bigcup[\epsilon,B_2]} z^2\nu(dz)$, it is enough to consider a region at least as big as the standard deviation of the unrestricted jump process[6] i.e. $h_X\bar{L} \geq \sigma_J$, where $\sigma_J^2 = \int_{[-\infty,-\epsilon]\bigcup[\epsilon,\infty]} z^2\nu(dz)$. Putting all together, the relation becomes $\bar{L} \geq \frac{\sigma_J}{\sigma_X}\sqrt{N}$, and replacing the values $\sigma_X = 0.2024$ and $\sigma_J = 0.1916$ we get $\bar{L} \geq 0.94\sqrt{N}$.

In the Table 6.6 we present several option prices computed with different values of N, but with fixed $\bar{L}$. In the case of the VG process it is more difficult to analyze the convergence results. This is due to the approximation (6.3) introduced to replace the infinite activity jump component with a Brownian motion. All the parameters in (2.39) depend on ϵ, and consequently on N. Within our discretization ($N = 150$), we have $\sigma_\epsilon = 0.0654$, $\lambda_\epsilon = 10.01$. With the parameters under consideration, we obtain an ATM price for zero transaction costs of 1.9821, which is very close the the PIDE price in Tab. 6.2.

The convergence rate of the VG PIDE is quite low and this is reflected in our algorithm. We refer to [Cont and Voltchkova, 2005a] for a detailed error

[6]For small values of ϵ the value of σ_ϵ^2 is negligible. In this case it is possible to use the expression for the variance of the VG process and write $\sigma_X^2 = \bar{\sigma}^2 + \theta^2\kappa$.

Convergence table			
$N = \bar{M}$	λ_ϵ	Price	Execution time
50	4.73	1.910934	3.63 ± 0.16
100	7.82	1.957806	26.54 ± 0.26
150	10.01	1.982078	82.51 ± 0.20
200	11.73	1.996180	185.2 ± 0.81
250	13.14	2.004719	350.3 ± 4.5
300	14.35	2.008536	654.2 ± 7.3
350	15.40	2.009436	1236 ± 12

Table 6.6: Convergence of ATM VG prices with $\bar{L} = 43$, zero transaction costs.

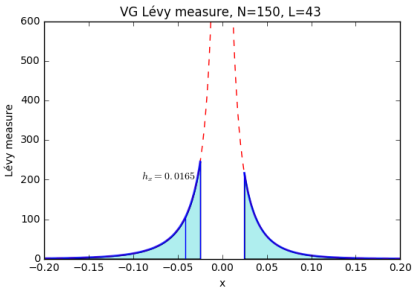

Figure 6.7: VG Lévy measure computed using parameters in Tab. 6.1, $N = 150$ and $\bar{L} = 43$. The domain $[-B_1, -\epsilon] \bigcup [\epsilon, B_2]$ has length $(\bar{L} - 3)h_x$. The highlighted area is 99.9% of the total area.

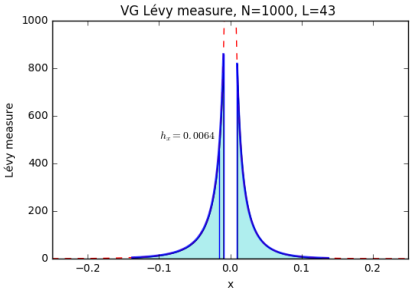

Figure 6.8: VG Lévy measure computed using parameters in Tab. 6.1, $N = 1000$ and $\bar{L} = 43$. The domain $[-B_1, -\epsilon] \bigcup [\epsilon, B_2]$ has length $(\bar{L} - 3)h_x$. The highlighted area is 98.9% of the total area.

analysis. In order to solve the PIDE (using an implicit-explicit scheme) we constructed a grid with 13000 space steps of size $\delta x = 0.0004$ and 7000 time steps, and obtained the price in Tab 6.2, with an approximated activity $\lambda_\epsilon = 75$. Consequently, we expect to have good convergence results in our algorithm when $N \sim 10^4$. All the presented prices (Figures 6.6) have thus a truncation error, which is adjusted by an accurate choice of the value of γ.

From Tab. 6.6 we can estimate the time complexity of this algorithm. The exponent is $\frac{\log(1236.0/654.2)}{\log(350/300)} = 4.12$, indeed very close to the theoretical $\mathcal{O}(N^4)$.

6.3.4 Numerical convergence analysis

In this section we want to analyze the convergence properties of the algorithm 1 considering the prices presented in tables 6.3, 6.5 and 6.6.

Let us first consider the prices in table 6.3, which are plotted in figure 6.9. We can see that for higher values of γ, not only the price increases, but the oscillations of the log-error increases as well. Let us assume the limit price V^* is represented by the price at $N = 3500$. For computational reasons, i.e. the

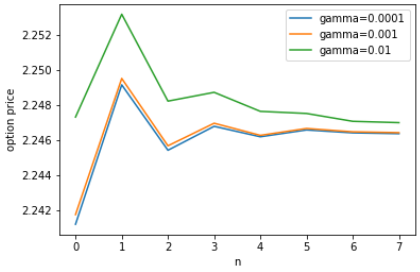

Figure 6.9: Prices in table 6.3 as function of $n = \log_2(N/50)$.

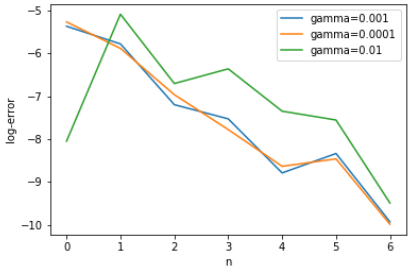

Figure 6.10: Plot of the log-errors as function of $n = \log_2(N/50)$.

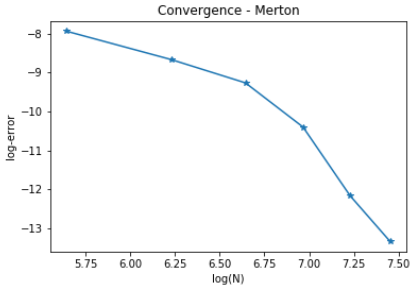

Figure 6.11: Plot of the log-errors as function of $n = \log_2(N/50)$, for Merton prices in 6.5

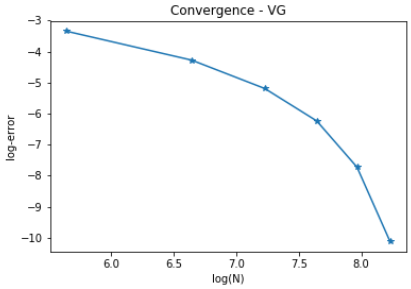

Figure 6.12: Plot of the log-errors as function of $n = \log_2(N/50)$, for VG prices in 6.6

very high computational time, it was not possible to consider higher values. Following the arguments in section 2.2.4, let us define the log-error $\log_2(\varepsilon_N)$ with $\varepsilon_N = |V^N - V^*|$. Then we consider $n = 0, 1, ..., 6$ such that $N = 50 \cdot 2^n$. In figure 6.10 we plot the log-error for each value of n. For $\gamma = 0.0001$ and $\gamma = 0.001$, even if the line is not straight, it is possible to recognize a linear behavior. In these cases it is possible to estimate the rate of convergence (the slope of the line is about -1 i.e. linear convergence). For $\gamma = 0.01$, given the irregular shape, it is hard to understand the functional behavior. The value at $n = 0$ can be excluded because it is probably an outlier originated by the rough grid resolution. In general, given the nonlinear nature of the optimization problem, we expect a nonlinear behavior of the convergence functional form.

This is confirmed by the pictures 6.11 and 6.12, containing the prices obtained for the jump-diffusion Merton process in table 6.5, and the VG process in table 6.6. The shape is more regular (here the number of steps N is much smaller), but it is not a linear function. The algorithm converges faster for higher values of N.

	cost = 0	cost = 0.01	cost = 0.02	cost = 0.03	cost = 0.04
Merton	3.4771	3.6400	3.8212	4.0054	4.1864
VG	1.9821	2.0921	2.1870	2.2568	2.3131

Table 6.7: Merton and VG writer prices for different transaction costs, with parameters as in Tab. (6.1).

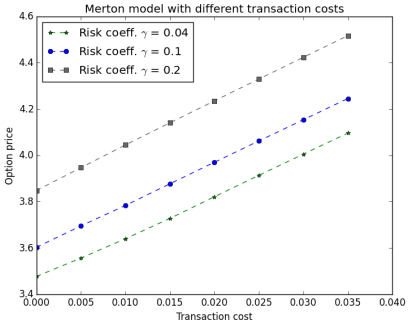

Figure 6.13: Merton option prices for the writer as function of the transaction cost, with different values of γ.

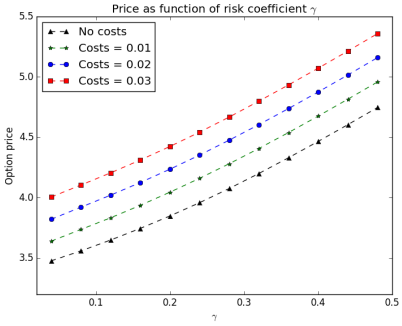

Figure 6.14: Merton option prices for the writer as function of γ, with different values of transaction costs.

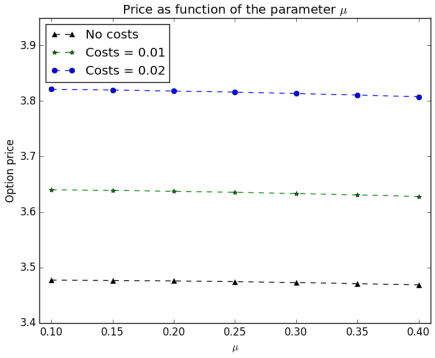

Figure 6.15: Merton option prices for the writer as function of μ, with different values of transaction costs.

6.3.5 Properties of the model

In this section we want to analyze the properties of the model and how the option price depends on the level of transaction costs θ_b, θ_s, the risk aversion parameter γ and the drift μ. In this numerical experiment, we use the Merton model with parameters of Tab. 6.1. In Tab. 6.7 we show the writer ATM option values for different transaction costs.

In Fig. 6.13 we can see better how the price for the writer is affected by the change of the transaction costs. The picture shows prices for different values of risk coefficient. The risk profile of the investor also plays an important role. As

already shown in [Hodges and Neuberger, 1989], the writer price is an increasing function of the risk aversion coefficient. Figure 6.14 confirms their results.

In all the previous computations we always used the drift term μ equal to the risk free interest rate r. This is the choice made in [Hodges and Neuberger, 1989], following the common rule of the standard no-arbitrage theory. The option price has to be independent of the expected return of the underlying asset. As we can see in Fig. 6.15, the numerical experiment shows that also in this model, the option prices do not depend on the drift μ.

6.4 Solution of the 4-dimensional problem

This section focuses on the original HJB equation 5.21, were no variables reduction is considered. A few authors have already presented some results for the diffusion case, using different approaches. In [Palczewsky et al., 2015] the authors propose a method to improve the performances of the Markov chain approximation method for the diffusion case. In [Wang and Li, 2014] the author propose a penalty method for the diffusion case of the 4 dimensional HJB equation, but the numerical results they present are not very clear.

If we want to solve the problem (5.20), we have to deal with a three dimensional state. In the variable reduction section 5.1.3 we assumed a high value of credit availability C, such that the default probability can be ignored. Here we do not ignore the default case, and compute option prices for investor with low credit availability. We work with the original problem (5.20) and derive a discrete time dynamic programming equation as done for (6.7).

Let us indicate with $B_n := B(t_n^-)$ the value of cash immediately before the possible transaction. Let us define $\Sigma_b := \{-K_5 h_b, ..., -h_b, 0, h_b, ..., +K_6 h_b\}$, where $h_b > 0$ is the discrete cash step and $K_5, K_6 \in \mathbb{N}$. Its dimension is $\bar{B} = \#(\Sigma_b) = K_5 + K_6 + 1$. The discretized SDE for the cash account process $\{B_t\}_{t\in[t_0,T]}$ in (5.18), with solution (5.25), is:

$$B_{n+1} = e^{r\Delta t}\Big(B_n - (1+\theta_b)e^{X_n}\Delta L_n + (1-\theta_s)e^{X_n}\Delta M_n\Big) \tag{6.22}$$

Let us derive the backward algorithm for computing the value function using the DPP, as we did for (6.7). We obtain the discrete DPE:

$$\begin{aligned} V(t_n, B_n, Y_n, X_n) = \max\Big\{ & \mathbb{E}_n\Big[V\big(t_{n+1}, e^{r\Delta t}B_n, Y_n, X_n + \Delta X_n\big)\Big], \\ & \max_{\Delta L_n} \mathbb{E}_n\Big[V\big(t_{n+1}, e^{r\Delta t}(B_n - e^{X_n}(1+\sigma_b)\Delta L_n), Y_n + \Delta L_n, X_n + \Delta X_n\big)\Big], \\ & \max_{\Delta M_n} \mathbb{E}_n\Big[V\big(t_{n+1}, e^{r\Delta t}(B_n + e^{X_n}(1-\sigma_s)\Delta M_n), Y_n - \Delta M_n, X_n + \Delta X_n\big)\Big]\Big\}, \end{aligned} \tag{6.23}$$

where all the expectations are conditioned on the current state (B_n, Y_n, X_n). We use the notation $V(t_n, b_h, y_j, x_i) = V^n_{h,j,i}$

Following the arguments in section 6.1.2, the computational complexity is

$$\mathcal{O}\Big((N+1)\big[\frac{N(\bar{L}-1)}{2}+1\big] \times \bar{M} \times \bar{B} \times \min\{\bar{M}, \bar{B}\}\Big).$$

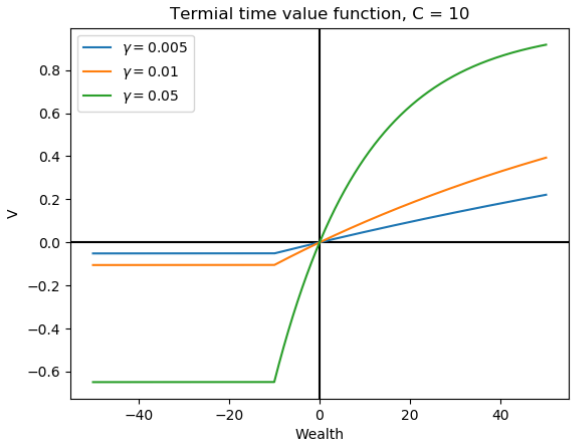

Figure 6.16: Terminal time value functions with $\mathcal{U}(w) = -C$ for $w < -C$.

The term $\min\{\bar{M}, \bar{B}\}$ is the computational complexity of the minimum search. It is performed for each (t_n, b_h, y_j, x_i) such that of all $l, m \in \mathbb{N}$:

$$y_j + lh_y \in \Sigma_y \bigcap (b_h - e^{x_i}(1 + \sigma_b)lh_y) \geq -K_5 h_b$$

and

$$y_j - mh_y \in \Sigma_y \bigcap (b_h + e^{x_i}(1 - \sigma_s)mh_y) \leq K_6 h_b.$$

If we set $\bar{L} = \sqrt{N}$ and $\bar{B} = \bar{M} = N$ we have total computational complexity $\mathcal{O}(N^{5.5})$.

In the following analysis of the problem (5.20), we assume that $\mathcal{U}(w) = -C$ for $w < -C$ and $\beta = r$. In figure (6.16) it is possible to see the shape of the value function at terminal time for different values of γ. In the points (b, y, x) such that $\mathcal{W}(b, y, x) = -C$ the function is not differentiable, and this may create some instabilities in the numerical computations.

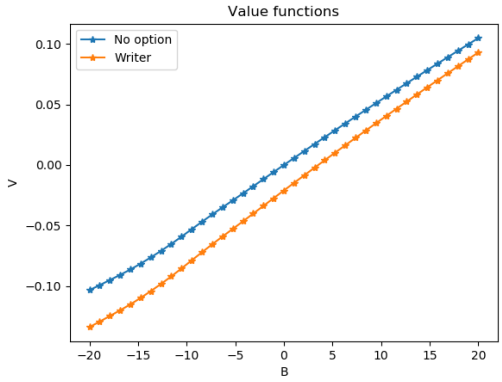

Figure 6.17: Value function for the writer and value function with no option at time t_0 with $Y_0 = 0$ and $S_0 = 15$. The option price at $B_0 = 0$ is $p^w = 3.48303858$.

Algorithm 2 Backward algorithm

Input: $r, (b, \sigma, \nu), X_0, K, T, \theta_b, \theta_s, \gamma, N, \bar{L}, \bar{M}, \bar{B}$.
Output: $V^j(t_0, b, y, X_0)$ for $j = 0, w, b$

1: Create the lattice for (6.4) and (6.22) with appropriate discrete steps $\Delta t, h_y, h_x, h_b$.
2: Create the vector of probabilities p_k as defined in 6.19.
3: Use (5.17) to initialize a $\bar{B} \times \bar{M} \times \big(N(\bar{L} - 1) + 1\big)$ grid for $V^N_{h,j,i}$.
4: **for** n = N-1 to 0 **do**
5: $W_{h,j,i} = \sum_{k=-K_1}^{K_2} p_k \, V^{n+1}_{h,j,i+k}$
6: Interpolate the value of W in the points:
- $W_1(n, h, j, i)$ at $(t_n, e^{r\Delta t} b_h, y_j, x_i)$.
- $W_2(n, h, j, i, l)$ at $(t_n, e^{r\Delta t}(b_h - e^{x_i}(1 + \sigma_b) l h_y), y_j + l h_y, x_i)$ for each l.
- $W_3(n, h, j, i, m)$ at $(t_n, e^{r\Delta t}(b_h + e^{x_i}(1 - \sigma_s) m h_y), y_j - m h_y, x_i)$ for each m.
7: $V^n_{h,j,i} = \max\Big\{ W_1(n, h, j, i),\ \max_l W_2(n, h, j, i, l),\ \max_m W_3(n, h, j, i, m) \Big\}$
8: **end for**

We present numerical solution for a Merton process with values in table 2.2. The cash vector is is chosen such that $-20 \leq B_0 \leq 20$, and consider values of $N = \bar{M} = \bar{B} = 25$ and $\bar{L} = 11$, but even with these small values, the algorithm takes about 2 hours to run. The algorithm is written in Python and run on a Linux machine with a i7 processor. In order to increase the speed of the program it is necessary to write the program in a low level language such as C of Fortran.

In the figure 6.17 we computed the value functions and the option prices for $C = 5000$. The value functions are smooth and, as expected, the option price (defined in (5.13)) corresponding to the horizontal distance between the value functions, is not affected too much by the initial wealth. The high value of the credit availability C has to be intended as a low default probability. Therefore the problem leads back to the problem with 3 variables. And the numerical results confirm it.

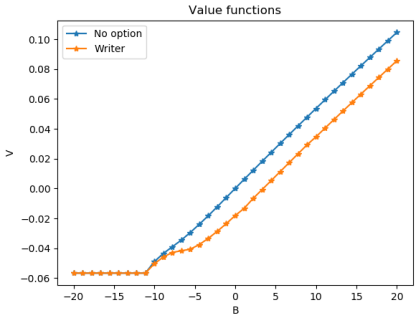

Figure 6.18: Writer and no option value functions at t_0 with $C = 10$, $Y_0 = 0$ and $S_0 = 15$.

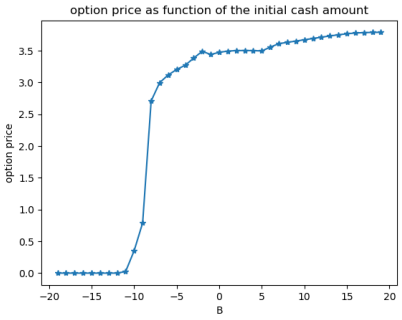

Figure 6.19: Option price as function of B_0 at t_0 with $C = 10$, $Y_0 = 0$ and $S_0 = 15$.

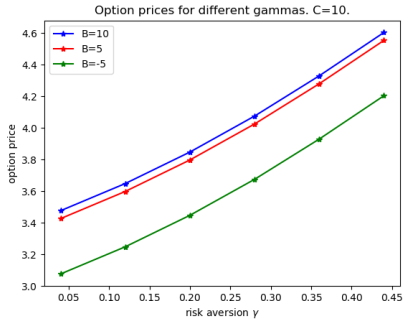

Figure 6.20: Option price as function of risk aversion, for several values of initial B_0. We set $C = 10$, $Y_0 = 0$ and $S_0 = 15$.

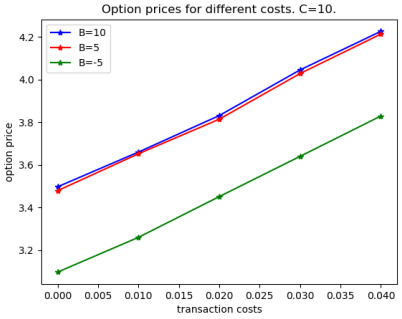

Figure 6.21: Option price as function of transaction costs, for several values of initial B_0. We set $C = 10$, $Y_0 = 0$ and $S_0 = 15$.

A different story happens when we choose a small credit availability. For example $C = 10$.
In figure 6.18 we can see that the two value functions have the same value for $B_0 < -C$, because in this region the value function corresponds to the boundary conditions (remember that we are considering $Y_0 = 0$). For higher values of B_0, the influence of the boundary conditions decreases and the value functions look like those in figure 6.17.

It is important to stress that the grid resolution we used in these example is quite rough. The numerical results are still good, but it is not possible at the moment to study the convergence properties of the algorithm. Furthermore, although the value function is highly non-linear, we used linear interpolation to interpolate the missing points in the grid, and this may create further errors in the shape of the function and in the final result.

We conclude this section by testing the model properties as we did in Section

Convergence table	
$N = \bar{M}$	Price
50	1.96121076540
100	1.96889900730
200	1.97154200723
400	1.97288296067
800	1.97354995910
1000	1.97368292970
1600	1.97388226442
2000	1.97394872091

Table 6.8: Convergence table for ATM VG prices with parameters in Tab 2.3, calculated with the multinomial method.

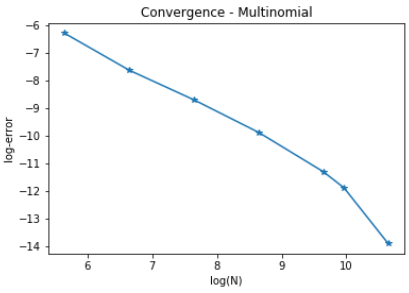

Table 6.9: Multinomial VG writer and buyer prices for different transaction costs.

6.3.5. In figures 6.20 and 6.21 we present several values of option prices as function of the risk aversion and transaction costs respectively. We can see that the value of B_0 does not affect the shape of the curve, but only its height. As we saw in figures 6.13 and 6.14, the option price is an increasing function of the risk aversion and of the transaction costs.

6.5 Multinomial method applied to the reduced problem

In section 6.1 we have seen that a possible technique to construct the Markov chain approximation is to discretize the infinitesimal generator of the continuous process by an explicit finite difference scheme. For Lévy processes with infinite activity, however, this technique cannot be used directly, and we need to consider the infinitesimal generator of the approximated jump-diffusion process. In section 2.2.4, we have seen that with this approximation the convergence is very slowly.

In this section we use the multinomial approximation method presented in chapter 3, in place of the discretization proposed in this chapter. With this method the total computational complexity is reduced because the number of branches is kept fixed. Although the method still relies on an approximation (the VG process is approximated by a general jump process with only the first four moments equal), the number of branches is fixed to $\bar{L} = 5$, and therefore the computational complexity is reduced by a factor $\sqrt{N}$. Recall that the complexity algorithm 1 is

$$\mathcal{O}\left((N+1)\Big[\frac{N(\bar{L}-1)}{2}+1\Big]\times\bar{M}\times\bar{M}\right).$$

Assuming $\bar{M} = N$ and $\bar{L} = 5$, the computational time is reduced to $\mathcal{O}(N^4)$. We use the values in table 2.3 and follow the discretization scheme proposed in chapter 3. In table 6.8 we reported prices for different values of N. In figure 6.9 we plot the log-error as a function of the logarithm of N. We can see that there is a linear relation $\log_2(\varepsilon) \sim p \log_2(N)$ between them with a rate of convergence

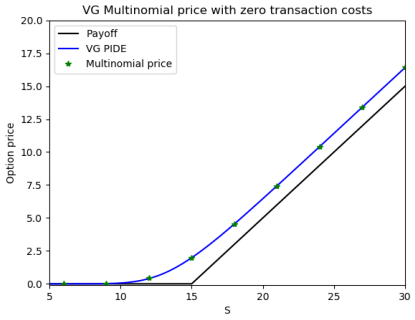

Figure 6.22: The multinomial VG prices for zero transaction costs agree with the solution of the VG PIDE (continuous line).

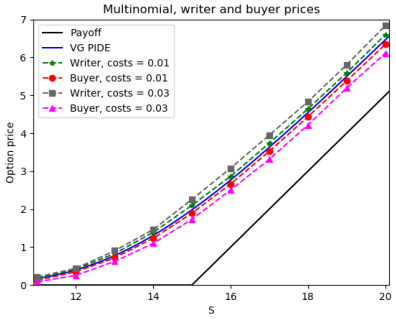

Figure 6.23: Multinomial VG writer and buyer prices for different transaction costs.

$p \approx 1.6$. In the figures 6.22 and 6.23 we show how good are the prices computed with the multinomial method.

6.6 Final conclusions

The main objective of the thesis is to develop a new model for pricing options in presence of proportional transaction costs. The main difference between this work and the previous works in the literature is that we considered a stock dynamics that follows a generic exponential Lévy process.

In chapter 5 we derived the general HJB equation of the model, eq. (5.21), which is a complicated equation with three state variables and a time variable. Therefore we opted to simplify the problem by reducing the number of variables, obtaining the simpler HJB Eq. (5.34). In the end of chapter 6 we developed the algorithm [2] to solve the maximization problem (5.20) associated to the HJB eq. (5.21) and presented numerical results.

However, more emphasis has been given to the reduced problem (5.27), which is also the main topic of the paper [Cantarutti et al., 2018]. We proposed a monotone, stable and consistent numerical scheme and proved that its solution converges to the viscosity solution of the HJB equation (5.34). Numerical results of the equation (5.34) are obtained for the particular cases of diffusion, Merton jump-diffusion and Variance Gamma process, although any Lévy process satisfying the conditions **EM** can be used. The transition probabilities in the Markov chain approximation are obtained by explicit finite difference discretization of the infinitesimal generator of the process. The Brownian motion and the Merton process can be discretized directly, while the VG process needs to be approximated to remove the infinite activity jump component. Due to this approximation, the algorithm [1] has slower convergence when applied to the VG dynamics. Using numerical experiments, we confirmed some features of the model such as the dependency degree of the price on the transaction costs, the risk aversion and the drift. In the end of chapter 6 we also computed the

option prices by using the multinomial method, i.e. we derived the transition probabilities from the approximation introduced in chapter 3.

To conclude, we want to mention some possible future improvements in this area of research. An interesting direction can be the development of a more efficient numerical method for the HJB equations (5.34) and in particular (5.21), which has a hige time complexity. There are several approaches in the literature to solve variational inequalities, such as the policy iteration method of [Forsyth and Huang, 2012b], or the penalty method of [Forsyth and Huang, 2012a], [Wang and Li, 2014]. We argue that an implicit/explicit (IMEX), with the possible help of the Fast-Fourier-Transform for evaluating the integral term, (as in [Andersen and Andreasen, 2000] for instance) can increase the efficiency of the numerical method. Also, using a non-uniform grid as in [Haentjens, 2013], can help to improve the efficiency and reduce the computational cost for both the differential and the integral part.

Acknowledgements

First of all, I want to thank my PhD supervisors Prof. João Guerra and Prof. Manuel Guerra for the patience they had with me, for the useful discussions and the many emails we have exchanged in recent years. I would like to express my thanks to Prof. Maria do Rosário Grossinho, who gave me the opportunity to start this doctoral program together with the European network STRIKE `http://www.itn-strike.eu/` and who has always given me positive energy. Last, but not least, I want to thank my family and my friends Pedro Polvora, Yaser Faghan Kord, Radoslav Vulkov, Daniele Polencic and Alessandro Scalia for their never-ending support.

This research was supported by the European Union in the FP7-PEOPLE-2012-ITN project STRIKE - Novel Methods in Computational Finance (304617), and by CEMAPRE MULTI/00491, financed by FCT/MEC through Portuguese national funds.

Bibliography

[Ait-Sahalia and Jacod, 2012] Ait-Sahalia, Y. and Jacod, J. (2012). Analyzing the spectrum of asset returns: Jump and volatility components in high frequency data. *Journal of Economic literature*, 50(4):1007–1050.

[Almendral, 2005] Almendral, A. (2005). Numerical valuation of american options under the CGMY process. *Eds., Exotic Option Pricing and Advanced Levy Models, Wiley, Hoboken.*

[Almendral and Oosterlee, 2007] Almendral, A. and Oosterlee, C. W. (2007). On American options under the Variance Gamma process. *Applied Mathematical Finance*, 14(2):131–152.

[Andersen and Andreasen, 2000] Andersen, L. and Andreasen, J. (2000). Jump-diffusion processes: Volatility smile fitting and numerical methods for pricing. *Rev. Derivatives Research*, 4:231 – 262.

[Applebaum, 2009] Applebaum, D. (2009). *Lévy Processes and Stochastic Calculus*. Cambridge University Press; 2nd edition.

[Barles and Imbert, 2008] Barles, G. and Imbert, C. (2008). Second-order elliptic integro-differential equations: viscosity solutions' theory revisited. *Annales de l'Institut Henri Poincare (C) Non Linear Analysis*, 25(3):567–585.

[Barles and Soner, 1998] Barles, G. and Soner, H. M. (1998). Option pricing with transaction costs and a nonlinear Black-Scholes equation. *Finance and Stochastics*, 2(4):369–397.

[Barles and Souganidis, 1991] Barles, G. and Souganidis, P. (1991). Convergence of approximation schemes for fully nonlinear second order equations. *J. Asymptotic Analysis*, 4:271–283.

[Barndorff-Nielsen, 1997] Barndorff-Nielsen (1997). Processes of Normal inverse Gaussian type. *Finance and Stochastics*, 2:41–68.

[Bellman, 1957] Bellman, R. (1957). *Dynamic Programming*. Princeton University Press.

[Benth et al., 2001] Benth, F., Karlsen, K., and Reikvam, K. (2001). Optimal portfolio selection with consumption and nonlinear integro-differential equations with gradient constraint: a viscosity solution approach. *Finance and Stochastics*, 5(3):275–303.

[Benth et al., 2002] Benth, F., Karlsen, K., and Reikvam, K. (2002). Portfolio optimization in a Lévy market with intertemporal substitution and transaction costs. *Stochastics and Stochastics Report*, 74(3-4):517–569.

[Björk, 2009] Björk, T. (2009). *Arbitrage theory in continuous time.* Oxford University Press; 3 edition.

[Black and Scholes, 1973] Black, F. and Scholes, M. (1973). The pricing of options and corporate liabilities. *The Journal of Political Economy*, 81(3):637–654.

[Boyle and Vorst, 1992] Boyle, P. and Vorst, T. (1992). Option replication in discrete time with transaction costs. *The Journal of Finance*, 47(1):271–293.

[Cantarutti and Guerra, 2018] Cantarutti, N. and Guerra, J. (2018). Multinomial method for option pricing under Variance Gamma. *International Journal of Computer Mathematics.*

[Cantarutti et al., 2018] Cantarutti, N., Guerra, J., Guerra, M., and Grossinho, M. (2018). Option pricing in exponential Lévy models with transaction costs. *Arxiv.*

[Carmona (Editor), 2009] Carmona (Editor), R. (2009). *Indifference Pricing.* Princeton University Press.

[Carr et al., 2002] Carr, P., Geman, H., D.B., M., and M., Y. (2002). The fine structure of asset returns: An empirical investigation. *Journal of Business*, 75(2):305–333.

[Carr and Madan, 1998] Carr, P. and Madan, D. (1998). Option valuation using the Fast Fourier transform. *Journal of Computational Finance*, 2:61–73.

[Cherny and Shiryaev, 2002] Cherny, A. and Shiryaev, A. (2002). Vector stochastic integrals and the fundamental theorems of asset pricing. *Proceedings of Steklov Institute of mathematics*, 237:6–49.

[Clewlow and Hodges, 1997] Clewlow, L. and Hodges, S. (1997). Optimal delta hedging under transaction costs. *Jornal of Economic Dynamics and Control*, 21:1353–1376.

[Constantinides, 1986] Constantinides, G. (1986). Capital market equilibrium with transaction costs. *Journal of Political Economy*, 94:842 – 862.

[Cont, 2001] Cont, R. (2001). Empirical properties of asset returns: stylized facts and statistical issues. *Quantitative Finance*, 1(2):223–236.

[Cont et al., 1997] Cont, R., Potters, M., and Bouchaud, J. (1997). Scaling in stock market data: stable laws and beyond. *Scale invariance and beyond, Springer.*

[Cont and Tankov, 2003] Cont, R. and Tankov, P. (2003). *Financial Modelling with Jump Processes*. Chapman and Hall/CRC; 1 edition.

[Cont and Voltchkova, 2005a] Cont, R. and Voltchkova, E. (2005a). A finite difference scheme for option pricing in jump diffusion and exponential Lévy models. *SIAM Journal of numerical analysis*, 43(4):1596–1626.

[Cont and Voltchkova, 2005b] Cont, R. and Voltchkova, E. (2005b). Integro-differential equations for option prices in exponential Lèvy models. *Finance and Stochastics*, 9:299–325.

[Cox et al., 1979] Cox, J., Ross, S., and Rubinstein, M. (1979). Option pricing: A simplified approach. *Journal of financial economics*, 7:229–263.

[Crandall et al., 1992] Crandall, M., Ishii, H., and Lions, P.-L. (1992). User's guide to viscosity solutions of second order partial differential equations. *Bulletin of the american mathematical society*, 27(1).

[Crandall and Lions, 1983] Crandall, M. and Lions, P. (1983). Viscosity solutions of Hamilton-Jacobi equations. *Transactions of the AMS*, pages 1–42.

[Dai et al., 2010] Dai, M., Xu, Z. Q., and Zhou, X. Y. (2010). Continuous-time markowitz's model with transaction costs. *SIAM Journal of Financial Mathematics*, 1(1):96–125.

[Dai and Yi, 2009] Dai, M. and Yi, F. H. (2009). Finite horizon optimal investment with transaction cost: A parabolic double obstacle problem. *Journal of Differential Equation*, 246:1445–1469.

[Damgaard, 1998] Damgaard, A. (1998). *Optimal Portfolio Choice and Utility Based Option Pricing in Markets with Transaction Costs*. PhD thesis, School of Business and Economics, Odense University.

[Davis and Norman, 1990] Davis, M. H. A. and Norman, A. R. (1990). Portfolio selection with transaction costs. *Math. Oper. Res.*, 15(4):676–713.

[Davis and Panas, 1994] Davis, M. H. A. and Panas, V. G. (1994). The writing price of a European contingent claim under proportional transaction costs. *Computational and Applied Mathematics*, 13(2):0101–8205.

[Davis et al., 1993] Davis, M. H. A., Panas, V. G., and Zariphopoulou, T. (1993). European option pricing with transaction costs. *SIAM J. Control Optim.*, 31(2):470–493.

[De Vallière et al., 2016] De Vallière, D., Kabanov, Y., and Lépinette, E. (2016). Consumption-investment problem with transaction costs for lévy-driven price processes. *Finance and Stochastics*, 20(3):705–740.

[Delbaen and Schachermayer, 1998] Delbaen, F. and Schachermayer, W. (1998). The fundamental theorem of asset pricing for unbounded stochastic processes. *Mathematische Annalen*, 312(2):215–250.

[Eberlein and Keller, 1995] Eberlein, E. and Keller, U. (1995). Hyperbolic distributions in finance. *Bernoulli*, 1(3):281–299.

[Ehrhardt et al., 2017] Ehrhardt, M., Günther, M., and Ter Maten, J. (2017). *Novel Methods in Computational Finance*. Springer.

[Ethier and Kurtz, 2005] Ethier, S. and Kurtz, T. (2005). *Markov Processes: Characterization And Convergence*. Wiley-Interscience; 2nd Revised edition.

[Fleming and Soner, 2005] Fleming, W. H. and Soner, M. H. (2005). *Controlled Markov Processes and Viscosity Solutions*. Springer; 2nd edition.

[Florescu et al., 2014] Florescu, I., Mariani, M. C., and Sengupta, I. (2014). Option pricing with transaction costs and stochastic volatility. *Electronic Journal of Differential Equations*, 2014(165):1–19.

[Forsyth and Huang, 2012a] Forsyth, P. and Huang, Y. (2012a). Analysis of a penalty method for pricing a guaranteed minimum withdrawal benefit (gmwb). *IMA, Journal of Numerical Analysis*, 32:320–351.

[Forsyth and Huang, 2012b] Forsyth, P. and Huang, Y. (2012b). Iterative methods for solution of the singular control formulation of a gmwb pricing problem. *Numerische Mathematik*, 122:133–167.

[Framstad et al., 1999] Framstad, N., Øksendal, B., and Sulem, A. (1999). Optimal consumption and portfolio in a jump diffusion market with proportional transaction costs. *Journal of Mathematical Economics*, 35:233–257.

[Fu, 2000] Fu, M. C. (2000). Variance-Gamma and Monte Carlo. *Advances in Mathematical Finance.*, pages 21–34.

[Gihman and Skorohod, 1979] Gihman, I. and Skorohod, A. (1979). *Controlled stochastic processes*. New York, Springer Verlang.

[Goldys and Wu, 2016] Goldys, B. and Wu, W. (2016). Dynamic programming principle for stochastic control problems driven by general Lévy noise. *Stochastic Analysis and Applications*, 34(6):1083 – 1093.

[Grinold and Kahn, 1999] Grinold, R. C. and Kahn, R. N. (1999). *Active Portfolio Management*. McGraw-Hill Education; 2 edition.

[Haentjens, 2013] Haentjens, T. (2013). Efficient and stable numerical solution of the heston–cox–ingersoll–ross partial differential equation by alternating direction implicit finite difference schemes. *International Journal of Computer Mathematics*, 90(11):2409–2430.

[Hainaut and MacGilchrist, 2010] Hainaut, D. and MacGilchrist, R. (Winter 2010). An interest tree driven by a lévy process. *The Journal of Derivatives*, 18(2):33–45.

[Harrison and Kreps, 1979] Harrison, J. and Kreps, D. (1979). Martingales and arbitrage in multiperiod security markets. *Journal of Economic theory*, 2:381–408.

[Harrison and Pliska, 1981] Harrison, J. and Pliska, S. R. (1981). Martingales and stochastic integrals in the theory of continuous trading. *Stochastic processes and their applications*, 11(3):215–260.

[Hirsa and Madan, 2001] Hirsa, A. and Madan, D. (2001). Pricing American options under Variance Gamma. *Journal of Computational Finance*, 7.

[Hodges and Neuberger, 1989] Hodges, S. D. and Neuberger, A. (1989). Optimal replication of contingent claims under transaction costs. *The Review of Future Markets*, 8(2):222–239.

[Ishii, 1989] Ishii, H. (1989). On uniqueness and existence of viscosity solutions of fully nonlinear second-order elliptic pdes. *Comm. Pure and Appl. Math.*, 42(1):15–45.

[Ishii and P.L., 1990] Ishii, H. and P.L., L. (1990). Viscosity solutions of fully nonlinear second-order elliptic partial differential equations. *J. Differential Equations*, 83(1):26–78.

[Jacod and Shiryaev, 2002] Jacod, J. and Shiryaev, A. N. (2002). *Limit Theorems for Stochastic Processes*. Springer; 2nd ed.

[Jensen, 1988] Jensen, R. (1988). The maximum principle for viscosity solutions of fully nonlinear second-order partial differential equations. *Archive for Rational Mechanics and Analysis*, 101(1):1–27.

[Jouini and Pilska, 2001] Jouini, E. and Pilska, S. (2001). *Option pricing, Interest rates and risk Management*. Cambridge university press.

[Kabanov, 2001] Kabanov, Y. (2001). Arbitrage theory. *in Jouini and Pilska [Jouini and Pilska, 2001]*, pages 3–42.

[Kabasinskas et al., 2009] Kabasinskas, A., Rachev, S., Sakalauskas, L., Wei, S., and Belovas, I. (2009). Alpha-stable paradigm in financial markets. *Journal of Computational Analysis and Applications*, 11(4):641–669.

[Kellezi and Webber, 2006] Kellezi, E. and Webber, N. (2006). Valuing Bermudan options when asset returns are Lévy processes. *Quantitative Finance*, 4(1):87–100.

[Kou, 2002] Kou, S. (2002). A jump-diffusion model for option pricing. *Management Science*, 48(8):1086–1101.

[Kushner and Dupuis, 2001] Kushner, H. and Dupuis, P. G. (2001). *Numerical Methods for Stochastic Control Problems in Continuous Time*. Springer; 2nd ed.

[Kushner and Martins, 1991] Kushner, H. J. and Martins, F. L. (1991). Numerical methods for stochastic singular control problems. *SIAM Journal of Control and Optimization*, 29(6):1443–1475.

[Leland, 1985] Leland, H. (1985). Option pricing and replication with transaction costs. *The Journal of Finance*, 40(5):1283–1301.

[Lions, 1983] Lions, P. L. (1983). Optimal control for diffusion processes and Hamilton-Jacobi-Bellman equations. *Comm. PDE*, 8:1101–1174.

[Liu and M., 2002] Liu, H. and M., L. (2002). Optimal portfolio selection with transaction costs and finite horizons. *Review of Financial Studies*, 15:805–835.

[Liu and M., 2007] Liu, H. and M., L. (2007). Optimal portfolio selection with transaction costs and event risk. *AFA 2008 New Orleans*.

[Madan et al., 1998] Madan, D., Carr, P., and Chang, E. (1998). The Variance Gamma process and option pricing. *European Finance Review*, 2:79–105.

[Madan and Milne, 1991] Madan, D. and Milne, F. (1991). Option pricing with V.G. martingale components. *Mathematical Finance*, 1:39–55.

[Madan and Seneta, 1990] Madan, D. and Seneta, E. (1990). The Variance Gamma (V.G.) model for share market returns. *The journal of Business*, 63(4):511–524.

[Maller et al., 2006] Maller, R., Solomon, D., and Szimayer, A. (2006). A multinomial approximation for american option prices in levy process models. *Mathematical Finance*, 16(4):613–633.

[Mandelbrot, 1963] Mandelbrot, B. (1963). Modeling financial data with stable distributions. *Journal of Business*, XXXVI(1):392–417.

[Merton, 1976] Merton, R. (1976). Option pricing when underlying stock returns are discontinuous. *Journal of Financial Economics*, 3:125–144.

[Merton, 1969] Merton, R. C. (1969). Lifetime portfolio selection under uncertainty: The continuous-time case. *The Review of Economics and Statistics*, 51(3):247–257.

[Merton, 1971] Merton, R. C. (1971). Optimum consumption and portfolio rules in a continuous time model. *Journal of economic theory*, 3:373–413.

[Mocioalca, 2007] Mocioalca, O. (2007). Jump diffusion options with transaction costs. *Revue Roumaine de Mathematiques Pures et Appliquees*, 52(3):349–366.

[Monoyios, 2003] Monoyios, M. (2003). Efficient option pricing with transaction costs. *Journal of Computational Finance*, 7:107–128.

[Monoyios, 2004] Monoyios, M. (2004). Option pricing with transaction costs using a Markov chain approximation. *Jornal of Economic Dynamics and Control*, 28:889–913.

[Musiela and Rutkowski, 2005] Musiela, M. and Rutkowski, M. (2005). *Martingale Methods in Financial Modeling*. Springer.

[Øksendal and Sulem, 2007] Øksendal, B. and Sulem, A. (2007). *Applied Stochastic Control of Jump Diffusions*. Springer; 2nd edition.

[Palczewsky et al., 2015] Palczewsky, J., Poulsen, R., Schenk-Hoppé, K. R., and Wang, H. (2015). Dynamic portfolio optimization with transaction costs and state-dependent drift. *European journal of operational research*, 243:921–931.

[Pham, 1998] Pham, H. (1998). Optimal stopping of controlled jump diffusion processes a viscosity solution approach. *Journal of mathematical system, Estimation and Control*, 8(1):1–27.

[Pham, 2009] Pham, H. (2009). *Continuous-time Stochastic Control and Optimization with Financial Applications*. Springer; 1st ed.

[Protter, 2004] Protter, P. E. (2004). *Stochastic integration and differential equations*. Springer-Verlag, Berlin, second edition.

[Sato, 1999] Sato, K. I. (1999). *Lévy processes and infinitely divisible distributions*. Cambridge University Press.

[Sayah, 1991] Sayah, A. (1991). Equations d'Hamilton-Jacobi du primer ordre avec termes intégro-differentiels: Parties 1 et 2. *Comm. PDE*, 10:1057 – 1093.

[Schachermayer, 2002] Schachermayer, W. (2002). No arbitrage: on the work of david kreps. *Positivity*, 6:1057–1093.

[Schoutens, 2003] Schoutens, W. (2003). *Lévy processes in finance*. Wiley, First Edition.

[Seneta, 2004] Seneta, E. (2004). Fitting the Variance-Gamma model to financial data. *Journal of Applied Probability*, 41:177–187.

[Sengupta, 2014] Sengupta, I. (2014). Option pricing with transaction costs and stochastic interest rate. *Applied Mathematical Finance*, 21(5):399–416.

[Shreve and Soner, 1994] Shreve, S. E. and Soner, H. M. (1994). Optimal investment and consumption with transaction costs. *Annals of Applied Probability*, 4(3):609–692.

[Soner, 1986a] Soner, H. (1986a). Jump Markov processes and viscosity solutions. *IMA Vols in Math. and Applic.*, 10:501 – 511.

[Soner, 1986b] Soner, H. M. (1986b). Optimal control with state space constraint. *Siam Journal of Control and Optimization*, 24:1110–122.

[Ssebugenyi and Konlack, 2013] Ssebugenyi, C.S. Mwaniki, I. and Konlack, V. (2013). On the minimal entropy martingale measure and multinomial lattices with cumulants. *Applied Mathematical Finance*, 20(4):359–379.

[Wang and Li, 2014] Wang, S. and Li, W. (2014). A numerical method for pricing european options with proportional transaction costs. *Journal of Global Optimization*, 60(1):59–78.

[Whalley and Wilmott, 1997] Whalley, A. E. and Wilmott, P. (1997). An asymptotic analysis of an optimal hedging model for option pricing with transaction costs. *Mathematical Finance*, 7(3):307–324.

[Yamada and Primbs, 2001] Yamada, Y. and Primbs, J. (2001). Construction of multinomial lattice random walks for optimal hedges. *Proc. of the 2001 international conference of Computational Science*, pages 579–588.

[Yamada and Primbs, 2003] Yamada, Y. and Primbs, J. (2003). Mean square optimal hedges using higher order moments. *Proc of the 2003 international conference on Computational Intelligence for Financial Engineering*.

[Yamada and Primbs, 2004] Yamada, Y. and Primbs, J. (2004). Properties of multinomial lattices with cumulants for option pricing and hedging. *Asia-Pacific Financial Markets*, 11:335–365.