OPTIMAL CONTROL WITH JUMP-DIFFUSIONS

CARLOS RODRIGUEZ-VEGA

Dissertation submitted for the MSc in Mathematical Finance

Department of Mathematics
University of York

Supervisor: Dr. Maciej Capiński

To Lana.

Genius is eternal patience.
Michelangelo

... never forget that until the day when God shall deign to reveal the future to man, all human wisdom is summed up in these two words, –'Wait and hope.'
Alexandre Dumas - The Count of Monte Cristo

First of all, I would like to express my gratitude to my supervisor, Maciej Capiński, for his support and feedback on my drafts. Thanks as well to Marek Capiński for suggesting me such an interesting and challenging topic. Similarly as with my previous two dissertations, I would not have been able to complete this one without the support of my parents and brothers. I am counting on you for dissertations to come in the future! Thanks also to Graham, Razvan and Cookie for giving me some distraction when I was not able to make sense of the math. And finally, words can barely describe how much in debt I feel to my girlfriend, Lana. You have been the wind pushing this boat to safe coast; you are where I found my heaven. Thank you so much for your patience, your understanding, and for putting up with me when I tried to optimise even the Saturday morning shopping.

Contents

1. Introduction 6
2. Stochastic analysis 8
2.1. Basic concepts of stochastic processes 8
2.2. Stochastic integration 10
2.3. Stochastic calculus 14
3. Jump processes 16
3.1. Pure jump processes 16
3.2. Jump-diffusion processes 24
3.3. Stochastic calculus for jump-diffusion processes 26
3.4. Stochastic exponential 31
4. Stochastic control 37
4.1. Elements of a stochastic control problem 37
4.2. The Hamilton-Jacobi-Bellman equation 39
5. Risk-sensitive stochastic control 41
5.1. Analytical setting 43
5.2. Problem setup 47
5.3. Analysis of the solution 52
6. Numerical approximation 56
6.1. Discretisation of the Hamilton-Jacobi-Bellman equation 56
6.2. Discretisation of the integral operator 58
6.3. Terminal and boundary conditions 60
6.4. The complete numerical scheme 60
6.5. Convergence of the numerical scheme 61
6.6. Numerical results with two factor processes 65
6.7. Numerical results with one factor process 68
Appendix A. MATLAB code 71
A.1. Two-factor implementation 71
A.2. One-factor implementation 79
References 84

1. Introduction

One of the central problems in mathematical finance concerns portfolio optimisation: given a set of available investment opportunities with different risk characteristics, how should we allocate capital in an optimal way? This question poses also the challenge of defining what optimality means for each investor and how it can be represented mathematically. Since risk apetite and investment goals may vary among investors, a suitable criterion –describing the preferences of the investor and the optimisation goal– must be carefully chosen.

The first contribution to the area of portfolio optimisation was done by Markowitz [21]. He suggested that investors should consider not only the expected return of the assets, but also the volatility of the asset prices. The volatility accounts for the risk associated with the assets and it is represented as the standard deviation of the asset returns. Markowitz came to the conclusion that capital is allocated optimaly when the proportions invested in the assets are such that the expected portfolio return is maximised for a given amount of risk.

The framework developed by Markowitz –known as the *mean-variance model*– stands out because it is both sufficiently general for a considerable range of practical situations and simple enough to perform theoretical analysis and numerical calculations. However, it has not been exempt from criticism since its inception. In Markowitz's original setup the portfolio weights are chosen assuming a single investment period. This is not practical for long-term investment planning, where investors are able to re-balance their portfolio frequently –considerable effort has been put on multi-period [20] and continuous-time [6] extensions, though–. The mean-variance model also assumes that investors' goal is to maximise the mean return and minimise the variance of the return. However, this assumption was challenged by the work of von Neumann and Morgenstern and their theory of *expected utility* [25], which deals with choice under uncertainty. Under some conditions, they argue that investors' preferences can be characterised by maximising the expected value of a function called *utility*. The utility function describes the investor's degree of satisfaction with the outcome of terminal wealth. It has two key properties: an upward slope, since a larger expected value of the utility function is always preferable, and a curvature, which measures the investor's attitude towards risk. In general, the more concave the utility function, the more risk-averse the investor will be; and the more convex the utility function, the more risk-seeking behavior the investor will show. The intermediate case is that of a linear utility function. Here the investor is risk-neutral and she does not care about the riskiness of her portfolio, only about its final expected value. Thus, our optimisation problem would ideally incorporate a multi-period setting, which allows asset positions to change as new information is available, and a criterion reflecting the investors' different degrees of risk aversion.

In general, maximising the expected utility of terminal wealth by choosing the portfolio weights is a complicated problem. However, by borrowing from the field of control engineering, a way forward is possible. Control engineering is the discipline that focuses on the mathematical modelling of dynamic systems and uses control theory to make the systems behave in a desired way. This definition fits well the problem of multi-period portfolio management, where we wish to make asset allocation decisions so that the portfolio behaves in a wanted manner with regard to risk and return. Since asset price dynamics are usually described by stochastic

differential equations, we capitalise on the approach of *stochastic control*, which is a subfield of control theory that deals with the existence of uncertainty either in observations of the state of the system or in the underlying process driving the evolution of the state. Stochastic control states the mathematical conditions of optimality given the model that describes the dynamic system and the criterion we want to optimise. When the conditions for optimality are met, a control is found that reacts optimally with respect to the criterion, new available information, the remaining investment horizon, and the assumed stochastic dynamics of our portfolio and assets.

This kind of optimisation problem with a continuous-time model driven by Brownian motion was first studied by Merton [22, 23], who provided one of the few explicit solutions of a Hamilton-Jacobi-Bellman equation in stochastic control. Since Merton's work there has been substantial research that relaxes some of his original assumptions in order to make the model more realistic: usage of stochastic interest rates, stochastic asset volatility and growth rate, transaction costs, etc. Remarkable contributions have been done by Bielecki and Pliska [4, 5], whose work has served as basis for Davis and Lleo [12, 13]. Using, in turn, the work carried out by Davis and Lleo in [12] as the main reference of this dissertation, we will study a portfolio optimisation problem where asset price dynamics are modelled as jump-diffusions. The uncertainty of asset returns is taken into account by including an external factor that may describe macroeconomic circumstances. As opposed to classical expected utility theory –where risk tolerance depends implicitly on the curvature of the utility function–, we use a criterion that includes a risk-sensitivity parameter that influences directly the outcome of the optimisation. The investor's goal is then to maximise her risk-adjusted portfolio return by maximising this criterion. Since the wealth process is modelled as a jump-diffusion, this would result in a Hamilton-Jacobi-Bellman partial integro-differential equation. However, applying a change measure will allow us to obtain an equivalent optimal control problem in the auxiliary factor process, which is a diffusion process and has no jumps, deriving therefore a Hamilton-Jacobi-Bellman partial differential equation instead.

The dissertation is organised as follows. Immediately after the introduction the reader will be introduced to some of the theory needed for the mathematical developments in this dissertation, with emphasis on random processes and stochastic calculus. The experienced reader may want to skip it, while others may read it as an introduction to new topics or use it for reference.

Section 3 deals with the topic of processes with jumps. Pure jump processes, such as the Poisson and compound Poisson processes, will be introduced first, and their relationship with Poisson random measures explained in detail. The characteristics of jump-diffusions are analysed within the framework of Lévy processes; and the applicable stochastic calculus theory is presented within the broader category of semimartingales.

In Section 3 we also cover the theory of stochastic exponential processes and their application in defining equivalent probability measures. The stochastic exponential process can be used as a Radon–Nikodým density process in Girsanov Theorem, but that requires its product with another Lévy-Itô process to be a local martingale in the original probability space. We present the condition that must be satisfied in order to fulfill such requirement. To our knowledge, this proof can not be found in the literature looked up for this dissertation.

Section 4 defines the most important concepts in stochastic control and derives the Hamilton-Jacobi-Bellman equation for a simple optimisation problem with a difussion process.

Section 5 contains the main results developed in this dissertation. We derive solutions to the stochastic differential equations that model the dynamics of the factor process, the money market account, the risky securities and the wealth of the portfolio. Since the driving stochastic process of the portfolio optimisation problem is a Lévy process, the Hamilton-Jacobi-Bellman equation arising from the problem would be a partial integro-differential equation. In order to avoid the extra complexity of an integral term involving the value function, we will show how, by a change of measure, Davis and Lleo obtain an alternative criterion in the factor process –and thus without jumps–, and find an equivalent optimal control problem. The change of measure relies upon the stochastic exponential and, as a necessary condition to be able to apply Girsanov theorem, Davis and Lleo show that it is a martingale.

Although not strictly needed –since Davis and Lleo prove the existence of a classical solution in [12] and every classical solution is a viscosity solution–, we prove that the optimisation problem actually admits a unique viscosity solution by establishing a comparison theorem, where the key is to find a modulus of continuity.

In Section 6 we construct two finite difference numerical schemes for the risk-sensitive optimisation problem. The numerical schemes approximate the solution of the Hamilton-Jacobi-Bellman equation over a reduced domain. We consider two cases: one factor process and one risky security; and two factor processes and two risky securities. We prove that the numerical schemes are consistent, stable and monotone, which is enough to ensure that the numerical schemes converge uniformly to the viscosity solution of the risk-sensitive optimisation problem on compact subsets. Both numerical schemes were implemented in MATLAB and the results of a set of numerical experiments are presented and discussed. Code is listed in Appendix A.

Apart from the proof of the Hamilton-Jacobi-Bellman equation in Subsection 4.2, none of the proofs given in this dissertation are direct copies of proofs found elsewhere, although many of them are inspired by similar proofs found in other papers and books.

2. Stochastic analysis

2.1. Basic concepts of stochastic processes.

We introduce in this section some basic concepts, starting with the definition of *stochastic process*, which is the most common descriptor of any financial product whose value varies with time.

Definition 2.1. (Stochastic process) A stochastic process $(X(t), t \geq 0)$ taking values in $\mathbb{R}$ is a family of $\mathbb{R}$-valued random variables indexed by time defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$.

A stochastic process can also be seen as a function $X : [0, \infty) \times \Omega \to \mathbb{R}$. For every $t \geq 0$, the mapping $X(t) : \Omega \to \mathbb{R}$ is a random variable; and for each ω, the trajectory $X(\omega) : [0, \infty) \to \mathbb{R}$ defines a function called the *sample path* of the process. For convenience, the stochastic process $X(t, \omega)$ will be written as $X(t)$ or $(X(t), t \geq 0)$.

The interpretation of the index t as a time variable causes that, as time goes, more information is progressively revealed to the observer. The information structure is

given by an increasing sequence of sub σ-algebras of $\mathcal{F}$. Thus, we need to add a time-dependent component to the probabability space $(\Omega, \mathcal{F}, \mathbb{P})$ to account for this additional feature. This is usually done using the concept of *filtration*.

Definition 2.2. (Filtration) Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and assume that for every $t \geq 0$ there is a σ-algebra of $\mathcal{F}$. A filtration on $(\Omega, \mathcal{F}, \mathbb{P})$ is an increasing family of sub σ-algebras $(\mathcal{F}_t, t \geq 0)$ such that $\mathcal{F}_s \subseteq \mathcal{F}_t$ whenever $s \leq t$.

Intuitively, a filtration guarantees that at a later time t the sub σ-algebra $\mathcal{F}_t$ contains at least as much information as there is in $\mathcal{F}_s$. A probability space $(\Omega, \mathcal{F}, \mathbb{P})$ that comes equipped with a filtration is called a *filtered probability space*. We will assume from now on that the probability space is always equipped with a filtration.

Definition 2.3. (Adapted process) A stochastic process $(X(t), t \geq 0)$ defined on a probability space is adapted to the filtration (or $\mathcal{F}_t$-adapted) if the random variable $X(t)$ is $\mathcal{F}_t$-measurable for each $t \geq 0$.

If $X(t)$ is $\mathcal{F}_t$-adapted, then $\mathcal{F}_t$ contains all the information needed to know the value of $X(t)$ at time t. Any process $X(t)$ is adapted to its own filtration $\mathcal{F}_t^X = \sigma\{X(s); s \leq t\}$, which is usually called the *natural filtration*.

Definition 2.4. (Progressively measurable process) A process $(X(t), t \geq 0)$ defined on a probability space is progressively measurable if, for every $t \geq 0$, the function $X(t, \omega)$ is $\mathcal{B}([0,t]) \otimes \mathcal{F}_t$-measurable, where $\mathcal{B}([0,t])$ denotes the Borel σ-algebra on $[0,t]$, and $\otimes$ stands for the tensor product.

In the context of gambling, *martingales* are the stochastic processes that capture the notion of a fair game where knowledge of past events never helps predict future winnings. Apart from gambling, they appear naturally in numerous applications of stochastic modelling.

Definition 2.5. (Martingale) Let $(X(t), t \geq 0)$ be an adapted stochastic process defined on a probability space that also satisfies the integrability condition $\mathbb{E}(|X(t)|) < \infty$ for all $t \geq 0$. We say it is a martingale with respect to $\mathcal{F}_t$ if, for all $0 \leq s < t < \infty$,

$$\mathbb{E}(X(t)|\mathcal{F}_s) = X(s), \ a.s.$$

In other words, the best prediction of a martingale's future value is its present value. A common example of a martingale is the Wiener process, which is the stochastic process that serves as source of randomness in the models for financial markets. It is called after mathematician Norbert Wiener, but it is also known as *Brownian motion*, in recognition to the work of botanist Robert Brown. It was first introduced into finance by Louis Bachelier.

Definition 2.6. (Wiener process) A stochastic process $(W(t), t \geq 0)$, defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, is called a Wiener process if it satisfies the following conditions:

(1) $W(0) = 0$, almost surely;
(2) for any $0 \leq s < t$, the random variable $W(t) - W(s)$ has normal distribution with mean 0 and variance $t - s$;
(3) $W(t)$ has independent increments: for any $0 \leq t_1 < t_2 < \cdots < t_n$, the random variables $W(t_1), W(t_2) - W(t_1), \ldots, W(t_n) - W(t_{n-1})$ are independent;

(4) the sample paths of $W(t)$ are continuous for almost all $\omega \in \Omega$.

Lemma 2.7. [8] *An $\mathcal{F}_t$-adapted Wiener process defined on a probability space is a martingale with respect to $\mathcal{F}_t$.*

Since they will be needed for later developments, we will mention some properties of the Wiener process regarding the irregularity of its sample paths. Before that, we need first the definition of *variation*.

Definition 2.8. (Variation) [8] For any process $X(t)$ and $t \in [0,T]$. Let $0 = t_0^n < t_1^n < \ldots < t_{k(n)}^n = t$ be a sequence of partitions of $[0,t]$ and

$$Y_n(t) = \sum_{i=0}^{k(n)-1} \mid X\left(t_{i+1}^n\right) - X\left(t_i^n\right) \mid .$$

The variation of $X(t)$ is the limit in probability of $Y_n(t)$ as $\max_{i=1,\ldots,k(n)-1} |t_{i+1}^n - t_i^n| \to 0$ when $n \to \infty$ for all t.here we could be a bit more precise. In fact, we have a sequance of sequences $t_1^n \ldots t_{k(n)}^n$ with $\max_{i=1,\ldots,k(n)-1} |t_{i+1}^n - t_i^n| \to 0$ as n goes to infinity.

Lemma 2.9. [8] *The sample paths of a Wiener process are almost surely not differentiable at any time t and almost surely of infinite variation.*

Related to the variation, we have also the *quadratic variation*.

Definition 2.10. (Quadratic variation) [8] For any process $X(t)$ and $t \in [0,T]$. Let $0 = t_0^n < t_1^n < \ldots < t_{k(n)}^n = t$ be a sequence of partitions of $[0,t]$ and

$$Y_n(t) = \sum_{i=0}^{k(n)-1} \left(X\left(t_{i+1}^n\right) - X\left(t_i^n\right)\right)^2 .$$

The quadratic variation of $X(t)$, denoted by $[X,X](t)$, is the limit in probability of $Y_n(t)$ as $\max_{i=1,\ldots,k(n)-1} |t_{i+1}^n - t_i^n| \to 0$ when $n \to \infty$ for all t.

Lemma 2.11. [8] *The quadratic variation of the Wiener process is $[W,W](t) = t$.*

The Wiener process is also the simplest and most fundamental *diffusion process*.

Definition 2.12. (Difussion process) Let $(X(t), t \geq 0)$ be a stochastic process defined on a probability space and adapted to the filtration $\mathcal{F}_t$. Then $X(t)$ is a diffusion process if the sample paths are almost surely continuous and it satisfies the Markov property with respect to $\mathcal{F}_t$, i.e. for each $x \in \mathbb{R}$ and any s, t with $s < t$,

$$P\left(X(t) \leq x | \mathcal{F}_s\right) = P\left(X(t) \leq x | X(s)\right).$$

2.2. **Stochastic integration.** Since integration is one of the main operations in calculus, we need a rigorous definition of the integrals of stochastic processes where the integrator is the Wiener process. K. Itô [17] developed the stochastic integration theory needed to give meaning to such integrals. The stochastic integral constructed here it is also known as the Itô integral, and it represents the basis of the Itô stochastic calculus.

Since the sample paths of the Wiener process are almost surely of unbounded variation, the classical Lebesgue-Stieltjes integration theory cannot be applied to define the integral of a stochastic process with respect to the Wiener process. However, since the quadratic variation of the Wiener process is finite, the integral can

be defined as the limit of suitable aproximating sums in $\mathcal{L}^2(\Omega)$ (the space of square integrable real-valued random variables on Ω) [15]. The stochastic integral will be defined for stochastic processes belonging to the space specified in the following Definition.

Definition 2.13. Let $\mathcal{M}^2([0,T])$ be the space of all stochastic processes $f(t)$ defined on a probability space satisfying:

(1) $f(t)$ is adapted to the filtration $\mathcal{F}_t$;
(2) $\mathbb{E}\left(\int_0^T f^2(t)dt\right) < \infty$.

These two conditions are necessary in order to have the martingale property for the stochastic integral [19]. The filtration $\mathcal{F}_t$ satisfies two conditions: for each t, $W(t)$ is $\mathcal{F}_t$-measurable; and for any $s \leq t$, the random variable $W(t) - W(s)$ is independent of the sub σ-algebra $\mathcal{F}_s$.

We will consider initially the interval $[0,T]$ and we will define first the integral for *simple processes* (i.e. processes that are constant on finitely many intervals). By the limiting procedure the integral is then defined for more general processes.

Definition 2.14. (Simple process) An $\mathbb{R}$-valued simple process on $[0,T]$ is a process $f(t)$ for which there is a partition $0 = t_0 < t_1 < \ldots < t_n = T$ and $\mathcal{F}_{t_i}$-measurable random variables ξ_i such that

$$f(t) = \sum_{i=0}^{n-1} \xi_i \mathbb{I}_{(t_i, t_{i+1}]}(t),\ t \in [0,T]\,.$$

Provided that $\mathbb{E}\left(\xi_i^2\right) < \infty$, then the stochastic integral over $[0,T]$ for any simple process $f(t) \in \mathcal{M}^2([0,T])$ is defined as

$$I(f) = \int_0^T f(t)dW(t) = \sum_{i=0}^{n-1} \xi_i \left(W\left(t_{i+1}\right) - W\left(t_i\right)\right), \tag{2.1}$$

which is a random variable in $\mathcal{L}^2(\Omega)$.

Theorem 2.15. [8] *Let $I(f)$ be defined as (2.1). Then $\mathbb{E}\left(I(f)\right) = 0$ and*

$$\mathbb{E}\left[\left(I(f)\right)^2\right] = \mathbb{E}\left[\int_0^T f^2(t)dt\right],$$

which receives the name of Itô isometry.

Equipping the space $\mathcal{M}^2([0,T])$ of integrands with the norm

$$\| f \|_{\mathcal{M}^2([0,T])} = \mathbb{E}\left[\int_0^T f^2(t)dt\right],$$

we can see that the stochastic integral of simple processes is an isometry, since

$$\| I(f) \|_{\mathcal{L}^2(\Omega)} = \mathbb{E}\left[\left(I(f)\right)^2\right] = \mathbb{E}\left[\int_0^T f^2(t)dt\right] = \| f \|_{\mathcal{M}^2([0,T])}\,.$$

The next step is to extend the definition of the stochastic integral to any stochastic process $f(t) \in \mathcal{M}^2([0,T])$, but first we need the following approximation Lemma.

Lemma 2.16. [8] *For every process $f(t) \in \mathcal{M}^2([0,T])$ there exists a sequence of simple processes $f_n(t) \in \mathcal{M}^2([0,T])$ such that*

$$\lim_{n\to\infty} \mathbb{E}\left(\int_0^T \left(f(t) - f_n(t)\right)^2 dt\right) \to 0,\ n \geq 1.$$

By Lemma 2.16 it is reasonable to define the stochastic integral of any stochastic process $f(t) \in \mathcal{M}^2([0,T])$ as

$$I\left(f(t)\right) = \lim_{n\to\infty} I\left(f_n(t)\right), \tag{2.2}$$

but we have to make sure first that the limit exists. In order to prove that the sequence $I\left(f_n(t)\right)$ converges in $\mathcal{L}^2(\Omega)$ we will use the Itô isometry property satisfied by simple processes. For each $n, m \geq 0$ we have

$$\begin{aligned} \| I\left(f_n\right) - I\left(f_m\right) \|_{\mathcal{L}^2(\Omega)} &= \mathbb{E}\left[\left(I\left(f_n\right) - I\left(f_m\right)\right)^2\right] \\ &= \mathbb{E}\left[\int_0^T \left(f_n(t) - f_m(t)\right)^2 dt\right] \to 0,\ n, m \to \infty. \end{aligned}$$

Hence the sequence $I\left(f_n\right)$ is Cauchy in $\mathcal{L}^2(\Omega)$ and thus the limit (2.2) does indeed exist in $\mathcal{L}^2(\Omega)$.

Definition 2.17. The limit defined in (2.2) for any stochastic process $f(t) \in \mathcal{M}^2([0,T])$ is called the stochastic integral of $f(t)$ and is denoted by

$$\int_0^T f(t) dW(t).$$

The extension to any stochastic process $f(t) \in \mathcal{M}^2([0,T])$ has the same properties as the integral of simple processes mentioned in Theorem 2.15.

So far we have considered a fixed terminal time T and defined the integral over the interval $[0,T]$. However, we would like to consider an interval of integration $[0,t]$ for any $t \in [0,T]$, so that the stochastic integral defines a stochastic process. For a stochastic process $f(t) \in \mathcal{M}^2([0,T])$ it can be shown [8] that we can find a sequence $f_n(t)$ of simple processes such that

$$X_n(t) = \int_0^T \left(\mathbb{I}_{[0,t]} f_n\right)(s) dW(s),\ n \geq 1$$

defines a sequence of stochastic processes that are a martingale with respect to $\mathcal{F}_t$ and has almost surely continuous sample paths. The sequence $X_n(t)$ converges in $\mathcal{L}^2(\Omega)$ to the stochastic process

$$X(t) = \int_0^T \left(\mathbb{I}_{[0,t]} f\right)(s) dW(s),$$

which, in turn, converges almost surely to a continuous martingale.

Theorem 2.18. [8] *If $f(t) \in \mathcal{M}^2([0,T])$, then there exists a continuous martingale $M(t)$ such that, for all $t \in [0,T]$,*

$$P\left(M(t) = \int_0^t f(s) dW(s) = \int_0^T \left(\mathbb{I}_{[0,t]} f\right)(s) dW(s)\right) = 1. \tag{2.3}$$

If we require the integrands to belong to $\mathcal{M}^2([0,T])$, then the stochastic integral

$$\int_0^T f(t)dW(t)$$

exists and it is a random variable in $\mathcal{L}^2(\Omega)$. However, in our intention of defining the integral as a stochastic process for every $t \in [0,T]$, we do not necessarily care whether the stochastic integral over $[0,T]$ actually exists (in other words, if the random variable $\int_0^T f(t)dW(t)$ is integrable and its expectation is well-defined). Thus, the condition that $f(t) \in \mathcal{M}^2([0,T])$ is very restrictive and we are now interested in extending the stochastic integral to a larger class of integrands, which is specified in the following Definition.

Definition 2.19. Let $\mathcal{L}^2([0,T])$ be the space of all stochastic processes $f(t)$ on a probability space satisfying:

(1) $f(t)$ is adapted to the filtration $\mathcal{F}_t$;
(2) $\int_0^T f^2(t)dt < \infty$ almost surely.

Among the processes $f(t) \in \mathcal{L}^2([0,T])$ that are not necessarily in the class $\mathcal{M}^2([0,T])$ are those of the form

$$f(t) = u\left(W(t), t\right),$$

where $u(x,t)$ is any function of $x \in \mathbb{R}$ and $t \geq 0$ that is jointly continuous in its two arguments. Such processes occur frequently in stochastic calculus and its applications. Using a method called *localisation* the stochastic integral can be defined for processes belonging to $\mathcal{L}^2([0,T])$. This technique consists of successive truncations at suitable *stopping times*.

Definition 2.20. (Stopping time) A random variable $\tau : \Omega \to [0,T]$ is called a stopping time with respect to filtration $\mathcal{F}_t$ if $\{\omega : \tau(\omega) \leq t\} \in \mathcal{F}_t$ for all $t \geq 0$.

We are interested in finding a sequence of stopping times with the properties listed in the following Definition.

Definition 2.21. [8] Let $f(t) \in \mathcal{L}^2([0,T])$. A sequence of stopping times is localising in $\mathcal{M}^2([0,T])$ for $f(t)$ if:

(1) it is increasing: $\tau_n \leq \tau_{n+1}$, for each $n \geq 1$;
(2) for each $n \geq 1$, $\left(\mathbb{I}_{\{t \leq \tau_n\}} f\right)(t) \in \mathcal{M}^2([0,T])$;
(3) $P\left(\bigcup_{n=1}^{\infty} \{\omega : \tau_n(\omega) = T\}\right) = 1$.

Proposition 2.22. [8] *If $f(t) \in \mathcal{L}^2([0,T])$, then the sequence of stopping times defined, for $n \geq 1$, as*

$$\tau_n = \inf\left\{t : \int_0^t f^2(s)ds \geq n\right\},$$

with the convention that $\inf\{\emptyset\} = T$, is a localising sequence for $f(t)$ in $\mathcal{M}^2([0,T])$.

Definition 2.23. An $\mathcal{F}_t$-adapted stochastic process $X(t)$ is called a local martingale if there is a localising sequence $(\tau_n, n \geq 1)$ such that for $n \geq 1$, the process $X\left(t \wedge \tau_n\right)$ is a martingale with respect to $\mathcal{F}_t$. The symbol $\wedge$ denotes the minimum of two values.

The sequence of adapted processes $\left(\mathbb{I}_{\{t\leq\tau_n\}}f\right)(t)$ converges to $f(t)$ with probability 1 for all sufficiently large n. Besides, each process $\left(\mathbb{I}_{\{t\leq\tau_n\}}f\right)(t)$ is a member of the space $\mathcal{M}^2([0,T])$, since

$$\mathbb{E}\left[\int_0^T \left(\mathbb{I}_{\{t\leq\tau_n\}}f\right)^2(s)ds\right] = \mathbb{E}\left[\int_0^{\tau_n} f^2(s)ds\right] \leq \mathbb{E}(n) = n < \infty,$$

and thus the stochastic integrals of $\left(\mathbb{I}_{\{t\leq\tau_n\}}f\right)(t)$ are well-defined over $[0,T]$. This allows us to define a sequence of continuous martingales with respect to $\mathcal{F}_t$

$$X_n(t) = \int_0^T \left(\mathbb{I}_{[0,t\wedge\tau_n]}f\right)(s)ds,$$

which converges to a *local martingale*.

Theorem 2.24. [8] *For $n \geq 1$, the process $X(t\wedge\tau_n)$ is a martingale with respect to $\mathcal{F}_t$. If $f(t) \in \mathcal{L}^2([0,T])$, then there exists a continuous local martingale $X(t)$ such that for all $t \in [0,T]$*

$$X(t) = \int_0^t f(s)dW(s) = \lim_{n\to\infty}\int_0^T \left(\mathbb{I}_{[0,t\wedge\tau_n]}f\right)(s)ds.$$

As opposed to the stochastic integral of processes $f(t) \in \mathcal{M}^2([0,T])$ over an interval $[0,T]$, whose expectation is equal to 0, the expectation of the stochastic integral of processes $f(t) \in \mathcal{L}^2([0,T])$ does not necessarily even exist.

2.3. **Stochastic calculus.** We are interested in studying general processes that include a Lebesgue-type integral in addition to a stochastic integral.

Definition 2.25. (Itô process) An $\mathbb{R}$-valued Itô process is a stochastic process of the form

$$X(t) = X(0) + \int_0^t \alpha(s)ds + \int_0^t \beta(s)dW(s),\ 0 \leq t \leq T, \tag{2.4}$$

where the stochastic processes $\alpha(t)$ and $\beta(t)$ satisfy almost surely $\int_0^T |\alpha(s)|ds < \infty$ and $\int_0^T \beta^2(s)ds < \infty$.

It is accepted practice to write equation (2.4) in differential form as

$$dX(t) = \alpha(t)dt + \beta(t)dW(t).$$

However, it must be noted that this expression has no meaning by itself since the paths of the Wiener process are nowhere differentiable. Since $\beta(t) \in \mathcal{L}^2([0,T])$ the Itô process is a local martingale.

In order to understand the volatility associated with Itô processes, we must determine the rate at which they accumulate quadratic varition.

Lemma 2.26. [31] *The quadratic variation of the Itô process is*

$$[X,X](t) = \int_0^t \beta^2(s)ds.$$

This means that, at each time t, the process $X(t)$ is accumulating quadratic variation at a rate $\beta^2(t)$ per unit of time. The quadratic variation is due solely to the quadratic variation of the stochastic integral.

In standard calculus the chain rule of differentation gives the derivative of a composite function $f(g(t))$. If $f(t)$ and $g(t)$ are differentiable, then $f(g(t))$ is also differentiable and has derivative

$$\frac{d}{dt}f(g(t)) = \dot{f}(g(t))\dot{g}(t).$$

The stochastic calculus theory developed by K. Itô provides the equivalent to the standard chain rule of differentiation when dealing with functions of stochastic processes. For a twice differentiable function $f(x)$ applied to an Itô process $X(t)$, we would have

$$df(X(t)) = \dot{f}(X(t))dX(t) + \frac{1}{2}\ddot{f}(X(t))d\left[X,X\right](t), \tag{2.5}$$

where $\left[X,X\right](t)$ is the quadratic variation of the process $X(t)$. The generalisation of this formula answers the question whether, given an Itô process and a smooth function $f(t,x)$, $f(t,X(t))$ is again an Itô process.

Theorem 2.27. (Itô formula) [8] *Let $(X(t), t \in [0,T])$ be an Itô process given by equation (2.4). Assume that $F : [0,T] \times \mathbb{R} \to \mathbb{R}$ is a continuous function that is $\mathcal{C}^1([0,T])$ and $\mathcal{C}^2(\mathbb{R})$, i.e. with continuous partial derivatives $F_t(t,x)$, $F_x(t,x)$ and $F_{xx}(t,x)$. Then $F(t,X(t))$ is also an Itô process and*

$$\begin{aligned} F(t,X(t)) - F(0,X(0)) &= \int_0^t F_t(s,X(s))\,ds + \int_0^t F_x(s,X(s))\,dX(s) \\ &\quad + \frac{1}{2}\int_0^t F_{xx}(s,X(s))\,d\left[X,X\right](s) \\ &= \int_0^t F_t(s,X(s))\,ds + \int_0^t F_x(s,X(s))\,\alpha(s)ds \\ &\quad + \frac{1}{2}\int_0^t F_{xx}(s,X(s))\,\beta^2(s)ds \\ &\quad + \int_0^t F_x(s,X(s))\,\beta(s)dW(s), \end{aligned}$$

where $\left[X,X\right](t) = \int_0^t \beta^2(s)ds$.

It follows that $F(t,X(t))$ is a local martingale because $\beta(t) \in \mathcal{L}^2([0,T])$ and $f_x(s,X(s))$ is continuous.

For our purposes it will be useful to extend the Itô formula to its multidimensional version. Let $W_1(t), W_2(t), \ldots, W_m(t)$ be m independent Wiener processes and $X_1(t), X_2(t), \ldots, X_n(t)$ be n Itô processes given by

$$X_i(t) = X_i(0) + \int_0^t \alpha_i(s)ds + \sum_{j=1}^m \int_0^t \beta_{ij}(s)dW_j(s),\ 1 \le i \le n,\ 0 \le t \le T \tag{2.6}$$

where stochastic processes $\alpha_i(t)$ satisfy $\int_0^T |\alpha_i(s)|ds < \infty$ and $\beta_{ij}(t) \in \mathcal{L}^2([0,T])$. If we introduce the vectors and matrix

$$W(t) = \begin{bmatrix} W_1(t) \\ \vdots \\ W_m(t) \end{bmatrix},\ X(t) = \begin{bmatrix} X_1(t) \\ \vdots \\ X_n(t) \end{bmatrix},$$

$$\alpha(t) = \begin{bmatrix} \alpha_1(t) \\ \vdots \\ \alpha_n(t) \end{bmatrix}, \ \beta(t) = \begin{bmatrix} \beta_{11}(t) & \cdots & \beta_{1m}(t) \\ \vdots & \ddots & \vdots \\ \beta_{n1}(t) & \cdots & \beta_{nm}(t) \end{bmatrix},$$

then equation (2.6) can be written in matrix form resulting in the same expression as equation (2.4).

Theorem 2.28. (Multidimensional Itô formula) [19] *Assume that* $F : [0,T] \times \mathbb{R}^n \to \mathbb{R}$ *is a continuous function with continuous partial derivatives* $F_t(t,x)$, $F_{x_i}(t,x)$, *and* $F_{x_i x_j}(t,x)$ *for* $1 \leq i,j \leq n$,. *Then* $F(t,X(t))$ *is also an Itô process and*

$$\begin{aligned} F(t,X(t)) &= F(0,X(0)) + \int_0^t F_t(s,X(s))\,ds \\ &+ \int_0^t \sum_{i=1}^n F_{x_i}(s,X(s))\,\alpha_i(s)ds \\ &+ \int_0^t \sum_{i=1}^n \sum_{j=1}^m F_{x_i}(s,X(s))\,\beta_{ij}(s)dW_j(s) \\ &+ \frac{1}{2}\int_0^t \sum_{k=1}^m \sum_{i,j=1}^n F_{x_i x_j}(s,X(s))\,\beta_{ik}(s)\beta_{jk}(s)ds. \end{aligned} \tag{2.7}$$

Unless mentioned otherwise vectors throughout the dissertation are given in column form.

3. Jump processes

3.1. **Pure jump processes.** Observed asset prices usually present jumps or spikes and the empirical distribution of returns exhibits fat tails and skewness. These features are not taken into account in models based on diffusion processes driven by the Wiener process. Such models lack rigorous descriptive power and they are likely to induce mispricing of financial products. One way of improvement is to build models as a combination of diffusion and jump processes, both of which belong to a broader family of stochastic processes that can describe the observed reality of financial markets in a more accurate way.

Definition 3.1. (Lévy process) [1] An $\mathbb{R}^d$-valued stochastic process $(X(t), t \geq 0)$ defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ is called a Lévy process if:

(1) $X(0) = 0$ almost surely;
(2) $X(t)$ has independent and stationary increments;
(3) $X(t)$ is *stochastically continuous*, i.e. for all $a > 0$ and for all $s \geq 0$

$$\lim_{t \to s} P(|X(t) - X(s)| > a) = 0.$$

The third condition does not imply that the sample paths are continuous. It means that, for any given time t, the probability of seeing a jump at t is zero, and thus discontinuities occur at random times. We need therefore a space that allows for discontinuous functions. The class of *càdlàg* functions happens to be a convenient class of discontinuous functions.

Definition 3.2. (Càdlàg function) [11] A function $f : [0,T] \to \mathbb{R}^d$ is said to be càdlàg if $f(t)$ is right-continuous with left limits: for all $t \in [0,T]$ the limits $f(t-) = \lim_{s\to t, s<t} f(s)$ and $f(t+) = \lim_{s\to t, s>t} f(s)$ exist and $f(t) = f(t+)$.

A càglàd function (i.e. one that is left-continuous with right limits) is defined similarly. Assuming a jump time at $t \in [0,T]$, for a càdlàg function we define its value at t to be the value after the jump, thus the value at t is not foreseeable by following the trajectory up to time t. For a càglàd function, on the contrary, we define the value at t to be the value before the jump, thus the value at t could be predicted when approaching t along the sample path. This distinction is important in the context of financial modelling: jumps represent sudden, unforeseeable events, so the choice of right-continuity is natural. On the other hand, if we want to model a discontinuous process whose values are predictable we should use a càglàd process.

Theorem 3.3. [1] *Every Lévy process has a unique càdlàg version that is itself a Lévy process.*

We will assume from now on, without loss of generality, that the sample paths of the Lévy processes are càdlàg.

The class of Lévy processes is very rich and includes, for example, the Wiener process, which is the only non-deterministic Lévy process with continuous sample paths [1]. Next we will also introduce examples of pure jump Lévy processes.

Definition 3.4. (Poisson process) Let $(\tau_n, n \geq 1)$ be a sequence of independent exponential random variables with parameter λ and a sequence $(T_n, n \geq 1)$ such that

$$T_n = \sum_{i=1}^{n} \tau_i.$$

The process $(M(t), t \geq 0)$[1] defined by

$$M(t) = \sum_{n=1}^{\infty} \mathbb{I}_{[T_n,\infty]}(t)$$

is a integer-valued process taking values in $\mathbb{N}$ called Poisson process with intensity λ.

The Poisson process is the most elementary example of a jump process. The random variables $(\tau_n, n \geq 1)$ are called the *interarrival times*, since they determine the time elapsed between jumps, and the random variables $(T_n, n \geq 1)$ are called the *arrival times.* Since the expected value of each exponential random variable is $1/\lambda$, the jumps occur at an average rate of λ per unit of time. The Poisson process is therefore a stochastic process with discontinuous paths that counts the number of jumps that occur before or at time t. Some important properties of the Poisson process are discussed in the following Proposition.

Proposition 3.5. [11] *Let $(M(t), t \geq 0)$ be a Poisson process, then:*

(1) $M(0) = 0$;

[1]In order to preserve the notation used in [12], we will use the letter N to refer to Poisson random measures, and to avoid confusion with the associated Poisson process, we will use the letter M to refer to it.

(2) *the sample paths are càdlàg*[2]*, piecesewise constant and increase with jumps of size* 1*;*
(3) *for any* $0 \leq s < t$*,* $M(t) - M(s)$ *follows a Poisson distribution with parameter* $\lambda(t-s)$*, i.e.*

$$\mathbb{P}\left((M(t) - M(s)) = n\right) = e^{-\lambda(t-s)} \frac{\left(\lambda\left(t-s\right)\right)^n}{n!}, \ \forall n \in \mathbb{N};$$

(4) *for any* $0 \leq t_1 < t_2 < \cdots < t_n$*, the random variables* $M(t_1), M\left(t_2\right) - M\left(t_1\right), \ldots, M\left(t_n\right) - M\left(t_{n-1}\right)$ *are stationary and independent.*

The Poisson process lacks, however, the property of being a martingale. But we can define the compensated Poisson process as

$$\tilde{M}(t) = M(t) - \lambda t,$$

which is a martingale due to the deterministic compensator λt.

The Poisson and compensated Poisson processes have the limitation of allowing only jumps of size 1. Models for financial markets need, however, to allow the jump size to be random. This is achieved by means of the compound Poisson process.

Definition 3.6. (Compound Poisson process) Let $M(t)$ be a Poisson process with intensity λ and a sequence $(Y_n, n \geq 1)$ of i.i.d. random variables with $\mathbb{E}\left(Y_n\right) = \beta$ that are also independent from $M(t)$. The process $(X(t), t \geq 0)$ defined as

$$X(t) = \sum_{i=1}^{M(t)} Y_i$$

is an $\mathbb{R}^+$-valued process called compound Poisson process with intensity λ.

The jumps in $X(t)$ occur at the same times T_i as the jumps in $M(t)$, and their size is given by Y_i. The expected value of the compound Poisson process is $\mathbb{E}\left(X(t)\right) = \lambda\beta t$, which means that, on average, there are λt jumps in the interval $[0, t]$ with average jump size equal to β. If $\beta \neq 0$, the compound Poisson process is not a martingale, but, as done for the Poisson process, we can define its compensated version as

$$\widetilde{X}(t) = X(t) - \lambda\beta t,$$

which turns out to be a martingale.

Poisson processes have a close relationship with the study of *random measures*, as we will see next.

Definition 3.7. (Random measure) [9] Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and $(\mathbf{Z}, \mathcal{E})$ be a measurable space, where $\mathbf{Z} \subset \mathbb{R}^d$ and $\mathcal{E}$ is a σ-algebra on $\mathbf{Z}$. A random measure $N(\omega, U)$ on $(\mathbf{Z}, \mathcal{E})$ is a mapping $N : \Omega \times \mathcal{E} \to \bar{\mathbb{R}}^+$, where $\bar{\mathbb{R}}^+$ is the non-negative extended real line. For convenience, the random measure $N(\omega, U)$ will be written as $N(U)$ or simply N.

The mapping N is called a *transition kernel* from $(\Omega, \mathcal{F}, \mathbb{P})$ into $(\mathbf{Z}, \mathcal{E})$ because it satisfies the conditions in the following Definition.

[2]The choice of càdlàg sample paths is motivated by the inclusion of unpredictable jumps in the models of price dynamics. Another possible choice would have been càglàd sample paths, but then the values of M in the near future are foreseeable.

Definition 3.8. (Transition kernel) [9] Let $(\mathbf{F},\mathcal{F})$ and $(\mathbf{E},\mathcal{E})$ be measurable spaces. Let N be a mapping from $\mathbf{F}\times\mathcal{E}$ into $\bar{\mathbb{R}}^{+}$. Then N is called a transition kernel from $(\mathbf{F},\mathcal{F})$ into $(\mathbf{E},\mathcal{E})$ if:

(1) for every set $U\in\mathcal{E}$, $N(\cdot,U)$ is a $\mathcal{F}$-measurable random variable;
(2) for every $\omega\in\mathbf{F}$, $N(\omega,\cdot)$ is a measure on $(\mathbf{E},\mathcal{E})$.

Given any partition $(U_n, n\in\mathbb{N})$ of mutually exclusive sets in $\mathcal{E}$, the random measure N is a σ-finite mapping because $N(U_n)<\infty$ almost surely for all n. Some other important properties of the random measure are listed in the following Proposition.

Proposition 3.9. [9] *Let N be a random measure defined on $(\mathbf{Z},\mathcal{E})$, then:*

(1) $N(\emptyset)=0$;
(2) *given any sequence $(U_n, n\in\mathbb{N})$ of mutually exclusive sets in $\mathcal{E}$,*

$$N\left(\bigcup_{n\in\mathbb{N}} U_n\right)=\sum_{n\in\mathbb{N}} N(U_n);$$

(3) *for any family of measurable disjoint sets $(U_n, n\in\mathbb{N})$ in $\mathcal{E}$, the random variables $(N(U_n), n\in\mathbb{N})$ are independent.*

A random measure can be seen is a collection of random variables that assigns a measure to every outcome ω in Ω.

Definition 3.10. (Intensity measure) The *intensity measure* μ of a random measure N is the measure μ on $(\mathbf{Z},\mathcal{E})$ defined as

$$\mu(U)=\mathbb{E}(N(U))=\int_{\Omega} N(U)\mathbb{P}(d\omega),\ U\in\mathcal{E}.$$

The intensity measure is nothing more than the average value of the random measure N on any set $U\in\mathcal{E}$.

Definition 3.11. (Random counting measure) Let $(M(t), t\geq 0)$ be a Poisson process and let $(T_n, n\geq 1)$ be the arrival times of the Poisson process. A random counting measure on $\mathbb{R}^{+}$ is a positive, integer-valued measure N with values in $\mathbb{N}$ defined as

$$N(U)=\#\{n\geq 1, T_n(\omega)\in U\},\ U\in\mathcal{B}(\mathbb{R}^{+}),$$

where $\mathcal{B}(\mathbb{R}^{+})$ denotes the Borel σ-algebra on $\mathbb{R}^{+}$.

The Poisson process was formulated in Definition 3.4 as a counting process: the sequence of jump times T_n can be regarded as a random countable set of points on $\mathbb{R}^{+}$, and the Poisson process simply counts the number of such points in the interval $[0,t]$ of the real line. The intensity λ of the Poisson process determines the intensity measure of the random counting measure N. From Definition 3.10, we can see that [31]

$$\begin{aligned}\mu(U) &= \mathbb{E}(N(U))=\sum_{n=0}^{\infty} n\frac{\left(\lambda\int_U dm\right)^n}{n!}e^{-\lambda\int_U dm}\\ &= \lambda\int_U dm=\lambda m(U),\ U\in\mathcal{B}(\mathbb{R}_{+}),\end{aligned}$$

where m represents the Lebesgue measure. Since $\mu(U)$ is the average value of the random counting measure on any set $U\in\mathcal{B}(\mathbb{R}^{+})$, then $N(U)$ is a Poisson random variable with intensity $\mu(U)$.

Definition 3.12. (Radon measure) [11] Let $(\mathbf{Z}, \mathcal{E})$ be a measurable space. A Radon measure on $(\mathbf{Z}, \mathcal{E})$ is a measure μ such that, for every compact measurable set $U \in \mathcal{E}$, $\mu(U) < \infty$.

The construction of the random counting measure N in Definition 3.11 can be extended to more general settings, such as the one described in Definition 3.7. If we replace $\mathbb{R}^+$ by a space $\mathbf{Z} \subset \mathbb{R}^d$ and the Lebesgue measure by any Radon measure μ on $\mathbf{Z}$, then the Poisson process distributes elements over some d-dimensional space in a random way. We can therefore define the Poisson random measure as in the following Definition.

Definition 3.13. (Poisson random measure) [11] Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, $(\mathbf{Z}, \mathcal{E})$ be a measurable space and μ be a Radon measure on it. A Poisson random measure N on $(\mathbf{Z}, \mathcal{E})$ with intensity measure μ is an integer-valued random measure $N : \Omega \times \mathcal{E} \to \mathbb{N}$, such that for every set $U \in \mathcal{E}$, $N(U)$ is a $\mathcal{F}$-measurable Poisson random variable with parameter $\mu(U)$, i.e.

$$\mathbb{P}\left(N(U) = n\right) = e^{-\mu(U)} \frac{\left(\mu(U)\right)^n}{n!}, \ \forall n \in \mathbb{N}.$$

The Poisson random measure also satisfies the properties mentioned in Proposition 3.9 for random measures.

Proposition 3.14. [11] *Let $(\mathbf{Z}, \mathcal{E})$ be a measurable space. For any Radon measure μ on $(\mathbf{Z}, \mathcal{E})$, there exists a Poisson random measure N on $(\mathbf{Z}, \mathcal{E})$ with intensity measure μ.*

We can further generalise the concept of Poisson random measure by studying the behavior of the jumps of a compound Poisson process, or, in the general case, of a Lévy process. The sizes of the jumps of the process are distributed randomly over the d-dimensional space as time progresses.

Definition 3.15. (Lévy measure) [1] Let $X(t)$ be a Lévy process on $\mathbb{R}^d$. The measure ν on $(\mathbf{Z}, \mathcal{E})$ defined as

$$\nu(U) = \mathbb{E}\left[\#\left\{\Delta X(t) \neq 0, \Delta X(t) \in U : t \in [0,1]\right\}\right], \ U \in \mathcal{E},$$

is called the *Lévy measure* of $X(t)$.

In other words, $\nu(U)$ is the expected number, per unit of time, of jumps whose size belongs to U. The Lévy measure describes the sizes and rate with which jumps of a Lévy process occur.

Assumption 3.16. *Since jumps of size 0 are excluded from the definition of the Lévy measure, we have that $\nu(\{0\}) = 0$. We will assume from now that $\mathbf{Z} = \mathbb{R}^d$ and $\mathcal{E}$ is the Borel σ-algebra $\mathcal{B}(\mathbf{Z}\backslash\{0\})$. Besides, we will assume that every set $U \in \mathcal{E}$ satisfies that its closure does not contain $\{0\}$.*

Without loss of generality, let $(X(t), t \geq 0)$ be an $\mathbb{R}^d$-valued compound Poisson process. We can define the jump process $(\Delta X(t), t \geq 0)$ associated with it as

$$\Delta X(t) = X(t) - X(t-),$$

and the random measure describing the jumps of $X(t)$ as

$$N(t, U) = \#\left\{\Delta X(s) \neq 0, \Delta X(s) \in U : s \in [0,t]\right\}, \ U \in \mathcal{E}.$$

For every measurable set $U \in \mathcal{E}$, $N(t, U)$ counts the number of jumps sized in U that occur in the interval $[0, t]$.

Proposition 3.17. [1] *Let $N(t,U)$ be a Poisson random measure on $\mathbb{R}^+ \times \mathbf{Z}$, then:*

1. *for every $t \geq 0$ and $\omega \in \Omega$, the function $N(t,\cdot)$ is a counting measure on $(\mathbf{Z}, \mathcal{E})$;*
2. *for any sequence $(U_n, n \in \mathbb{N})$ of mutually exclusive sets in $\mathcal{E}$, the series $(N(t,U_n), t \geq 0)$ is comprised of independent stochastic processes;*
3. *for every $U \in \mathcal{E}$, $(N(\cdot,U), t \geq 0)$ is a Poisson process with intensity $\nu(U) = \mathbb{E}(N(1,U))$;*
4. *for a compound Poisson process with intensity λ and jump size distribution f, the Poisson random measure $N(t,U)$ has intensity measure $\mu(t \times U) = \nu(U)t = \lambda f(U)\, t$, where ν is the Lévy measure.*

For fixed $t \geq 0$ and $U \in \mathcal{E}$, $N(t,U)$ is a Poisson random variable with parameter $\mu(t \times U)$.

Lemma 3.18. [14] *For any Radon measure ν defined on $\mathbf{Z}$ with total mass $\lambda = \nu(\mathbf{Z})$ there exists a sequence $(Y_n, n \geq 1)$ of i.i.d. random variables such that*

$$P(Y_n \in U) = \frac{\nu(U)}{\lambda},\ U \in \mathcal{E},$$

and an independent Poisson process $(M(t), t \geq 0)$ with intensity λ, such that almost surely

$$X(t) = \sum_{i=1}^{M(t)} Y_i$$

is a compound Poisson process defined on $(\Omega, \mathcal{F}, \mathbb{P})$ with Lévy measure ν.

The Poisson random measure N can be described as the counting measure associated with the random sequence of points that represent the information about the jumps of the compound Poisson process. If a path of the process given by ω has a jump of size $\Delta X(s) = z$ at time t, then the random measure N places a Dirac mass at the point (t,z) in $\mathbb{R}^+ \times \mathbf{Z}$. Therefore, for a time interval $[0,t]$ and a set $U \in \mathcal{E}$, the Poisson random measure counts how many jumps of size within U occur for this particular path ω from time 0 to t.

Definition 3.19. (Poisson point process) Let $X(t)$ be a compound Poisson process taking values in $\mathbb{R}^d$, let $(T_n, n \geq 1)$ be the sequence of random times describing the occurence of jumps in $X(t)$ and let $(Y_n, n \geq 1)$ be a sequence of i.i.d. random variables representing the jump size at time T_n. Then a Poisson random measure on $\mathbf{Z} \times [0,t]$ is defined as

$$N = \sum_{n \geq 1} \delta_{(T_n, Y_n)} = \sum_{t \in [0,T]}^{\Delta X(t) \neq 0} \delta_{(t, \Delta X(t))}.$$

The pair (T_n, Y_n) produces a random configuration of points in $\mathbf{Z} \times [0,t]$ that receives the name of Poisson point process and will be denoted by $\mathbf{p}$. The Poisson point process contains all the information about the jumps of the process $X(t)$.

Proposition 3.20. [11] *Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. The random measure N is a Poisson random measure adapted to $\mathcal{F}_t$ if:*

1. *$(T_n, n \geq 1)$ is an increasing sequence of non-anticipating random times, i.e. T_n is adapted to $\mathcal{F}_t$;*
2. *Y_n is adapted to $\mathcal{F}_{T_n}$.*

If the Poisson random measure N is $\mathcal{F}_t$-adapted, then the associated Poisson point process is also $\mathcal{F}_t$-adapted.

Definition 3.21. (σ-finite Poisson point process) [16] A Poisson point process $\mathbf{p}$ with associated Poisson random measure $N(t,U)$ is σ-finite if there exists a sequence of sets $(U_n, n \in \mathbb{N})$ in $\mathcal{E}$ such that in $\bigcup_{n\in\mathbb{N}} U_n = \mathcal{E}$ and $\mathbb{E}\left(N(t,U_n)\right) < \infty$ for all $t > 0$ and $n \in \mathbb{N}$.

For a given σ-finite Poisson point process $\mathbf{p}$ we define the set

$$\mathcal{E}_{\mathbf{p}} = \{U \in \mathcal{E}, \mathbb{E}\left(N(t,U)\right) < \infty : t > 0\}.$$

Definition 3.22. ((QL) Poisson point process) [16] An $\mathcal{F}_t$-adapted Poisson point process $\mathbf{p}$ with associated Poisson random measure $N(t,U)$ is said to be of class (QL) (meaning Quasi-Left continuous) if it is σ-finite and there exists a random measure $\widehat{N}(t,U)$ such that:

(1) for every $U \in \mathcal{E}_{\mathbf{p}}$, $(N(\cdot,U), t \geq 0)$ is a continuous $\mathcal{F}_t$-adapted increasing process;
(2) for every $t \geq 0$ and $\omega \in \Omega$, the function $\widehat{N}(t,\cdot)$ is a σ-finite measure on $(\mathbf{Z}, \mathcal{E})$;
(3) for every $U \in \mathcal{E}_{\mathbf{p}}$, the random measure $\widetilde{N}(t,U) = N(t,U) - \widehat{N}(t,U)$ is a martingale with respect to $\mathcal{F}_t$.

The random measure $\widehat{N}(t,U)$ is called the compensator of the Poisson point process.

Since we consider only Poisson point processes that are stationary, then their compensators take the form of the intensity measure of N, i.e. $\widehat{N}(t,U) = \nu(U)t$. The compensated Poisson random measure $\widetilde{N}$ has then the form $\widetilde{N}(t,U) = N(t,U) - \nu(U)t$ for every $U \in \mathcal{E}$. In this case we say that $\widetilde{N}$ is a *martingale-valued measure* with respect to $\mathcal{F}_t$.

Similarly as done with other measures –such as the Lebesgue measure–, we can define the integral with respect to the Poisson random measure.

Definition 3.23. (Poisson integration) [1] Let N be a Poisson random measure defined on $\mathbf{Z} \times [0,t]$ and let $f : \mathbf{Z} \to \mathbb{R}^d$ be a measurable function. We can define the Poisson integral of $f(z)$ over a set $U \in \mathcal{E}$ as

$$\int_U f(z)N(t,dz) = \sum_{z\in U} f(z)N\left(t,\{z\}\right),$$

which is a random variable in $\mathbb{R}^d$ with compound Poisson distribution.

For any value $z \in \mathbf{Z}$, $N\left(t,\{z\}\right) \neq 0$ only if the underlying compound Poisson process associated with it has a jump of size z, i.e. $\Delta X(s) = z$ for at least one $0 \leq s \leq t$. Then an equivalent expression for the Poisson integral is

$$\int_U f(z)N(t,dz) = \sum_{0\leq s\leq t} f\left(\Delta X(s)\right)\mathbb{I}_U\left(\Delta X(s)\right),$$

where $\mathbb{I}_U$ is the indicator function of the set $U \in \mathcal{E}$. As t varies the Poisson integral gives rise to a compound Poisson process with intensity $\nu(U)$ and jump size distribution

$$P\left(Y_n \in V\right) = \frac{\nu\left(U \cap f^{-1}\left(V\right)\right)}{\nu(U)},\ V \in \mathcal{E},$$

where $(Y_n, n \geq 1)$ is a sequence of i.i.d. random variables [1]. As we can see, the Poisson integral gives a systematic way of generating jump processes from Poisson random measures.

Theorem 3.24. [1] *Let N be a Poisson random measure on $(\mathbf{Z}, \mathcal{E})$ and let $f : \mathbf{Z} \to \mathbb{R}^d$ be a measurable function. For every set $U \in \mathcal{E}$:*

(1) *if $f \in \mathcal{L}^1(U)$, we have*

$$\mathbb{E}\left[\int_U f(z)N(t,dz)\right] = t\int_U f(z)\nu(dz);$$

(2) *if $f \in \mathcal{L}^2(U)$, we have*

$$Var\left[\int_U f(z)N(t,dz)\right] = t\int_U f^2(z)\nu(dz).$$

If we choose the measurable function $f(z)$ to be equal to 1, we can see from Definition 3.23 that the connection between the Poisson process $M(t)$ and the Poisson random measure N is given by

$$M(t) = \int_0^t \int_U N(ds,dz),$$

for every $U \in \mathcal{E}$ and $t \in [0,T]$. The integral over U counts the number of jumps whose size is in U, and since this value varies with time, the integral over $[0,t]$ produces a Poisson process. Furthermore, if the integrand is a measurable function $f : \mathbf{Z} \times [0,T] \to \mathbb{R}^d$ satisfying $f \in \mathcal{L}^1(\mathbf{Z} \times [0,T])$, then the Poisson integral over any set $U \in \mathcal{E}$ can be similarly defined as

$$\int_0^t \int_U f(s,z)N(ds,dz) = \sum_{0\leq s\leq t} f\left(s, \Delta X(s)\right)\mathbb{I}_U\left(\Delta X(s)\right).$$

Proposition 3.25. *Let $\widetilde{N} = N - \nu t$ be the compensated Poisson random measure defined on $(\mathbf{Z}, \mathcal{E})$ and let $f : \mathbf{Z} \times [0,T] \to \mathbb{R}^d$ be a measurable function such that $f \in \mathcal{L}^1(\mathbf{Z} \times [0,T])$. The Poisson integral of $f(s,z)$ with respect to $\widetilde{N}$ over any set $U \in \mathcal{E}$ results in a stochastic process defined as*

$$\begin{aligned} X(t) &= \int_0^t \int_U f(s,z)\widetilde{N}(ds,dz) \\ &= \int_0^t \int_U f(s,z)N(ds,dz) - \int_0^t \int_U f(s,z)\nu(dz)ds. \end{aligned} \tag{3.1}$$

Since the integrator $\widetilde{N}$ in (3.1) is a martingale, we can study under which conditions the process $X(t)$ is a martingale.

Proposition 3.26. [26] *If $f \in \mathcal{L}^2(\mathbf{Z} \times [0,T])$, then the stochastic process $X(t)$ defined in (3.1) is:*

(1) *a martingale if $\mathbb{E}\left[\int_0^T \int_U f^2(s,z)\nu(dz)ds\right] < \infty$;*

(2) *a local martingale if $\int_0^T \int_U f^2(s,z)\nu(dz)ds < \infty$ almost surely.*

Poisson and compound Poisson processes have σ-finite Lévy measures defined on $\mathbf{Z}$ and thus they are examples of pure jump Lévy processes that have finitely many jumps in a finite time interval. However, it is possible to chose a Lévy measure that is not σ-finite on $\mathbf{Z}$ and that it might diverge too strongly towards the origin. If

this is the case, then the associated Lévy process may have an infinite number of small jumps in a finite time interval. In fact, the Lévy measure does only need to satisfy the condition in Proposition 3.27.

Proposition 3.27. [11] *Let $(X(t), t \geq 0)$ be an $\mathbb{R}^d$-valued Lévy process. The Lévy measure ν associated with $X(t)$ is defined on $\mathbf{Z}$ and satisfies the integrability condition*

$$\int_{\mathbf{Z}} \left(1 \wedge z^2\right) \nu(dz) < \infty.$$

This conditions guarantess that it is not possible to have an infinite number of big jumps (of size ≥ 1) in a finite time interval. Depending on the σ-finiteness of the Lévy measure, Lévy processes can be grouped in two categories.

Proposition 3.28. [1] *Let $X(t)$ be a Lévy process with Lévy measure ν. We say that the process has finite activity if $\nu(\mathbf{Z}) < \infty$, and then almost all paths have only a finite number of jumps along any time interval of finite length. If $\nu(\mathbf{Z}) = \infty$, then $X(t)$ has infinite activity and almost all paths have infinitely many jumps along any time interval of finite length.*

3.2. **Jump-diffusion processes.** Jump-diffusion processes mix a jump process and a diffusion process, and their use in financial modelling has grown in recent years. We will introduce some of their basic concepts in the framework of Lévy processes. These were presented in Definition 3.1, but the following Theorem is even more interesting, since it reveals the building blocks of a Lévy process.

Theorem 3.29. (Lévy-Itô decomposition) [26, 1] *Let $(\mathbf{Z}, \mathcal{E})$ be a measurable space satisfying Assumption 3.16. If $X(t)$ is an $\mathbb{R}^d$-valued Lévy process, then there exists $b \in \mathbb{R}^d$, a d-dimensional Wiener process $W(t)$ with covariance matrix Σ and an independent Poisson random measure N defined on $\mathbb{R}^+ \times \mathbf{Z}$ generated by the jumps of $X(t)$ such that, for each $t \geq 0$,*

$$X(t) = bt + \Sigma W(t) + \int_0^t \int_{\mathbf{Z}\setminus\mathbf{Z}_0} zN(ds, dz) + \int_0^t \int_{\mathbf{Z}_0} z\widetilde{N}(ds, dz), \tag{3.2}$$

where the set $\mathbf{Z}_0 \in \mathcal{B}(\mathbf{Z}\setminus\{0\})$, fixed around the origin, is such that $z \in \mathbf{Z}_0$ if $|z| < R$, for any arbitrarily chosen positive constant R in $\mathbb{R}^d$. Furthermore,

$$b = \mathbb{E}\left(X(1) - \int_{\mathbf{Z}\setminus\mathbf{Z}_0} zN(1, dz)\right).$$

The value of R determines the separation between small and big jumps. The Lévy-Itô decomposition shows that every Lévy process can be described in terms of three independent auxiliary Lévy processes, each with different types of path behaviour.

The term $bt + \Sigma W(t)$ of (3.2) represents a linear Wiener process with drift. The term $\int_0^t \int_{\mathbf{Z}\setminus\mathbf{Z}_0} zN(ds, dz)$ represents the sum of jumps with absolute value larger than R. Since the paths of a Lévy process are càdlàg, the Lévy measure satisfies that $\int_{\mathbf{Z}\setminus\mathbf{Z}_0} \nu(dz) < \infty$, which means that over any bounded interval $[0, t]$ there can be only a finite number of jumps with absolute value larger that R [28]. This guarantees that the sum of jumps is a finite series and then the third term of (3.2) is well defined as a compound Poisson process with intensity $\nu(\mathbf{Z} \setminus \mathbf{Z}_0)$ and jump

distribution equal to

$$\frac{\mathbb{I}_{\{|z|>R\}}(z)\nu\left(dz\right)}{\nu\left(\mathbf{Z}\setminus\mathbf{Z}_0\right)},$$

unless $\nu\left(\mathbf{Z}\setminus\mathbf{Z}_0\right)=0$, in which case the third term is equal to zero.

Not all Lévy processes have finite Lévy measures (i.e. $\nu\left(\mathbf{Z}\right)<\infty$). However, since Lévy measures must satisfy that $\int_{\mathbf{Z}\setminus\mathbf{Z}_0}\nu(dz)<\infty$, the finiteness of $\nu\left(\mathbf{Z}\right)$ is subjected to the finiteness of $\int_{\mathbf{Z}_0}\nu(dz)$. By Proposition 3.27 the Lévy measure ν has no mass at 0, but it is possible it has singularities around 0 (i.e. there can be infinitely many jumps around the origin) and then the sum of small jumps has to be interpreted as the limit

$$\lim_{\epsilon\to 0}\int_0^t\int_{\epsilon<|z|<R} zN(ds,dz),\ \epsilon>0,$$

which, unfortunately, does not necessarily converge. However, it can be forced to converge by compensating it (i.e. subtracting the corresponding average increase of the process along $[0,t]$) [30]. We can define a sequence of processes [1]

$$\int_0^t\int_{\frac{1}{k}<|z|<R} z\widetilde{N}(ds,dz),\ k=1,2,\ldots$$

that converge in $\mathcal{L}^2(\Omega)$ to a martingale denoted by

$$\int_0^t\int_{\mathbf{Z}_0} z\widetilde{N}(ds,dz),$$

which is the fourth term in (3.2) and it represents the compensated sum of small jumps.

Definition 3.30. (Semimartingale) [16] Let $(S(t),t\geq 0)$ be an $\mathcal{F}_t$-adapted stochastic process defined on a probability space $(\Omega,\mathcal{F},\mathbb{P})$. We say that $S(t)$ is a semimartingale if, for each $t\geq 0$, it can be decomposed as

$$S(t)=S(0)+M(t)+A(t), \tag{3.3}$$

where $(M(t),t\geq 0)$ is a local martingale and $(A(t),t\geq 0)$ is a càdlàg $\mathcal{F}_t$-adapted process of finite variation.

The Lévy-Itô decomposition shows that every Lévy process can be represented as the sum of a Brownian motion with drift, a compound Poisson process and a pure jump martingale. It also implies that the paths of a Lévy process are almost surely of infinite variation in the case that a Wiener process is present. However, if we consider the case $\Sigma=0$, then the fourth term in (3.2) decides whether the Lévy process is of finite or infinite variation, since the third term in (3.2) is a compound Poisson process and therefore of finite variation.

Proposition 3.31. [11] *An $\mathbb{R}^d$-valued Lévy process $X(t)$ with characteristic triplet (b,Σ,ν) is almost surely of finite variation if and only if $\Sigma=0$ and $\int_{\mathbf{Z}\setminus\mathbf{Z}_0}|z|\nu(dz)<\infty$.*

By Lévy-Itô decomposition we have

$$M(t)=\Sigma W(t)+\int_0^t\int_{\mathbf{Z}_0} z\widetilde{N}(ds,dz)$$

$$A(t) = bt + \int_0^t \int_{\mathbf{Z}\setminus\mathbf{Z}_0} zN(ds, dz).$$

Since the Wiener process is a martingale and, by Proposition 3.26, the integral with respect to the compensated Poisson process is also a martingale, then the process $M(t)$ is a martingale. Process $A(t)$ is of finite variation since the number of jumps of absolute size $\geq R$ is locally finite, as follows from the fact that $\int_{\mathbf{Z}\setminus\mathbf{Z}_0} \nu(dz) < \infty$. Thus, we can conclude that every Lévy process is a *semimartingale*.

In financial applications it is often important to check whether a given Lévy process is a martingale. We will give the necessary and sufficient conditions in the following Proposition.

Proposition 3.32. [11] *An $\mathbb{R}^d$-valued Lévy process $X(t)$ with characteristic triplet (b, Σ, ν) is a martingale if and only if $\int_{\mathbf{Z}\setminus\mathbf{Z}_0} |z|\nu(dz) < \infty$ and $b + \int_{\mathbf{Z}\setminus\mathbf{Z}_0} z\nu(dz) = 0$ almost surely.*

3.3. **Stochastic calculus for jump-diffusion processes.** Following a similar motivation as in Subsections 2.2 and 2.3, we would like to have at our disposal a rigorous definition for integrals of stochastic processes with respect to Lévy processes; and also an equivalent Itô formula to describe the time evolution of a function that depends on a Lévy process. However, the class of Lévy processes is not stable under stochastic integration or non-linear transformations (i.e. the result of the operations are, in general, not Lévy process), and thus we will introduce first the main results of stochastic calculus for processes with jumps in the general framework of semimartingales, which is stable under the operations considered above.

Due to the representation of a semimartingale in the form of (3.3), integration with respect to semimartingales involves two integrals: one with respect to a local martingale and another one with respect to a finite variation process. The integral with respect to $A(t)$ can be done path by path as a Lebesgue-Stieltjes integral, since $A(t)$, though random, is of finite variation. The integral with respect to $M(t)$ generalises the Itô integral, where integration is done with respect to Wiener process, since $M(t)$ is allowed now to have jumps. Because of this, an essential requirement for the integrand is its predictability, in order to compensate the jumps of the local martingale $M(t)$ [19].

Definition 3.33. (Predictable σ-algebra) [1] Let $\mathcal{F}_t$ be the filtration of the probability space $(\Omega, \mathcal{F}, \mathbb{P})$. We denote by $\mathcal{P}$ the smallest σ-algebra of subsets of $\mathbb{R}^+ \times \Omega$ generated by all left-continuous and $\mathcal{F}_t$-adapted stochastic processes.

The previous definition applies naturally to the case where $\mathbb{R}^+$ is replaced by $[0, T]$.

Definition 3.34. (Predictable process) [11] We say that the stochastic process $(H(t), t \in [0, T])$ is predictable if the mapping $H : [0, T] \times \Omega \to \mathbb{R}^d$ is $\mathcal{P}$-measurable.

In other words, the process $H(t)$ is predictable if the random variable $H(t)$ is measurable with respect to the σ-algebra $\mathcal{F}_{t-}$ for each time t. Any left-continuous process $H(t)$ is by definition predictable: if the limit $\lim_{s\to t, s<t} H(s) = H(t-)$ exists and $H(t) = H(t-)$ then the value of $H(t)$ is predicted by the values at preceding instants.

In the context of financial modelling the requirement of predictability is fundamental to calculate the capital gain of a trading strategy. The positions in a

portfolio are based on information available at time t, so when the decision to change the portfolio is made the previous position is still valid. Thus, the portfolio takes the new value right after t.

Besides predictability, the integrands of the stochastic integral with respect to a semimartingale must satisfy additional conditions, which are listed in the definition of the space $\mathcal{L}^2_{pred}([0,T])$.

Definition 3.35. [14, 28] Let $S(t)$ be a semimartingale that admits the decomposition (3.3). Let $\mathcal{L}^2_{pred}([0,T])$ be the space of all stochastic processes $H(t)$ on a probability space satisfying:

(1) $H(t)$ is predictable;
(2) $\int_0^T H^2(t)d\left[M,M\right](t) < \infty$ almost surely;
(3) $\int_0^T |H(t)||dA(t)| < \infty$ almost surely.

The integrals with respect to $[M,M](t)$ and $A(t)$ are well-defined because both are processes of finite variation [28]. Considering simple processes that approximate the integrand by a limiting procedure, we can now define the stochastic integral.

Lemma 3.36. [28] *Let the stochastic process $(S(t), t \in [0,T])$ be a semimartingale and let $H(t) \in \mathcal{L}^2_{pred}([0,T])$. Let $\pi^n = (0 = t_0^n < t_1^n < \ldots < t_n^n = T)$ be a sequence of random partitions of $[0,T]$ such that $|\pi^n| = \sup_i |t_i^n - t_{i-1}^n| \to 0$ almost surely when $n \to \infty$. We define the process $H(t)$ sampled at π^n to be*

$$H_n(t) = \sum_{i=0}^{n-1} H\left(t_i^n\right) \mathbb{I}_{\left(t_i^n, t_{i+1}^n\right]}(t).$$

Then there exists a sequence of simple processes $H_n(t) \in \mathcal{L}^2_{pred}([0,T])$ such that

$$\lim_{n\to\infty} \mathbb{E}\left(\int_0^T \left(H(t) - H_n(t)\right)^2 d\left[M,M\right]t + \int_0^T |H(t) - H_n(t)||dA(t)|\right) \to 0, \ n \geq 1.$$

Definition 3.37. [28] Let the stochastic process $(S(t), t \in [0,T])$ be a semimartingale and let $H(t) \in \mathcal{L}^2_{pred}([0,T])$. Then the limit defined as

$$\lim_{n\to\infty} \left[\sum_{i=0}^{n-1} H\left(t_i^n\right)\left[S\left(t_{i+1}^n \wedge t\right) - S\left(t_i^n \wedge t\right)\right]\right]$$

converges in $\mathcal{L}^2(\Omega)$ for $t \in [0,T]$ to the stochastic integral of $H(t)$ with respect to $S(t)$, which is denoted by

$$\int_0^t H(s)dS(s).$$

It is logical to wonder now whether, given a semimartingale $X(t)$ and a smooth function $F(t,x)$, $F(t,X(t))$ is again a semimartingale. This question finds answer in the following Theorem.

Theorem 3.38. (Itô formula for semimartingales) [28] *Let the $\mathbb{R}$-valued process $(X(t), t \in [0,T])$ be a semimartingale and let $F : [0,T] \times \mathbb{R} \to \mathbb{R}$ be a continuous*

function that is $\mathcal{C}^1([0,T])$ *and* $\mathcal{C}^2(\mathbb{R})$. *Then* $F(t,X(t))$ *defined as*

$$\begin{aligned} F(t,X(t)) &= F(0,X(0)) + \int_0^t F_t(s,X(s))\,ds + \int_0^t F_x(s,X(s-))dX(s) \\ &+\frac{1}{2}\int_0^t F_{xx}(s,X(s-))d[X,X]^c(s) \\ &+\sum_{0<s\leq t}\Big[F(s,X(s)) - F(s,X(s-)) - F_x(s,X(s-))\Delta X(s)\Big] \end{aligned} \tag{3.4}$$

is a semimartingale, where $X(t-)$ *is the value of the process immediately before a jump,* $[X,X]^c(t)$ *is the continuous part of the quadratic variation of* $X(t)$ *and* $\Delta X(t)$ *is the size of the jump at time* t.

Since a semimartingale need not be continuous, Itô formula contains an additional term to ensure that the jumps of the process are correctly dealt with. Using the chain rule stated in equation (2.5) and considering the function $F(x) = x^2$, we can derive the form of the quadratic variation of a semimartingale $X(t)$.

Theorem 3.39. (Quadratic variation of semimartingale) [11] *Let* $(X(t), t \geq 0)$ *be a semimartingale. By definition* $X(t)$ *is an adapted càdlàg process and we can define the process* $X_-(t) = X(t-)$ *for* $t \geq 0$, *which is adapted and càglàd, and thus also predictable. The quadratic variation process of* $X(t)$ *is the adapted càdlàg process defined as*

$$[X,X](t) = X^2(t) - 2\int_0^t X(s-)dX(s).$$

The quadratic variation is a non-decreasing process of finite variation [14] and thus it admits the alternative representation

$$[X,X](t) = [X,X]^c(t) + \sum_{s\leq t}(\triangle X(s))^2. \tag{3.5}$$

In view of the Lévy-Itô decomposition, it is natural to consider stochastic processes of the following form.

Definition 3.40. (Lévy-Itô process) [26] Let $(\mathbf{Z},\mathcal{E})$ be a measurable space satisfying Assumption 3.16 with $d=1$. Let

$$\bar{N}(ds,dz) = \begin{cases} \widetilde{N}(ds,dz) & z \in \mathbf{Z}_0 \\ N(ds,dz) & z \in \mathbf{Z}\setminus\mathbf{Z}_0 \end{cases}.$$

Let $\alpha : [0,T]\times\Omega \to \mathbb{R}$ and $\beta : [0,T]\times\Omega \to \mathbb{R}$ be $\mathcal{F}_t$-adapted processes and let $\gamma : [0,T]\times\mathbf{Z}\times\Omega \to \mathbb{R}$ be an $\mathcal{F}_t$-predictable process satisfying the growth condition

$$\int_0^t \left\{ \mid \alpha(s) \mid + \beta^2(s) + \int_{\mathbf{Z}} \gamma^2(s,z)\nu(dz)\right\} ds < \infty \ a.s.$$

for all $t > 0$. Then the stochastic process

$$X(t) = X(0) + \int_0^t \alpha(s)ds + \int_0^t \beta(s)dW(s) + \int_0^t\int_{\mathbf{Z}} \gamma(s,z)\bar{N}(ds,dz), \tag{3.6}$$

for $0 \leq t \leq T$ receives the name of Lévy-Itô process.

We will also find convenient to say that the Lévy-Itô process $X(t)$ satisfies the stochastic differential equation

$$dX(t) = \alpha(t)dt + \beta(t)dW(t) + \int_{\mathbf{Z}} \gamma(t,z)\bar{N}(dt,dz). \tag{3.7}$$

In the paper by Davis and Lleo [12] the dynamics of risky securities are modelled using an Lévy-Itô process, thus we will find useful to have a version of formula (3.4) when the process $X(t)$ is a Lévy-Itô process.

Theorem 3.41. (Itô formula for Lévy-Itô processes) [26] *Let $(X(t), t \in [0,T])$ be an $\mathbb{R}$-valued Lévy-Itô process and let $F : [0,T] \times \mathbb{R} \to \mathbb{R}$ be a continuous function that is $\mathcal{C}^1([0,T])$ and $\mathcal{C}^2(\mathbb{R})$. Then $F(t,X(t))$ is also an Lévy-Itô process and*

$$\begin{aligned} F(t,X(t)) &= F(0,X(0)) + \int_0^t F_t(s,X(s))ds + \int_0^t F_x(s,X(s-))\alpha(s)ds \\ &\quad + \frac{1}{2}\int_0^t F_{xx}(s,X(s-))\beta^2(s)ds + \int_0^t F_x(s,X(s-))\beta(s)dW(s) \\ &\quad + \int_0^t \int_{\mathbf{Z}_0} \Big[F\left(s,X(s-)+\gamma(s,z)\right) - F(s,X(s-)) \\ &\quad - F_x(s,X(s-))\gamma(s,z)\Big]\nu(dz)ds \\ &\quad + \int_0^t \int_{\mathbf{Z}} \Big[F\left(s,X(s-)+\gamma(s,z)\right) - F(s,X(s-))\Big]\bar{N}(ds,dz), \end{aligned} \tag{3.8}$$

for $t \in [0,T]$.

Proof. The derivation starts by replacing $\Delta X(s)$ with $\gamma(s,z)$ –which determines the size of the jumps–, and $dX(s)$ with $\alpha(s)ds + \beta(s)dW(s) + \int_{\mathbf{Z}} \gamma(s,z)\bar{N}(ds,dz)$ in equation (3.4). We then obtain

$$\begin{aligned} F(t,X(t)) &= F(0,X(0)) + \int_0^t F_t\left(s,X(s)\right)ds + \frac{1}{2}\int_0^t F_{xx}(s,X(s-))\beta^2(s)ds \\ &\quad + \int_0^t F_x(s,X(s-))\alpha(s)ds + \int_0^t F_x(s,X(s-))\beta(s)dW(s) \\ &\quad + \int_0^t \int_{\mathbf{Z}} F_x(s,X(s-))\gamma(s,z)\bar{N}(ds,dz) \\ &\quad + \sum_{0<s\le t} \Big[F(s,X(s-)+\gamma(s,z)) - F(s,X(s-)) \\ &\quad - F_x(s,X(s-))\gamma(s,z)\Big]. \end{aligned}$$

By Definition 3.23 we can replace the sum by an integral, and we can also expand the integral with respect to $\bar{N}(ds,dz)$ to have integrals with respect to $N(ds,dz)$

and $\nu(ds)dz$. The result is

$$\begin{aligned} F(t,X(t)) &= F(0,X(0)) + \int_0^t F_t\left(s,X(s)\right)ds + \frac{1}{2}\int_0^t F_{xx}(s,X(s-))\beta^2(s)ds \\ &+ \int_0^t F_x(s,X(s-))\alpha(s)ds + \int_0^t F_x(s,X(s-))\beta(s)dW(s) \\ &+ \int_0^t \int_{\mathbf{Z}_0} F_x(s,X(s-))\gamma(s,z)N(ds,dz) \\ &- \int_0^t \int_{\mathbf{Z}_0} F_x(s,X(s-))\gamma(s,z)\nu(ds)dz \\ &+ \int_0^t \int_{\mathbf{Z}\setminus\mathbf{Z}_0} F_x(s,X(s-))\gamma(s,z)N(ds,dz) \\ &+ \int_0^t \int_{\mathbf{Z}} \Big[F(s,X(s-)+\gamma(s,z)) - F(s,X(s-)) \\ &- F_x(s,X(s-))\gamma(s,z)\Big]N(ds,dz). \end{aligned}$$

We now split the integral in $\mathbf{Z}$ into two: one integral in $\mathbf{Z}$ and another one in $\mathbf{Z}\setminus\mathbf{Z}_0$. Afterwards, we group integrals with respect to $N(ds,dz)$ and we obtain

$$\begin{aligned} F(t,X(t)) &= F(0,X(0)) + \int_0^t F_t\left(s,X(s)\right)ds + \frac{1}{2}\int_0^t F_{xx}(s,X(s-))\beta^2(s)ds \\ &+ \int_0^t F_x(s,X(s-))\alpha(s)ds + \int_0^t F_x(s,X(s-))\beta(s)dW(s) \\ &- \int_0^t \int_{\mathbf{Z}_0} F_x(s,X(s-))\gamma(s,z)\nu(ds)dz \\ &+ \int_0^t \int_{\mathbf{Z}_0} \Big[F(s,X(s-)+\gamma(s,z)) - F(s,X(s-))\Big]N(ds,dz) \\ &+ \int_0^t \int_{\mathbf{Z}\setminus\mathbf{Z}_0} \Big[F(s,X(s-)+\gamma(s,z)) - F(s,X(s-))\Big]N(ds,dz). \end{aligned}$$

We add and substract the term $\int_0^t \int_{\mathbf{Z}_0} \big[F(s,X(s-)+\gamma(s,z)) - F(s,X(s-))\big]\nu(ds)dz$ and we finally reach to the final formulation. □

If we consider now an $\mathbb{R}^n$-valued Lévy-Itô process $X(t)$, we would have in (3.6) that $\alpha : [0,T]\times\Omega \to \mathbb{R}^n$, $\beta : [0,T]\times\Omega \to \mathbb{R}^{n\times m}$ and $\gamma : [0,T]\times\mathbf{Z}\times\Omega \to \mathbb{R}^{n\times d}$ are $\mathcal{F}_t$-predictable stochastic processes; $W(t)$ is an m-dimensional Wiener process; and $\bar{N}(ds,dz)$ is a vector of d independent Poisson random measures

$$\bar{N}(ds,dz) = \begin{bmatrix} \bar{N}_1\left(ds,dz_1\right) \\ \vdots \\ \bar{N}_d\left(ds,dz_d\right) \end{bmatrix},$$

whose Lévy measure ν is likewise a vector of Lévy measures coming from d independent one-dimensional Lévy processes

$$\nu(dz) = \begin{bmatrix} \nu_1(dz_1) \\ \vdots \\ \nu_d(dz_d) \end{bmatrix}.$$

Then each component of the Lévy-Itô process may be written in the form

$$\begin{aligned} X_i(t) &= X_i(0) + \int_0^t \alpha_i(s)ds + \int_0^t \sum_{j=1}^m \beta_{ij}(s)dW_j(s) \\ &\quad + \int_0^t \int_{Z^{(i)}} \sum_{j=1}^d \gamma_{ij}(s,z)\bar{N}_j(ds,dz_j),\ 1 \le i \le n,\ 0 \le t \le T, \end{aligned}$$

where $Z^{(i)} = \mathbb{R}$ is the i-th component of $\mathbf{Z}$. Thus, the multidimensional version of the Itô formula can be expressed as follows.

Theorem 3.42. (Multidimensional Itô formula for Lévy-Itô processes) [26] *Let $(X(t), t \in [0,T])$ be an $\mathbb{R}^n$-valued Lévy-Itô process and let $F : [0,T] \times \mathbb{R}^n \to \mathbb{R}$ be a continuous function that is $\mathcal{C}^1([0,T])$ and $\mathcal{C}^2(\mathbb{R}^n)$. Then $F(t, X(t))$ is also a Lévy-Itô process and*

$$\begin{aligned} F(t,X(t)) &= F(0,X(0)) + \int_0^t F_t(s,X(s))ds + \int_0^t \sum_{i=1}^n F_{x_i}(s,X(s-))\alpha_i(s)ds \\ &\quad + \frac{1}{2}\int_0^t \sum_{k=1}^m \sum_{i,j=1}^n F_{x_i x_j}(s,X(s-))\beta_{ik}(s)\beta_{jk}(s)ds \\ &\quad + \int_0^t \sum_{i=1}^n \sum_{j=1}^m F_{x_i}(s,X(s-))\,\beta_{ij}(s)dW_j(s) \\ &\quad + \sum_{p=1}^d \int_0^t \int_{Z_0^{(p)}} \Big[F\left(s, X(s-) + \gamma^{(p)}(s,z)\right) - F(s,X(s-)) \\ &\quad - \sum_{i=1}^n \gamma_{ip}(s,z)F_{x_i}(s,X(s-)) \Big] \nu_p(dz_p)\,ds \\ &\quad + \sum_{p=1}^d \int_0^t \int_{Z^{(p)}} \Big[F\left(s, X(s-) + \gamma^{(p)}(s,z)\right) \\ &\quad - F(s,X(s-)) \Big] \bar{N}_p(ds,dz_p), \end{aligned} \tag{3.9}$$

for $t \in [0,T]$, where $\gamma^{(k)} \in \mathbb{R}^n$ is the k column of the $n \times d$ matrix γ, and $Z_0^{(p)} \in \mathcal{B}\left(Z^{(p)} \backslash \{0\}\right)$ is the p-th component of $\mathbf{Z}_0$, such that $z \in Z_0^{(p)}$ if $|z| < R_p$, for constant $R_p \in \mathbb{R}$.

3.4. **Stochastic exponential.** The *stochastic exponential* (also known as *Doléans-Dade exponential*) is the stochastic analogue of the exponential function. In the deterministic case, if $f(t)$ is a smooth function, then $g(t) = e^{f(t)}$ is the solution to the differential equation $\dot{g}(t) = g(t)\dot{f}(t)$. The stochastic exponential is defined as a solution to a similar stochastic equation.

Definition 3.43. (Stochastic exponential) [28] Let $(X(t), t \geq 0)$ be a semimartingale with $X(0) = 0$. The stochastic exponential of $X(t)$, written $\mathcal{E}_X(t)$, is the (unique) semimartingale $(Z(t), t \geq 0)$ that satisfies the equation

$$Z(t) = 1 + \int_0^t Z(s-)dX(s). \tag{3.10}$$

Theorem 3.44. ***(Doléans-Dade)*** [28] *Let $(X(t), t \geq 0)$ be a semimartingale with $X(0) = 0$. The solution to (3.10) is given by*

$$\mathcal{E}_X(t) = e^{\left(X(t) - \frac{1}{2}[X,X](t)\right)} \prod_{0<s\leq t} (1 + \triangle X(s))\, e^{\left(-\triangle X(s) + \frac{1}{2}(\triangle X(s))^2\right)}, \tag{3.11}$$

where the infinite product converges.

Using (3.5) the stochastic exponential can also be expressed as

$$\mathcal{E}_X(t) = e^{\left(X(t) - \frac{1}{2}[X,X]^c(t)\right)} \prod_{0<s\leq t} (1 + \triangle X(s))\, e^{(-\triangle X(s))}.$$

When $X(t)$ is a continuous semimartingale, equation (3.11) simplifies considerably to

$$\mathcal{E}_X(t) = e^{\left(X(t) - \frac{1}{2}[X,X](t)\right)}.$$

As it can be noticed, the stochastic exponential of $X(t)$ produces a different process to the ordinary exponential of $X(t)$, $e^{X(t)}$. In fact, contrarily to the ordinary exponential, which is obviously a positive process, the stochastic exponential is not necessarily positive. It is easy to see from (3.11) that the stochastic exponential is always non-negative if all jumps of $X(t)$ are greater than -1. Assuming now that $X(t)$ is a Lévy process and that $\mathcal{E}_X(t) > 0$, then it can be shown that the stochastic exponential of $X(t)$ can be rewritten as the ordinary exponential of another Lévy process, and vice versa [1].

Proposition 3.45. [1] *Let $(\mathbf{Z}, \mathcal{E})$ be a measurable space satisfying Assumption 3.16 with $d = 1$. Let $(X(t), t \geq 0)$ be an unidimensional Lévy-Itô process of the form (3.6) satisfying that $\nu((-\infty, -1]) = 0$ so that $\mathcal{E}_X(t) > 0$. Then the stochastic exponential of $X(t)$ has the alternative form of $\mathcal{E}_X(t) = e^{S_X(t)}$, where $(S_X(t), t \geq 0)$ is a Lévy process satisfying the equation*

$$\begin{aligned} dS_X(t) &= \beta(t)dW(t) + \left(\alpha(t) - \frac{1}{2}\beta^2(t)\right)dt \\ &\quad + \int_{\mathbf{Z}} \ln(1 + \gamma(t,z))\, \bar{N}(dt, dz) \\ &\quad + \int_{\mathbf{Z}_0} \Big[\ln(1 + \gamma(t,z)) - \gamma(t,z)\Big]\nu(dz)dt. \end{aligned} \tag{3.12}$$

Proof. Applying Itô formula to (3.12) with $F(x) = e^x$, we obtain

$$\begin{aligned} e^{S_X(t)} &= 1 + \int_0^t e^{S_X(s-)} \Big[\alpha(s) - \frac{1}{2}\beta^2(s) \\ &\quad + \int_{\mathbf{Z}_0} \left(\ln\left(1+\gamma(s,z)\right) - \gamma(s,z) \right) \nu(dz) \Big] ds \\ &\quad + \int_0^t e^{S_X(s-)} \beta(s) dW(s) + \int_0^t \frac{1}{2} e^{S_X(s-)} \beta^2(s) ds \\ &\quad + \int_0^t \int_{\mathbf{Z}_0} \Big[e^{S_X(s-)+\ln(1+\gamma(s,z))} - e^{S_X(s-)} \\ &\quad - e^{S_X(s-)} \ln\left(1+\gamma(s,z)\right) \Big] \nu(dz) ds \\ &\quad + \int_0^t \int_{\mathbf{Z}} \Big[e^{S_X(s-)+\ln(1+\gamma(s,z))} - e^{S_X(s-)} \Big] \bar{N}(ds,dz), \end{aligned}$$

which can be simplified to

$$\begin{aligned} e^{S_X(t)} &= 1 + \int_0^t e^{S_X(s-)} \Big[\alpha(s) + \int_{\mathbf{Z}_0} \left(\ln\left(1+\gamma(s,z)\right) - \gamma(s,z) \right) \nu(dz) \Big] ds \\ &\quad + \int_0^t e^{S_X(s-)} \beta(s) dW(s) + \int_0^t \int_{\mathbf{Z}} e^{S_X(s-)} \gamma(s,z) \bar{N}(ds,dz) \\ &\quad + \int_0^t \int_{\mathbf{Z}_0} e^{S_X(s-)} \Big[\gamma(s,z) - \ln\left(1+\gamma(s,z)\right) \Big] \nu(dz) ds. \end{aligned}$$

Removing the elements that cancel each other out, we obtain

$$\begin{aligned} e^{S_X(t)} &= 1 + \int_0^t e^{S_X(s-)} \alpha(s) ds + \int_0^t e^{S_X(s-)} \beta(s) dW(s) \\ &\quad + \int_0^t \int_{\mathbf{Z}} e^{S_X(s-)} \gamma(s,z) \bar{N}(ds,dz), \end{aligned}$$

which can also be expressed as

$$e^{S_X(t)} = 1 + \int_0^t e^{S_X(s-)} \left[\alpha(s) ds + \beta(s) dW(s) + \int_{\mathbf{Z}} \gamma(s,z) \bar{N}(ds,dz) \right].$$

Since $X(t)$ satisfies the stochastic differential equation (3.7) we can write

$$e^{S_X(t)} = 1 + \int_0^t e^{S_X(s-)} dX(s),$$

which is identical to (3.10) if we denote $e^{S_X(t)}$ by $Z(t)$. Therefore, the ordinary exponential of $S_X(t)$ satisfies the same equation as the stochastic exponential of $X(t)$. □

Processes of the form $e^{S_X(t)}$, where $S_X(t)$ solves (3.12), are an important element in the *change of measure* technique. Given a probablity space, the change to an equivalent measure involves assigning new probabilities to the events on the original probability space; this procedure plays an important role in mathematical finance [11].

Definition 3.46. [24] Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space equipped with the filtration $(\mathcal{F}_t, t \in [0,T])$. Two probability measures $\mathbb{P}$ and $\mathbb{Q}$ on $(\Omega, \mathcal{F}_T)$ are said to be equivalent if, for any event $A \in \mathcal{F}_T$, the equality $\mathbb{P}(A) = 0$ holds if and only if $\mathbb{Q}(A) = 0$.

In other words, both probability measures are equivalent on $(\Omega, \mathcal{F}_T)$ if they have the same set of null events in the sub σ-algebra $\mathcal{F}_T$. Then $\mathbb{P}$ and $\mathbb{Q}$ also enjoy the property of being equivalent on any sub σ-algebra $(\mathcal{F}_t, t \in [0,T])$. One way of generating equivalent measures is given by the following Definition.

Definition 3.47. (Radon-Nikodým derivative) [24] Let $\mathbb{P}$ and $\mathbb{Q}$ be two equivalent probability measures on $(\Omega, \mathcal{F}_T)$. The Radon–Nikodým derivative of $\mathbb{Q}$ with respect to $\mathbb{P}$ is defined as the unique $\mathcal{F}_T$-measurable random variable $\eta(T)$ such that7

$$\frac{d\mathbb{Q}}{d\mathbb{P}} = \eta(T). \tag{3.13}$$

The Radon–Nikodým derivative satisfies that $\mathbb{E}_{\mathbb{P}}(\eta(T)) = 1$, where $\mathbb{E}_{\mathbb{P}}$ denotes the expectation under $\mathbb{P}$. Formula (3.13) defines then a probability measure $\mathbb{Q}$ on $(\Omega, \mathcal{F}_T)$ that is equivalent to $\mathbb{P}$. It also implies that for any $\mathbb{Q}$-integrable random variable ξ we have that $\mathbb{E}_{\mathbb{Q}}(\xi) = \mathbb{E}_{\mathbb{P}}(\xi\eta(T))$.

Definition 3.48. (Radon-Nikodým density process) [24] The Radon–Nikodým density process $(\eta(t), t \in [0,T])$ of $\mathbb{Q}$ with respect to $\mathbb{P}$ and the filtration $\mathcal{F}_t$ is defined, for $t \in [0,T]$, as

$$\eta(t) = \mathbb{E}_{\mathbb{P}}\left(\eta(T)\Big|_{\mathcal{F}_t}\right) = \mathbb{E}_{\mathbb{P}}\left(\frac{d\mathbb{Q}}{d\mathbb{P}}\Big|_{\mathcal{F}_t}\right).$$

The Radon-Nikodým density process is a strictly positive martingale under $\mathbb{P}$ verifying that $\mathbb{E}_{\mathbb{P}}(\eta(t)) = 1$.

Since processes of the form $e^{S_X(t)}$, with $S_X(t)$ solving (3.12), can be used as a source of Radon–Nikodým derivatives, we need to know under which conditions $e^{S_X(t)}$ is, first, a local martingale.

Theorem 3.49. ***(Preservation of local martingale property)*** [1, 28] *Let the process $(X(t), t \in [0,T])$ be an unidimensional Lévy-Itô process satisfying Theorem 3.50, and let $H(t) \in \mathcal{L}^2_{pred}([0,T])$. Then, for $t \in [0,T]$, the stochastic integral $\int_0^t H(s)dX(s)$ is a local martingale.*

Theorem 3.50. [1] *Let $(\mathbf{Z}, \mathcal{E})$ be a measurable space satisfying Assumption 3.16 with $d = 1$. Let $(X(t), t \in [0,T])$ be an unidimensional Lévy-Itô process of the form (3.6) such that*

$$\int_0^t \int_{\mathbf{Z}\setminus\mathbf{Z}_0} \gamma(s,z)\nu(dz)ds < \infty,$$

then $X(t)$ and $\mathcal{E}_X(t)$ are local martingales if and only if

$$\alpha(t) + \int_{\mathbf{Z}\setminus\mathbf{Z}_0} \gamma(t,z)\nu(dz) = 0$$

almost surely for $t \in [0,T]$.

Proof. The process $Z(t)$ (or equivalently, $e^{S_X(t)}$) satisfies equation

$$dZ(t) = Z(t-)dX(t), \tag{3.14}$$

where $X(t)$ satisfies (3.7). By Theorem 3.49 $Z(t)$ is a local martingale if and only if

$$\alpha(t) + \int_{\mathbf{Z}\setminus\mathbf{Z}_0} \gamma(t,z)\nu(dz) = 0$$

almost surely for $t \geq 0$. □

Therefore, given a Lévy-Itô process $X(t)$, we have found the condition under which $\mathcal{E}_X(t)$ (or equivalently, $e^{S_X(t)}$) is a local martingale. The conditions required for $\mathcal{E}_X(t)$ to be in fact a martingale are given in the following Theorem.

Theorem 3.51. [1] *If $(X(t), t \geq 0)$ is an unidimensional Lévy-Itô process such that $\mathcal{E}_X(t)$ is a local martingale, then $\mathcal{E}_X(t)$ is a martingale if and only if $\mathbb{E}(\mathcal{E}_X(t)) = 1$ for all $t \in [0,T]$.*

Lemma 3.52. [28] *Let $\mathbb{Q}$ and $\mathbb{P}$ be equivalent probability measures on $(\Omega, \mathcal{F}_T)$, and let $(\eta(t), t \in [0,T])$ be the Radon–Nikodým density process. An adapted, càdlàg process $(X(t), t \in [0,T])$ is a $\mathbb{Q}$-local martingale if and only if $(\eta(t)X(t), t \in [0,T])$ is a $\mathbb{P}$-local martingale.*

Definition 3.53. [26] Let $X(t), Y(t) \in \mathbb{R}$ be two càdlàg semimartingales. The quadratic covariation of $X(t)$ and $Y(t)$ is the unique semimartingale such that

$$X(t)Y(t) = X(0)Y(0) + \int_0^t X(s-)dY(s) + \int_0^t Y(s-)dX(s) + [X,Y](t).$$

If we assume now that we have a process $Z(t)$ of the form (3.14) satisfying Theorems 3.50 and 3.51, i.e.

$$dZ(t) = Z(t-)\left[\beta_1(t)dW(t) + \int_{\mathbf{Z}} \gamma_1(t,z)\widetilde{N}(dt,dz)\right], \tag{3.15}$$

and a Lévy-Itô process $Y(t)$ of the form (3.6), i.e.

$$dY(t) = \alpha_2(t)dt + \beta_2(t)dW(t) + \int_{\mathbf{Z}} \gamma_2(t,z)\bar{N}(dt,dz), \tag{3.16}$$

we can find the conditions that the processes $\beta_1(t)$ and $\gamma_1(t,z)$ must satisfy so that $Z(t)$ can be use as a Radon-Nikodým density process. The process $Z(t)$ is the stochastic exponential of a certain process satisfying the stochastic differential equation

$$dU(t) = \beta_1(t)dW(t) + \int_{\mathbf{Z}} \gamma_1(t,z)\widetilde{N}(dt,dz),$$

so that if $Z(t)$ is to be used as a Radon-Nikodým density process, we need to show that Lemma 3.52 is satisfied and that the process $Z(t)Y(t)$ is a local martingale in the original probability space.

Definition 3.54. [1] Let $X(t), Y(t) \in \mathbb{R}$ be two Lévy-Itô processes, then the quadratic covariation of $X(t)$ and $Y(t)$ is defined as

$$[X,Y](t) = \int_0^t \beta_1(s)\beta_2(s)ds + \int_0^t \int_{\mathbf{Z}} \gamma_1(s)\gamma_2(s)N(ds,dz).$$

Using Definition 3.53 we find the expression for the stochastic differential equation satisfied by $Z(t)Y(t)$, i.e.

$$\begin{aligned} d\left(Z(t)Y(t)\right) &= Z(t-)dY(t) + Y(t-)dZ(t) + \beta_1(t)\beta_2(t)Z(t-)dt \\ &\quad + \int_{\mathbf{Z}} \gamma_1(t,z)\gamma_2(t,z)Z(t-)N(dt,dz). \end{aligned}$$

We now replace $dZ(t)$ and $dY(t)$ with the right-hand side of equations (3.15) and (3.16), respectively. We will also use the equality $N(dt,dz) = \widetilde{N}(dt,dz) + \nu(dz)ds$, so that

$$\begin{aligned} d\left(Z(t)Y(t)\right) &= Z(t-)\left[\alpha_2(t)dt + \beta_2(t)dW(t) + \int_{\mathbf{Z}} \gamma_2(t,z)\bar{N}(dt,dz)\right] \\ &\quad + Y(t-)Z(t-)\left[\beta_1(t)dW(t) + \int_{\mathbf{Z}} \gamma_1(t,z)\widetilde{N}(dt,dz)\right] \\ &\quad + \int_{\mathbf{Z}} \gamma_1(t,z)\gamma_2(t,z)Z(t-)\widetilde{N}(dt,dz) \\ &\quad + \int_{\mathbf{Z}} \gamma_1(t,z)\gamma_2(t,z)Z(t-)\nu(dz)dt + \beta_1(t)\beta_2(t)Z(t-)dt. \end{aligned}$$

Grouping all terms that are local martingales first, we obtain

$$\begin{aligned} d\left(Z(t)Y(t)\right) &= Z(t-)\Bigg[\beta_2(t)dW(t) + \int_{\mathbf{Z}_0} \gamma_2(t,z)\widetilde{N}(dt,dz) \\ &\quad + Y(t-)\left(\beta_1(t)dW(t) + \int_{\mathbf{Z}} \gamma_1(t,z)\widetilde{N}(dt,dz)\right) \\ &\quad + \int_{\mathbf{Z}} \gamma_1(t,z)\gamma_2(t,z)\widetilde{N}(dt,dz) + \int_{\mathbf{Z}\setminus\mathbf{Z}_0} \gamma_2(t,z)\widetilde{N}(dt,dz) \\ &\quad + \int_{\mathbf{Z}} \gamma_1(t,z)\gamma_2(t,z)\nu(dz)dt + \int_{\mathbf{Z}\setminus\mathbf{Z}_0} \gamma_2(t,z)\nu(dz)dt \\ &\quad + \beta_1(t)\beta_2(t)dt + \alpha_2(t)dt\Bigg], \end{aligned}$$

which, for our purposes, can be expressed as

$$\begin{aligned} d\left(Z(t)Y(t)\right) &= Z(t-)\Bigg[\alpha_2(t)dt + \beta_1(t)\beta_2(t)dt \\ &\quad + \int_{\mathbf{Z}} \gamma_1(t,z)\gamma_2(t,z)\nu(dz)dt + \int_{\mathbf{Z}\setminus\mathbf{Z}_0} \gamma_2(t,z)\nu(dz)dt\Bigg] \\ &\quad + (local)\ martingale. \end{aligned}$$

Thus, the process $Z(t)Y(t)$ will be a local martingale in the original probability space if and only if

$$\alpha_2(t) = -\beta_1(t)\beta_2(t) - \int_{\mathbf{Z}} \gamma_1(t,z)\gamma_2(t,z)\nu(dz) - \int_{\mathbf{Z}\setminus\mathbf{Z}_0} \gamma_2(t,z)\nu(dz).$$

Then by Lemma 3.52 the process $Z(t)$ defines an equivalent probability measure under which $Y(t)$ is a local martingale. We are now in position to formulate the following Theorem, which will describe the change of measure by means of the Radon-Nikodým density process.

Theorem 3.55. ***(Girsanov theorem for Lévy-Itô processes)*** [26] *Let* $(\mathbf{Z}, \mathcal{E})$ *be a measurable space satisfying Assumption 3.16. Let* $(X(t), t \in [0,T])$ *be an* $\mathbb{R}^n$*-valued Lévy-Itô process of the form given by (3.6) such that*

$$\int_{\mathbf{Z}\setminus\mathbf{Z}_0} \gamma(t,z)\nu(dz) = 0.$$

Assume that there exist $\mathcal{F}_t$*-predictabe processes* $u(t) \in \mathbb{R}^m$ *and* $\theta(t,z) \in \mathbb{R}^d$ *such that*

$$\alpha(t) = \beta(t)u(t) + \int_{\mathbf{Z}} \gamma(t,z)\theta(t,z)\nu\left(dz\right), \ \forall t \in [0,T],$$

and such that the process

$$\begin{aligned} Z(t) \quad = \quad & \exp\Bigg\{ - \int_0^t u(s)dW(s) - \frac{1}{2}\int_0^t u^2(s)ds \\ & + \int_0^t \int_{\mathbf{Z}} \ln\left(1-\theta(s,z)\right)\widetilde{N}\left(ds,dz\right) \\ & + \int_0^t \int_{\mathbf{Z}} \Big[\ln\left(1-\theta(s,z)\right) + \theta(s,z)\Big]\nu\left(dz\right)ds \Bigg\}, \end{aligned} \tag{3.17}$$

is well-defined for $t \in [0,T]$*. If we assume that* $\mathbb{E}\left(Z(T)\right) = 1$*, then it is possible to define the probability measure* $\mathbb{Q}$ *on* $\mathcal{F}_T$ *as* $d\mathbb{Q} = Z(T)d\mathbb{P}$*, such that* $W_{\mathbb{Q}}(t) = u(t)dt + W(t)$ *is a* $(\mathcal{F}_t, \mathbb{Q})$*-standard Wiener process and* $\widetilde{N}_{\mathbb{Q}}(dt,dz) = \theta(t,z)\nu(dz)dt + \widetilde{N}(dt,dz)$ *is a* $(\mathcal{F}_t, \mathbb{Q})$*-compensated Poisson random measure. Moreover, then* $X(t)$ *is a* $\mathbb{Q}$*-local martingale.*

Notice that formula (3.17) is the solution to the stochastic differential equation

$$dZ(t) = Z(t-)\left[-u(t)dW(t) - \int_{\mathbf{Z}} \theta(t,z)\widetilde{N}(dt,dz)\right].$$

4. Stochastic control

This section is an introduction to the mathematical methods for optimisation of dynamic stochastic systems and their application to problems arising in economics or finance. Typical problems targeted by these techniques are, for example, portfolio allocation and production-consumption problems [27]. When analysing problems of these types, we observe the current state of the system and study how a control policy can be used to move the system from its present status to a future more desirable state. These tools give analysts a better understanding of dynamic systems under uncertainty.

4.1. Elements of a stochastic control problem.

Stochastic optimisation problems can be subdivided into finite and infinite horizon problems, where we optimise the objective either over a given finite or over an infinite time interval. Since the paper by Davis and Lleo [12] deals with a portfolio optimisation problem within a finite time interval, we will only treat finite horizon problems in this section.

In general terms, an optimal control problem consists of the following elements.

Definition 4.1. (State process) The state process is a stochastic process that describes the state of the system of interest. We will assume that the dynamics

of the state process are given by an $\mathbb{R}^n$-valued controlled stochastic differential equation of the form

$$dX(t) = \alpha\left(t, X(t), h(t)\right) dt + \beta\left(t, X(t), h(t)\right) dW(t), \tag{4.1}$$

where $W(t)$ is a m-dimensional Wiener process on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, and $h(t)$ is a control policy valued in U, a subset of $\mathbb{R}^k$. The measurable functions $\alpha : [0,T] \times \mathbb{R}^n \times U \to \mathbb{R}^n$ and $\beta : [0,T] \times \mathbb{R}^n \times U \to \mathbb{R}^{n \times m}$ satisfy a uniform Lipschitz condition in U, i.e. for $t \in [0,T]$, $x, y \in \mathbb{R}^n$ and $h \in U$, there exists a constant K such that

$$|\alpha(t,x,h) - \alpha(t,y,h)| + |\beta(t,x,h) - \beta(t,y,h)| \leq K|x-y|, \tag{4.2}$$

and have linear growth, i.e. there exists a constant C such that

$$|\alpha(t,x,h)| + |\beta(t,x,h)| \leq C\left(1 + |h| + |x|\right). \tag{4.3}$$

Definition 4.2. (Control process) The evolution of the state process is influenced by the action of a control policy modelled as a process $(h(t), t \geq 0)$, whose value depends on the available information at any time t. The process $h(t)$ is progressively measurable with respect to $\mathcal{B}\left([0,t]\right) \otimes \mathcal{F}_t$.

We will restrict our attention to Markov controls, which should satisfy additionally some requirements, as given by the following Definition.

Definition 4.3. (Admissible controls) Given a finite time horizon $0 < T < \infty$, we denote by $\mathcal{A}$ the set of control processes $h(t)$ such that

$$\mathbb{E}\left[\int_0^T \left|\alpha\left(s, x, h(s)\right)\right| + \beta^2\left(s, x, h(s)\right) ds\right] < \infty \tag{4.4}$$

for all $(t,x) \in [0,T] \times \mathbb{R}^n$.

The conditions (4.2), (4.3) and (4.4) ensure for all $h(t) \in \mathcal{A}$ and for any initial condition $(t,x) \in [0,T] \times \mathbb{R}^n$ the existence and uniquess of a strong solution to the controlled stochastic differential equation (4.1). The set of all admissible controls $\mathcal{A}$ may depend on the initial value of the state process, thus we have the following Definition.

Definition 4.4. Let $f : [0,T] \times \mathbb{R}^n \times U \to \mathbb{R}$ and $g : \mathbb{R}^n \to \mathbb{R}$ be two measurable functions satisfying the quadratic growth condition

$$|f(t,x,h)| + |g(x)| \leq L\left(1 + |h| + x^2\right)$$

for some constant L and for all $(t,x,h) \in [0,T] \times \mathbb{R}^n \times U$. For $(t,x) \in [0,T] \times \mathbb{R}^n$ we denote by $\mathcal{A}(t,x)$ the subset of controls $h(t)$ in $\mathcal{A}$ such that

$$\mathbb{E}\left[\int_t^T \left|f\left(s, X_{t,x}(s), h(s)\right)\right| ds\right] < \infty,$$

where $(X_{t,x}(s), s \geq t)$ denotes the solution to (4.1) with control process $h(s)$ and initial condition (t,x) (i.e. at time t the process takes value x).

Associated with the system there is some objective function that may depend on the system state itself and on the control used. The objective function can be seen as a reward or a cost. Depending on the control problem, the optimisation entails maximising the reward or minimising the cost. We will focus on the maximisation problem, but the same concepts can be applied for a minimisation one.

Definition 4.5. (Functional objective) We define the functional objective J as

$$J(t,x,h) = \mathbb{E}\left[\int_t^T f\left(s, X_{t,x}(s), h(s)\right) ds + g\left(X_{t,x}(T)\right)\right], \tag{4.5}$$

for all $(t,x) \in [0,T] \times \mathbb{R}^n$ and $h(t) \in \mathcal{A}(t,x)$.

The functional objective $J(t,x,h)$ can be interpreted as the expected reward obtained in interval $[t,T]$ when control h is implemented and given that the control strategy in operation in the interval $[0,t]$ has left the system in the state x. Function $f(t,x,h)$ receives the name of *running profit function* and $g(x)$ is usually referred as *terminal reward function*. The goal of a stochastic control problem is then to maximise the functional objective over all admissible control processes.

Definition 4.6. (Value function) The value function $\Phi : [0,T] \times \mathbb{R}^n \to \mathbb{R}$ is given by

$$\Phi(t,x) = \sup_{h \in \mathcal{A}(t,x)} J\left(t,x,h\right). \tag{4.6}$$

Give an initial state $(t,x) \in [0,T] \times \mathbb{R}^n$ we say that $h^*(t) \in \mathcal{A}(t,x)$ is an optimal control if $\Phi(t,x) = J\left(t,x,h^*\right)$. Assuming that such optimal strategy $h^*(t)$ exists, we would like to find a way to actually calculate what it is. The idea behind the procedure is the well-known *dynamic programming principle* due to Bellman [3].

Theorem 4.7. ***(Dynamic programming principle)*** *Let $(t,x) \in [0,T] \times \mathbb{R}^n$, then we have*

$$\Phi(t,x) = \sup_{h \in \mathcal{A}(t,x)} \mathbb{E}\left[\int_t^\tau f\left(s, X_{t,x}(s), h(s)\right) ds + \Phi\left(\tau, X_{t,x}(\tau)\right)\right] \tag{4.7}$$

for any stopping time $\tau \in [t,T]$, or equivalently,

$$\Phi(t,x) \geq \mathbb{E}\left[\int_t^\tau f\left(s, X_{t,x}(s), h(s)\right) ds + \Phi\left(\tau, X_{t,x}(\tau)\right)\right]$$

for all $h(t) \in \mathcal{A}(t,x)$ and $\tau \in [t,T]$.

The premise of the dynamic programming principle is that an optimisation problem on the whole time interval $[0,T]$ can be split in smaller parts: first we search for an optimal control from time τ given that the state value is $X_{t,x}(\tau)$ (i.e. compute $\Phi\left(\tau, X_{t,x}(\tau)\right)$), and then maximise over controls on $[t,\tau]$ the quantity

$$\mathbb{E}\left[\int_t^\tau f\left(s, X_{t,x}(s), h(s)\right) ds + \Phi\left(\tau, X_{t,x}(\tau)\right)\right].$$

4.2. **The Hamilton-Jacobi-Bellman equation.** The Hamilton-Jacobi-Bellman (HJB) equation is the infinitesimal version of the dynamic programming principle; it describes the local behavior of the value function when we take the limit $\tau \to t$ in (4.7).

Theorem 4.8. ***(Hamilton-Jacobi-Bellman equation)*** [27] *Let $(X(t), t \in [0,T])$ be a state process satisfying (4.1). For a stochastic optimisation problem involving a functional objective of the form (4.5), the value function $\Phi(t,x)$, as defined in (4.6), satisfies the partial differential equation*

$$-\Phi_t\left(t,x\right) - H\left(t,x,\Phi_x\left(t,x\right), \Phi_{xx}\left(t,x\right)\right) = 0,$$

where, for all $(t,x,r,p) \in [0,T] \times \mathbb{R}^n \times \mathbb{R}^n \times \mathcal{S}^n$*, the function*

$$H(t,x,p,D) = \sup_{h \in \mathcal{A}(t,x)} \left[\alpha(t,x,h)p + \frac{1}{2} tr\left(\beta(t,x,h)\beta'(t,x,h)D\right) + f(t,x,h) \right]$$

is called the Hamiltonian of the associated control problem. The symbol tr *denotes the trace of a matrix and* $\mathcal{S}^n$ *is the set of symmetric* $n \times n$ *matrices.*

Proof. We start from (4.7) with $\tau = t + \Delta t$ and an arbitrary constant control $h \in \mathcal{A}(t,x)$. Since the control is not optimal, then the value function satisfies

$$\Phi(t,x) \geq \mathbb{E}\left[\int_t^{t+\Delta t} f(s, X_{t,x}(s), h)\, ds + \Phi\left(t + \Delta t, X_{t,x}(t+\Delta t)\right) \right]. \tag{4.8}$$

If we assume now that the value function Φ is smooth enough (i.e. $\mathcal{C}^1([0,T])$ and $\mathcal{C}^2(\mathbb{R}^n)$), we may then apply Itô formula (2.7) between t and $t+\Delta t$ and obtain

$$\begin{aligned}
\Phi\left(t+\Delta t, X_{t,x}(t+\Delta t)\right) &= \Phi(t,x) + \int_t^{t+\Delta t} \Phi_t\left(s, X_{t,x}(s)\right) ds \\
&\quad + \int_t^{t+\Delta t} \sum_{i=1}^n \Phi_{x_i}\left(s, X_{t,x}(s)\right) \alpha_i\left(s, X_{t,x}(s), h\right) ds \\
&\quad + \int_t^{t+\Delta t} \left[\sum_{i=1}^n \sum_{j=1}^m \Phi_{x_i}\left(s, X_{t,x}(s)\right) \right. \\
&\quad \left. \beta_{ij}\left(s, X_{t,x}(s), h\right) dW_j(s) \right] \\
&\quad + \frac{1}{2} \int_t^{t+\Delta t} \left[\sum_{k=1}^m \sum_{i,j=1}^n \Phi_{x_i x_j}\left(s, X_{t,x}(s)\right) \right. \\
&\quad \left. \beta_{ik}\left(s, X_{t,x}(s), h\right) \beta_{jk}\left(s, X_{t,x}(s), h\right) ds \right],
\end{aligned}$$

which can be written as

$$\begin{aligned}
\Phi\left(t+\Delta t, X_{t,x}(t+\Delta t)\right) &= \Phi(t,x) + \int_t^{t+\Delta t} \Phi_t\left(s, X_{t,x}(s)\right) ds \\
&\quad + \int_t^{t+\Delta t} \sum_{i=1}^n \Phi_{x_i}\left(s, X_{t,x}(s)\right) \alpha_i\left(s, X_{t,x}(s), h\right) ds \\
&\quad + \frac{1}{2} \int_t^{t+\Delta t} \left[\sum_{i,j=1}^n \left(\beta\left(s, X_{t,x}(s), h\right) \beta'\left(s, X_{t,x}(s), h\right)\right)_{ij} \right. \\
&\quad \left. \Phi_{x_i x_j}\left(s, X_{t,x}(s)\right) ds \right] \\
&\quad + (local)\ martingale,
\end{aligned} \tag{4.9}$$

where $\beta(t,x,h)\beta'(t,x,h)$ is an $n \times n$ matrix with components

$$\left(\beta(t,x,h)\beta'(t,x,h)\right)_{ij} = \sum_{k=1}^m \beta_{ik}(t,x,h)\beta_{jk}(t,x,h).$$

We define now the differential operator of second order L_t^h as

$$L_t^h \Phi(t,x) = \alpha(t,x,h)\Phi_x(t,x) + \frac{1}{2} tr\left(\beta(t,x,h)\beta'(t,x,h)\Phi_{xx}(t,x)\right),$$

which is called the infinitesimal generator associated to $X(t)$ with control h. Formula (4.9) can then be written as

$$\begin{aligned} \Phi(t+\Delta t, X_{t,x}(t+\Delta t)) &= \Phi(t,x) \\ &+ \int_t^{t+\Delta t} \Big[\Phi_t(s, X_{t,x}(s)) + L_t^h \Phi(s, X_{t,x}(s))\Big] ds \\ &+ (local)\ martingale, \end{aligned} \tag{4.10}$$

If we assume now that the local martingale in (4.10) is, in fact, a martingale, then we substitute (4.10) in (4.8) to obtain

$$0 \geq \mathbb{E}\left[\int_t^{t+\Delta t} \Big[\Phi_t(s, X_{t,x}(s)) + L_t^h \Phi(s, X_{t,x}(s)) + f(s, X_{t,x}(s), h)\Big] ds\right].$$

Dividing by Δt, and according to the mean-value theorem, when taking the limit $\Delta t \to 0$ we get

$$0 \geq \Phi_t(t,x) + L_t^h \Phi(t,x) + f(t,x,h),$$

which holds true for any arbitrary constant control $h \in \mathcal{A}(t,x)$ and thus we may obtain the inequality

$$-\Phi_t(t,x) - \sup_{h \in \mathcal{A}(t,x)} \left[L_t^h \Phi(t,x) + f(t,x,h)\right] \geq 0. \tag{4.11}$$

If we now assume that $h^*(t)$ is an optimal control, we would have in the preceding equations equality signs instead of inequality ones. By similar arguments we would find that

$$-\Phi_t(t,x) - L_t^{h^*} \Phi(t,x) - f(t,x,h^*(t)) = 0,$$

which combined with (4.11) suggests that the value function should satisfy

$$-\Phi_t(t,x) - \sup_{h \in \mathcal{A}(t,x)} \left[L_t^h \Phi(t,x) + f(t,x,h)\right] = 0,\ \forall (t,x) \in (0,T) \times \mathbb{R}^n, \tag{4.12}$$

if the supremum in h is finite. We express the Hamilton-Jacobi-Bellman equation (4.12) in the form

$$-\Phi_t(t,x) - H\left(t,x,\Phi_x(t,x), \Phi_{xx}(t,x)\right) = 0,$$

where the function

$$H(t,x,p,D) = \sup_{h \in \mathcal{A}(t,x)} \left[\alpha(t,x,h)p + \frac{1}{2} tr\left(\beta(t,x,h)\beta'(t,x,h)D\right) + f(t,x,h)\right],$$

and then the proof is completed. □

5. Risk-sensitive stochastic control

We take as starting point for the work presented in the rest of this dissertation the paper [12], where Davis and Lleo consider a dynamic portfolio optimisation problem on a finite time horizon, where infinite activity jumps in risky asset prices and correlation between factor risks and asset risks are allowed. The goal is to maximise, by means of risk-sensitive stochastic control, the expected growth rate of a portfolio of assets subject to a constraint on variance.

Risk-sensitive control generalises classical stochastic control by explicitly including the degree of risk aversion or risk tolerance of the optimising agent in the functional objective, and thus influencing directly the outcome of the optimisation. In contrast, classical stochastic control characterises the investor's risk aversion implictly through a utility function, as mentioned in Section 1. In risk-sensitive control, the decision maker's objective is to select a control policy $h(t)$ to maximise the criterion

$$J(t,x,h;\theta) = -\frac{1}{\theta}\ln\mathbb{E}\left(e^{-\theta F(t,x,h)}\right), \tag{5.1}$$

where $t \in [0,T]$ represents time, x is the state variable, F is a given reward function, and the risk sensitivity $\theta \in (-1,0)\cup(0,\infty)$ is an exogenous parameter representing the degree of risk aversion of the decision maker. The meaning of θ becomes evident by looking at the Taylor expansion of (5.1) around $\theta = 0$.

Proposition 5.1. *The Taylor expansion of (5.1) around $\theta = 0$ is*

$$J(t,x,h;\theta) = \mathbb{E}\left(F(t,x,h)\right) - \frac{\theta}{2}Var\left(F(t,x,h)\right) + O\left(\theta^2\right).$$

Proof. We follow here the same steps as in [33]. For a function $f(x) = e^{-\theta x}$ the expansion of $f\left(F(t,x,h)\right) = e^{-\theta F(t,x,h)}$ in powers of $F(t,x,h) - \mathbb{E}\left(F(t,x,h)\right)$ leads to

$$\begin{aligned}
\mathbb{E}\left(f\left(F(t,x,h)\right)\right) &= f\left(\mathbb{E}\left(F(t,x,h)\right)\right) + \frac{1}{2}\ddot{f}\left(\mathbb{E}\left(F(t,x,h)\right)\right)Var\left(F(t,x,h)\right) \\
&\quad + O\left(Var\left(F(t,x,h)\right)\right) \\
&= e^{-\theta\mathbb{E}(F(t,x,h))} + \frac{\theta^2}{2}e^{-\theta\mathbb{E}(F(t,x,h))}Var\left(F(t,x,h)\right) \\
&\quad + O\left(Var\left(F(t,x,h)\right)\right) \\
&= \mathbb{E}\left(e^{-\theta F(t,x,h)}\right).
\end{aligned}$$

This implies then that

$$\begin{aligned}
f^{-1}\left(\mathbb{E}\left(f\left(F(t,x,h)\right)\right)\right) &= \mathbb{E}\left(F(t,x,h)\right) + \frac{Var\left(F(t,x,h)\right)}{2}\frac{\ddot{f}\left(\mathbb{E}\left(F(t,x,h)\right)\right)}{\dot{f}\left(\mathbb{E}\left(F(t,x,h)\right)\right)} \\
&\quad + O\left(Var\left(F(t,x,h)\right)\right) \\
&= \mathbb{E}\left(F(t,x,h)\right) - \frac{\theta}{2}Var\left(F(t,x,h)\right) + O\left(\theta^2\right) \\
&= -\frac{1}{\theta}\ln\mathbb{E}\left(e^{-\theta F(t,x,h)}\right).
\end{aligned}$$

□

We say that the optimising agent is risk-averse if $\theta > 0$; on the other hand, a value of $\theta < 0$ would be chosen by a risk-seeking investor. The functional objective $J(t,x,h;\theta)$ can then be interpreted as the expected value of the reward function minus a penalty for variance and with an error proportional to θ^2. The notation $O(\theta)$ refers to a term whose absolute value is bounded by a constant multiple of $|\theta|$ as $\theta \to 0$.

In optimal investment problems we take $F(t,x,h(t)) = \ln V(t,x,h(t))$, where $V(t,x,h(t))$ is the value of the portfolio resulting from the implementation of an asset allocation strategy $h(t)$. Risk-sensitive control can be seen as a dynamic

Markowitz portfolio optimisation: we maximise the expected return subject to a constraint on the variance of returns.

5.1. **Analytical setting.** In [12] a market of assets comprised of one money market account $S_0(t)$ and m risky securities $S_1(t), S_2(t), \ldots, S_m(t)$ is considered. The dynamics of the price process of the risky securities are modelled as jump-diffusions whose growth rates depend on an auxiliary n-dimensional factor process that satisfies the stochastic differential equation

$$dX(t) = (b + BX(t))\,dt + \Lambda dW(t),\ X(0) = x, \tag{5.2}$$

where $b \in \mathbb{R}^n$, $B \in \mathbb{R}^{n\times n}$, $\Lambda \in \mathbb{R}^{n\times M}$ and $W(t)$ is an $\mathbb{R}^M$-valued Wiener process with $M = n+m$. The components $(W_i(t), 1 \le i \le M)$ of the M-dimensional Wiener process are defined on the underlying probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and adapted to the filtration $\mathcal{F}_t$.

Assumption 5.2. [12] *The matrix $\Lambda\Lambda'$ is positive definite.*

The components of the factor process $X(t)$ can be interpreted either as macroeconomic factors (GDP growth, inflation, unemployment rates, etc.) or simply as a statistical representation of the uncertainty of asset returns. We will assume that the factor process is directly observed, although it would be possible to use $X(t)$ as a latent factor to model the volatility of returns, in which case only the prices, and not $X(t)$, will be observed.

Lemma 5.3. *Let $(W(t), t \ge 0)$ be an $\mathbb{R}^M$-valued Wiener process and let $(X(t), t \ge 0)$ be an $\mathbb{R}$-valued stochastic process satisfying equation (5.2) for $n = 1$. Then the solution to (5.2) is given by*

$$X(t) \;=\; e^{Bt}\left[X(0) + \frac{b}{B}\left(1 - e^{-Bt}\right) + \int_0^t e^{-Bs}\Lambda dW(s)\right],$$

where $\Lambda \in \mathbb{R}^M$ and the stochastic integral $\int_0^t e^{-Bs}\Lambda dW(s)$ is a martingale.

Proof. We first consider the *homogeneous* differential equation

$$dX(t) = BX(t)dt,\ X(0) = x,$$

whose solution is $X(t) = X(0)e^{Bt}$. If we now take the function $F(t,x) = xe^{-Bt}$ (which is $\mathcal{C}^1$ in t and $\mathcal{C}^2$ in x) and apply Itô Formula with respect to equation (5.2), we obtain

$$\begin{aligned}
F(t, X(t)) \;&=\; F(0, X(0)) - \int_0^t BX(s)e^{-Bs}ds \\
&\quad + \int_0^t (b + BX(s))e^{-Bs}ds + \sum_{i=1}^M \int_0^t e^{-Bs}\Lambda_i dW_i(s) \\
&=\; F(0, X(0)) + \int_0^t be^{-Bs}ds + \sum_{i=1}^M \int_0^t e^{-Bs}\Lambda_i dW_i(s) \\
&=\; F(0, X(0)) + \frac{b}{B}\left(1 - e^{-Bt}\right) + \sum_{i=1}^M \int_0^t e^{-Bs}\Lambda_i dW_i(s).
\end{aligned}$$

And since $F(t, X(t)) = X(t)e^{-Bt}$, then

$$\begin{aligned} X(t) &= e^{Bt}\left[X(0) + \frac{b}{B}\left(1 - e^{-Bt}\right) + \sum_{i=1}^{M}\int_0^t e^{-Bs}\Lambda_i dW_i(s)\right] \\ &= e^{Bt}\left[X(0) + \frac{b}{B}\left(1 - e^{-Bt}\right) + \int_0^t e^{-Bs}\Lambda dW(s)\right]. \end{aligned}$$

□

The process $S_0(t)$ representing the wealth invested in the money market account is assumed to fulfill the dynamics given by the equation

$$dS_0(t) = S_0(t)\left(a_0 + A_0' X(t)\right)dt, \; S_0(0) = s_0, \tag{5.3}$$

where $a_0 \in \mathbb{R}$ is a scalar constant, $A_0 \in \mathbb{R}^n$ and $X(t)$ is the n-dimensional factor process. Throughout the dissertation x' denotes the transpose of the matrix or vector x.

Lemma 5.4. *Let $(S_0(t), t \geq 0)$ be an $\mathbb{R}$-valued stochastic process satisfying equation (5.3) for $n = 1$. Then the solution to equation (5.3) is given by*

$$S_0(t) = S_0(0)e^{\left(a_0 + A_0' X(t)\right)t},$$

where $A_0' = A_0$ since $n = 1$.

In [12] the price process of the risky assets is represented by a stochastic differential equation driven by a Wiener process and a Poisson random measure, with drifts that are functions of the factor process $X(t)$. The dynamics of the m-column vector of risky securities satisfy the stochastic differential equation

$$dS(t) = S(t-)\left[\left(a + AX(t)\right)dt + \Sigma dW(t) + \int_{\mathbf{Z}} \gamma(z)\bar{N}(dt, dz)\right], \; S(0) = s, \tag{5.4}$$

where $a \in \mathbb{R}^m$, $A \in \mathbb{R}^{m\times n}$, $\Sigma \in \mathbb{R}^{m\times M}$ is the matrix of the volatility of the logarithmic stock's returns, and $\gamma(z) \in \mathbb{R}^m$ satisfies Assumption 5.5. As mentioned before, the mean returns of the securities are explicitly affected by the factor process.

Assumption 5.5. *The function $\gamma(z) \in \mathbb{R}^m$ with components $\gamma_i(z) \in \mathbb{R}$ satisfies $-1 < \gamma_i^{\min} \leq \gamma_i(z) \leq \gamma_i^{\max} < \infty$ and $-1 < \gamma_i^{\min} \leq 0 \leq \gamma_i^{\max} < \infty$ for $1 \leq i \leq m$. Furthermore, the function $\gamma(z)$ satisfies*

$$\int_{\mathbf{Z}_0} \gamma^2(z)\nu(dz) < \infty.$$

By Assumption 5.5 every risky asset has, with positive probability, both upward and downward jumps. Furthermore, it also ensures that the jumps of $S(t)$ are greater than -1.

Assumption 5.6. [12] *The matrix $\Sigma\Sigma'$ is positive definite.*

Lemma 5.7. *Let $(S(t), t \geq 0)$ be the $\mathbb{R}$-valued stochastic process satisfying equation (5.4) for $m = 1$, i. e.*

$$dS(t) = S(t-)\left[\left(a + AX(t)\right)dt + \sum_{i=1}^{M}\sigma_i dW_i(t) + \int_{\mathbf{Z}} \gamma(z)\bar{N}(dt, dz)\right],$$

where $\Sigma \in \mathbb{R}^M$ and $\sigma_i \in \mathbb{R}$ is the i-th value of vector Σ. Then the solution to the equation is given by the process

$$\begin{aligned} S(t) &= S(0)\exp\Bigg\{\int_0^t AX(s)ds + \left(a - \frac{1}{2}\sum_{i=1}^M \sigma_i^2\right)t + \sum_{i=1}^M \sigma_i W_i(t) \\ &\quad + \int_{\mathbf{Z}_0}\Big[\ln\left(1+\gamma(z)\right) - \gamma(z)\Big]\nu(dz)t \\ &\quad + \int_0^t\int_{\mathbf{Z}} \ln\left(1+\gamma(z)\right)\bar{N}(ds,dz)\Bigg\}, \end{aligned}$$

which receives the name of geometric Lévy-Itô process.

Proof. Applying (3.9) with $F(x) = \ln(x)$ we obtain

$$\begin{aligned} F(S(t)) &= F(S(0)) + \int_0^t \left(a + AX(s)\right)ds + \sum_{i=1}^M \int_0^t \sigma_i dW_i(s) - \frac{1}{2}\sum_{i=1}^M \int_0^t \sigma_i^2 ds \\ &\quad + \int_0^t\int_{\mathbf{Z}_0}\Big[\ln\left(S(s-)+\gamma(z)S(s-)\right) - \ln\left(S(s-)\right) - \gamma(z)\Big]\nu(dz)ds \\ &\quad + \int_0^t\int_{\mathbf{Z}}\Big[\ln\left(S(s-)+\gamma(z)S(s-)\right) - \ln\left(S(s-)\right)\Big]\bar{N}(ds,dz). \end{aligned}$$

After rearranging we get the expression

$$\begin{aligned} F(S(t)) &= F(S(0)) + \int_0^t \left(a + AX(s) - \frac{1}{2}\sum_{i=1}^M \sigma_i^2\right)ds + \sum_{i=1}^M \int_0^t \sigma_i dW_i(s) \\ &\quad + \int_0^t\int_{\mathbf{Z}_0}\Big[\ln\left(1+\gamma(z)\right) - \gamma(z)\Big]\nu(dz)ds \\ &\quad + \int_0^t\int_{\mathbf{Z}} \ln\left(1+\gamma(z)\right)\bar{N}(ds,dz), \end{aligned}$$

which can be simplified to

$$\begin{aligned} F(S(t)) &= F(S(0)) + \int_0^t AX(s)ds + \left(a - \frac{1}{2}\sum_{i=1}^M \sigma_i^2\right)t + \sum_{i=1}^M \sigma_i W_i(t) \\ &\quad + \int_{\mathbf{Z}_0}\Big[\ln\left(1+\gamma(z)\right) - \gamma(z)\Big]\nu(dz)t \\ &\quad + \int_0^t\int_{\mathbf{Z}} \ln\left(1+\gamma(z)\right)\bar{N}(ds,dz) \end{aligned}$$

and after applying the exponential to both sides we obtain the final formula. Note that by Assumption 5.5 $\ln\left(1+\gamma(z)\right)$ is well-defined. □

Our investment strategy is steered by the control process $h(t) \in \mathbb{R}^m$. Each component $(h_i(t), 1 \le i \le m)$ represents the fraction of the current value of the portfolio invested in the i-th asset. Thus, the fraction invested in the money market account is then $h_0(t) = 1 - \sum_{i=1}^m h_i(t)$. The wealth of the investor in response to

an investment strategy $h(t)$ follows then the dynamics

$$\begin{aligned} dV(t) &= V(t-)\Big[(a_0 + A_0'X(t))\,dt + h'(t)\left(\hat{a} + \hat{A}X(t)\right)dt \\ &\quad + h'(t)\Sigma dW(t) + \int_{\mathbf{Z}} h'(t)\gamma(z)\bar{N}(dt,dz)\Big],\ V(0) = v, \end{aligned} \tag{5.5}$$

where $\hat{a} = a - a_0\mathbf{1}$ and $\hat{A} = A - A_0\mathbf{1}$, and where $\mathbf{1} \in \mathbb{R}^m$ denotes the m-dimensional unit vector.

Lemma 5.8. *Let $(V(t), t \geq 0)$ be the $\mathbb{R}$-valued stochastic process satisfying equation (5.5). Then $V(t)$ is given by the expression*

$$\begin{aligned} V(t) &= V(0)\exp\Bigg\{\int_0^t \Big[a_0 + A_0'X(s) + h'(s)\left(\hat{a} + \hat{A}X(s)\right) \\ &\quad - \frac{1}{2}h'(s)\Sigma\Sigma' h'(s)\Big]ds + \int_0^t h'(s)\Sigma dW(s) \\ &\quad + \int_0^t \int_{\mathbf{Z}_0} \Big[\ln\left(1 + h'(s)\gamma(z)\right) - h'(s)\gamma(z)\Big]\nu(dz)ds \\ &\quad + \int_0^t \int_{\mathbf{Z}} \ln\left(1 + h'(s)\gamma(z)\right)\bar{N}(ds,dz)\Bigg\}. \end{aligned} \tag{5.6}$$

Proof. Apply (3.8) with $F(x) = \ln(x)$. □

For the problem at hand we are interested in control processes that satisfy certain conditions. Definitions 5.9 and 5.11 specify which ones.

Definition 5.9. [12] Let $\mathcal{G}_t = \sigma\left((X(s), S(s)), 0 \leq s \leq t\right)$ for $t \in [0,T]$ be the filtration generated by the factor process and the price process of the risky securities. An $\mathbb{R}^m$-valued control process $h(t)$ belongs to class $\mathcal{H}([0,T])$ if:

(1) $h(t)$ is progressively measurable with respect to $\mathcal{B}([0,t]) \otimes \mathcal{G}_t$ and is càdlàg;
(2) $P\left(\int_0^T h^2(t)dt < \infty\right) = 1$, for all $T > 0$;
(3) $h'(t)\gamma(z) > -1$, for $t \in [0,T]$ and $z \in \mathbf{Z}$.

The third condition in Definition 5.9 guarantees that the jumps of $V(t)$ are greater than -1 and that $\ln\left(1 + h'(s)\gamma(z)\right)$ is well-defined. This condition is endogenous to the control problem and can be interpreted as a risk management safeguard preventing the investor from investing in some of the portfolios if the jump component of these portfolios could result in the investor's bankruptcy [12].

Lemma 5.10. *Under Assumption 5.5 a control process $h(t)$ satisfying the third condition in Definition 5.9 is bounded.*

Proof. The components $\gamma_i(z)$ of the process $\gamma(z)$ are bounded from below by -1. In order then to satisfy that $h'(t)\gamma(z) > -1$ the components $h_i(t)$ of the control process $h(t)$ are bounded from above by 1. In the case where only one component $\gamma_i(z)$ takes the value $\gamma_i^{\min}$ and the rest are zero, then the component $h_i(t)$ can take 1 as the maximum value, so that $h'(t)\gamma(z) > -1$. Similarly, the components $h_i(t)$ are bounded from below by $\frac{-1}{\gamma_i^{\max}}$, so that $h'(t)\gamma(z) > -1$ even when $\gamma_i(z)$ takes the value $\gamma_i^{\max}$. Again, when one of the components $\gamma_i(z)$ takes the value $\gamma_i^{\max}$ and the

rest are zero, then $h_i(t)$ can take the closest value to $\frac{-1}{\gamma_i^{\max}}$, so that $h'(t)\gamma(z) > -1$. In short, each component $h_i(t)$ thus satisfies

$$\frac{-1}{\gamma_i^{\max}} < h_i(t) \leq 1.$$

□

Definition 5.11. [12] A control process $h(t)$ is in class $\mathcal{A}(T)$ for all $T > 0$ if:

(1) $h(t) \in \mathcal{H}([0,T])$;
(2) $\mathbb{E}(\mathcal{E}_U(T)) = 1$, where $(\mathcal{E}_U(t), t \in [0,T])$, defined as

$$\begin{aligned}\mathcal{E}_U(t) &= \exp\Bigg\{-\int_0^t \theta h'(s)\Sigma dW(s) - \frac{1}{2}\int_0^t \theta^2 h'(s)\Sigma\Sigma' h(s) ds \\ &\quad + \int_0^t \int_{\mathbf{Z}} \ln\left(1 - G(z, h(s); \theta)\right) \widetilde{N}(ds, dz) \\ &\quad + \int_0^t \int_{\mathbf{Z}} \Big[\ln\left(1 - G(z, h(s); \theta)\right) + G(z, h(s); \theta)\Big] \nu(dz)\, ds\Bigg\},\end{aligned} \tag{5.7}$$

is the stochastic exponential (or Doléans-Dade exponential) of the process $(U(t), t \in [0,T])$ satisfying the equation

$$dU(t) = -\theta h'(t)\Sigma dW(t) - \int_{\mathbf{Z}} G(z, h(s); \theta)\, \widetilde{N}(dt, dz),$$

with $G(z, h; \theta) = 1 - (1 + h'\gamma(z))^{-\theta}$.

We say that a control process $h(t)$ is admissible if $h(t) \in \mathcal{A}(T)$.

The third condition in Definition 5.9 ensures that $G(z, h; \theta) < 1$ and therefore the jumps of $U(t)$ are greater than -1 and $\mathcal{E}_U(t) > 0$. We will see next that the second condition in Definition 5.11 plays a crucial role.

5.2. **Problem setup.** The goal of our investment strategy is to maximise the risk-adjusted growth of the portfolio over a finite time horizon. Keeping this in mind, the objective of the risk-sensitive management problem is to find a control process $h^*(t) \in \mathcal{A}(T)$ that maximises the functional objective (5.1). The greatest difficulty we have to deal with in the control problem is the presence of jumps in the process $V(t)$. However, extending Kuroda and Nagai's change of measure technique [18], Davis and Lleo remove the jumps and manage to turn the control problem into a more tractable challenge.

Theorem 5.12. [12] *Let $(V(t), t \in [0,T])$ be a wealth process of the form (5.6). Functional objective*

$$J(t, x, h; \theta) = -\frac{1}{\theta} \ln \mathbb{E}\left(e^{-\theta \ln V(t,x,h)}\right) \tag{5.8}$$

can be expressed as

$$J(t, x, h; \theta) = -\frac{1}{\theta} \ln \mathbb{E}\left[v^{-\theta} \exp\left\{\theta \int_0^t g(X(s), h(s); \theta)\, ds\right\} \mathcal{E}_U(t)\right],$$

where

$$\begin{aligned} g(x,h;\theta) &= \frac{1}{2}(\theta+1)\,h'\Sigma\Sigma'h - a_0 - A_0'x - h'\left(\hat{a}+\hat{A}x\right) \\ &\quad + \int_{\mathbf{Z}} \frac{1}{\theta}\left[(1+h'\gamma(z))^{-\theta}-1\right]\nu(dz) + \int_{\mathbf{Z}_0} h'\gamma(z)\nu(dz), \end{aligned}$$

and $\mathcal{E}_U(t)$ is the Doléans-Dade exponential given by (5.7).

Proof. We start by applying (3.8) to equation (5.5) with $F(x) = \ln(x)$ so that we obtain

$$\begin{aligned} \ln V(t) &= \ln v + \int_0^t \Big[a_0 + A_0'X(s) + h'(s)\left(\hat{a}+\hat{A}X(s)\right)\Big]ds \\ &\quad -\frac{1}{2}\int_0^t h'(s)\Sigma\Sigma'h(s)ds + \int_0^t h'(s)\Sigma dW(s) \\ &\quad + \int_0^t\int_{\mathbf{Z}_0}\Big[\ln\left(1+h'(s)\gamma(z)\right) - h'(s)\gamma(z)\Big]\nu(dz)ds \\ &\quad + \int_0^t\int_{\mathbf{Z}}\ln\left(1+h'(s)\gamma(z)\right)\bar{N}(ds,dz). \end{aligned}$$

We can thus define the process $e^{-\theta\ln V(t)}$ as

$$\begin{aligned} e^{-\theta\ln V(t)} &= v^{-\theta}\exp\Bigg\{-\int_0^t \theta\Big[a_0 + A_0'X(s) + h'(s)\left(\hat{a}+\hat{A}X(s)\right)\Big]ds \\ &\quad +\frac{\theta}{2}\int_0^t h'(s)\Sigma\Sigma'h(s)ds - \int_0^t \theta h'(s)\Sigma dW(s) \\ &\quad -\int_0^t\int_{\mathbf{Z}_0}\theta\Big[\ln\left(1+h'(s)\gamma(z)\right) - h'(s)\gamma(z)\Big]\nu(dz)ds \\ &\quad -\int_0^t\int_{\mathbf{Z}}\theta\ln\left(1+h'(s)\gamma(z)\right)\bar{N}(ds,dz)\Bigg\}. \end{aligned}$$

Multiplying and dividing at the same time $e^{-\theta\ln V(t)}$ by the terms

$$\exp\left\{\int_0^t\int_{\mathbf{Z}}\left[1-(1+h'(s)\gamma(z))^{-\theta}\right]\nu(dz)ds\right\},$$

$$\exp\left\{\int_0^t\int_{\mathbf{Z}\setminus\mathbf{Z}_0}\theta\ln\left(1+h'(s)\gamma(z)\right)\nu(dz)ds\right\},$$

and

$$\exp\left\{\frac{\theta^2}{2}\int_0^t h'(s)\Sigma\Sigma'h(s)ds\right\},$$

we obtain

$$\begin{aligned} e^{-\theta \ln V(t)} &= v^{-\theta} \exp\Bigg\{ - \int_0^t \theta \Big[a_0 + A_0' X(s) + h'(s) \left(\hat{a} + \hat{A} X(s) \right) \Big] ds \\ &\quad + \frac{1}{2} \int_0^t \theta \left(\theta + 1 \right) h'(s) \Sigma \Sigma' h(s) ds \\ &\quad + \int_0^t \int_{\mathbf{Z}} \left[\left(1 + h'(s) \gamma(z) \right)^{-\theta} - 1 \right] \nu(dz) ds \\ &\quad + \int_0^t \int_{\mathbf{Z}_0} \theta h'(s) \gamma(z) \nu(dz) ds \Bigg\} \\ &\quad \exp\Bigg\{ - \int_0^t \theta h'(s) \Sigma dW(s) - \frac{\theta^2}{2} \int_0^t h'(s) \Sigma \Sigma' h(s) ds \\ &\quad - \int_0^t \int_{\mathbf{Z}} \theta \ln \left(1 + h'(s) \gamma(z) \right) \widetilde{N}(ds, dz) \\ &\quad - \int_0^t \int_{\mathbf{Z}} \Big[\theta \ln \left(1 + h'(s) \gamma(z) \right) - \left(1 - \left(1 + h'(s) \gamma(z) \right)^{-\theta} \right) \Big] \nu(dz) ds \Bigg\}. \end{aligned}$$

Since $G\left(z, h; \theta\right) = 1 - \left(1 + h' \gamma(z)\right)^{-\theta}$, then.

$$e^{-\theta \ln V(t)} = v^{-\theta} \exp\left\{ \theta \int_0^t g\left(X(s), h(s); \theta \right) ds \right\} \mathcal{E}_U(t).$$

□

We have found a way to express $e^{-\theta \ln V(t,x,h)}$ as the product of two processes: $v^{-\theta} \exp\left\{ \theta \int_0^t g\left(X(s), h(s); \theta \right) ds \right\}$ and $\mathcal{E}_U(t)$. Since $\mathcal{E}_U(t)$ is a stochastic exponential process, we would like to use it as a Radon–Nikodým density process, so that we can perform a change of measure and replace $\mathbb{E}\left(e^{-\theta \ln V(t,x,h)} \right)$ by the expectaton of $v^{-\theta} \exp\left\{ \theta \int_0^t g\left(X(s), h(s); \theta \right) ds \right\}$ under the equivalent probability measure. In order to achieve this, we need $\mathcal{E}_U(t)$ to be a martingale. Furthermore, as stated in Definition 5.11, the admissibility of the control process $h(t)$ also depends on $\mathcal{E}_U(t)$ being a martingale.

Lemma 5.13. [12] *The Doléans-Dade exponential $\mathcal{E}_U(t)$ is a martingale.*

Since $\mathcal{E}_U(t)$ is a martingale we can define a probability measure $\mathbb{Q}$ on $(\Omega, \mathcal{F}_t)$ that is equivalent to $\mathbb{P}$. By Definition 3.47 we can define $\mathbb{Q}$ via the Radon-Nikodým derivative

$$\frac{d\mathbb{Q}}{d\mathbb{P}} = \mathcal{E}_U(t).$$

By Theorem 3.55 we can then define

$$W_{\mathbb{Q}}(t) = \int_0^t \theta \Sigma' h(s) ds + W(t),$$

which is a $(\mathcal{F}_t, \mathbb{Q})$-standard Wiener process and

$$\begin{aligned}\int_0^t \int_{\mathbf{Z}} \widetilde{N}_{\mathbb{Q}}(ds,dz) &= \int_0^t \int_{\mathbf{Z}} G\left(z,h(s);\theta\right)\nu(dz)ds + \int_0^t \int_{\mathbf{Z}} \widetilde{N}(ds,dz) \\ &= \int_0^t \int_{\mathbf{Z}} G\left(z,h(s);\theta\right)\nu(dz)ds + \int_0^t \int_{\mathbf{Z}} N(ds,dz) \\ &\quad - \int_0^t \int_{\mathbf{Z}} \nu(dz)ds \\ &= \int_0^t \int_{\mathbf{Z}} N(ds,dz) - \int_0^t \int_{\mathbf{Z}} \Big[1 - G\left(z,h(s);\theta\right)\Big]\nu(dz)ds \\ &= \int_0^t \int_{\mathbf{Z}} N(ds,dz) - \int_0^t \int_{\mathbf{Z}} \left(1 + h'(s)\gamma(z)\right)^{-\theta}\nu(dz)ds,\end{aligned}$$

where $\widetilde{N}_{\mathbb{Q}}$ is a $(\mathcal{F}_t, \mathbb{Q})$-compensated Poisson random measure. As a result, the factor process satisfies on the probability space $(\Omega, \mathcal{F}_t, \mathbb{Q})$ the stochastic differential equation

$$dX(t) = \left(b + BX(t) - \theta\Lambda\Sigma' h(t)\right)dt + \Lambda dW_{\mathbb{Q}}(t),\ t \in [0,T]\,. \tag{5.9}$$

Furthermore, the functional objective (5.8) can be expressed as

$$J(t,x,h;\theta) = -\frac{1}{\theta}\ln \mathbb{E}_{\mathbb{Q}}\left[v^{-\theta}\exp\left\{\theta\int_0^t g\left(X(s),h(s);\theta\right)ds\right\}\right],$$

where $\mathbb{E}_{\mathbb{Q}}$ denotes the expectation taken with respect to the probability measure $\mathbb{Q}$. The functional objective comprises the function $g\left(x,h;\theta\right)$, which depends on the factor process and the control. We have thus reduced the risk-sensitive control problem to a stochastic control problem in the factor process, which does not have jumps. Furthermore, we can derive the auxiliary functional objective directly associated with the risk-sensitive control problem

$$I(t,x,h;\theta,T,v) = -\frac{1}{\theta}\ln \mathbb{E}_{\mathbb{Q}}\left[\exp\left\{\theta\int_t^T g\left(X_{t,x}(s),h(s);\theta\right)ds - \theta\ln v\right\}\right],$$

and an exponentially transformed functional criterion

$$\widetilde{I}(t,x,h;\theta,T,v) = \mathbb{E}_{\mathbb{Q}}\left[\exp\left\{\theta\int_t^T g\left(X_{t,x}(s),h(s);\theta\right)ds - \theta\ln v\right\}\right],$$

which we will find helpful to use in further derivations. The subscript t,x indicate the initial conditions for the factor process.

Theorem 5.14. ***(Feynman-Kac formula)*** [24] *Let $(X(t), t \in [0,T])$ be the $\mathbb{R}^n$-valued process that satisfy the stochastic differential equation*

$$dX(t) = \alpha\left(t,X(t)\right)dt + \beta\left(t,X(t)\right)dW(t),\ X(0) = x,$$

where $W(t)$ is a d-dimensional Wiener process and the measurable functions $\alpha : [0,T]\times\mathbb{R} \to \mathbb{R}$ and $\beta : [0,T]\times\mathbb{R}^n \to \mathbb{R}^{n\times d}$ satisfy the usual uniform Lipschitz and linear growth conditions. Let $f : [0,T]\times\mathbb{R}^n \to \mathbb{R}$ and $g : \mathbb{R}^n \to \mathbb{R}$ be continuous functions such that $g(x) \geq 0$ for all $x \in \mathbb{R}^n$ be a function bounded from below.

Assume that $v : [0,T] \times \mathbb{R}^n \to \mathbb{R}$ *is a function* $\mathcal{C}^1((0,T))$ *and* $\mathcal{C}^2(\mathbb{R}^n)$ *that is the unique bounded solution to the partial differential equation*

$$v_t(t,x) + \alpha(t,x)v_x(t,x) + \frac{1}{2}tr\left(\beta(t,x)\beta'(t,x)v_{xx}(t,x)\right) = f(t,x)v(t,x),$$

for all $(t,x) \in (0,T) \times \mathbb{R}^n$ *and subject to the terminal condition* $v(T,x) = g(x)$ *for all* $x \in \mathbb{R}^n$. *Then* $v(t,x)$ *has the following representation*

$$v(t,x) = \mathbb{E}\left[\exp\left\{-\int_t^T f(s,X(s))\,ds\right\} g(X(T))\Big| X(t) = x\right],$$

for all $(t,x) \in (0,T) \times \mathbb{R}^n$.

The Feynman-Kac formula expresses the solution of a parabolic partial differential equation as the expected value of a function of a stochastic processs driven by the Wiener process.

Theorem 5.15. [12] *Let* $\Phi(t,x)$ *be the value function for the auxiliary functional objective* $I(t,x,h;\theta,T,v)$. *Then* $\Phi(t,x)$ *is defined as*

$$\Phi(t,x) = \sup_{h\in\mathcal{A}(T)} I(t,x,h;\theta,T,v), \tag{5.10}$$

and it satisfies the Hamilton-Jacobi-Bellman equation

$$\Phi_t(t,x) + \sup_{h\in\mathcal{A}(T)} L_t^h\Phi(t,x) = 0,\ \forall\,(t,x) \in (0,T)\times\mathbb{R}^n,$$

where

$$\begin{aligned} L_t^h\Phi(t,x) &= \left(b + Bx - \theta\Lambda\Sigma' h(t)\right)'\Phi_x(t,x) + \frac{1}{2}tr\left(\Lambda\Lambda'\Phi_{xx}(t,x)\right) \\ &\quad -\frac{\theta}{2}\Phi_x'(t,x)\Lambda\Lambda'\Phi_x(t,x) - g(x,h(t);\theta), \end{aligned}$$

and subject to terminal condition $\Phi(T,x) = \ln v$ *for all* $x \in \mathbb{R}^n$.

Proof. We first define the auxiliary value function $\widetilde{\Phi}(t,x) = \exp(-\theta\Phi(t,x))$ and we multiple $I(t,x,h;\theta,T,v)$ by $-\theta$. We then apply an exponential transformation to the result, so that we obtain the auxiliary criterion

$$\widetilde{I}(t,x,h;\theta,T,v) = \mathbb{E}_{\mathbb{Q}}\left[\exp\left\{\theta\int_t^T g(X_{t,x}(s),h(s);\theta)\,ds\right\} v^{-\theta}\right].$$

If we now assume that $h^*(t)$ is an optimal control, then the value function can be defined as

$$\widetilde{\Phi}(t,x) = \mathbb{E}_{\mathbb{Q}}\left[\exp\left\{\theta\int_t^T g(X_{t,x}(s),h^*(s);\theta)\,ds\right\} v^{-\theta}\right],$$

Applying Theorem 5.14 we obtain that the value function $\widetilde{\Phi}(t,x)$ satisfies the partial differential equation

$$\begin{aligned} &\widetilde{\Phi}_t(t,x) + \left(b + BX(t) - \theta\Lambda\Sigma' h^*(t)\right)'\widetilde{\Phi}_x(t,x) \\ &+\frac{1}{2}tr\left(\Lambda\Lambda'\widetilde{\Phi}_{xx}(t,x)\right) + \theta g(x,h^*(t);\theta)\,\widetilde{\Phi}(t,x) \quad = \quad 0 \end{aligned}$$

with the terminal condition $\widetilde{\Phi}(T,x) = v^{-\theta}$. We now replace $\widetilde{\Phi}(t,x)$ by $\exp(-\theta\Phi(t,x))$ and then we divide by $-\theta\widetilde{\Phi}(t,x)$. After rearranging we obtain the equation

$$\begin{aligned} &\Phi_t(t,x) + \left(b + BX(t) - \theta\Lambda\Sigma' h^*(t)\right)' \Phi_x(t,x) \\ +\frac{1}{2}tr\left(\Lambda\Lambda'\Phi_{xx}(t,x)\right) - \frac{\theta}{2}\Phi_x'(t,x)\Lambda\Lambda'\Phi_x(t,x) - g\left(x,h^*(t);\theta\right) &= 0, \end{aligned}$$

which can be written as

$$\Phi_t\left(t,x\right) + L_t^{h^*}\Phi\left(t,x\right) = 0.$$

For any other control process $h(t) \in \mathcal{A}(T)$ that is not optimal, the value function will then satisfy

$$\Phi_t\left(t,x\right) + \sup_{h\in\mathcal{A}(T)} L_t^h\Phi\left(t,x\right) = 0,$$

as stated in the Theorem. □

As part of the proof we have seen that the value function for the exponentially transformed functional objective satisfies

$$\widetilde{\Phi}_t\left(t,x\right) + \frac{1}{2}tr\left(\Lambda\Lambda'\widetilde{\Phi}_{xx}(t,x)\right) + H\left(x,\widetilde{\Phi}(t,x),\widetilde{\Phi}_x(t,x)\right) = 0$$

subject to the terminal condition $\widetilde{\Phi}\left(T,x\right) = v^{-\theta}$, where

$$H\left(x,r,p\right) = \inf_{h\in\mathcal{A}(T)}\left[\left(b + Bx - \theta\Lambda\Sigma' h\right)' p + \theta g\left(x,h;\theta\right) r\right].$$

Since $\Phi(t,x)$ and $\widetilde{\Phi}(t,x)$ are related through a strictly monotone continuous transformation, an admissible (or optimal) control policy for the exponentially transformed problem is also admissible (or optimal) for the risk-sensitive problem. Grouping all the terms where the control process $h(t)$ appears, we can write the Hamilton-Jacobi-Bellman equation for the risk-sensitive problem as

$$\begin{aligned} \Phi_t\left(t,x\right) + \left(b + Bx\right)' \Phi_x\left(t,x\right) + \frac{1}{2}tr\left(\Lambda\Lambda'\Phi_{xx}\left(t,x\right)\right) + a_0 + A_0'x \\ -\frac{\theta}{2}\Phi_x'\left(t,x\right)\Lambda\Lambda'\Phi_x\left(t,x\right) + \sup_{h\in\mathcal{A}(T)}\Bigg[-\frac{1}{2}\left(\theta+1\right)h'\Sigma\Sigma'h \\ -\theta h'\Sigma\Lambda'\Phi_x\left(t,x\right) + h'\left(\hat{a} + \hat{A}x\right) \\ -\int_{\mathbf{Z}}\frac{1}{\theta}\left[\left(1 + h'\gamma(z)\right)^{-\theta} - 1\right]\nu(dz) - \int_{\mathbf{Z}_0} h'\gamma(z)\nu(dz)\Bigg] &= 0, \end{aligned}$$

for all $(t,x) \in (0,T)\times\mathbb{R}^n$. Please note that the dynamics of the state variable x satisfy the stochastic differential equation (5.9).

Corollary 5.16. [12] *Value function $\Phi(t,x)$ is convex in its second argument, x.*

5.3. **Analysis of the solution.** The Hamilton-Jacobi-Bellman equation for our risk-sensitive optimisation problem (including its corresponding terminal condition) can be expressed as

$$\begin{cases} \Phi_t\left(t,x\right) + G\left(x,\Phi_x(t,x),\Phi_{xx}(t,x)\right) = 0, & \forall\left(t,x\right) \in (0,T)\times\mathbb{R}^n, \\ \Phi\left(T,x\right) = \ln v, & \forall x \in \mathbb{R}^n, \end{cases} \tag{5.11}$$

where

$$G(x,p,D) = \sup_{h\in\mathcal{A}(T)}\left[\frac{1}{2}tr\left(\Lambda\Lambda' D\right) - \frac{\theta}{2}p'\Lambda\Lambda' p + \left(b + Bx - \theta\Lambda\Sigma' h\right)' p - g(x,h;\theta)\right]. \tag{5.12}$$

In [12] Davis and Lleo prove via a verification procedure that if problem (5.11) has a sufficiently smooth solution (i.e. it is, at least, C^1 in time and C^2 in the state space), then that solution is equal to the value function Φ defined by (5.10) and the control $h^*(t)$ is optimal. The verification theorem in [12] implicitly tells us that the value function is the unique solution to the problem (5.11) within the class of functions having the specified properties. However, the validity of the verification procedure depends completely on the existence of a solution to (5.11).

Definition 5.17. (Classical solutions) [11] A function $\Phi \in C^{1,2}\left([0,T]\times\mathbb{R}^n\right)$ is a classical subsolution (respectively, supersolution) of (5.11) if, for all $(t,x) \in [0,T]\times\mathbb{R}^n$, we have

$$\Phi_t(t,x) + G\left(x,\Phi_x(t,x),\Phi_{xx}(t,x)\right) \leq 0 \ \ (resp. \ \geq 0).$$

A function $\Phi \in C^{1,2}\left([0,T]\times\mathbb{R}^n\right)$ is a classical solution of (5.11) if it is both a classical subsolution and a classical supersolution of (5.11).

Theorem 5.18. [12] *The problem (5.11) has a solution* $\Phi(t,x) \in C^{1,2}\left([0,T]\times\mathbb{R}^n\right)$ *with* Φ *continuous in* $[0,T]\times\mathbb{R}^n$.

Theorem 5.18 guarantees that there exists a solution to problem (5.11) in the classical sense given by Definition 5.17, and the verification procedure ensures that the solution is equal to the value function.

Definition 5.19. Let $f:\mathbb{R}^n\to\mathbb{R}$ and let $x_0\in\mathbb{R}^n$.

(1) The function $f(x)$ is upper semicontinuous at x_0 if $\lim_{x\to x_0} f(x) \leq f(x_0)$.
(2) The function $f(x)$ is lower semicontinuous at x_0 if $\lim_{x\to x_0} f(x) \geq f(x_0)$.

If $f(x)$ is both upper and lower semicontinuous, then it is continuous. We also denote by

$$USC\left([0,T]\times\mathbb{R}^n\right) = \{f:[0,T]\times\mathbb{R}^n\to\mathbb{R} \mid f \text{ is upper semicontinuous on } [0,T]\times\mathbb{R}^n\},$$

$$LSC\left([0,T]\times\mathbb{R}^n\right) = \{f:[0,T]\times\mathbb{R}^n\to\mathbb{R} \mid f \text{ is lower semicontinuous on } [0,T]\times\mathbb{R}^n\},$$

the sets of upper and lower semicontinuous functions on $[0,T]\times\mathbb{R}^n$.

Definition 5.20. (Viscosity solutions) [11] A function $\Phi\in USC\left([0,T]\times\mathbb{R}^n\right)$ (respectively, $\Phi\in LSC\left([0,T]\times\mathbb{R}^n\right)$) is a viscosity subsolution (respectively, supersolution) of (5.11) if, for all $(t,x)\in[0,T]\times\mathbb{R}^n$ and $\phi\in C^{1,2}\left([0,T]\times\mathbb{R}^n\right)$ for which $\Phi-\phi$ has a global maximum (respectively, minimum) at (t,x), we have

$$\Phi_t(t,x) + G\left(x,\Phi_x(t,x),\Phi_{xx}(t,x)\right) \leq 0 \ \ (resp. \ \geq 0).$$

A function $\Phi\in C\left([0,T]\times\mathbb{R}^n\right)$ is a viscosity solution of (5.11) if it is both a viscosity subsolution and a viscosity supersolution of (5.11).

However, there are often situations where no solution in the classical sense can be shown to exist for a Hamilton-Jacobi-Bellman equation, which is the case for many stochastic control problems. In those situations, we have to address the problem by means of *viscosity solutions*, which have in recent years proved particularly useful for obtaining weak solutions to a Hamilton-Jacobi-Bellman equation. The viscosity solution approach overcomes the problem when the value function is not sufficiently smooth to satisfy the Hamilton-Jacobi-Bellman equation in a classical sense (e.g. there may be points where the value function is not differentiable). A viscosity solution does not need to be differentiable everywhere, but it still satisfies the equation in an appropriate sense.

The theory of viscosity solutions applies to parabolic partial differential equations of the form

$$v_t(t,x) + G(t,x,v,v_x,v_{xx}) = 0,$$

where the G must satisfy the conditions in the following Definition.

Definition 5.21. [10] A functional $G(t,x,r,p,D)$ is said to be *proper* if it satisfies the following two properties:

(1) The functional G is degenerate elliptic, i.e.

$$G(t,x,r,p,D) \leq G(t,x,r,p,E)$$

for all $D, E \in \mathcal{S}^n$ such that $E \leq D$.

(2) The functional G is monotone, i.e.

$$G(t,x,r,p,D) \leq G(t,x,s,p,D)$$

for all $r, s \in \mathbb{R}^n$ such that $r \leq s$.

With regards to our problem, we can express the Hamilton-Jacobi-Bellman partial differential equation in (5.11) as

$$-\Phi_t(t,x) + G(x, \Phi_x(t,x), \Phi_{xx}(t,x)) = 0,$$

where the functional defined as

$$\begin{aligned} G(x,p,D) &= -\sup_{h \in \mathcal{A}(T)} \Big[\frac{1}{2} tr(\Lambda\Lambda' D) - \frac{\theta}{2} p'\Lambda\Lambda' p \\ &\quad + (b + Bx - \theta\Lambda\Sigma' h)' p - g(x,h;\theta) \Big] \end{aligned}$$

is indeed proper because of Assumption 5.2. The following Theorem shows that the notions of viscosity and classical solution coincide when the value function is sufficiently smooth. Thus, the classical solution to (5.11), whose existence is proved in [12], is also a viscosity solution.

Theorem 5.22. [29] *Any classical solution of (5.11) is a viscosity solution of (5.11). If Φ is a viscosity solution of (5.11) and $\Phi \in C^{1,2}([0,T] \times \mathbb{R}^n)$, then Φ is a classical solution of (5.11).*

Since viscosity solutions represent a weaker type of solution, it is easier to find a solution in the viscosity sense. Thus, the issue of existence is usually not a problem and the difficulty shifts from proving existence of a solution to proving uniqueness of the solution. In order for the Hamilton-Jacobi-Bellman equation to be an effective characterisation of the value function we must guarantee that the solution is unique for a certain class of functions. This is achieved by obtaining a *comparison principle*.

Lemma 5.23. ***(Modulus of continuity)*** [32] *Let $G(x,p,D)$ be the functional defined as (5.12). Then there exists a function $\omega:\mathbb{R}^+\to\mathbb{R}^+$ with $\omega(0)=0$ such that*

$$G(y,\alpha(x-y),E)-G(x,\alpha(x-y),D)\leq\omega\left(\alpha|x-y|^2+|x-y|\right),$$

for all $x,y\in\mathbb{R}^n$, $\alpha>0$ and symmetric matrices D,E satisfying

$$-3\alpha\begin{pmatrix} I & 0\\ 0 & I\end{pmatrix}\leq\begin{pmatrix} D & 0\\ 0 & -E\end{pmatrix}\leq 3\alpha\begin{pmatrix} I & -I\\ -I & I\end{pmatrix}. \tag{5.13}$$

Proof. Following a similar argument as in [13], we take $m=G(y,\alpha(x-y),E)-G(x,\alpha(x-y),D)$, so that we have

$$\begin{aligned} m &= \sup_{h\in\mathcal{A}(T)}\Big[\alpha\left(b+By-\theta\Lambda\Sigma'h\right)'(x-y)+\frac{1}{2}tr\left(\Lambda\Lambda'E\right)\\ &\quad-\frac{\theta\alpha^2}{2}(x-y)'\Lambda\Lambda'(x-y)-g(y,h;\theta)\Big]\\ &\quad-\sup_{h\in\mathcal{A}(T)}\Big[\alpha\left(b+Bx-\theta\Lambda\Sigma'h\right)'(x-y)+\frac{1}{2}tr\left(\Lambda\Lambda'D\right)\\ &\quad-\frac{\theta\alpha^2}{2}(x-y)'\Lambda\Lambda'(x-y)-g(x,h;\theta)\Big]. \end{aligned}$$

Removing the terms that cancel each other out, we obtain

$$\begin{aligned} m &= \frac{1}{2}tr\left(\Lambda\Lambda'E-\Lambda\Lambda'D\right)+\alpha\left(B(y-x)\right)'(x-y)\\ &\quad+\sup_{h\in\mathcal{A}(T)}\left[g(x,h;\theta)-g(y,h;\theta)\right]\\ &\leq \frac{1}{2}|tr\left(\Lambda\Lambda'D-\Lambda\Lambda'E\right)|+\alpha|B||x-y|^2\\ &\quad+\sup_{h\in\mathcal{A}(T)}\left[|A_0'+h'\hat{A}||x-y|\right]. \end{aligned}$$

On the other hand, symmetric matrices D and E satisfy condition (5.13), and therefore

$$\begin{aligned} tr\left(\Lambda\Lambda'D-\Lambda\Lambda'E\right) &= tr\left(\begin{bmatrix}\Lambda\Lambda' & \Lambda\Lambda'\\ \Lambda\Lambda' & \Lambda\Lambda'\end{bmatrix}\begin{bmatrix} D & 0\\ 0 & -E\end{bmatrix}\right)\\ &\leq 3\alpha tr\left(\begin{bmatrix}\Lambda\Lambda' & \Lambda\Lambda'\\ \Lambda\Lambda' & \Lambda\Lambda'\end{bmatrix}\begin{bmatrix} I & -I\\ -I & I\end{bmatrix}\right)\\ &= 0. \end{aligned}$$

As a result, we have

$$\begin{aligned} m &\leq \alpha|B||x-y|^2+\sup_{h\in\mathcal{A}(T)}\left[|A_0'+h'\hat{A}||x-y|\right]\\ &\leq \alpha J|x-y|^2+K|x-y|, \end{aligned}$$

for some constants $J,K>0$. Finally, we can choose a function $\omega(r)=Cr$, with $C=J\vee K$, where $\vee$ denotes the maximum of two values, such that

$$m\leq\omega\left(\alpha|x-y|^2+|x-y|\right).$$

The function $\omega:\mathbb{R}^+\to\mathbb{R}^+$ is called a *modulus of continuity*. □

Remark 5.24. [27] Since we consider the entire space $\mathbb{R}^n$ for the state values, the boundary conditions for any viscosity solution are growth conditions at infinity in x (e.g. polynomial growth in x).

Theorem 5.25. ***(Comparison principle)*** [27, 32] *Let the functional $G(x,p,D)$ defined as (5.12) be uniformly continuous and proper; let $G(x,p,D)$ also satisfy Lemma 5.23. Let $u \in USC([0,T]\times\mathbb{R}^n)$ and $v \in LSC([0,T]\times\mathbb{R}^n)$ be bounded viscosity sub- and supersolutions of (5.11) with polynomial growth condition such that $u \leq v$ for $t = T$. Then $u \leq v$ for all $(t,x) \in [0,T)\times\mathbb{R}^n$.*

The comparison principle allows us to compare a viscosity sub- and supersolution on $[0,T)\times\mathbb{R}^n$ from the comparison on the boundary of the domain. In particular, it proves uniqueness of the viscosity solution of (5.11) from a terminal condition at $t = T$.

6. Numerical approximation

In general, it is not possible to find the analytical solution of the Hamilton-Jacobi-Bellman equation that arises in most stochastic control problems. We thus have to find the value function and the optimal control policy numerically. Since classical solutions to the Hamilton-Jacobi-Bellman equations may not exist, we must instead try to find the viscosity solution. It is important then to ensure that we generate a discrete solution that converges to the viscosity solution as the mesh size and timestep are reduced. We will carry a numerical analysis of the problem (5.11) that consists of two steps:

(1) Discretisation of the Hamilton-Jacobi-Bellman equation using a *finite difference method*.
(2) Solution of the discrete equation by means of an iterative method and proof of converge of the algorithm.

Among the finite difference methods, we will apply the explicit method to numerically solve the problem (5.11) because it is relatively easy to implement and reasonably effective, although it is less accurate than the implicit or Crank-Nicolson method and it can also be unstable. As part of our numerical experiments we implemented two numerical schemes: one that approximates the solution of the Hamilton-Jacobi-Bellman equation when there is only one auxiliary factor process and another scheme that approximates it when two factor processes are allowed to exist. Results obtained with both are presented in Subsections 6.7 and 6.6, although only details of the latter numerical scheme will be explained in the following subsections. We will, however, make use of the one-factor numerical scheme in a number of proofs that will justify the convergence of the numerical approximation to the viscosity solution.

6.1. Discretisation of the Hamilton-Jacobi-Bellman equation.

The first step for obtaining a numerical solution is to discretise the given domain into a network of points called *grid*. At each grid point the Hamilton-Jacobi-Bellman equation is approximated by replacing the partial derivatives by their corresponding difference approximations. This results in algebraic equations for each grid point, in which the value of Φ at that point is unknown, but will be calculated from the values of Φ known at a certain set of neighbouring points.

We would like to find $\Phi(t,x)$ on the whole domain $[0,T]\times\mathbb{R}^n$, but since numerical calculations can only be performed on finite domains, we must reduce the problem

(5.11) to a bounded domain in order to obtain a system of equations of finite dimension. We will then restrict ourselves to the domain

$$Q_T = [0,T] \times Q \subset [0,T] \times \mathbb{R}^n, \tag{6.1}$$

where $Q = \left[x_1^{min}, x_1^{max}\right] \times \left[x_2^{min}, x_2^{max}\right]$ for $x_1^{min}, x_1^{max}, x_2^{min}, x_2^{max} \in \mathbb{R}$ such that $x_1^{min} < x_1^{max}$ and $x_2^{min} < x_2^{max}$. By choosing a domain $Q \in \mathbb{R}^2$ where we will approximate the solution to (5.11), we implicitly consider the situation where we have two auxiliary factors and thus $n = 2$ in equation (5.9).

For some constants $N_t, N_{x_1}, N_{x_2} \in \mathbb{N}$ we define

$$\triangle t = \frac{T}{N_t}, \triangle x_1 = \frac{x_1^{max} - x_1^{min}}{N_{x_1}}, \triangle x_2 = \frac{x_2^{max} - x_2^{min}}{N_{x_2}}$$

to be the timestep size and the mesh spacing for the state space. We will then construct a discretisation of the domain Q_T defined by the grid points

$$I_t = \{m \in \mathbb{N} : 0 \leq m \leq N_t\},$$

$$I_x = \left\{(i,j) \in \mathbb{N}^2 : 0 \leq i \leq N_{x_1}, 0 \leq j \leq N_{x_2}\right\},$$

$$Q^d = \{(x_1,x_2) \in Q : x_1 = i\triangle x_1, x_2 = j\triangle x_2, \forall (i,j) \in I_x\},$$

$$Q_T^d = \{(t,x_1,x_2) \in Q_T : t = m\triangle t, x_1 = i\triangle x_1, x_2 = j\triangle x_2, \forall m \in I_t, \forall (i,j) \in I_x\}.$$

For convenience, we will use indeces (m,i,j) as a shorthand notation when refering to any point $(m\triangle t, i\triangle x_1, j\triangle x_2)$ of Q_T^d. The discretisation produces a three-dimensional grid of points, which can be seen as a set of rectangular grids over the state domain indexed by time. We recall that the state domain represents the range of possible values that the factors satisfying equation (5.9) may take.

Since the key idea of the finite difference methods is to replace the partial derivatives with finite differences, we need to establish some expressions to approximate the partial derivatives on the grid points. First we will approximate the spatial partial derivatives of $\Phi(t,x) = \Phi(t,x_1,x_2)$, where $x \in \mathbb{R}^2$ has components $(x_i, 1 \leq i \leq 2)$ and t is fixed. Applying Taylor expansions, we find

$$\Phi_{x_1}(t,x_1,x_2) = \frac{\Phi(t,x_1+\triangle x_1,x_2) - \Phi(t,x_1-\triangle x_1,x_2)}{2\triangle x_1} + O\left((\triangle x_1)^2\right),$$

$$\Phi_{x_2}(t,x_1,x_2) = \frac{\Phi(t,x_1,x_2+\triangle x_2) - \Phi(t,x_1,x_2-\triangle x_2)}{2\triangle x_2} + O\left((\triangle x_2)^2\right),$$

$$\begin{aligned}\Phi_{x_1x_1}(t,x_1,x_2) &= \frac{\Phi(t,x_1+\triangle x_1,x_2) - 2\Phi(t,x_1,x_2) + \Phi(t,x_1-\triangle x_1,x_2)}{(\triangle x_1)^2} \\ &\quad + O\left((\triangle x_1)^2\right),\end{aligned}$$

$$\begin{aligned}\Phi_{x_2x_2}(t,x_1,x_2) &= \frac{\Phi(t,x_1,x_2+\triangle x_2) - 2\Phi(t,x_1,x_2) + \Phi(t,x_1,x_2-\triangle x_2)}{(\triangle x_2)^2} \\ &\quad + O\left((\triangle x_2)^2\right),\end{aligned}$$

$$\begin{aligned}\Phi_{x_1x_2}(t,x_1,x_2) &= \frac{\Phi(t,x_1+\triangle x_1,x_2+\triangle x_2)-\Phi(t,x_1+\triangle x_1,x_2-\triangle x_2)}{4\triangle x_1\triangle x_2}\\ &\quad -\frac{\Phi(t,x_1-\triangle x_1,x_2+\triangle x_2)-\Phi(t,x_1-\triangle x_1,x_2-\triangle x_2)}{4\triangle x_1\triangle x_2}\\ &\quad +O(\triangle x_1\triangle x_2)=\Phi_{x_2x_1}(t,x_1,x_2),\end{aligned}$$

which are known as central difference approximations. Similarly, we can find an approximation of $\Phi_t(t,x)$, known as the backward difference approximation

$$\Phi_t(t,x_1,x_2)=\frac{\Phi(t,x_1,x_2)-\Phi(t-\triangle t,x_1,x_2)}{\triangle t}+O(\triangle t).$$

Using the notation $\Phi^m_{i,j}=\Phi(m\triangle t,i\triangle x,j\triangle y)$ to denote the value of $\Phi(t,x_1,x_2)$ at the node (m,i,j), we define now the discrete numerical approximation $\hat{\Phi}:Q^d_T\to\mathbb{R}$ that results from our numerical scheme as

$$\hat{\Phi}^m_{i,j}\simeq\Phi^m_{i,j},$$

for all $m\in I_t,(i,j)\in I_x$. And we will use the following finite differences

$$\frac{\partial\hat{\Phi}^m_{i,j}}{\partial x_1}\simeq\frac{\hat{\Phi}^m_{i+1,j}-\hat{\Phi}^m_{i-1,j}}{2\triangle x_1},$$

$$\frac{\partial\hat{\Phi}^m_{i,j}}{\partial x_2}\simeq\frac{\hat{\Phi}^m_{i,j+1}-\hat{\Phi}^m_{i,j-1}}{2\triangle x_2},$$

$$\frac{\partial\hat{\Phi}^m_{i,j}}{\partial x_1^2}\simeq\frac{\hat{\Phi}^m_{i+1,j}-2\hat{\Phi}^m_{i,j}+\hat{\Phi}^m_{i-1,j}}{(\triangle x_1)^2},$$

$$\frac{\partial\hat{\Phi}^m_{i,j}}{\partial x_2^2}\simeq\frac{\hat{\Phi}^m_{i,j+1}-2\hat{\Phi}^m_{i,j}+\hat{\Phi}^m_{i,j-1}}{(\triangle x_2)^2},$$

$$\frac{\partial\hat{\Phi}^m_{i,j}}{\partial x_1x_2}=\frac{\partial\hat{\Phi}^m_{i,j}}{\partial x_2x_1}\simeq\frac{\hat{\Phi}^m_{i+1,j+1}-\hat{\Phi}^m_{i+1,j-1}-\hat{\Phi}^m_{i-1,j+1}+\hat{\Phi}^m_{i-1,j-1}}{4\triangle x_1\triangle x_2},$$

$$\frac{\partial\hat{\Phi}^m_{i,j}}{\partial t}\simeq\frac{\hat{\Phi}^m_{i,j}-\hat{\Phi}^{m-1}_{i,j}}{\triangle t}$$

as approximations to its partial derivatives.

6.2. **Discretisation of the integral operator.** The numerical scheme approximating the solution to (5.11) must also incorporate a discretisation of the integral with respect to the Lévy measure included in $g(x,h;\theta)$. This translates in a reduction of the region of integration to a bounded interval, which, regarding the jump term of (5.4), amounts to the truncation of large jumps. In our numerical scheme we are going to consider a finite Lévy measure. In particular, we will assume that the jumps of the Lévy-Itô process satisfying (5.4) are caused by a compound Poisson process. In this case the Lévy measure does not have a singularity around zero. Similarly as done for the discretisation of the state space, we will first then restrict ourselves to the two-dimensional domain $\mathbf{Z}_t\subset\mathbf{Z}=\mathbb{R}^d$, where $\mathbf{Z}_t=\left[z_1^{min},z_1^{max}\right]\times\left[z_2^{min},z_2^{max}\right]$ for $z_1^{min},z_1^{max},z_2^{min},z_2^{max}\in\mathbb{R}$ such that

$z_1^{min} < z_1^{max}$ and $z_2^{min} < z_2^{max}$. The domain $\mathbf{Z}_t$ represents the space of possible jump sizes. For some constants $N_{z_1}, N_{z_2} \in \mathbb{N}$ we define as well

$$\triangle z_1 = \frac{z_1^{max} - z_1^{min}}{N_{z_1}}, \triangle z_2 = \frac{z_2^{max} - z_2^{min}}{N_{z_2}}$$

to be the mesh spacing in the domain $\mathbf{Z}_t$. We will then construct a discretisation of the domain $\mathbf{Z}_t$ defined by the grid points

$$I_z = \left\{(l,n) \in \mathbb{N}^2 : 0 \leq m \leq N_{z_1}, 0 \leq i \leq N_{z_2}\right\},$$

$$\mathbf{Z}_t^d = \left\{(z_1, z_2) \in \mathbf{Z}_t : z_1 = l\triangle z_1, z_2 = n\triangle z_2, \forall (l,n) \in I_z\right\},$$

where, again, we will use indeces (l,n) as abbreviation for any point $(l\triangle z_1, n\triangle z_2)$ of $\mathbf{Z}_t^d$. For our simulations we choose a Lévy measure that has gaussian density

$$\nu(z) = \frac{\lambda}{\sqrt{(2\pi)^2 \det(\Sigma)}} \exp\left(-\frac{1}{2}(z-\mu)'\Sigma^{-1}(z-\mu)\right), \tag{6.2}$$

where z is a two-dimensional column vector with components $(z_i, 1 \leq i \leq 2)$, $\mu \in \mathbb{R}^2$ is the mean vector, $\Sigma \in \mathbb{R}^{2\times 2}$ is the covariance matrix, and $\lambda \in \mathbb{R}$ is the intensity, which is the arrival rate of jumps of the underlying Poisson process. The values of μ and Σ determine the average jump size and the variation of jump sizes. The Lévy measure satisfies $\nu(\mathbf{Z}_t) < \infty$ and thus the geometric Lévy-Itô process solving equation (5.4) has finite activity and a finite number of jumps on any time interval of finite length.

We recall that the integral operator in $g(x,h;\theta)$ has the form

$$\mathcal{I}(h;\theta) = \int_{\mathbf{Z}} \left\{\left[(1+h'\gamma(z))^{-\theta} - 1\right] + \theta h'\gamma(z)\mathbb{I}_{\mathbf{Z}_0}(z)\right\}\nu(dz),$$

and with our choice of Lévy measure we can write it as

$$\mathcal{I}(h;\theta) = \int_{\mathbf{Z}} f(z,h;\theta)\,dz,$$

where

$$f(z,h;\theta) = \left\{\left[(1+h'\gamma(z))^{-\theta} - 1\right] + \theta h'\gamma(z)\mathbb{I}_{\mathbf{Z}_0}(z)\right\}\nu(z).$$

We can then write the integral in the truncated domain $\mathbf{Z}_t$ as

$$\mathcal{I}_t(h;\theta) = \int_{\mathbf{Z}_t} f(z,h;\theta)\,dz = \int_{z_2}\int_{z_1} f(z_1,z_2,h;\theta)\,dz_1 dz_2,$$

where we have just simply expanded the two-dimensional vector z into its components in order to express the integral over $\mathbf{Z}_t$ as a double integral. We introduce the notation $f_{l,n}(h;\theta) = f(l\triangle z_1, n\triangle z_2, h;\theta)$ to denote the value of the function at the point $(l\triangle z_1, n\triangle z_2)$, so that the next step is to approximate the integral numerically as a finite Riemann sum on $\mathbf{Z}_t^d$. At every point (l,n) of $\mathbf{Z}_t^d$ for $0 \leq l < N_{z_1}, 0 \leq n < N_{z_2}$ we evaluate $f_{l,n}(h;\theta)$ at points (l,n), $(l+1,n)$, $(l,n+1)$ and $(l+1,n+1)$, and we take the average. The integral is then approximated by the sum

$$\begin{aligned}\mathcal{I}_t(h;\theta) \quad \simeq \quad \mathcal{I}_t^d(h;\theta) = \sum_{l,n=0,0}^{N_{z_1}-1,N_{z_2}-1} \frac{1}{4}\Big(& f_{l,n}(h;\theta) + f_{l+1,n}(h;\theta) \\ & + f_{l,n+1}(h;\theta) + f_{l+1,n+1}(h;\theta)\Big)\triangle z_1 \triangle z_2,\end{aligned}$$

which converges to $\mathcal{I}_t(h;\theta)$ as $\triangle z_1, \triangle z_2 \to 0$.

6.3. **Terminal and boundary conditions.** Our risk-sensitive optimisation problem (5.11) has a concise terminal condition because it is defined over a finite time interval. However, it is also defined over an infinite state space, and thus needs to be truncated if we aim to simulate it on a computer. This is one of the disadvantages of finite difference methods: they can only handle bounded state domains; the state space must be truncated if it is naturally infinite, as it is in our case.

The truncation defined in (6.1) introduced artificial boundaries, and therefore we must specify boundary conditions at $i = 0, i = N_{x_1}$ and $j = 0, j = N_{x_2}$. One possibility is to impose *Dirilecht conditions*, which means that we specify the value of $\hat{\Phi}^m_{i,j}$ at the extremes of the bounded state domain. However, due to the difficulty of choosing sensible values, we are inclined to enforce *Neumann conditions*, where we influence the value of the first and second order partial derivatives of $\hat{\Phi}^m_{i,j}$ with respect to x_1 and x_2. In order to be able to calculate the finite differences that approximate the partial derivatives at the extremes of the domain, we will artificially extend the rectangular grid and extrapolate the value function on those points. We obtain by extrapolation, for all $0 \leq m \leq N_t$, the values $\hat{\Phi}^m_{-1,j}$ for $-1 \leq j \leq N_{x_2}+1$, $\hat{\Phi}^m_{i,-1}$ for $-1 \leq i \leq N_{x_1}+1$, $\hat{\Phi}^m_{N_{x_1}+1,j}$ for $-1 \leq j \leq N_{x_2}+1$ and $\hat{\Phi}^m_{i,N_{x_2}+1}$ for $-1 \leq i \leq N_{x_1}+1$.

6.4. **The complete numerical scheme.** The numerical scheme that will be implemented in MATLAB can be written as

$$\begin{aligned} \frac{\hat{\Phi}^m_{i,j} - \Phi^{m-1}_{i,j}}{\triangle t} + (b + Bx)' D\hat{\Phi}^m_{i,j} + \frac{1}{2} tr\left(\Lambda\Lambda' D^2 \hat{\Phi}^m_{i,j}\right) \\ + a_0 + A_0' x - \frac{\theta}{2}\left(D\hat{\Phi}^m_{i,j}\right)' \Lambda\Lambda' D\hat{\Phi}^m_{i,j} + \sup_{\hat{h}\in\mathcal{A}(T)}\left[-\frac{1}{2}(\theta+1)\,\hat{h}'\Sigma\Sigma'\hat{h}\right. \\ \left. -\theta\hat{h}'\Sigma\Lambda' D\hat{\Phi}^m_{i,j} + \hat{h}'\left(\hat{a} + \hat{A}x\right) - \frac{1}{\theta}\mathcal{I}^d_t\left(\hat{h};\theta\right)\right] = 0, \end{aligned} \tag{6.3}$$

for all $1 \leq m \leq N_t, (i,j) \in I_x$ and with terminal condition

$$\hat{\Phi}^{N_t}_{i,j} = \ln v,$$

for $m = N_t$ and all $(i,j) \in I_x$, where $\theta \in \mathbb{R}$, $b \in \mathbb{R}^2$, $B \in \mathbb{R}^{2\times 2}$, $\Lambda \in \mathbb{R}^{2\times 2}$, $\Sigma \in \mathbb{R}^{2\times 2}$, $a_0 \in \mathbb{R}$, $A_0 \in \mathbb{R}^2$, $\hat{a} \in \mathbb{R}^2$, $\hat{A} \in \mathbb{R}^2$ and $v > 0$. The numerical scheme assumes that there exist two risky securities available for investment. We use the notation $h^m = h(m\triangle t)$ to denote the value of the control policy $h(t)$ at time m with initial condition $(i,j) \in I_x$. The discrete numerical approximation obtained from the simulation of the numerical scheme can be defined as

$$\hat{h}^m \simeq h^m,$$

and it is also subject to the bounds found in the proof of Lemma 5.10. Similarly, we denote by $\gamma_{l,n}$ the value of the function $\gamma(z)$ present in $\mathcal{I}^d_t\left(\hat{h};\theta\right)$ at the point $(l\triangle z_1, n\triangle z_2)$ for $(l,n) \in I_z$. Note that both the control policy and the value of $\gamma_{l,n}$ for any $(l,n) \in I_z$ are two-dimensional vectors in our numerical scheme.

The column vector $x \in Q^d$ represents the value of the state variables x_1, x_2 at the coordinates (i,j). Finally, the finite differences approximating the partial

derivatives at time m with respect to x_1 and x_2 are

$$D\hat{\Phi}^m_{i,j} = \begin{bmatrix} \frac{\hat{\Phi}^m_{i+1,j} - \hat{\Phi}^m_{i-1,j}}{2\triangle x_1} \\ \frac{\hat{\Phi}^m_{i,j+1} - \hat{\Phi}^m_{i,j-1}}{2\triangle x_2} \end{bmatrix},$$

$$D^2\hat{\Phi}^m_{i,j} = \begin{bmatrix} \frac{\hat{\Phi}^m_{i+1,j} - 2\hat{\Phi}^m_{i,j} + \hat{\Phi}^m_{i-1,j}}{(\triangle x_1)^2} & \frac{\hat{\Phi}^m_{i+1,j+1} - \hat{\Phi}^m_{i+1,j-1} - \hat{\Phi}^m_{i-1,j+1} + \hat{\Phi}^m_{i-1,j-1}}{4\triangle x_1 \triangle x_2} \\ \frac{\hat{\Phi}^m_{i+1,j+1} - \hat{\Phi}^m_{i+1,j-1} - \hat{\Phi}^m_{i-1,j+1} + \hat{\Phi}^m_{i-1,j-1}}{4\triangle x_1 \triangle x_2} & \frac{\hat{\Phi}^m_{i,j+1} - 2\hat{\Phi}^m_{i,j} + \hat{\Phi}^m_{i,j-1}}{(\triangle x_2)^2} \end{bmatrix}.$$

The numerical scheme is purely explicit, since the value of $\hat{\Phi}^m_{i,j}$ can be determined for all $0 \leq m < N_t, (i,j) \in I_x$ from the values of $\hat{\Phi}$ at time $m+1$. The problem is thus solved backwards in time. The simulation of the numerical scheme runs the following algorithm:

(1) Set $\hat{\Phi}^{N_t}_{i,j}$ for all $(i,j) \in I_x$ using the given terminal condition. Set $m = N_t - 1$.
(2) Calculate the optimal control policy $\hat{h}^{*,m}$ for any given initial condition $(i,j) \in I_x$.
(3) Calculate $\hat{\Phi}^m_{i,j}$ for $(i,j) \in I_x$ using the values of $\hat{\Phi}$ at time $m+1$. Reduce m by 1, go to point 2 and use the new computed values of $\hat{\Phi}$ at time m as input while $m > 0$.

The values of $\mathcal{I}^d_t\left(\hat{h};\theta\right)$ for the admissible discrete control policies are pre-computed before the start of the iterative algorithm, since the admissible discrete controls $\hat{h}$ and $\gamma(z)$ are independent of time and the state variable.

6.5. **Convergence of the numerical scheme.** The devised numerical scheme will only be useful if we can ensure that its solution converges to the solution of (5.11) as the discretisation parameters go to 0. Since the solution to many stochastic control problems may not be sufficiently smooth, we have to guarantee that the solution of the numerical scheme converges to the relevent solution, which is the viscosity solution. As we showed in Theorem 5.25, the partial differential equation in (5.11) verifies a comparison principle. Thus, by the convergence argument developed in [2, 7], any finite difference scheme that is *consistent*, *stable* (in the maximum norm) and *monotone* converges uniformly on each compact subset of $[0,T] \times \mathbb{R}^n$ to the unique viscosity solution, even when solutions are not smooth [11]. We will first define what we mean by consistency, stability and monotonicity.

Our numerical scheme is said to be consistent if the finite difference operator defined in (6.3) converges to the partial differential equation in (5.11) as the timestep size and mesh spacing tend to 0. The difference between the discretised and the exact Hamilton-Jacobi-Bellman equation is called the *truncation error*, and it is caused by replacing the partial derivatives in (5.11) with finite difference approximations. The truncation error is usually estimated by substituting in (6.3) the numerical approximation $\hat{\Phi}^m_{i,j}$ at all grid points with a Taylor expansion of the exact solution $\Phi^m_{i,j}$.

Lemma 6.1. *The numerical scheme (6.3) is consistent with the partial differential equation in (5.11) if* $\triangle t, \triangle x_1, \triangle x_2 \to 0$.

Proof. For the sake of simplicity, we will prove this result by dealing with the one-factor finite difference operator

$$\begin{aligned}(6.4)\qquad &\frac{\hat{\Phi}_i^m-\hat{\Phi}_i^{m-1}}{\triangle t}+(b+Bx_1)\frac{\hat{\Phi}_{i+1}^m-\hat{\Phi}_{i-1}^m}{2\triangle x_1}\\ &+\frac{\Lambda^2}{2}\left(\frac{\hat{\Phi}_{i+1}^m-2\hat{\Phi}_i^m+\hat{\Phi}_{i-1}^m}{(\triangle x_1)^2}\right)-\frac{\theta}{2}\left(\Lambda\frac{\hat{\Phi}_{i+1}^m-\hat{\Phi}_{i-1}^m}{2\triangle x_1}\right)^2\\ &+\sup_{\hat{h}\in\mathcal{A}(T)}\left[-\theta\hat{h}\Sigma\Lambda\frac{\hat{\Phi}_{i+1}^m-\hat{\Phi}_{i-1}^m}{2\triangle x_1}-g\left(x_1,\hat{h};\theta\right)\right] \;=\; 0,\end{aligned}$$

for all $1\leq m\leq N_t, 0\leq i\leq N_{x_1}$, where $\theta,b,B,\Lambda,\Sigma,a_0,A_0,\hat{a},\hat{A}\in\mathbb{R}$. Assuming that the high order derivatives of the value function exists, we replace the values of the numerical approximations $\hat{\Phi}_i^m$ by the Taylor expansions of Φ_i^m for points $i+1$ and $i-1$ at time m

$$\Phi_{i+1}^m=\Phi_i^m+\triangle x_1\frac{\partial\Phi_i^m}{\partial x_1}+\frac{(\triangle x_1)^2}{2!}\frac{\partial\Phi_i^m}{\partial x_1^2}+\frac{(\triangle x_1)^3}{3!}\frac{\partial\Phi_i^m}{\partial x_1^3}+\frac{(\triangle x_1)^4}{4!}\frac{\partial\Phi_i^m}{\partial x_1^4}+\ldots,$$

$$\Phi_{i-1}^m=\Phi_i^m-\triangle x_1\frac{\partial\Phi_i^m}{\partial x_1}+\frac{(\triangle x_1)^2}{2!}\frac{\partial\Phi_i^m}{\partial x_1^2}-\frac{(\triangle x_1)^3}{3!}\frac{\partial\Phi_i^m}{\partial x_1^3}+\frac{(\triangle x_1)^4}{4!}\frac{\partial\Phi_i^m}{\partial x_1^4}+\ldots$$

and at time m for point i

$$\Phi_i^m=\Phi_i^{m-1}+\triangle t\frac{\partial\Phi_i^m}{\partial t}+\frac{(\triangle t)^2}{2!}\frac{\partial\Phi_i^m}{\partial t^2}+\frac{(\triangle t)^3}{3!}\frac{\partial\Phi_i^m}{\partial t^3}+\frac{(\triangle t)^4}{4!}\frac{\partial\Phi_i^m}{\partial t^4}+\ldots$$

in equation (6.4), so that we obtain the truncation error

$$\begin{aligned}T_i^m \;=\;& \frac{\partial\Phi_{i,j}^m}{\partial t}+O\left(\triangle t\right)+(b+Bx_1)\left(\frac{\partial\Phi_{i,j}^m}{\partial x_1^2}+O\left((\triangle x_1)^2\right)\right)\\ &+\frac{\Lambda^2}{2}\left(\frac{\partial\Phi_{i,j}^m}{\partial x_1^2}+O\left((\triangle x_1)^2\right)\right)-\frac{\theta}{2}\Lambda^2\left(\frac{\partial\Phi_{i,j}^m}{\partial x_1}+O\left((\triangle x_1)^2\right)\right)^2\\ &+\sup_{\hat{h}\in\mathcal{A}(T)}\left[-\theta\hat{h}\Sigma\Lambda\left(\frac{\partial\Phi_{i,j}^m}{\partial x_1^2}+O\left((\triangle x_1)^2\right)\right)-g\left(x_1,\hat{h};\theta\right)\right],\end{aligned}$$

for all $0<m\leq N_t, 0\leq i\leq N_{x_1}$. By definition of the risk-sensitive optimisation problem we have

$$\begin{aligned}&\frac{\partial\Phi_{i,j}^m}{\partial t}+(b+Bx_1)\left(\frac{\partial\Phi_{i,j}^m}{\partial x_1}\right)+\frac{\Lambda^2}{2}\left(\frac{\partial\Phi_{i,j}^m}{\partial x_1^2}\right)\\ -\frac{\theta}{2}&\left(\Lambda\frac{\partial\Phi_{i,j}^m}{\partial x_1}\right)^2+\sup_{\hat{h}\in\mathcal{A}(T)}\left[-\theta\hat{h}\Sigma\Lambda\left(\frac{\partial\Phi_{i,j}^m}{\partial x_1}\right)-g\left(x_1,\hat{h};\theta\right)\right] \;=\; 0,\end{aligned}$$

and thus the truncation error becomes

$$\begin{aligned}T_i^m \;=\;& O\left(\triangle t\right)+\left(b+Bx_1-\theta\hat{h}^{*,m}\Sigma\Lambda\right)\left(O\left((\triangle x_1)^2\right)\right)\\ &+\frac{\Lambda^2}{2}O\left((\triangle x_1)^2\right)-\frac{\theta}{2}\Lambda^2\left(O\left((\triangle x_1)^2\right)\right)^2,\end{aligned}$$

where $\hat{h}^{*,m}$ is the optimal discrete control strategy with initial condition i. Thus, the numerical scheme is consistent as $\triangle t,\triangle x_1\to 0$. In fact, it is consistent with order 1 in t and order 2 in x_1. This proof can be extended to the two-factor

numerical scheme (6.3) to show that it is as well consistent when $\triangle x_2 \to 0$. Please note that we do not check consistency of the numerical scheme at the boundaries of the state domain. The boundaries were introduced only because we needed to truncate the space $\mathbb{R}^n$, thus the numerical scheme will be consistent indenpendently of the boundaries and their conditions when the limits of the truncated domain go to ∞ and $\triangle t, \triangle x_1, \triangle x_2 \to 0$. □

Stability problems are due to the rounding errors introduced into the numerical solution when using finite precision computer arithmetic to solve the difference equations. We say that our numerical scheme is stable if it does not amplify the errors that appear in the course of the numerical algorithm. Stability ensures that the numerical scheme produces a bounded solution whenever the solution of the exact Hamilton-Jacobi-Bellman equation is bounded.

Lemma 6.2. *The numerical scheme (6.3) is stable if there exists a constant $M \in \mathbb{R}$ such that $\| \hat{\Phi}^m \|_\infty \leq M$ for all $m \in I_t$, where $\| \cdot \|_\infty$ is the maximum norm defined as $\| \hat{\Phi}^m \|_\infty = \max_{(i,j)\in I_x} |\hat{\Phi}^m_{i,j}|$.*

Proof. As in the proof of Lemma 6.1, we will also prove this result in the context of the one-factor finite difference operator (6.4). For $m = N_t$ we have

$$\begin{aligned} &\frac{\hat{\Phi}_i^{N_t} - \hat{\Phi}_i^{N_t-1}}{\triangle t} + (b + Bx_1)\frac{\hat{\Phi}_{i+1}^{N_t} - \hat{\Phi}_{i-1}^{N_t}}{2\triangle x_1} \\ &+\frac{\Lambda^2}{2}\left(\frac{\hat{\Phi}_{i+1}^{N_t} - 2\hat{\Phi}_i^{N_t} + \hat{\Phi}_{i-1}^{N_t}}{(\triangle x_1)^2}\right) - \frac{\theta}{2}\left(\Lambda\frac{\hat{\Phi}_{i+1}^{N_t} - \hat{\Phi}_{i-1}^{N_t}}{2\triangle x_1}\right)^2 \\ &+ \sup_{\hat{h}\in\mathcal{A}(T)}\left[-\theta\hat{h}\Sigma\Lambda\frac{\hat{\Phi}_{i+1}^{N_t} - \hat{\Phi}_{i-1}^{N_t}}{2\triangle x_1} - g\left(x_1, \hat{h}; \theta\right)\right] = 0. \end{aligned}$$

Since $\hat{\Phi}_i^{N_t} = \ln v$ for $v > 0$ and all $0 \leq i \leq N_{x_1}$, then we have

$$\frac{\ln v - \hat{\Phi}_i^{N_t-1}}{\triangle t} + \sup_{\hat{h}\in\mathcal{A}(T)}\left[-g\left(x_1, \hat{h}; \theta\right)\right] = 0,$$

which becomes

$$\hat{\Phi}_i^{N_t-1} = \triangle t\left[\ln v - g\left(x_1, \hat{h}^{*,N_t-1}; \theta\right)\right]$$

if we take the optimal control $\hat{h}^{*,N_t-1}$. The maximum norm of $\hat{\Phi}_i^{N_t-1}$ is defined as

$$\| \hat{\Phi}^{N_t-1} \|_\infty = \max_{0\leq i\leq N_{x_1}} |\hat{\Phi}_i^{N_t-1}|,$$

and if we denote by C the constant

$$C = \max_{0\leq i\leq N_{x_1}}\left[-g\left(x_1, \hat{h}^{*,N_t-1}; \theta\right)\right] < \infty,$$

then we can write

$$\| \hat{\Phi}^{N_t-1} \|_\infty \leq \triangle t\,(\ln v + C) < \infty,$$

and thus we have shown that $\| \hat{\Phi}^{N_t-1} \|_\infty$ is finite. If we now take any $m < N_t - 1$ and we assume $\| \hat{\Phi}^m \|_\infty < \infty$, we wish to show that $\| \hat{\Phi}^{m-1} \|_\infty < \infty$. For $m-1$ we

have basically equation (6.4). Selecting the optimal control $\hat{h}^{*,m-1}$ and maximum values for the finite differences, we obtain the inequality

$$\begin{aligned}\hat{\Phi}_i^{m-1} \quad \leq \quad & \triangle t\Bigg\{\parallel \hat{\Phi}^m \parallel_\infty + (b+Bx_1)\frac{\parallel \hat{\Phi}^m \parallel_\infty}{\triangle x_1} + 2\Lambda^2\frac{\parallel \hat{\Phi}^m \parallel_\infty}{(\triangle x_1)^2} \\ & -\frac{\theta}{2}\left(\Lambda\frac{\parallel \hat{\Phi}^m \parallel_\infty}{\triangle x_1}\right)^2 - \theta\hat{h}^{*,m-1}\Sigma\Lambda\frac{\parallel \hat{\Phi}^m \parallel_\infty}{\triangle x_1} - g\left(x_1,\hat{h}^{*,m-1};\theta\right)\Bigg\}.\end{aligned}$$

In order to find an upper bound for the maximum norm of $\hat{\Phi}_i^{m-1}$ we need to find the value of i that maximises the terms depending on the state variable, x_1. Thus, we have

$$\begin{aligned}\parallel \hat{\Phi}^{m-1} \parallel_\infty \quad \leq \quad & \triangle t\Bigg\{\parallel \hat{\Phi}^m \parallel_\infty \left(1+\frac{b}{\triangle x_1}+\frac{2\Lambda^2}{(\triangle x_1)^2}\right. \\ & \left.-\frac{\theta\parallel \hat{\Phi}^m \parallel_\infty}{2}\left(\frac{\Lambda}{\triangle x_1}\right)^2 - \frac{\theta\hat{h}^{*,m-1}\Sigma\Lambda}{\triangle x_1}\right) \\ & + \max_{0\leq i\leq N_{x_1}}\left[Bx_1\frac{\parallel \hat{\Phi}^m \parallel_\infty}{\triangle x_1} - g\left(x_1,\hat{h}^{*,m-1};\theta\right)\right]\Bigg\} < \infty.\end{aligned}$$

Therefore, it is possible to find a constant $M \in \mathbb{R}$ such that $\parallel \hat{\Phi}^m \parallel_\infty \leq M$ for all $m \in I_t$. We can reach to a similar conclusion for the two-factor finite difference operator (6.3). □

Since stability can be difficult to investigate, it is usually checked using numerical experiments: i.e. repeating simulations on series of successively refined grids. If the numerical scheme is stable and if all approximations used in the discretisation process are consistent, it is usually found that the solution converges to a grid-independent solution.

Lemma 6.3. *The numerical scheme (6.3) is monotone if the discrete functional defined as*

$$\begin{aligned}G(x,p,D) \quad = \quad & -\sup_{\hat{h}\in\mathcal{A}(T)}\left[\frac{1}{2}tr\left(\Lambda\Lambda' D\right) - \frac{\theta}{2}p'\Lambda\Lambda' p\right. \\ & \left.+\left(b+Bx-\theta\Lambda\Sigma'\hat{h}\right)' p - g\left(x,\hat{h};\theta\right)\right]\end{aligned}$$

satisfies that $G(x,p,D) \leq G(x,p,E)$ for all symmetric matrices D,E satisfying (5.13) for $\alpha > 0$ such that $E \leq D$.

Proof. We will again resort to the one-factor finite difference operator defined in (6.4) to prove this result first. For the discrete functional defined as

$$G(x_1,p,D) = -\sup_{\hat{h}\in\mathcal{A}(T)}\left[\frac{1}{2}\left(\Lambda^2 D\right) - \frac{\theta}{2}\left(\Lambda p\right)^2 + \left(b+Bx_1-\theta\Lambda\Sigma\hat{h}\right)p - g\left(x_1,\hat{h};\theta\right)\right]$$

we would like to prove that $G(x_1, p, D) - G(x_1, p, E) \leq 0$ for $E, D \in \mathbb{R}$ such that $E \leq D$. We take the difference

$$\begin{aligned} G(x_1, p, D) - G(x_1, p, E) &= -\frac{1}{2}\Lambda^2 D + \frac{1}{2}\Lambda^2 E \\ &= -\frac{1}{2}\Lambda^2 (D - E) \\ &\leq 0, \end{aligned}$$

since $E \leq D$, which completes the proof. The proof can be easily extended to the two-factor finite difference operator since

$$\begin{aligned} G(x, p, D) - G(x, p, E) &= -\frac{1}{2} tr(\Lambda\Lambda' D) + \frac{1}{2} tr(\Lambda\Lambda' E) \\ &= -\frac{1}{2} tr(\Lambda\Lambda' D - \Lambda\Lambda' E) \\ &= -\frac{1}{2} tr\left(\begin{bmatrix} \Lambda\Lambda' & \Lambda\Lambda' \\ \Lambda\Lambda' & \Lambda\Lambda' \end{bmatrix} \begin{bmatrix} D & 0 \\ 0 & -E \end{bmatrix} \right) \\ &\leq -\frac{3}{2}\alpha \begin{pmatrix} I & -I \\ -I & I \end{pmatrix} = 0. \end{aligned}$$

□

Theorem 6.4. ***(Convergence to the viscosity solution)*** [2] *Provided that the Hamilton-Jacobi-Bellman equation of (5.11) satisfies Theorem 5.25 and the numerical scheme (6.3) satisfies Lemmas (6.1), (6.2) and (6.3), then the solution of the numerical scheme converges uniformly on each compact subset of* $[0, T] \times \mathbb{R}^n$ *to the unique viscosity solution of (5.11) for* $n = 2$.

6.6. **Numerical results with two factor processes.** The numerical scheme (6.3) was implemented in MATLAB and we present below the results obtained with its execution. The code provided consists of two files: file *riskSensitive-HJB2factor.m* performs the calculations to simulate the risk-sensitive optimisation problem described by equation (5.11); and file *example2factor.m* serves as an example of how to build the optimisation problem by setting the input parameters, solve it and display the results. As mentioned above the software is designed to deal with a two-dimensional state domain, and thus we have two factor processes. It is also our choice to limit the number of risky securities to two.

We include in this subsection the graphs generated by running the file *example2factor.m*. The values of all input parameters were kept unchanged as they are in the provided file, i.e.

$$v = 1000,\ b = \begin{bmatrix} 0.3 \\ 0.7 \end{bmatrix},\ B = \begin{bmatrix} 0.5 & -0.3 \\ -0.3 & 0.2 \end{bmatrix},\ \Lambda = \begin{bmatrix} 0.2 & -0.2 \\ -0.2 & 0.5 \end{bmatrix},$$

$$a_0 = 0.05,\ A_0 = \begin{bmatrix} -0.3 \\ -0.4 \end{bmatrix},\ a = \begin{bmatrix} 0.2 \\ 0.4 \end{bmatrix},\ A = \begin{bmatrix} 0.4 & -0.3 \\ -0.2 & 0.5 \end{bmatrix},$$

$$\Sigma = \begin{bmatrix} 0.3 & -0.2 \\ -0.2 & 0.4 \end{bmatrix},\ \gamma_{min} = -0.9,\ \gamma_{max} = 10,\ R = \begin{bmatrix} 1 \\ 2 \end{bmatrix},$$

and we only varied the value of the risk-sensitivity parameter θ, which represents the investor's degree of risk aversion, and the intensity λ of the density function of the Lévy measure. We recall here that $\theta \in (-1, 0) \cup (0, \infty)$ and that risk-averse investors choose a positive value for θ. In the risk-sensitive optimisation

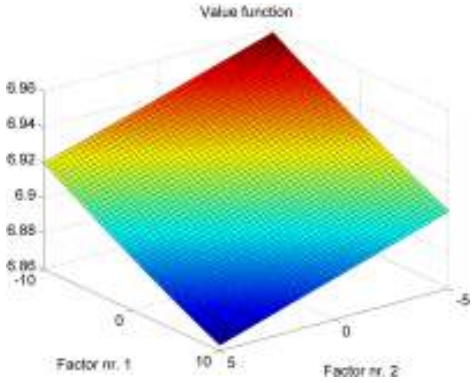

FIGURE 6.1. Value function for $\theta = 2$ and $\lambda = 5$.

FIGURE 6.2. Weight for money market account for $\theta = 2$ and $\lambda = 5$.

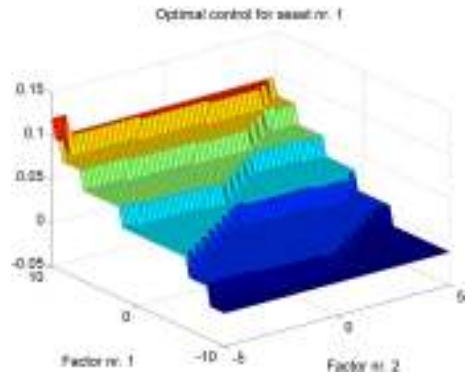

FIGURE 6.3. Optimal control for first asset for $\theta = 2$ and $\lambda = 5$.

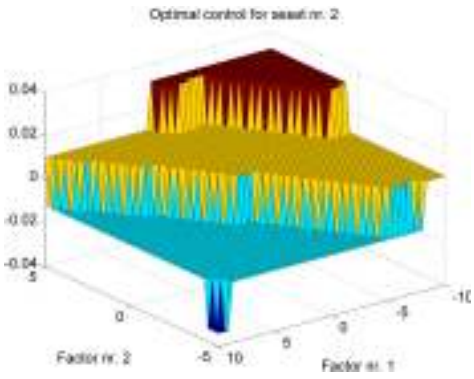

FIGURE 6.4. Optimal control for second asset for $\theta = 2$ and $\lambda = 5$.

problems that we have simulated we considered the time interval $[0, 1]$ and the state domain $(x_1, x_2) \in [-10, 10] \times [-5, 5]$. Besides, the space $\mathbf{Z}$, which represents the space of possible jump sizes, was truncated to the domain $[-5, 5] \times [-5, 5]$. In order to discretise all of them, we chose the values $N_t = 100$, $N_{x_1} = N_{x_2} = 50$, $N_{z_1} = N_{z_2} = 20$. As mentioned at the beginning of Subsection 5.1, the state domain was chosen in such a way that its values could represent, for instance, rates of GDP growth and inflation.

Figures 6.1, 6.2, 6.3 and 6.4 were generated for values $\theta = 2$ and $\lambda = 5$, and they show the value function and the distribution of capital among the available assets one timestep before maturity, i.e. at time $t = 0.99$ (or equivalently, $m = 99$). Being the value of $\theta > 0$ we know that our investor is risk-averse, and this is evident from the graphs: most capital is allocated in the money market account, a safe investment.

Running the software with values $\theta = -0.9$ and $\lambda = 5$ produced Figures 6.5, 6.6, 6.7 and 6.8. First thing we can observe respect to the previous simulation is that

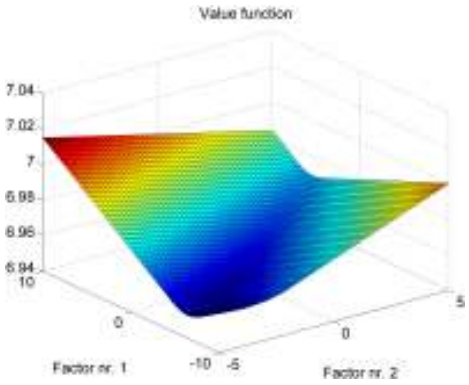

FIGURE 6.5. Value function for $\theta = -0.9$ and $\lambda = 5$.

FIGURE 6.6. Weight for money market account for $\theta = -0.9$ and $\lambda = 5$.

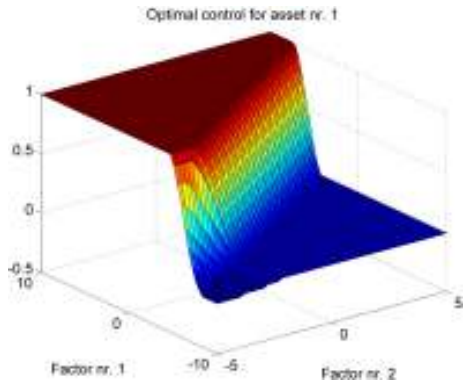

FIGURE 6.7. Optimal control for first asset for $\theta = -0.9$ and $\lambda = 5$.

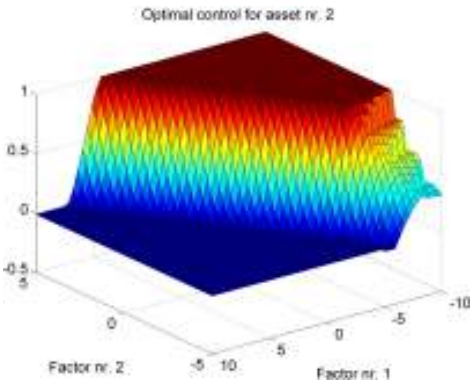

FIGURE 6.8. Optimal control for second asset for $\theta = -0.9$ and $\lambda = 5$.

capital is invested in the risk-free asset only for lower values of the state domain. Low values of x_1 and x_2 may suggest that the economy is performing poorly, and thus it is less uncertain to invest in the risk-free asset. On the contrary, as x_1 and x_2 move away from the minimums, our investor assigns incrementally a larger fraction of the portfolio value to risky securities. This agrees with our choice of θ, which makes our investor a risk-seeker. Since

$$A = \begin{bmatrix} 0.4 & -0.3 \\ -0.2 & 0.5 \end{bmatrix}$$

the first risky asset is positively correlated with the first factor process, and thus, when x_1 takes high values, the largest share of the current value of the portfolio corresponds to the investment in the first risky asset. Likewise, the second risky asset is positively correlated with the second factor process, and when x_2 takes high values, the second risky asset becomes the main investment.

In Figures 6.9, 6.10, 6.11 and 6.12 we can appreciate the effect that the value of λ has in the investment strategy. The value of θ is equal to 2, and thus our

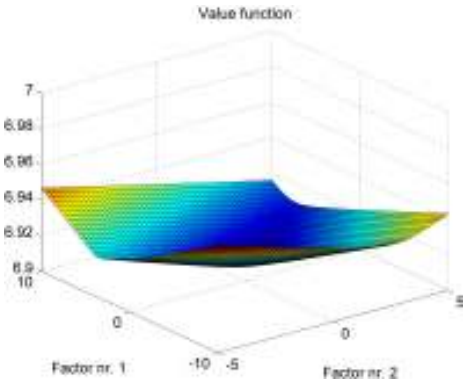

FIGURE 6.9. Value funtion for $\theta = 2$ and $\lambda = 0.5$.

FIGURE 6.10. Weight for money market account for $\theta = 2$ and $\lambda = 0.5$.

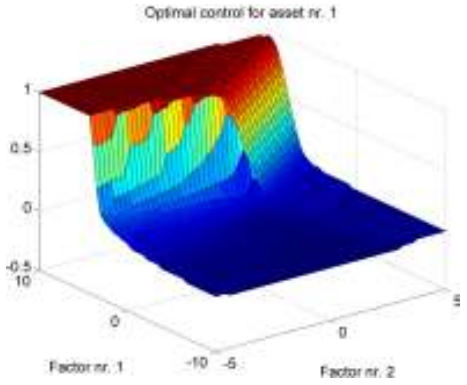

FIGURE 6.11. Optimal control for first asset for $\theta = 2$ and $\lambda = 0.5$.

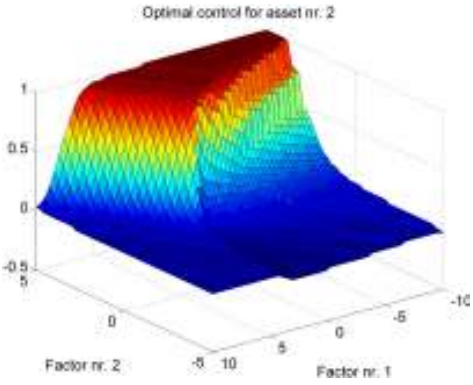

FIGURE 6.12. Optimal control for second asset for $\theta = 2$ and $\lambda = 0.5$.

investor shows the same degree of risk aversion as in the first simulation of this subsection. However, λ is 10 times smaller than before, and this means that, on average, there are 0.5 jumps per unit of time instead of 5. Therefore, the value of the risky securities does not jump so often and investing in them is seen less unsafe. Still, the portfolio wealth is predominantly due to the risk-free asset, and capital is only allocated in the risky securities for high values of the state domain.

In all three simulations it is worth noticing that the value function generated by the numerical scheme satisfies Corollary 5.16: we can see that at $t = 0.99$ the numerical approximation $\hat{\Phi}^m_{i,j}$ is convex with respect to the state variable.

6.7. **Numerical results with one factor process.** We also implemented in MATLAB the numerical scheme (6.4). The code is provided in files *riskSensitive-HJB1factor.m* and *example1factor.m*. Similarly as with the files for the two-factor numerical scheme, the former performs the calculations to simulate the risk-sensitive optimisation problem; and the latter provides an example of how to set the input

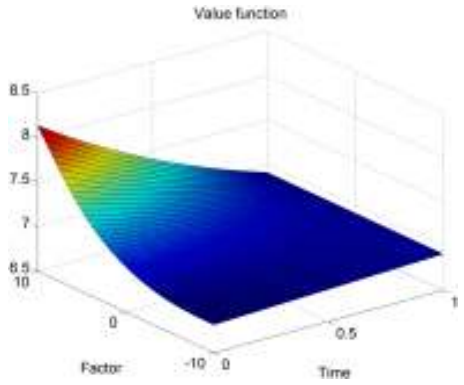

FIGURE 6.13. Value funtion for $\theta = 0.05$ and $\lambda = 5$.

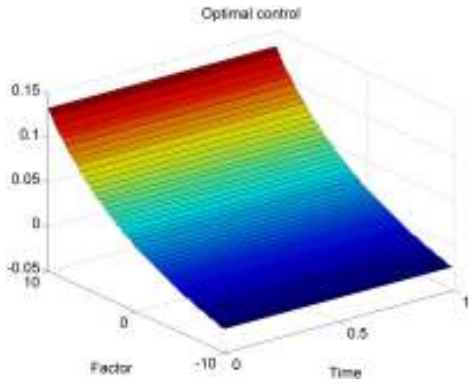

FIGURE 6.14. Optimal control for $\theta = 0.05$ and $\lambda = 5$.

parameters, solve the optimisation problem and display the results. The software is designed to deal with a one-dimensional state domain, and thus we have one factor process and one risky asset. We present below the results obtained for a number of simulations.

All the graphs included in this subsection have been produced running the file *example1factor.m*. Apart from the parameters θ and λ, the values of the other parameters were not modified and they kept the values shown in the file throughout our simulations, i.e.

$$v = 1000,\ b = 0.7,\ B = 0.5,\ \Lambda = 0.2,\ a_0 = 0.05,\ A_0 = 0.02,$$
$$a = 0.4,\ A = 0.8,\ \Sigma = 0.3,\ \gamma_{min} = -0.9,\ \gamma_{max} = 10,\ R = 2.$$

We simulated risk-sensitive optimisation problems where we considered the time interval $[0, 1]$, the state domain $[-10, 10]$ and the jump-size domain $[-5, 5]$, choosing values $N_t = 100$, $N_{x_1} = 50$, $N_{z_1} = 100$ to carried out the discretisation. Similarly as in Subsection 6.6 the state domain could represent in this case rates of GDP growth.

Figures 6.13 and 6.14 were obtained for values $\theta = 0.05$ and $\lambda = 5$. We will use these two graphs as reference against which to compare the graphs obtained with lower and higher values of θ. It can be seen that a higher fraction of the portfolio wealth is invested in the risky asset as the state variable takes higher values. If we assume that the state variable represents rates of GDP growth, we can see that this investment strategy makes sense: the higher the GDP growth rate, the more the economy grows, and it is usually more profitable to invest in risky assets. However, the fraction invested in the risky security never exceeds 15% and thus most of the capital is allocated in the risk-free asset.

If value $\theta = 2$ and $\lambda = 5$ are chosen, Figures 6.15 and 6.16 may be obtained. If we compare these to Figures 6.13 and 6.14, we can observe that the investment in the risky asset represents even a smaller portion of the portfolio wealth. Since the value of θ is higher than before, the investor is more risk-averse, and thus prefers to invest more heavily in the money market account, which is a safer investment.

Figures 6.17 and 6.18 were generated after choosing values $\theta = -0.9$ and $\lambda = 5$. It can be noticed –comparing with respect to Figures 6.13 and 6.14– that a higher

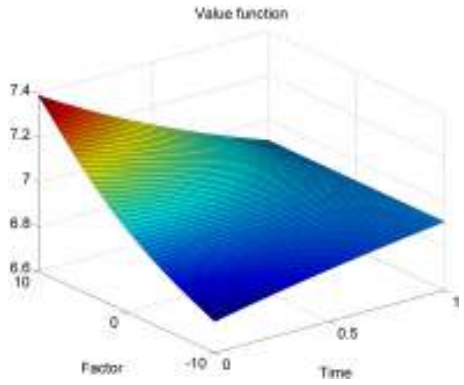

FIGURE 6.15. Value funtion for $\theta = 2$ and $\lambda = 5$.

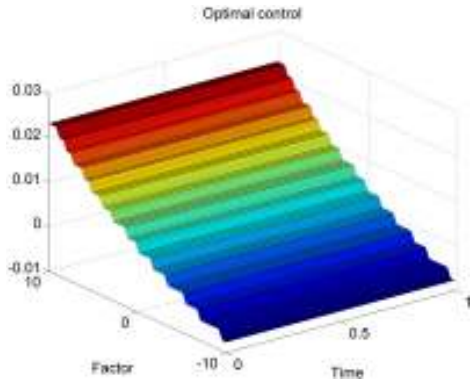

FIGURE 6.16. Optimal control for $\theta = 2$ and $\lambda = 5$.

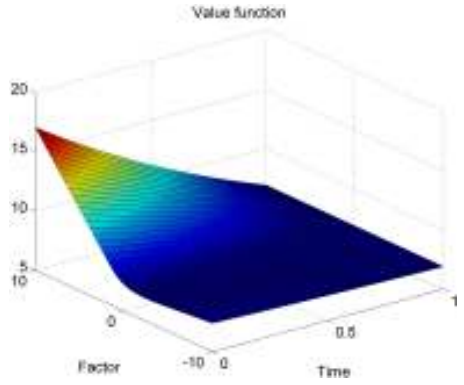

FIGURE 6.17. Value funtion for $\theta = -0.9$ and $\lambda = 5$.

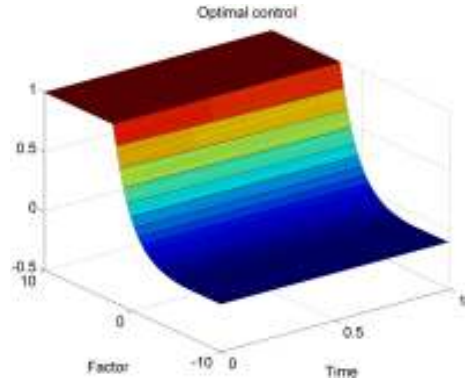

FIGURE 6.18. Optimal control for $\theta = -0.9$ and $\lambda = 5$.

fraction of capital is invested in the risky asset. In fact, when the state variable takes values close to the maximum, the portfolio consists uniquely of the investment in the risky security and no capital is allocated in the money market account. Since $\theta < 0$, the investor is risk-seeker and pursues a more aggressive investment strategy as the state variable takes higher values.

Finally, Figures 6.19 and 6.20 help us examine again the effect that the value of λ has in the investment strategy. They were generated for values $\theta = 2$ and $\lambda = 0.5$, and thus the value of the risky security jumps, on average, 10 times less often per unit of time. We can see how this translates into a greater fraction of the portfolio wealth being invested in the risky security, in particular as the state variable takes higher values. Even though these figures were generated with $\theta = 2$ (and simulate, thus, the behavior of a risk-averse investor), they look more similar to Figures 6.17 and 6.18, which were generated with $\theta = -0.9$ (i.e. a risk-seeking investor). This shows how the value of λ influences the result of the optimisation problem, and

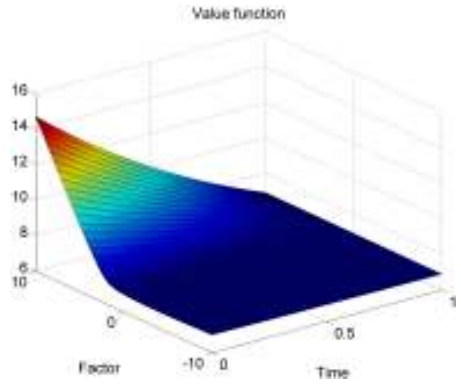

FIGURE 6.19. Value funtion for $\theta = 2$ and $\lambda = 0.5$.

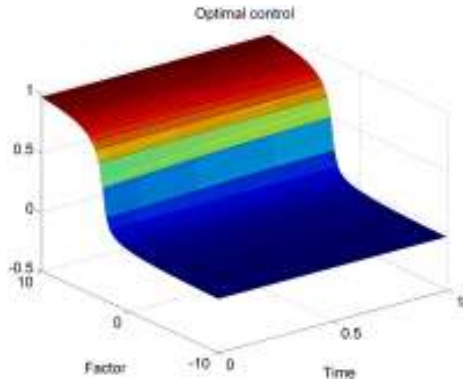

FIGURE 6.20. Optimal control for $\theta = 2$ and $\lambda = 0.5$.

how an investment in a risky security might be perceived less uncertain if its value jumps less often, and thus it is less volatile.

It is again worth noticing that, in all four examples, the value function generated by the numerical scheme satisfies Corollary 5.16, i.e. for any fixed $m \in I_t$, $\hat{\Phi}_i^m$ is convex with respect to the state variable.

APPENDIX A. MATLAB CODE

A.1. **Two-factor implementation.** The numerical scheme (6.3) is implemented in the file *riskSensitiveHJB2factor.m* as a class. Thus, we need first to instantiate an object of type *riskSensitiveHJB2factor* and set the following input parameters before trying to solve the optimisation problem:

- Parameters *t*, *x1* and *x2* are one-dimensional vectors representing the possible values of t, x_1 and x_2 in domain Q_T^d.
- Parameter v corresponds to v in equation (5.11).
- Parameter *numberOfControls* stores the number of possible control strategies.
- Parameters b, B and *lambda* correspond, respectively, to b, B, and Λ in equation (5.2) for $n, M = 2$.
- Parameters *a0* and *A0* correspond, respectively, to a_0 and A_0 in equation (5.3) for $n = 2$.
- Parameters a, A and *sigma* correspond, respectively, to a, A and Σ in equation (5.4) for $n, m, M = 2$.
- Parameters *z1* and *z2* are one-dimensional vectors representing the set of values for both coordinates of the domain $\mathbf{Z}_t^d$.
- Parameters *gamma_ min* and *gamma_ max* correspond, respectively, to the values $\gamma_i^{\min}$ and $\gamma_i^{\max}$ from Assumption 5.5 for $i = 1, 2$.
- Parameter *gamma* is a three-dimensional matrix that corresponds to function $\gamma(z)$ in equation (5.4) for $m = 2$ and $z \in \mathbf{Z}_t^d$.
- Parameter R correspond to constant R in equation (3.2) for $d = 2$.
- Parameters *intensity*, *mean* and *covariance* correspond, respectively, to λ, μ, and Σ in equation (6.2).

The class *riskSensitiveHJB2factor* exposes a public method called *solve* that returns the following output parameters:

- Output V is a three-dimensional array that corresponds to $\hat{\Phi}^m_{i,j}$ in (6.3). First two dimensions of V correspond to coordinates i, j of the state variables and the third dimension corresponds to time m.
- Output h is a four-dimensional array that corresponds to the optimal control process $\hat{h}^m$ in (6.3). First two dimensions of h indicate the indeces i, j of the state variables; third dimension represents the risky security index; and fourth dimension corresponds to time m.

```
clear;

test = riskSensitiveHJB2factor();
test.t = linspace(1 - 1/100, 1, 2);
test.x1 = linspace(-10, 10, 50);
test.x2 = linspace(-5, 5, 50);
test.v = 1000;
test.numberOfControls = 50;
test.theta = 2;
test.b = [0.3
          0.7];
test.B = [0.5 -0.3
          -0.3 0.2];
test.lambda = [0.2 -0.2
               -0.2 0.5];
test.a0 = 0.05;
test.A0 = [-0.3
           -0.4];
test.a = [0.2
          0.4];
test.A = [0.4 -0.3
          -0.2 0.5];
test.sigma = [0.3 -0.2
              -0.2 0.4];
test.z1 = linspace(-5, 5, 20);
test.z2 = linspace(-5, 5, 20);
test.gamma_min = -0.9;
test.gamma_max = 10;
test.gamma = (test.gamma_max - test.gamma_min) ...
             .*rand(length(test.z1), length(test.z2), 2) + test.
               gamma_min;
test.R = [1
          2];
test.intensity = 5;
test.mean = [0.5
             0.8];
test.covariance = [0.8 0.5
                   0.5 0.8];

% Find solution to HJB problem
[V, h] = test.solve();

axis tight
[X, Y] = meshgrid(test.x2, test.x1);
surf(X, Y, V(:, :, 1));
xlabel('Factor nr. 2');
```

```
ylabel('Factor nr. 1');
title('Value function');

figure;
surf(X, Y, h(:, :, 1, 1));
xlabel('Factor nr. 2');
ylabel('Factor nr. 1');
title('Optimal control for asset nr. 1');

figure;
surf(X, Y, h(:, :, 1, 2));
xlabel('Factor nr. 2');
ylabel('Factor nr. 1');
title('Optimal control for asset nr. 2');

figure;
surf(X, Y, ones(length(X), length(Y)) - (h(:, :, 1, 1) + h(:, :, 1, 2)
  ));
xlabel('Factor nr. 2');
ylabel('Factor nr. 1');
title('Weight for money market account');
```

LISTING 1. File example2factor.m

```
% riskSensitiveHJB2factor approximates solution to risk-sensitive
% HJB equation with two factor process
%
% Reference: M.H.A. Davis and S. Lleo. Jump-diffusion risk-sensitive
% asset management I: Diffusion factor model. SIAM Journal on
% Financial Mathematics, 2:22-54, 2011.

classdef (Sealed = true) riskSensitiveHJB2factor
properties (Constant)
  one = [1
         1];
end
properties (GetAccess = private, SetAccess = private)
  % Value function
  V
  % Optimal control
  h
  % Wealth dynamics
  a_tilde
  A_tilde
  % Control policy
  weights
  isControlAdmissible
  % Results from integration with respect to Levy measure
  integrals
end
properties
  t         % Time range
  x1        % Factor 1 value range
  x2        % Factor 2 value range
  numberOfControls
  % Risk sensitivity
  theta
  % Factor dynamics
```

```
  b
  B
  lambda
  % Asset market dynamics
  a0
  A0
  % Risky security dynamics
  a
  A
  sigma
  z1
  z2
  gamma_min
  gamma_max
  gamma
  % Small/big jump border
  R
  % HJB problem
  v
  % Levy measure with Gaussian density
  intensity
  mean
  covariance
end

methods
function obj = riskSensitiveHJB2factor()
end

function [V, h] = solve(obj)
% Check if input values are correct
obj.checkParameters();

dt = diff(obj.t(1:2));
dx1 = diff(obj.x1(1:2));
dx2 = diff(obj.x2(1:2));

obj.a_tilde = obj.a - obj.a0.*riskSensitiveHJB2factor.one;
obj.A_tilde = obj.A - riskSensitiveHJB2factor.one*obj.A0';

% Preallocate matrices
% Value function
obj.V = zeros(length(obj.x1) + 2, length(obj.x2) + 2, length(obj.t));
% Optimal control
obj.h = zeros(length(obj.x1) + 2, length(obj.x2) + 2, length(obj.t) -
  1, 2);

% Calculate asset weights
obj.weights = linspace(-1/obj.gamma_max, 1, obj.numberOfControls + 1);
obj.weights = obj.weights(2:end);

% Find what controls are admissible
obj.isControlAdmissible = obj.findAdmissibleControls();

% Precompute integrals with respect to Levy measure
obj.integrals = obj.calculateIntegralsWithRespectToLevyMeasure();

% Apply terminal condition to value function
```

```
obj.V(:, :, length(obj.t)) = log(obj.v);

for m = length(obj.t):-1:2          % time
  if m < length(obj.t)
    obj.V(:, :, m) = obj.extrapolateValueFunctionBeyondBorders(dx1,
      dx2, m);
  end

  for i = 2:1:length(obj.x1) + 1    % factor 1
    for j = 2:1:length(obj.x2) + 1  % factor 2
      x = [obj.x1(i - 1)
           obj.x2(j - 1)];

      [DV, D2V] = obj.partialDerivatives(dx1, dx2, m, i, j);
      [sup, optimal_h] = obj.supOperatorL(x, DV);
      obj.h(i, j, m - 1, :) = optimal_h;
      obj.V(i, j, m - 1) = obj.V(i, j, m) ...
        + dt*((obj.b + obj.B*x)'*(DV) ...
        + 0.5*trace(obj.lambda*(obj.lambda'*D2V)) ...
        - (obj.theta/2)*(DV'*(obj.lambda*(obj.lambda'*DV))) ...
        + (obj.a0 + obj.A0'*x + sup));

      fprintf('Value function for state (%.2f, %.2f) calculated at t =
        %.4f\n', obj.x1(i - 1), obj.x2(j - 1), obj.t(m - 1));
    end
  end
end
V = obj.V(2:end - 1, 2:end - 1, :);
h = obj.h(2:end - 1, 2:end - 1, :, :);
end
end % public methods

methods (Access = private)
%CHECKPARAMETERS Check that some of the input parameters are valid
% in order to solve the HJB equation.
function checkParameters(obj)
% Check that lambda*lambda' is positive definite
[R, p] = chol(obj.lambda*obj.lambda');
if p > 0
  error('riskSensitiveHJB2factor:invalidInputs', 'lambda*lambda'' is
    not positive definite.');
end
% Check that sigma*sigma' is positive definite
[R, p] = chol(obj.sigma*obj.sigma');
if p > 0
  error('riskSensitiveHJB2factor:invalidInputs', 'sigma*sigma'' is not
     positive definite.');
end
if obj.gamma_min <= -1 || obj.gamma_min >= 0
  error('riskSensitiveHJB2factor:invalidInputs', 'gamma_min must be >
    -1 and < 0.');
end
if obj.gamma_max <= 0
  error('riskSensitiveHJB2factor:invalidInputs', 'gamma_max must be >
    0.');
end
if obj.theta == 0 || obj.theta <= -1
```

```
  error('riskSensitiveHJB2factor:invalidInputs', 'theta cannot be 0 or
     <= -1.');
end
end

%EXTRAPOLATEVALUEFUNCTIONBEYONDBORDERS Extrapolates one extra value
% around the borders of the value function
function [val] = extrapolateValueFunctionBeyondBorders(obj, dx1, dx2,
  m)
[X, Y] = meshgrid(obj.x2, obj.x1);
[Xq, Yq] = meshgrid(obj.x2(1) - dx2:dx2:obj.x2(end) + dx2, obj.x1(1) -
   dx1:dx1:obj.x1(end) + dx1);
val = interp2(X, Y, obj.V(2:end - 1, 2:end - 1, m), Xq, Yq, 'spline');
end

%PARTIALDERIVATIVES Return first and second order partial derivatives
% with respect to the state variable
function [DV, D2V] = partialDerivatives(obj, dx1, dx2, m, i, j)
% First order partial derivative of value function
DV = [(obj.V(i + 1, j, m) - obj.V(i - 1, j, m))/(2*dx1)
      (obj.V(i, j + 1, m) - obj.V(i, j - 1, m))/(2*dx2)];

% Second order partial derivative of value function (here x=x1, y=x2)
fxx = (obj.V(i + 1, j, m) - 2*obj.V(i, j, m) ...
  + obj.V(i - 1, j, m))/(dx1^2);
fyy = (obj.V(i, j + 1, m) - 2*obj.V(i, j, m) ...
  + obj.V(i, j - 1, m))/(dx2^2);
fxy = (obj.V(i + 1, j + 1, m) - obj.V(i + 1, j - 1, m) ...
  - obj.V(i - 1, j + 1, m) + obj.V(i - 1, j - 1, m))/(4*dx1*dx2);
fyx = fxy;
D2V = [fxx fxy
       fyx fyy];
end

%FINDADMISSIBLECONTROLS Find what control policies are admissible
function [isControlAdmissible] = findAdmissibleControls(obj)
isControlAdmissible = zeros(length(obj.weights), length(obj.weights));
for i = 1:1:length(obj.weights)   % Weight for asset 1
  for j = 1:1:length(obj.weights) % Weight for asset 2
    h = [obj.weights(i)
         obj.weights(j)];

    if obj.checkControlIsAdmissible(h) == true
      isControlAdmissible(i, j) = true;
    else
      isControlAdmissible(i, j) = false;
    end
  end
end
end

%CHECKCONTROLISADMISSIBLE Check if input control policy is admissible
function [isAdmissible] = checkControlIsAdmissible(obj, h)
isAdmissible = true;
for i = 1:1:length(obj.gamma(:, 1, 1))
  if isAdmissible == false
    break;
  end
```

```
  for j = 1:1:length(obj.gamma(1, :, 1))
    % Condition for admissibility
    if [obj.gamma(i, j, 1) obj.gamma(i, j, 2)]*h <= -1 || sum(h) > 1
      isAdmissible = false;
      break;
    end
  end
end
end

%SUPOPERATORL Calculate the maximum value and the control policy
% that produces it
function [val, optimal_h] = supOperatorL(obj, x, DV)
f = zeros(length(obj.weights), length(obj.weights));
for i = 1:1:length(obj.weights)   % Weight for asset 1
  for j = 1:1:length(obj.weights) % Weight for asset 2
    if obj.isControlAdmissible(i, j) == true
      h = [obj.weights(i)
           obj.weights(j)];

      f(i, j) = -0.5*(obj.theta + 1)*(h'*obj.sigma)*(obj.sigma'*h) ...
                - obj.theta*(h'*(obj.sigma*(obj.lambda'*DV))) ...
                + h'*(obj.a_tilde + obj.A_tilde*x) ...
                - (1/obj.theta)*obj.integrals(i, j);
    else
      f(i, j) = nan;
    end
  end
end

% Find maximum value
[val, ind] = max(f(:));
[row, column] = ind2sub(size(f), ind);
% Control policy that produces the maximum
optimal_h = [obj.weights(row)
             obj.weights(column)];
end

%CALCULATEINTEGRALWITHRESPECTTOLEVYMEASURE Calculate integral with
% respect to Levy measure for every admissible control policy
function [integrals] = calculateIntegralsWithRespectToLevyMeasure(obj)
fprintf('Calculating integrals with respect to Levy measure');

integrals = zeros(length(obj.weights), length(obj.weights));
for i = 1:1:length(obj.weights)   % Weight for asset 1
  for j = 1:1:length(obj.weights) % Weight for asset 2
    if obj.isControlAdmissible(i, j) == true
      h = [obj.weights(i)
           obj.weights(j)];

      integrals(i, j) = obj.integralWithRespectToLevyMeasure(h);
    else
      integrals(i, j) = nan;
    end
  end
  fprintf('.');
end
```

```
fprintf('\n');
end

%INTEGRALWITHRESPECTTOLEVYMEASURE Approximate integral with respect
% to Levy measure
function val = integralWithRespectToLevyMeasure(obj, h)
val = 0;
dz1 = diff(obj.z1(1:2));
dz2 = diff(obj.z2(1:2));

for i = 1:1:length(obj.z1) - 1
  for j = 1:1:length(obj.z2) - 1
    % Evaluate the integrand at every corner of the rectangle
    corner1 = obj.evaluateIntegrand(i, j, h);
    corner2 = obj.evaluateIntegrand(i + 1, j, h);
    corner3 = obj.evaluateIntegrand(i, j + 1, h);
    corner4 = obj.evaluateIntegrand(i + 1, j + 1, h);
    val = val + 0.25*(corner1 + corner2 + corner3 + corner4)*dz1*dz2;
  end
end
end

%EVALUATEINTEGRAND Evaluate integrand of integral with respect to
% Levy measure at one point of the domain
function val = evaluateIntegrand(obj, i, j, h)
aux1 = (power(1 + [obj.gamma(i, j, 1) obj.gamma(i, j, 2)]*h, -1*obj.
  theta) - 1);
aux2 = 0;
if (abs(obj.z1(i)) <= obj.R(1) && abs(obj.z2(j)) <= obj.R(2))
  aux2 = obj.theta*[obj.gamma(i, j, 1) obj.gamma(i, j, 2)]*h;
end
z = [obj.z1(i)
     obj.z2(j)];
val = (aux1 + aux2)*obj.gaussianDensity2D(z);
end

%GAUSSIANDENSITY2D Calculate the value of the 2D Gaussian density
% function at any input point
function val = gaussianDensity2D(obj, z)
if isequal(z, [0 0]')
  val = 0; % Levy measure has no mass at origin
else
  val = obj.intensity*(1/sqrt(power(2*pi, length(obj.mean))*det(obj.
    covariance)))*(exp(-0.5*(z - obj.mean)'*(obj.covariance\(z - obj.
    mean))));
end
end

end % private methods
end % classdef
```

LISTING 2. file riskSensitiveHJB2factor.m

A.2. **One-factor implementation.** The numerical scheme allowing one factor process is implemented in the file *riskSensitiveHJB1factor.m*. The code implements also a class, and thus we need first to instantiate an object of type *riskSensitiveHJB1factor* and set the following input parameters before trying to solve the optimisation problem:

- Parameters t, *x1* are one-dimensional vectors representing the possible values of t and x_1.
- Parameter v corresponds to v in equation (5.11).
- Parameter *numberOfControls* stores the number of possible control strategies.
- Parameters b, B and *lambda* correspond, respectively, to b, B, and Λ in equation (5.2) for $n, M = 1$.
- Parameters *a0* and *A0* correspond, respectively, to a_0 and A_0 in equation (5.3) for $n = 1$.
- Parameters a, A and *sigma* correspond, respectively, to a, A and Σ in equation (5.4) for $n, m, M = 1$.
- Parameter z is a one-dimensional vector representing the set of values of the domain $\mathbf{Z}_t^d$.
- Parameters *gamma_ min* and *gamma_ max* correspond, respectively, to the values $\gamma_i^{\min}$ and $\gamma_i^{\max}$ from Assumption 5.5 for $i = 1$.
- Parameter *gamma* is a one-dimensional vector that corresponds to function $\gamma(z)$ in equation (5.4) for $m = 1$ and $z \in \mathbf{Z}_t^d$.
- Parameter R correspond to constant R in equation (3.2) for $d = 1$.
- Parameters *intensity*, *mean* and *variance* correspond, respectively, to λ, μ, and Σ in the one-dimensional version of equation (6.2).

The class *riskSensitiveHJB1factor* exposes a public method called *solve* that returns the following output parameters:

- Output V is a two-dimensional array that corresponds to $\hat{\Phi}_i^m$ in (6.4). First dimension of V corresponds to coordinate i of the state variable and the second dimension corresponds to time m.
- Output h is a two-dimensional array that corresponds to the optimal control process $\hat{h}^m$ in (6.4). First dimension of h represents the index i of the state variable and the second dimension corresponds to time m.

```
clear;

test = riskSensitiveHJB1factor();
test.t = linspace(0, 1, 100);
test.x1 = linspace(-10, 10, 50);
test.v = 1000;
test.numberOfControls = 500;
test.theta = 2;
test.b = 0.7;
test.B = 0.5;
test.lambda = 0.2;
test.a0 = 0.05;
test.A0 = 0.02;
test.a = 0.4;
test.A = 0.8;
test.sigma = 0.3;
test.z = linspace(-5, 5, 100);
```

```
test.gamma_min = -0.9;
test.gamma_max = 10;
test.gamma = (test.gamma_max - test.gamma_min) ...
             .*rand(length(test.z), 1) + test.gamma_min;
test.R = 2;
test.intensity = 5;
test.mean = 1;
test.variance = 0.8;

% Find solution to HJB problem
[V, h] = test.solve();

axis tight
[t, X] = meshgrid(test.t, test.x1);
surf(t, X, V);
xlabel('Time');
ylabel('Factor');
title('Value function');

figure
[t, X] = meshgrid(test.t(1:end -1), test.x1);
surf(t, X, h);
xlabel('Time');
ylabel('Factor');
title('Optimal control');
```

LISTING 3. File example1factor.m

```
% riskSensitiveHJB1factor approximates solution to risk-sensitive
% HJB equation with one factor process
%
% Reference: M.H.A. Davis and S. Lleo. Jump-diffusion risk-sensitive
% asset management I: Diffusion factor model. SIAM Journal on
% Financial Mathematics, 2:22-54, 2011.

classdef (Sealed = true) riskSensitiveHJB1factor
properties (GetAccess = private, SetAccess = private)
  % Value function
  V
  h
  % Wealth dynamics
  a_tilde
  A_tilde
  % Control policy
  weights
  isControlAdmissible
  % Results from integration with respect to Levy measure
  integrals
end
properties
  t         % Time range
  x1        % Factor value range
  numberOfControls
  % Risk sensitivity
  theta
  % Factor dynamics
  b
  B
```

```
  lambda
  % Asset market dynamics
  a0
  A0
  % Risky security dynamics
  a
  A
  sigma
  z
  gamma_min
  gamma_max
  gamma
  % Small/big jump border
  R
  % HJB problem
  v
  % Levy measure with Gaussian density
  intensity
  mean
  variance
end

methods
function obj = riskSensitiveHJB1factor()
end

function [V, h] = solve(obj)
% Check if input values are correct
obj.checkParameters();

dt = diff(obj.t(1:2));
dx1 = diff(obj.x1(1:2));

obj.a_tilde = obj.a - obj.a0;
obj.A_tilde = obj.A - obj.A0;

% Preallocate matrices
% Value function
obj.V = zeros(length(obj.x1) + 2, length(obj.t));
% Optimal control
obj.h = zeros(length(obj.x1) + 2, length(obj.t) - 1);

% Calculate asset weights
obj.weights = linspace(-1/obj.gamma_max, 1, obj.numberOfControls + 1);
obj.weights = obj.weights(2:end);

% Find what controls are admissible
obj.isControlAdmissible = obj.findAdmissibleControls();

% Precompute integrals with respect to Levy measure
obj.integrals = obj.calculateIntegralsWithRespectToLevyMeasure();

% Apply terminal condition to value function
obj.V(:, length(obj.t)) = log(obj.v);

for m = length(obj.t):-1:2  % time
  if m < length(obj.t)
    obj.V(:, m) = obj.extrapolateValueFunctionBeyondBorders(dx1, m);
```

```
  end

  for i = 2:1:length(obj.x1) + 1  % factor
    [DV, D2V] = obj.partialDerivatives(dx1, m, i);
    [sup, optimal_h] = obj.supOperatorL(obj.x1(i - 1), DV);
    obj.h(i, m - 1, :) = optimal_h;
    obj.V(i, m - 1) = obj.V(i, m) ...
                      + dt*((obj.b + obj.B*obj.x1(i - 1))*(DV) ...
                      + 0.5*(obj.lambda^2)*(D2V) ...
                      - (obj.theta/2)*(obj.lambda^2)*(DV^2) ...
                      + obj.a0 + obj.A0*obj.x1(i - 1) + sup);

    fprintf('Value function for state %.2f calculated at t = %.4f\n',
      obj.x1(i - 1), obj.t(m - 1));
  end
end
V = obj.V(2:end - 1, :);
h = obj.h(2:end - 1, :, :);
end
end % public methods

methods (Access = private)
%CHECKPARAMETERS Check that some of the input parameters are valid
% in order to solve the HJB equation.
function checkParameters(obj)
if obj.gamma_min <= -1 || obj.gamma_min >= 0
  error('riskSensitiveHJB1factor:invalidInputs', 'gamma_min must be >
    -1 and < 0.');
end
if obj.gamma_max <= 0
  error('riskSensitiveHJB1factor:invalidInputs', 'gamma_max must be >
    0.');
end
if obj.theta == 0 || obj.theta <= -1
  error('riskSensitiveHJB1factor:invalidInputs', 'theta cannot be 0 or
     <= -1.');
end
end

%EXTRAPOLATEVALUEFUNCTIONBEYONDBORDERS Extrapolates one extra value
% around the borders of the value function
function [val] = extrapolateValueFunctionBeyondBorders(obj, dx1, m)
Xq = obj.x1(1) - dx1:dx1:obj.x1(end) + dx1;
val = interp1(obj.x1, obj.V(2:end - 1, m), Xq, 'linear', 'extrap');
end

%PARTIALDERIVATIVES Return first and second order partial derivatives
% with respect to the state variable
function [DV, D2V] = partialDerivatives(obj, dx1, m, i)
% First order partial derivative of value function
DV = (obj.V(i + 1, m) - obj.V(i - 1, m))/(2*dx1);

% Second order partial derivative of value function
D2V = (obj.V(i + 1, m) - 2*obj.V(i, m) + obj.V(i - 1, m))/(dx1^2);
end

%SUPOPERATORL Calculate the maximum value and the control policy that
% produces it
```

```
function [val, optimal_h] = supOperatorL(obj, x, DV)
f = zeros(length(obj.weights), 1);
for i = 1:1:length(obj.weights)   % Weight for risky asset
  if obj.isControlAdmissible(i) == true
    h = obj.weights(i);

    f(i) = -0.5*(obj.theta + 1)*(h^2)*(obj.sigma^2) ...
           - obj.theta*(h*obj.sigma*obj.lambda*DV) ...
           + h*(obj.a_tilde + obj.A_tilde*x) ...
           - (1/obj.theta)*obj.integrals(i);
  else
    f(i) = nan;
  end
end

% Find maximum value
[val, ind] = max(f(:));
[row, column] = ind2sub(size(f), ind);
% Control policy that produces the maximum
optimal_h = obj.weights(row);
end

%FINDADMISSIBLECONTROLS Find what control policies are admissible
function [isControlAdmissible] = findAdmissibleControls(obj)
isControlAdmissible = zeros(length(obj.weights), 1);
for i = 1:1:length(obj.weights)   % Weight for risky asset
  if obj.checkControlIsAdmissible(obj.weights(i)) == true
    isControlAdmissible(i) = true;
  else
    isControlAdmissible(i) = false;
  end
end
end

%CHECKCONTROLISADMISSIBLE Check if input control policy is admissible
function [isAdmissible] = checkControlIsAdmissible(obj, h)
isAdmissible = true;
for i = 1:1:length(obj.gamma)
  % Condition for admissibility
  if obj.gamma(i)*h <= -1 || h > 1
    isAdmissible = false;
    break;
  end
end
end

%CALCULATEINTEGRALWITHRESPECTTOLEVYMEASURE Calculate integral with
% respect to Levy measure for every admissible control policy
function [integrals] = calculateIntegralsWithRespectToLevyMeasure(obj)
fprintf('Calculating integrals with respect to Levy measure');

integrals = zeros(length(obj.weights), 1);
for i = 1:1:length(obj.weights)   % Weight for risky asset
  if obj.isControlAdmissible(i) == true
    integrals(i) = obj.integralWithRespectToLevyMeasure(obj.weights(i)
      );
  else
    integrals(i) = nan;
```

```
  end
  fprintf('.');
end

fprintf('\n');
end

%INTEGRALWITHRESPECTTOLEVYMEASURE Approximate integral with respect to
% Levy measure
function val = integralWithRespectToLevyMeasure(obj, h)
val = 0;
dz = diff(obj.z(1:2));

for i = 2:1:length(obj.z) - 1
  % Evaluate the integrand and accumulate result
  val = val + obj.evaluateIntegrand(i, h);
end
val = dz*(obj.evaluateIntegrand(1, h)/2 + val + obj.evaluateIntegrand(
  length(obj.z), h)/2);
end

%EVALUATEINTEGRAND Evaluate integrand of integral with respect to
% Levy measure at one point of the domain
function val = evaluateIntegrand(obj, i, h)
aux1 = (power(1 + obj.gamma(i)*h, -1*obj.theta) - 1);
aux2 = 0;
if abs(obj.z(i)) <= obj.R
  aux2 = obj.theta*obj.gamma(i)*h;
end
val = (aux1 + aux2)*obj.gaussianDensity1D(obj.z(i));
end

%GAUSSIANDENSITY1D Calculate the value of the 1D Gaussian density
% function at any input point
function val = gaussianDensity1D(obj, z)
if z == 0
  val = 0; % Levy measure has no mass at origin
else
  val = obj.intensity*normpdf(z, obj.mean, obj.variance);
end
end

end % private methods
end % classdef
```

LISTING 4. File riskSensitiveHJB1factor.m

REFERENCES

[1] Appelbaum, D., Lévy processes and stochastic calculus, 1st Edition, Cambridge University Press, 2009

[2] Barles, G., "Convergence of numerical schemes for degenerate parabolic equations arising in finance", In Rogers, L. C. G. and Talay, D., editors, *Numerical methods in finance*, pages 1-21, Cambridge University Press, 1997

[3] Bellman, R. E., Dynamic Programming. Princeton University Press, 1st Edition, 1957

[4] Bielecki, T. R. and Pliska, S. R., "Risk sensitive asset management with transaction costs", *Finance and Stochastics*, 2000, **4**, 1-33

[5] Bielecki, T. R., Pliska, S. R. and Sheu, S. J., "Risk sensitive portfolio management Cox-Ingersoll-Ross interest rates: the HJB equation", *SIAM Journal of Control and Optimization*, 2005, 44, 1811-1843
[6] Bielecki, T. R., Jin, H., Pliska, S. R. and Zhou, X. Y., "Continuous-time mean-variance portfolio selection with bankruptcy prohibition", *Mathematical Finance*, 2005, **15**, 213-244
[7] Barles, G., Souganidis, P. E., "Convergence of approximation schemes for fully nonlinear equations", *Asymptotic Analysis*, 1991, **4**, 271-283
[8] Capiński, M., Kopp E. and Traple J., Stochastic calculus for finance, 1st Edition, Cambridge University Press, 2012
[9] Çinlar, E., Probability and stochastics, 1st Edition, Springer, 2011
[10] Crandall, M., Ishii, H., Lions and P.-L., "User's guide to viscosity solutions of second order partial differential equations", *Bulletin of the American Mathematical Society*, 1992, **27**, 1-67
[11] Cont, R. and Tankov, P., Financial modelling with jump processes, 1st Edition, Chapman & Hall/CRC Financial Mathematical Series, 2004
[12] Davis, M. H. A. and Lleo, S., "Jump-diffusion risk-sensitive asset management I: Diffusion factor model", *SIAM Journal on Financial Mathematics*, 2011, **2**, 22–54
[13] Davis, M. H. A. and Lleo, S., "Jump-diffusion risk-sensitive asset management II: Jump-diffusion factor model", *SIAM Journal of Control and Optimization*, 2011, **51**, 1441-1480
[14] Eberle, A., Stochastic analysis, 2012 [cited 2013 Jun 20]. Available from: http://www.uni-bonn.de/~eberle/StoAn1112/StochasticAnalysisFinal.pdf
[15] Elliott, R. J. and Kopp, P. E., Mathematics of financial markets, 2nd Edition, Springer, 2005
[16] Ikeda, N. and Watanabe, S., Stochastic differential equations and diffusion processes, 1st Edition, Kodansha Ltd, 1981
[17] Itô, K., "On a stochastic integral equation", *Proc. Imp. Acad. Tokyo* 1944, **20**, 519-524
[18] Kuroda, K. and Nagai, H., "Risk-sensitive portfolio optimization on infinite time horizon", *Stochastics and Stochastics Reports*, 2002, **73**, 309-331
[19] Kuo, H., Introduction to Stochastic Integration, 1st Edition, Springer, 2006
[20] Li, D. and Ng, W.-L., "Optimal dynamic portfolio selection: Multiperiod mean-variance formulation", *Mathematical Finance*, 2000, **10**, 387-406
[21] Markowitz, H., "Portfolio selection", *The Journal of Finance*, 1952, **7**, 77-91
[22] Merton, R. C., "Lifetime portfolio selection under uncertainty: the continuous-time case", *Review of Economics and Statistics*, 1969, **51**, 247-257
[23] Merton R. C., "Optimal consumption and portfolio rules in a continuous-time model", *Journal of Economic Theory*, 1971, **3**, 373-413
[24] Musiela, M. and Rutkowski, M., Martingale methods in financial modelling, 2nd Edition, Springer, 2005
[25] von Neumann, J. and Morgenstern, O., Theory of games and economics behavior, 3rd Edition, Princeton University Press, 1953
[26] Øksendal, B. and Sulem, A., Applied stochastic control of jump diffusions, 3rd Edition, Springer, 2009
[27] Pham, H., Continuous-time stochastic control and optimization with financial applications, 1st, Springer, 2009
[28] Protter, P., Stochastic integration and differential equations, 2nd edition, Springer, 2004
[29] Ross, K., Stochastic control in continuous time [cited 2013 Sep 24]. Available from: http://www.swarthmore.edu/NatSci/kross1/Stat220notes.pdf
[30] Sato, K., Lévy processes and infinitely divisible distributions, 1st Edition, Cambridge University Press, 1999
[31] Shreve, S. E., Stochastic calculus for finance II: Continuous-time models, 1st Edition, Springer, 2004
[32] Touzi, N., Stochastic control and finance, 2010, [cited 2013 October 4]. Available from: http://www.fields.utoronto.ca/programs/scientific/09-10/finance/courses/touzi1.pdf

[33] Whittle, P., Optimal control: Basics and beyond, 1st Edition, Wiley, 1996