Universität Ulm
Institut für Versicherungswissenschaften

Portfolio Selection, Delta Hedging and Robustness in Brownian and Jump-Diffusion Models

Dissertation

zur Erlangung des Doktorgrades Dr. rer. nat. der Fakultät für Mathematik und Wirtschaftswissenschaften der Universität Ulm
vorgelegt von

Frank Bosserhoff
aus

Oberhausen
im Jahr

2019
Amtierender Dekan: Professor Dr. Martin Müller
1. Gutachter: Professor Dr. Mitja Stadje
2. Gutachter: Professor Dr. Robert Stelzer
Tag der Promotion: 17. Januar 2020
Acknowledgments

First of all, I would like to thank my advisor Prof. Dr. Mitja Stadje for his constant support throughout the previous years. His continuous stream of ideas, his curiosity in research questions and his in-depth knowledge, from which I benefited a lot, as well as his patience make him an outstanding supervisor. He also taught me how to write a proper scientific paper and improved earlier versions of this thesis through highly fruitful comments. It was my pleasure and honor to work with him.

Likewise, I am grateful to Prof. Dr. Robert Stelzer who agreed to act as referee for this thesis. Various of his lectures I attended as a master’s student prepared me very well for this work.

The financial support from Stiftung WiMa Ulm by providing a scholarship for the first two years I have been working on this thesis is highly acknowledged by me.

I also want to thank all the people I met at several conferences and symposiums for sharing their research ideas and providing valuable suggestions.

Furthermore, I am greatly indebted to the head of the Institute of Insurance Science, Prof. Dr. An Chen, for the pleasant working environment. Likewise, I would like to thank my colleagues at the institute for the good time and much fun that we had, their continued helpful advice when it came to research questions and, of course, many games of table football.

I greatly appreciate all the support I received from my family and friends. It is a particular please for me to thank Ivy for proofreading large parts of this thesis. My special thanks go to my brother Volker for his continuous encouragement over the last years as well as to my dear friends Haoyuan, Nici, Patrick and Wenjing; they all were very important persons to me during this time.

Last but not at least I want to express my eternal gratitude to my parents for their endless and unconditional support - be it school, university, Ph.D. studies or anything else. This thesis would not have been possible without them. Thank you very much!

Frank Bosserhoff
This thesis covers miscellaneous topics in financial and insurance mathematics. The first two chapters deal with generalizations of mean-variance portfolio selection and the final one with Delta hedging and robustness issues.

In Chapter 2, we consider a time-consistent mean-variance portfolio selection problem of an insurer and allow for the incorporation of basis (mortality) risk. The optimal solution is identified with a Nash subgame perfect equilibrium. We characterize an optimal strategy as solution of a system of partial integro-differential equations (PIDEs), a so called extended Hamilton-Jacobi-Bellman (HJB) system. We prove that the equilibrium is necessarily a solution of the extended HJB system. Under certain conditions we obtain an explicit solution to the extended HJB system and provide the optimal trading strategies in closed-form. A simulation shows that the previously found strategies yield payoffs whose expectations and variances are robust regarding the distribution of jump sizes of the stock. The same phenomenon is observed when the variance is correctly estimated, but erroneously ascribed to the diffusion components solely. Further, we show that differences in the insurance horizon and the time to maturity of a longevity asset do not add to the variance of the terminal wealth.

In Chapter 3, we first note that investors are particularly concerned about downside risk, which is not properly reflected by classical mean-variance portfolio selection due to the symmetry of the second central moment. We therefore consider the following extension: Denoting by $V_T$ the value of some portfolio consisting of a risky and a riskless asset at some time $T > 0$, Itô’s representation theorem implies the existence of a predictable square-integrable process $H^V = (H^V_t)_{t \in [0,T]}$ such that

$$V_T = \mathbb{E}(V_T) + \int_0^T H^V_t \, dW_t,$$

i.e., the terminal time portfolio value can be represented by its expected value plus some stochastic integral w.r.t. a standard Brownian motion $W$. The
process $H^V$ reflects the local volatility of the portfolio’s terminal payoff. Let $y = (y_t)_{t \in [0,T]}$ denote the absolute amount of risky assets held. We solve the following mean-local-volatility optimization problem:

$$\max_y \mathbb{E} \left( V_T - \gamma \int_0^T g(t, H^V_t) \, dt \right),$$

where $g$ is some possibly asymmetric penalty function satisfying certain conditions and $\gamma > 0$ is a risk-aversion parameter. As this optimization problem is time-inconsistent, Bellman’s dynamic programming principle is not applicable. We approach this time-inconsistent stochastic optimal control problem in continuous time by solving its discrete time analogue first. Thereby the discrete time optimal policies are represented in closed-form for arbitrary many trading time points per unit time. Subsequently, we employ a convergence argument to describe the continuous time optimal investment policy.

In Chapter 4, we consider an investor aiming at Delta hedging a European contingent claim $h(S(T))$ in a jump-diffusion model, but incorrectly specifying the stock price’s volatility and jump sensitivity, so that any hedging strategy is calculated under a misspecified model. When does the erroneously computed strategy super-replicate the true claim in an appropriate sense? If the misspecified volatility and jump sensitivity dominate the true ones, we show that following the misspecified Delta strategy does super-replicate $h(S(T))$ in expectation among a wide collection of models. We also show that if a robust pricing operator with a whole class of models is used, the corresponding hedge is dominating the contingent claim under each model in expectation. Our results rely on proving stochastic flow properties of the jump-diffusion and the convexity of the value function. In the pure Poisson case, we establish that an overestimation of the jump sensitivity results in an almost sure one-sided hedge. Moreover, in general the misspecified price of the option dominates the true one if the volatility and the jump sensitivity are overestimated.
TABLE OF CONTENTS

1 Introduction
   1.1 Motivation ........................................... 1
   1.2 Outline of the Thesis ............................... 7

2 Mean-Variance Hedging of Unit Linked Life Insurance Contracts in a Jump-Diffusion Model 11
   2.1 Introduction ........................................... 11
   2.2 Model Setup .......................................... 15
   2.3 Optimization Problem ................................. 19
   2.4 Sufficiency and Necessity ............................ 22
   2.5 Explicit Solution ...................................... 32
   2.6 Numerical Results .................................... 42
   2.A Appendix ............................................. 51

3 Extensions of Mean-Variance Portfolio Selection 55
   3.1 Introduction .......................................... 55
   3.2 Model Setup .......................................... 58
   3.3 Discrete Time Problem ............................... 62
   3.4 Convergence ......................................... 82
   3.5 Conclusion .......................................... 87
   3.A Appendix ............................................. 87
Chapter 1

Introduction

1.1 Motivation

How to optimally allocate money among a basket of assets? This question is faced by various private and institutional investors such as the asset side of insurance companies, pension as well as endowment funds and commercial banks. However, the interpretation of the term *optimal* is highly subjective as it depends on the investors’ preferences and the goals they seek to achieve. Moreover, the model an investor uses to determine her investment strategy needs to be chosen such that the developments at the stock market are mimicked appropriately. This thesis sheds some new light on widely known portfolio allocation criteria as well as model uncertainty combined with robustness properties in Brownian as well as general jump-diffusion settings.

Investors typically want to achieve a high return on their capital while keeping control of the risk induced. A fundamental problem thereby is the identification of the risk of an investment. Intuitively, risk is the deviation of some random outcome from its expected value. However, the quantification of this intuitive notion is by far not unique. Basically every researcher and practitioner being concerned with portfolio optimization problems has studied the contributions of Markowitz, known as the founder of modern portfolio theory, who assumed in his pioneering work (Markowitz (1952))
that the preferences of an agent are completely described by the expectation and the variance of her final income, i.e., the risk of an investment portfolio is entirely quantified by the variance of the terminal payoff. A portfolio allocation rule with respect to (w.r.t.) a mean-variance criterion is then called optimal if the variance of the final income is minimized for a given expected value. Markowitz has solved this problem in a single-period model and many distinguished researchers have built up on his ideas and extended the problem in various ways, see for instance [Hakansson (1971), Grauer and Hakansson (1993), Schweizer (1995), Li and Ng (2000), Li and Zhou (2000) and Lim (2005)]. The identification of risk with the variance of the terminal wealth leads to an analytically well tractable problem. Furthermore, as any gain in the expected terminal payoff can be measured in terms of an increase of the variance, the tradeoff between risk and return is explicit. Thus, it is an intuitively appealing and easily comprehensible portfolio allocation concept. For these reasons, mean-variance portfolio selection is extensively used by practitioners. However, there are at least two issues coming along with this approach. First of all, due to the non-recursiveness and the quadratic non-linearity of the variance part, the mean-variance portfolio allocation problem is time-inconsistent (cf. for instance the explanations in Li and Zhou (2000) or Björk and Murgoci (2010); Karnam et al. (2017) provide a short enlightening example on time-inconsistency of the mean-variance problem). Informally speaking, an optimization problem is time-inconsistent if a strategy that is optimal today may no longer be optimal later on. If an investor wants to circumvent this problem, she could for instance resort to maximize the expected utility function of some terminal wealth. Thereby a utility function is chosen such that it captures the risk attitude of an investor, the most common forms include an exponential, power, log or quadratic function. Utility maximization is a major stream in the literature, but not the main scope of this thesis, which is why we cut short by recommending Föllmer and Schied (2011) and references therein for a detailed discussion on the preferences of representative investors. Although time-inconsistency issues, when searching for a strategy maximizing the expected utility of the terminal wealth, can be circumvented, a drawback of the utility maximization approach is that
eliciting particular utility functions from investors is rather difficult. Moreover, compared to the mean-variance criterion, the tradeoff between risk and return is implicit, making this criterion less intuitive. In Chapter 2 and in Chapter 3 of this thesis we consider approaches combining advantages of both mean-variance optimization as well as expected utility maximization.

Presuming a jump-diffusion model, in Chapter 2 we reformulate the mean-variance problem in game-theoretic terms. The investment procedure is thereby modeled as a game where at each point in time an investor is the decision maker, that is, there are infinitely many players. Alternatively, one could think of the players as future incarnations of an investor or her preferences, respectively. In this setting a strategy is called optimal if it is a Nash equilibrium in each subgame (see Myerson (2013) for a general introduction to game theory; a philosophical view is found in Ross (2019)). This induces a dynamically optimal trading strategy that can be obtained as the solution to a system of partial integro-differential equations (PIDEs). In addition to enforcing time-consistency, the intuitive notion of the mean-variance criterion is preserved. In this setting we provide theoretical results and numerical illustrations of the optimal policies. The reformulation of time-inconsistent stochastic optimal control problems in game-theoretic terms goes back to Ekeland and Pirvu (2008) and Björk and Murgoci (2010). For a further discussion of literature related to this topic see the introduction to Chapter 2 below.

Another drawback going along with the mean-variance criterion is that it usually does not properly reflect investors’ concerns. Investors are mainly interested in the downside risk of an investment while unexpected gains are appreciated. The variance, however, is a symmetric criterion penalizing gains and losses equally. This gives rise to the usage of asymmetric risk measures such as the Value at Risk or the Tail-Risk (McNeil et al. (2005)). Following the illustrations in Chapter 9 of Cochrane (2005), investors tend to prefer those portfolios reacting inversely to good and bad news about stock market returns. The reason for this preference pattern is that the marginal utility of income and consumption is usually higher in recessions than in economically prospering times. The Value at Risk and the Tail-Risk, however, are law-
invariant: they only depend on the distribution of the final payoff, but do not take into account the correlation with the underlying financial market. Thus, when an investor wants to shape a portfolio reacting positively to bad news about stocks’ returns, those well known risk measures cannot be used just like that. In Chapter 3 below we present an alternative approach by solving a portfolio allocation problem with a risk measure taking the correlation with the general economic environment into account. Its existence is justified by the predictable representation property of certain random variables. To be more specific, we call this risk measure local volatility because it refers to the local variation of the final payoff w.r.t. the movements of the financial market. Hence, in the target functional the variance of the terminal income is substituted by the accumulated penalized local volatility. What is striking about this approach? For instance, the usage of the volatility parameter to measure risk is widely known: in the highly celebrated and extensively employed Black-Scholes model (Black and Scholes (1973)) the volatility parameter equals the standard deviation of the log-stock price per unit time and is used as such to quantify risk. Instead, the novelty in our approach is that by solving the mean-local-volatility problem for a relatively wide class of penalty functions, we are able to use well-known risk measures asymmetrically. Hence, this criterion is easily comprehensible as well and the tradeoff between risk and return is as explicit as before while investors’ preferences are simultaneously better reflected. We remark that this target accounts for the higher marginal utility in recessions without the need of knowing an investor’s explicit utility function. Note that the mean-local-volatility problem is a time-inconsistent optimal control problem. Thus, borrowing the language of Strotz (1955), an investor following the optimal strategy is said to pre-commit. Hence, while Chapter 2 tackles the time-inconsistency of the mean-variance problem through a reformulation in game-theoretic terms, Chapter 3 is complementary in the sense that an alternative to the symmetric mean-variance criterion is suggested while the time-inconsistency is accepted.

The above discussion illustrates among others that there is actually not the one perfect portfolio optimization criterion, instead each approach has its respective advantages and disadvantages. Once an optimization target
1.1. MOTIVATION

has been decided on, the investor needs to determine her respective strategy. In order for this strategy to lead as close to the outcome desired as possible, the computation should be based on a model that mirrors the developments at the real markets adequately. Actually, the choice of the market model is by no means less important than the optimization criterion itself. Chapter 4 of this thesis deals with misspecification and robustness issues, thereby complementing the topic of optimal allocation addressed in the first two parts.

Modeling stock prices based on the Brownian motion has been common practice for a long time and dates back to [Bachelier (1900)]. In this work, the stock price is modeled as arithmetic Brownian motion, thus, it allows for negative values. This problem is circumvented in the Black-Scholes model [Black and Scholes (1973)], where the stock price evolves as geometric Brownian motion with time-independent coefficients. An extension allowing for stochastic volatility is presented in [Heston (1993)]. A reason for modeling stock prices as a function of the Brownian motion is the availability of closed-form solutions to option prices, see [Schachermayer and Teichmann (2008)] for comparisons of prices among the aforementioned models. Another good reason is the general analytical and numerical tractability of Brownian settings. However, such models induce almost surely (a.s.) continuous paths of the asset price processes. There are situations where stock prices exhibit an abrupt stepwise change as illustrated in the following non-exhaustive list of (recent) examples:

- Brexit referendum: After the announcement of the turnout of the Brexit referendum in June 2016, the FTSE 100 decreased by almost 9 %, the Euro Stoxx 600 lost 7 % on one day while the Nikkei fell 7.9 % (see [Mackenzie and Platt (2016)] for further examples).

- SAP quarterly report: After the second quarter of 2019, the software vendor SAP announced to have suffered from pending trade conflicts and further problems. The previously well performing share price suddenly dropped by roughly 10 % and the trade shut down for a short period of time ([Ries (2019)]).
CHAPTER 1. INTRODUCTION

- Wirecard vs. Financial Times (FT): After arguing with the company Wirecard for several months, the British newspaper FT has published an article on October 15th, 2019, alleging that Wirecard has inflated sales and profits at businesses in Dubai and Ireland. As a consequence, the stock price decreased all of a sudden by about 23 % (Peitsmeier (2019)).

- Tesla quarterly report: After releasing highly positive news about the business in the third quarter of 2019 on October 23rd, 2019, the stock price of the electric car manufacturer Tesla increased immediately by roughly 20 % (Leisinger (2019)).

From these examples it becomes clear that stock prices exhibit significant, abrupt changes that can be either positive or negative, see Cont and Tankov (2012) and the references therein for further details. Moreover, these changes are usually not foreseeable, i.e., they occur at random times. Needless to say, models based on the Brownian motion do not capture these facts. Indeed, Cont and Voltchkova (2005) explain that any reasonable stock price model should be based on either a Brownian component and a finite activity jump process or, when dropping the diffusion part, on an infinite activity pure jump process (see also Madan (2001) and German (2002)). However, upon choosing one of these models, can the investor then be sure to achieve the desired hedging objective? The answer to this question is certainly not a genuine ‘yes’, because among others the model parameters need to be specified. If they are unrealistic, the hedging objective might not be attainable. Chapter 4 of this thesis treats the problem of parameter misspecification in general jump-diffusion models. To be more precise, we realistically presume the stock price develops as a jump-diffusion process, but the investor erroneously specifies the corresponding volatility and jump sensitivity. As jump-diffusion models induce an incomplete market anyway, what is the impact of an additional misspecification on the performance of certain hedging strategies? Since Delta hedging is known to yield a perfect replication in complete markets, many practitioners apply it against better knowledge in incomplete markets as well. For that reason, we study their robustness properties and
establish among others results where Delta strategies yield a one-sided hedge in expectation in a wide class of models.

1.2 Outline of the Thesis

As each of the subsequent chapters in this thesis is based on a different paper, they are essentially self-contained. Each chapter provides an introduction illustrating the specific topic treated and a classification into the scientific literature. This is followed by a remark on notation, which is only unified within each chapter. In the following we briefly summarize the contents of each paper.

Chapter 2 is based on Bosserhoff and Stadje (2019a). In this paper the investor is assumed to be an insurance company solving a time-consistent mean-variance portfolio selection problem. Additionally, we consider the incorporation of basis (mortality) risk. The insurer thereby models the dynamics of the stock prices as well as the force of mortality by jump-diffusions. A time-consistent version of the mean-variance problem is achieved by rephrasing the optimization problem in game-theoretic terms and identifying the optimal solution with a Nash subgame perfect equilibrium. The optimal strategy is actually given as the solution of a system of PIDEs, which we call extended Hamilton-Jacobi-Bellman (HJB) system. Our first main result is a theorem stating that the equilibrium necessarily solves the extended HJB system. We also explain that solving the extended HJB system is sufficient. Subsequently, under mild assumptions, we provide explicit solutions to the extended HJB system. In particular, the optimal investment policies are available in closed-form. The chapter closes with an extensive simulation investigating certain properties of the previously found strategies. For example, we numerically illustrate that the optimal trading strategies lead to terminal payoffs that are robust w.r.t. the distribution of the jump sizes of the stock. A similar robustness behavior is observed when the variance is precisely estimated, but erroneously thought to stem from the Brownian components solely. Moreover, a result being of high practical relevance is
that potential differences in the insurance horizon and the time to maturity of a longevity asset do not increase the variance of the final payout.

Chapter 3 is based on a working paper of the author and the first reviewer of this thesis. As previously noted, classical mean-variance portfolio selection does not adequately mimic investors’ concerns due to the symmetry of the second central moment. For that reason we suggest an extension inspired by Itô’s representation theorem. Writing \( V_T \) for the terminal payoff of some portfolio process, this random variable can be represented by its expected value plus some Itô integral w.r.t. the driving Brownian motion \( W \):

\[
V_T = \mathbb{E}[V_T] + \int_0^T H_t^V \, dW_t.
\]

The process \( H = (H_t^V)_{t \in [0,T]} \) reflects the local volatility of the portfolio’s final wealth per unit time. In particular, positive values of the local volatility indicate a positive correlation of the terminal payoff with the underlying financial market and negative values show an inverse relation. In the optimization criterion, we substitute the variance by the accumulated local volatility and by allowing for a large class of asymmetric penalty functions, this approach reflects investors’ concerns generally better than the classical mean-variance one. As the mean-local-volatility problem is time-inconsistent, Bellman’s dynamic programming principle is not applicable. Thus, we approach this time-inconsistent control problem in continuous time by solving its discrete time counterpart in a first step because in a finite filtration setting we are able to make more explicit calculations and optimize over the whole space. In a second step, we characterize the continuous time optimal policy as the limit of the discrete time one in a suitable sense.

Chapter 4 is based on Bosserhoff and Stadje (2019b). In this paper we consider an investor modeling the price of a stock traded at the stock exchange by a jump-diffusion, but incorrectly specifying the stock price’s volatility and jump sensitivity. If the investor’s goal is a Delta hedge of some European contingent claim \( h(S(T)) \), the hedging strategy is consequently
determined under a misspecified model. A natural question to ask is: under which conditions does the wrongly calculated strategy super-replicate the true claim suitably? Our first main result shows that in case the volatility and the jump sensitivity dominate the true ones, following the misspecified Delta hedging strategy yields a super-replication of $h(S(T))$ in expectation in a large collection of models. The second main theorem establishes that if a robust pricing operator with a whole class of models is applied, the calculated hedge dominates the physical contingent claim under each model in expectation. We provide a large list of practically highly relevant examples for the latter result. The fact that the domination holds in expectation and not almost surely obviously leaves some risk to the investor. On the other hand, superhedges are extremely expensive and therefore typically not used in practice. Mathematically, our proofs are essentially based on using stochastic flow properties in the jump-diffusion case. A pure Poisson market is known to be complete, i.e., every contingent claim can be perfectly replicated by following its Delta strategy. In this case we show that an overestimation of the jump sensitivity yields an almost sure one-sided hedge. We also argue that the misspecified price of the option dominates the true one in case the volatility and the jump sensitivity are systematically overestimated.
Chapter 2

Mean-Variance Hedging of Unit Linked Life Insurance Contracts in a Jump-Diffusion Model

2.1 Introduction

This chapter is based on Bosserhoff and Stadje (2019a). Two major risks faced by life insurance companies are longevity and asset risks. Longevity risk refers to the risk that the future changes in the mortality rates are incorrectly estimated while asset risk refers to the possibility of a future loss in the investment portfolio. Increases in the life expectancy might among others stem from sudden changes in environmental or medical conditions that are not foreseeable upon the contract initiation. Clearly, these changes need to be accounted for by the mortality model. An adequate way to do so is the modeling of the force of mortality with a diffusion process supplemented by jumps, see Luciano and Vigna (2008) and Cairns et al. (2008) for a detailed discussion. Another practical challenge of hedging longevity risk is basis risk. The payoff of for instance a longevity bond depends on a particular mortality rate that is related to but certainly not identical with the insurer’s portfolio,
see Coughlan et al. (2011) and Li and Hardy (2011) for empirical studies on this issue. Thus, buying a longevity asset can only provide a partial hedge against an insurance company’s mortality exposure.

Asset risk stems from the fact that the premiums paid by the insured are to be gainfully invested at the capital market inducing financial risk. Empirical evidence suggests that returns are non-gaussian and leptokurtic, see e.g. Cont and Tankov (2012), Schoutens (2003) and references therein. Thus, in order to properly capture this risk, an insurer should base a stock price model on a Brownian component and additionally allow for jumps. Such a financial market is known to be generically incomplete. Consequently, an insurance company facing the aforementioned risks cannot perfectly hedge its obligations. Hence, a way to quantify risk needs to be specified. In this paper we identify risk with the variance of the terminal wealth. The identification of risk with the variance of the terminal payoff has a long tradition in academia as well as industry and dates back already to Markowitz (1952). Mean-variance portfolio selection is intuitively appealing and analytically tractable. A major drawback, however, is time-inconsistency, which means that due to the non-linearity and non-recursiveness of the variance part, the dynamic programming approach fails. An investor might initiate a dynamic strategy because it is optimal at a particular point in time, knowing fully well that she will deviate from this strategy later on. Investors ignoring the sub-optimality of a previously found strategy are said to pre-commit, see Strotz (1955) for fundamentals on this problem. In Li and Zhou (2000) and Lim and Zhou (2002), the pre-commitment version of a mean-variance portfolio selection problem in a continuous-time economy is solved. The question of dynamically optimal, i.e., time-consistent mean-variance policies has been addressed for example by Basak and Chabakauri (2010). Their solution approach is based on a recursive formulation allowing for the application of dynamic programming. The authors point out that the same solution could be found as the Nash subgame perfect equilibrium outcome whereby the investor is playing a game with a future incarnation of herself, that is, a game with infinitely many players. Thereby a strategy is a Nash subgame perfect equilibrium if at some given point in time an investor knows that
every future "player" will follow a certain strategy, then it is optimal for her not to deviate. For a general game-theoretic background we refer to Peters (2015) and references therein. The reformulation of time-inconsistent control problems in game-theoretic terms has been originally proposed by Ekeland and Pirvu (2008) and Björk and Murgoci (2010). This line of research has been followed by Basak and Chabakauri (2010), Wang and Forsyth (2011), Czichowsky (2013), Bensoussan et al. (2014) and Lindensjö (2016).

It is well known that the optimal value function of a standard time-consistent stochastic control problem can be characterized as the unique solution of a non-linear partial integro-differential equation (PIDE), see Øksendal and Sulem (2005), known as Hamilton-Jacobi-Bellman (HJB) equation. In Björk and Murgoci (2010) it is shown that the reformulation of a time-inconsistent control problem in game-theoretic terms leads to a system of nonlinear PIDEs, the so called extended HJB system. Further, they provide a verification result showing that solving the extended HJB system is a sufficient condition for being an equilibrium control law. Lindensjö (2016) proves that under certain regularity assumptions solving the extended HJB system is a necessary condition for being an equilibrium; however, the proof is restricted to the diffusion case.

We consider a continuous-time Markovian economy in which an insurer trades in an arbitrary quantity of risky financial assets (stocks), a zero-coupon longevity bond and a riskless asset in order to hedge some terminal payout with regard to mean-variance optimality. Thereby the underlying financial assets as well as the force of mortality are modeled by jump-diffusions. Our first main contribution is an extension of the work of Lindensjö (2016) by proving that an equilibrium necessarily solves the extended HJB system. Secondly, for the case that an insurer neglects the hedge of some terminal payoff, we are able to present explicit closed-form solutions for the optimal trading strategies, the equilibrium value function and the expected terminal wealth. Thirdly, we exemplify our findings along a tractable model and provide numerical as well as graphical illustrations. When the jumps are erroneously neglected while the expected values and variances of the stock price and the force of mortality are correctly determined, the expected optimal
terminal payoff and its variance hardly change. Thus, the time-consistent mean-variance optimal terminal wealth is robust regarding the consideration of jumps. A similar result is found when changing the distribution of the jump sizes of the stock. As the market for longevity assets is relatively illiquid, one can certainly not expect to find a hedging instrument whose time to maturity coincides with the insurance horizon. We find that this does not affect the expectation and variance of the optimal final payoff in our setup. Moreover, we find that our strategies significantly outperform the gain from investing in the riskless asset only.

The remainder of this paper is structured as follows: In Section 2 we introduce the financial and longevity markets under consideration and clarify what is meant by an admissible trading strategy. In the following section we turn to the optimization problem by gradually defining several auxiliary functions, operators and the notion of equilibrium control. With these in hand, Section 3 closes with the specification of the extended HJB system. In Section 4 our first main result (Theorem 2.4.4) is stated and proved. In Section 5 we neglect the hedge of a terminal payout and present an explicit equilibrium solution, the main result here is Theorem 2.5.1. The final section contains the numerical applications.

**Notation.** Denote by \( \mathbb{R}_+ \) the positive real numbers and by \( \mathbb{R}_+^m \) its \( m \)-fold Cartesian product. The zero vector in any Euclidean space \( \mathbb{R}^m \) is written as \( 0 \). For \( x \in \mathbb{R}^m \) we use \( \|x\|_2 := \sqrt{\sum_{i=1}^m |x_i|^2} \). The symbol \( \mathbb{R}^{m \times n} \) denotes the space of real-valued matrices with \( m \) rows and \( n \) columns. If \( x \in \mathbb{R}^m \), the matrix \( \text{Diag}(x) \in \mathbb{R}^{m \times m} \) is the square matrix with the entries of \( x \) on the diagonal and all off-diagonal elements being equal to zero. For \( A \in \mathbb{R}^{m \times n} \), the symbol \( A^\top \) means the transpose of \( A \). If \( A \) is a square matrix, we write \( \text{Tr}(A) \) for its trace. For any \( T > 0 \), \( t \in (0, T) \) and some function \( f : [0, T] \times \mathbb{R}^m \to \mathbb{R} \) such that \( f \in C^{1,2} \), we define \( \dot{f}(t, \cdot) := \frac{\partial}{\partial t} f(t, \cdot) \) and for any \( x \in \mathbb{R}^m \) we denote for arbitrary \( j \in \{1, \ldots, m\} \) by \( f_{x_j}(\cdot, x) \) the first order partial derivative of \( f \) w.r.t. the \( j \)th component of \( x \). Moreover, let the gradient of \( f \) w.r.t. \( x \) be denoted by \( \nabla_x f(\cdot, x) \) and the Hessian matrix by \( H_x f(\cdot, x) \).
2.2 Model Setup

Let $T > 0$ be the planning horizon. Consider the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ that is equipped with a standard $d + 1$-dimensional Brownian motion $\hat{W} := (W^1, \ldots, W^d, W^{d+1})^\top$, whereby we define $W := (W^1, \ldots, W^d)^\top$ and $\hat{W} := W^{d+1}$, and a Poisson random measure $J_{\hat{X}}(dt, d\hat{x})$ on $[0, T] \times \mathbb{R}^{k+1} \setminus \{0\}$, independent of $\hat{W}$, with respective intensity measure $\vartheta_{\hat{X}}(dx) dt$. Denote its compensated version by $\tilde{J}_{\hat{X}}(dt, d\hat{x}) = J_{\hat{X}}(dt, d\hat{x}) - \vartheta_{\hat{X}}(d\hat{x})dt$. Let $(\mathcal{F}_t)_{t \in [0, T]}$ be the right-continuous completion of the filtration generated by $\hat{W}$ and $J_{\hat{X}}$.

Throughout this paper we impose the following condition:

**Assumption 2.2.1.** The Lévy measure $\vartheta_{\hat{X}}$ is such that

\[
\int_{\mathbb{R}^{k+1} \setminus \{0\}} |\hat{x}|^2 \vartheta_{\hat{X}}(d\hat{x}) < \infty. \tag{2.2.1}
\]

Under (2.2.1), let $\hat{X} := (X^1, \ldots, X^k, X^{k+1})^\top$ be a vector of pure-jump independent $(\mathcal{F}_t)$-martingales with $X^j_t = \int_0^t \int_{\mathbb{R}^{k+1} \setminus \{0\}} \hat{x}^j \tilde{J}_{\hat{X}}(ds, d\hat{x})$, $j = 1, \ldots, k, k + 1$, whereby $\hat{x}^j$ is the $j$th coordinate of $\hat{x} \in \mathbb{R}^{k+1}$. We define $X := (X^1, \ldots, X^k)^\top$ and $\tilde{X} := X^{k+1}$, so $\hat{X} = (X, \tilde{X})^\top$. Further, we may write $J_{\hat{X}}(dt, d\hat{x}) = J_{X, \tilde{X}}(dt, dx, d\tilde{x}) = 1_{x=0} J_X(dt, dx) + 1_{\tilde{x}=0} J_{\tilde{X}}(dt, d\tilde{x})$. For any $E \subseteq \mathbb{R}^{k+1}$, the independence of $X$ and $\tilde{X}$ implies that (cf. Cont and Tankov (2012), Proposition 5.3)

\[
\vartheta_{X, \tilde{X}}(E) = \vartheta_X(E_X) + \vartheta_{\tilde{X}}(E_{\tilde{X}}),
\]

with

\[
E_X := \{x \in \mathbb{R}^k : (x, 0) \in E\}, \\
E_{\tilde{X}} := \{\tilde{x} \in \mathbb{R} : (0, \tilde{x}) \in E\}.
\]

Assume that $x^j > -1$ for all $j \in \{1, \ldots, k\}$. The financial market under consideration consists of a bank account paying interest at a deterministic rate $r \geq 0$ and $m$ risky stocks, $1 \leq m \leq \min\{d, k\}$, with price processes $S^i = (S^i_t)_{t \in [0, T]}$, $i = 1, \ldots, m$, satisfying the SDEs given by
\[ \frac{dS_i^t}{S_i^t} = \mu_i \, dt + \sum_{j=1}^{d} \sigma_{ij} \, dW_i^j + \sum_{j=1}^{k} \rho_{ij} \, dX_i^j, \]  

(2.2.2)

where \( S_i^0 = s_i \in \mathbb{R}_+ \), \( \mu_i \in \mathbb{R} \), \( \sigma_{ij} \in \mathbb{R}_+ \) and \( \rho_{ij} \in \mathbb{R}_+ \) such that \( \sum_{j=1}^{k} \rho_{ij} \leq 1 \) respectively denote the initial price of stock \( i \), the rate of appreciation, the volatilities and the jump-sensitivities. We assume that the financial market is free of arbitrage, i.e., there exists a measure \( Q \) that is equivalent to \( \mathbb{P} \) such that the discounted stock price processes \( \left( \frac{S_i^t}{e^{rt}} \right)_{t \in [0,T]} \), \( i = 1, \ldots, m \), are \( \mathcal{F}_t \)- martingales under \( Q \).

In addition to the financial market, we consider an arbitrage-free mortality market on which an investor can buy a zero-coupon longevity bond. We use the Brownian motion \( \bar{W} \) and the jump component \( \bar{X} \) to model the force of mortality. In particular, the force of mortality \( \lambda \) shall be given as the solution of the SDE

\[ d\lambda_t = \mu_{\lambda}(t, \lambda_t) \, dt + \sigma_{\lambda}(t, \lambda_t) \, d\bar{W}_t + \int_{\mathbb{R} \setminus \{0\}} \tilde{\sigma}_{\lambda}(t, \lambda_t, \bar{x}) \, \tilde{J}(dt, d\bar{x}), \]  

(2.2.3)

whereby \( \lambda_0 \in \mathbb{R}_+ \), and \( \mu_{\lambda}, \sigma_{\lambda} \) and \( \tilde{\sigma}_{\lambda} \) satisfy Assumption \textbf{2.2.2} below. We remark that the force of mortality can become negative with positive probability. In practical applications it is therefore common to chose \( \mu_{\lambda} \) high and \( \sigma_{\lambda} \) as well as \( \tilde{\sigma}_{\lambda} \) small enough, see Luciano and Vigna (2008) for a discussion on calibration. We take \( \lambda_t > 0 \) for all \( t \in [0,T] \).

\textbf{Assumption 2.2.2.} We assume that \( \mu_{\lambda}, \sigma_{\lambda} : [0,T] \times \mathbb{R}_+ \to \mathbb{R} \) and \( \tilde{\sigma}_{\lambda} : [0,T] \times \mathbb{R}_+ \times \mathbb{R} \setminus \{0\} \to \mathbb{R} \) satisfy the following conditions:

(i) \textbf{(At most linear growth)} There exists a constant \( B_1 < \infty \) such that for all \( a \in \mathbb{R}_+ \) it holds that

\[ |\mu_{\lambda}(t,a)|^2 + |\sigma_{\lambda}(t,a)|^2 + |\tilde{\sigma}_{\lambda}(t,a,\bar{x})|^2 \leq B_1(1 + |a|^2). \]

(ii) \textbf{(Uniform Lipschitz continuity)} There exists a constant \( C_1 < \infty \) such
that for all \( a, b \in \mathbb{R}_+ \) it holds that

\[
|\mu_\lambda(t, a) - \mu_\lambda(t, b)|^2 + |\sigma_\lambda(t, a) - \sigma_\lambda(t, b)|^2
+ \int_{\mathbb{R}\setminus\{0\}} |\tilde{\sigma}_\lambda(t, a, \bar{x}) - \tilde{\sigma}_\lambda(t, b, \bar{x})|^2 \vartheta_X(d\bar{x}) \leq C_1|b - a|^2.
\]

We consider a longevity bond where the reference cohort is assumed to satisfy the following:

- at time \( t = 0 \), all members of the cohort are of the same age,
- the force of mortality of the cohort is entirely described by \( \lambda \),
- the cohort is sufficiently large such that the idiosyncratic risk is pooled away.

In addition, we assume that the insurance’s planning horizon \( T \) and the time to maturity of the longevity bond coincide. An investor who has bought the zero-coupon longevity bond at time \( 0 \leq t_1 \leq T \) paying \( L_\lambda(t_1, T) \) receives \( \exp \left( -\int_{t_1}^{T} \lambda_s \, ds \right) \) at time \( T \). Consequently, the price \( L_\lambda(t_1, T) \) is given by

\[
L_\lambda(t_1, T) = \mathbb{E}_Q \left[ e^{-\int_{t_1}^{T} (\lambda_s + r) \, ds} \bigg| \mathcal{F}_{t_1} \right]. \tag{2.2.4}
\]

Let \( 0 \leq t_1 < t_2 \leq T \), suppose investor \( A \) has bought the longevity bond at time \( t_1 \) at price \( L_\lambda(t_1, T) \) and there is a second investor, say \( B \), who has bought the bond at time point \( t_2 \) paying \( L_\lambda(t_2, T) \). As the final payoff depends on the length of the holding period, it is clear that investor \( A \) would not have sold her bond to \( B \) at price \( L_\lambda(t_2, T) \) at time \( t_2 \), but she would have demanded a price of \( \exp \left( -\int_{t_1}^{t_2} \lambda_s \, ds \right) L_\lambda(t_2, T) \). Therefore, if an investor has bought the longevity asset at time \( t_1 \), the dollar value of her investment at any time \( t_2 > t_1 \) is given by \( Y_{t_2} := \exp \left( -\int_{t_1}^{t_2} \lambda_s \, ds \right) L_\lambda(t_2, T) \). We name the dollar value process \( Y \) from now on; the discounted version of \( Y \) should be a martingale under the same risk-neutral measure \( Q \).
CHAPTER 2. MV HEDGING OF UNIT LINKED LIFE INSURANCE

Assumption 2.2.3. We assume that \((\lambda, Y)\) is a Markovian Itô jump-diffusion satisfying

\[
\frac{dY_t}{Y_t} = (r + \nu_L(t, \lambda_t, Y_t)) \, dt + \sigma_L(t, \lambda_t, Y_t) \, d\bar{W}_t + \int_{\mathbb{R} \setminus \{0\}} \eta_L(t, \lambda_t, Y_t, \bar{x}) \, \tilde{J}_X(dt, d\bar{x}),
\]

with \(Y_0 = L_\lambda(0, T)\) and deterministic functions \(\nu_L, \sigma_L, \eta_L\).

Note that the function \(\nu_L\) in (2.2.5) is the market price of longevity risk.

We further need the following assumption:

Assumption 2.2.4. We assume that \(\nu_L, \sigma_L : [0, T] \times \mathbb{R}_+ \times \mathbb{R}_+ \to \mathbb{R}\) and \(\eta_L : [0, T] \times \mathbb{R}_+ \times \mathbb{R}_+ \times \mathbb{R} \setminus \{0\} \to \mathbb{R}\) satisfy the following conditions:

(i) (At most linear growth) There exists a constant \(B_2 < \infty\) such that for all \(a, b \in \mathbb{R}_+\) it holds that

\[
|\nu_L(t, a, b)|^2 + |\sigma_L(t, a, b)|^2 + \int_{\mathbb{R} \setminus \{0\}} |\eta_L(t, a, b, \bar{x})|^2 \vartheta_X(d\bar{x}) \leq B_2(1 + |a|^2 + |b|^2).
\]

(ii) (Uniform Lipschitz continuity) There exists a constant \(C_2 < \infty\) such that for all \(a_1, a_2, b_1, b_2 \in \mathbb{R}_+\) it holds that

\[
|b_1 \nu_L(t, a_1, b_1) - b_2 \nu_L(t, a_2, b_2)| + |b_1 \sigma_L(t, a_1, b_1) - b_2 \sigma_L(t, a_2, b_2)| + \int_{\mathbb{R} \setminus \{0\}} |b_1 \eta_L(t, a_1, b_1, \bar{x}) - b_2 \eta_L(t, a_2, b_2, \bar{x})|^2 \vartheta_X(d\bar{x}) \leq C_2 \|(a_1, b_1) - (a_2, b_2)\|_2.
\]

We now consider an insurer who can invest in the \(m\) risky stocks, deposit money in the bank account and use the longevity asset to partially hedge against its mortality exposure. Let \(U \subseteq \mathbb{R}^{m+1}\). An allocation rule is a predictable function \(u : [0, T] \to U, \ t \mapsto (u_S(t), u_Y(t))^T\), whereby \(u_S = (u_{S1}, \ldots, u_{Sm})^T\) denotes the dynamic allocation process that indicates the total wealth that is invested in the stocks \(1, \ldots, m\), and \(u_Y\) the total wealth invested in the longevity asset. The portfolio process of the insurance
company using the allocation rule \( u \) is denoted by \( P^u = (P^u_t)_{t \in [0, T]} \) and fulfills the SDE

\[
dP^u_t = u^T_S(t) \frac{dS_t}{S_t} + u^T_Y(t) \frac{dY_t}{Y_t} + (P^u_t - u^T_S(t) \mathbf{1} - u^T_Y(t)) r \, dt,
\]

with initial wealth \( P^u_0 = p > 0 \) and \( \mathbf{1} \in \mathbb{R}^m \) denotes a column vector of ones. Observe that \( P^u \) as defined in (2.2.6) is self-financing.

**Definition 2.2.5.** An allocation rule \( u \) is admissible if for any point \( (t, p) \in [0, T) \times \mathbb{R}_+ \) there exists a unique càdlàg adapted solution \( P^u \) to (2.2.6) such that \( \mathbb{E}[|P^u_t|^2] < \infty \) for all \( t \). We denote by \( \mathcal{U} \) the set of admissible allocation rules.

### 2.3 Optimization Problem

Classical mean-variance portfolio selection aims at finding a strategy that simultaneously maximizes the expected terminal payoff of a portfolio while minimizing its variance. We first consider the more general case where an insurance company trades in the financial and longevity markets in order to hedge a terminal condition. Before rigorously defining what is meant by an equilibrium control in a stochastic optimization problem, we need some more notation and a target functional. Let \( Z := (S^1, \ldots, S^m, \lambda, Y) \in \mathbb{R}^{m+2}_+ \) be the vector containing the traded assets as well as the force of mortality \( \lambda \). Let \( H = (H_t)_{t \in [0, T]} \) be an \( l \)-dimensional Markovian jump-diffusion adapted to \( (\mathcal{F}_t) \) and \( D : \mathbb{R}^l \to \mathbb{R} \) some function. The goal is a mean-variance optimal hedge of \( D(H_T) \) using \( P^u \).

**Example 2.3.1.** Consider a process \( \hat{\lambda} = (\hat{\lambda}_t)_{t \in [0, T]} \) solving the SDE

\[
d\hat{\lambda}_t = \mu_\lambda(t, \hat{\lambda}_t) \, dt + \sigma_\lambda(t, \hat{\lambda}_t) \, d\bar{W}_t + \int_{\mathbb{R}_+ \setminus \{0\}} \tilde{\sigma}_\lambda(t, \hat{\lambda}_t, \bar{x}) \, \tilde{J}_\lambda(dt, d\bar{x}),
\]

with \( \hat{\lambda}_0 > 0 \). Suppose \( \hat{\lambda} \) describes the force of mortality of the pool of insured persons, let \( m = 1 \) for ease of exposition. If the insurance needs to deliver
one share of $S$ to each person in its pool that has survived until the terminal
time $T$, then the obligation $D(H_T) = S_T e^{-\int_0^T \lambda_s \, ds}$ is to be hedged.

We write $E_{t,p,z,h}[\cdot] = E[\cdot|P_t = p, Z_t = z, H_t = h]$ for the conditional
expectation given $(t, p, z, h) \in [0, T) \times \mathbb{R}_+ \times \mathbb{R}_m^{+2} \times \mathbb{R}_l$ and $\text{Var}_{t,p,z,h}$ denotes
the conditional variance accordingly. Let $\gamma > 0$ be a risk-aversion parameter.

**Definition 2.3.2.** For each $u \in \mathcal{U}$ and $\gamma > 0$, we define the functions $F_u, g_u : [0, T] \times \mathbb{R}_+ \times \mathbb{R}_m^{+2} \times \mathbb{R}_l \rightarrow \mathbb{R}$ by

$$g_u(t, p, z, h) = E_{t,p,z,h}[P_T^u - D(H_T)],$$

$$F_u(t, p, z, h) = E_{t,p,z,h} \left[ P_T^u - \frac{\gamma}{2} (P_T^u)^2 + \gamma P_T^u D(H_T) - D(H_T) - \frac{\gamma}{2} D(H_T)^2 \right].$$

(2.3.1)

We also need to define the following differential operator:

**Definition 2.3.3.** To any vector $u \in \mathcal{U}$ we associate the operator $A^u : f \mapsto A^uf$ mapping $C^{1,2,2,2}([0, T] \times \mathbb{R}_+ \times \mathbb{R}_m^{+} \times \mathbb{R}_l, \mathbb{R})$ to $C^{0,0,0,0}([0, T] \times \mathbb{R}_+ \times \mathbb{R}_m \times \mathbb{R}_l, \mathbb{R})$ given by

$$A^u f(t, p, z, h) = \lim_{\epsilon \downarrow 0} \frac{E_{t,p,z,h}[f(t + \epsilon, P_{t+\epsilon}^u, Z_{t+\epsilon}, H_{t+\epsilon})] - f(t, p, z, h)}{\epsilon},$$

(2.3.2)

if the limit exists.

The differential operator introduced in Definition [2.3.3] is known as *infinitesimal generator of the graph* of the process $(P^u, Z, H)$. Two further differential operators are needed; we presume them to act on suitably differentiable functions $f$:

- $\mathcal{L}^u f(t, p, z, h) := A^u f(t, p, z, h) - \dot{f}(t, p, z, h)$, which is called *infinitesimal generator* of the process $(P^u, Z, H)$,

- $\mathcal{G}^u f(t, p, z, h) := \gamma f(t, p, z, h) \mathcal{L}^u f(t, p, z, h) - \frac{\gamma}{2} \mathcal{L}^u f^2(t, p, z, h)$.
2.3. OPTIMIZATION PROBLEM

**Definition 2.3.4.** We define the value function \( J : [0, T] \times \mathbb{R}_+ \times \mathbb{R}_+^{m+2} \times \mathbb{R}^l \times \mathcal{U} \to \mathbb{R} \) by

\[
J(t, p, z, h, u) := \mathbb{E}_{t, p, z, h}[P_t^u - D(H_T)] - \frac{\gamma}{2} \text{Var}_{t, p, z, h}[P_t^u - D(H_T)]
\]

\[
= F_u(t, p, z, h) + \frac{\gamma}{2} g_u^2(t, p, z, h).
\]

The second equality in Definition 2.3.4 easily follows from (2.3.1). Finding some \( u^* \in \mathcal{U} \) such that \( J(t, p, z, h, u) \) is maximal is a time-inconsistent control problem and induces a path an investor would not follow. Therefore we next introduce the concept of equilibrium control.

**Definition 2.3.5.**

- A trading strategy \( u^* \in \mathcal{U} \) is an equilibrium control if

\[
\liminf_{c \to 0} \frac{J(t, p, z, h, u^*) - J(t, p, z, h, u_{t+c})}{c} \geq 0,
\]

for any \((t, p, z, h) \in [0, T] \times \mathbb{R}_+ \times \mathbb{R}_+^{m+2} \times \mathbb{R}^l \) and for all

\[
u_{t+c} := \begin{cases} u, & \text{on } [t, t+c] \times B_p \times B_z \times B_h, \\ u^*, & \text{on } \{(t, t+c] \times B_p \times B_z \times B_h\}', \end{cases}
\]

\( t+c \leq T \), where \( u \in \mathcal{U} \) and \( B_p, B_z, B_h \) are some arbitrary balls centered at respectively \( p, z, h \).

- The equilibrium value function is defined by

\[
V(t, p, z, h) := J(t, p, z, h, u^*).
\]

- An equilibrium policy \( u^* \) is of feedback type if, for some feedback function

\[
u_\ast : [0, T] \times \mathbb{R}_+ \times \mathbb{R}_+^{m+2} \times \mathbb{R}^l \to \mathcal{U}, \text{ we have}
\]

\[
u_t^\ast = u_\ast(t, P_t^\ast, Z_t, H_t), \quad t \in [0, T],
\]

with \( P_0^\ast = p, Z_0 = Z_0 \) and \( H_0 = H_0 \).
We see from (2.3.3) that a strategy is an equilibrium if a deviation is sub-optimal given the knowledge that every future player will obey that strategy. In the sequel we will search for an equilibrium control law of feedback type. Recall that the optimal value function of a standard time-consistent stochastic optimal control problem is the solution of a partial integro-differential equation (PIDE) known as Hamilton-Jacobi-Bellman (HJB) equation. In Björk and Murgoci (2010) a similar approach for time-inconsistent stochastic optimal control problems is introduced leading to a system of PIDEs. The system to be solved is subsequently referred to as extended HJB system and reduces to the classical case for a time-consistent problem. We now specify the extended HJB system corresponding to the value function from Definition 2.3.4.

**Definition 2.3.6.** For \((t, p, z, h) \in [0, T] \times \mathbb{R}^+ \times \mathbb{R}^{m+2} \times \mathbb{R}^l\), the extended HJB system is given by

\[
\dot{V}(t, p, z, h) + \sup_{u \in U} \{\mathcal{L}^u V(t, p, z, h) + \mathcal{G}^u g_u(t, p, z, h)\} = 0,
\]

\[
V(T, p, z, h) = p - D(h),
\]

\[
A^u g_u(t, p, z, h) = 0,
\]

\[
g_u(T, p, z, h) = \mathbb{E}_{T, p, z, h}[P_T^u - D(H_T)] = p - D(h),
\]

where \(\hat{u} = \arg \sup_{u \in U} \{\mathcal{L}^u V(t, p, z, h) + \mathcal{G}^u g_u(t, p, z, h)\}\).

A solution to the extended HJB system is the quadruple \((\hat{u}, V(t, p, h, z), F_\hat{u}(t, p, h, z), g_\hat{u}(t, p, h, z))\).

### 2.4 Sufficiency and Necessity

Before proving two verification results, we need the following assumption:

**Assumption 2.4.1.** The limit \(A^u V(t, p, z, h)\) defined in (2.3.2) exists.

**Definition 2.4.2.** A regular equilibrium is a quadruple \((u^*, V(t, p, z, h), F_{u^*}(t, p, z, h), g_{u^*}(t, p, z, h))\), with \(u^*\) being an equilibrium control of feedback type with corresponding equilibrium value function \(V\) (cf. Definition 2.3.5).
The first verification theorem says that if the extended HJB system given by Definition 2.3.6 has a solution, then it must be the equilibrium control law for the mean-variance hedge. In other words, the solvability of the extended HJB system is sufficient for the existence of an equilibrium control.

**Theorem 2.4.3.** Suppose $F_{u^*}(t, p, z, h), g_{u^*}(t, p, z, h) \in C^{1,2,2,2}(0, T) \times \mathbb{R}^+ \times \mathbb{R}_+ \times \mathbb{R}^l \times \mathbb{R}^l$ and $V, F_{u^*}, g_{u^*}$ solve the extended HJB system in Definition 2.3.6. Assume the control law $u^*$ realizes the supremum in the first row for every quadruple $(t, p, z, h) \in [0, T] \times \mathbb{R}_+ \times \mathbb{R}_+ \times \mathbb{R}^l$. Then there exists an equilibrium control law $u^*$ in the sense of Definition 2.3.5 and it is given by the optimal $u$ in the first row of (2.3.5). Moreover, $V$ is the corresponding equilibrium value function and $F_{u^*}$ and $g_{u^*}$ are given by (2.3.1).

**Proof.** The proof can be conducted similarly to the proof of Theorem 7.1 in Björk and Murgoci (2010) and is therefore omitted.

Next we show that an equilibrium control is necessarily a solution of the extended HJB system. Such a proof is provided in Lindensjö (2016) for a diffusion case and we extend it to the present jump-diffusion setting including the hedge of the terminal condition.

**Theorem 2.4.4.** A regular equilibrium $(u^*, V(t, p, z, h), F_{u^*}(t, p, z, h), g_{u^*}(t, p, z, h))$ in the sense of Definition 2.4.2 necessarily solves the extended HJB system (2.3.5) and $u^*$ realizes the supremum in the first row.

The proof is delivered in several steps. We start by introducing two sequences of stopping times that will be repeatedly needed in the sequel. Let $(c_k)_{k \in \mathbb{N}}$ be a strictly positive monotone sequence satisfying $\lim_{k \to \infty} c_k = 0$. Let $(t, p, z, h, u) \in [0, T] \times \mathbb{R}_+ \times \mathbb{R}_+ \times \mathbb{R}^l \times \mathcal{U}$ arbitrary and denote by $B_p, B_z, B_h$ balls centered at respectively $p, z, h$. Define the sequence of stopping times $(\sigma^u_k)$ by

$$\sigma^u_k := \inf \{ s > t : (s, P_s^u, Z_s, H_s) \notin [t, t + c_k) \times B_p \times B_z \times B_h \} \land T. \quad (2.4.1)$$
Proposition 2.4.5. Consider the sequence of stopping times \( (\sigma^u_k) \) with a typical element given by (2.4.1). It holds that

\[
\sigma^u_k > t \text{ a.s.}
\]

Proof. This is an immediate consequence of the càdlàg property of the mapping \( t \mapsto (t, P^u_t, Z_t, H_t) \) (cf. Applebaum (2009), p.106): let \( k \in \mathbb{N} \) and \( \omega \in \Omega \) arbitrary. For any \( \epsilon > 0 \) there exists some \( \delta > 0 \) such that for all \( s \in (t, t + \delta) \) it holds that

\[
\left\| (s, P^u_s, Z_s, H_s) - (t, P^u_t, Z_t, H_t) \right\|_2 < \epsilon,
\]

thus, \( \sigma^u_k \geq s > t \text{ a.s.} \)

Observe that \( \lim_{k \to \infty} \sigma^u_k = t \). Let \( (a_k)_{k \in \mathbb{N}} \) be another positive monotone sequence satisfying \( \lim_{k \to \infty} a_k = 0 \) such that the sequence of events \( (A_k)_{k \in \mathbb{N}} \) is characterized by

\[
A_k := \{ \omega \in \Omega : \sigma^u_k > t + a_k \},
\]

\[
P(A_k) \geq 1 - \frac{1}{k^2}.
\] (2.4.2)

Lemma 2.4.6. Consider the event \( A_k \) and its probability of occurrence defined by (2.4.2). Then it holds that

\[
\mathbb{1}_{A_k}(\omega) = 1 \text{ a.s.,}
\]

for all but finitely many \( k \).

Proof. Observe that \( P(A_k^c) \leq \frac{1}{k^2} \) and therefore \( \sum_{k=1}^{\infty} P(A_k^c) \leq \frac{\pi^2}{6} < \infty \). The Borel-Cantelli lemma implies that \( \mathbb{1}_{A_k^c}(\omega) = 0 \) for all but finitely many \( k \) and the claim follows.

Define a typical element of the sequence of stopping times \( (\tau^u_k)_{k \in \mathbb{N}} \) by

\[
\tau^u_k := \min\{\sigma^u_k, t + a_k\}.
\] (2.4.3)
Lemma 2.4.7. Let $u^*$ be an equilibrium control and consider the function $g_u$ defined by (2.3.1). Then it holds that

$$A^{u^*}g_{u^*}(t, p, z, h) = 0. \quad (2.4.4)$$

Proof. Using Dynkin’s formula (see e.g. Øksendal and Sulem (2005), p.12), we find that

$$g_{u^*}(t, p, z, h) = E_{t,p,z,h}\left[g_{u^*}(\tau_{i_k}^*, P_{i_k}^{u^*}, Z_{i_k}^{u^*}, H_{i_k}^{u^*}) - \int_{t}^{\tau_{i_k}^*} A^{u^*}g_{u^*}(s, P_{s}^{u^*}, Z_{s}, H_{s}) \, ds\right].$$

It is a simple consequence of the tower property that

$$E_{t,p,z,h}[g_{u^*}(\tau_{i_k}^*, P_{i_k}^{u^*}, Z_{i_k}^{u^*}, H_{i_k}^{u^*})] = E_{t,p,z,h}[P_{T_k}^{u^*} - D(H_{T_k})] = g_{u^*}(t, p, z, h).$$

Combining the previous two results, we find that

$$E_{t,p,z,h}\left[\frac{\int_{t}^{\tau_{i_k}^*} A^{u^*}g_{u^*}(s, P_{s}^{u^*}, Z_{s}, H_{s}) \, ds}{a_k}\right] = 0.$$

Consider the sequence of random variables

$$\left(\frac{\int_{t}^{\tau_{i_k}^*} A^{u^*}g_{u^*}(s, P_{s}^{u^*}, Z_{s}, H_{s}) \, ds}{a_k}\right)_{k \in \mathbb{N}},$$

and note that the integrand is bounded on the interval $[t, \tau_{i_k}^*]$, even if there is a large jump at $\tau_{i_k}^*$ since this point has Lebesgue measure zero. Therefore we can use dominated convergence to see that
According to Lemma 2.4.6 we have for arbitrary but fixed \( \omega \in \Omega \) that 
\[
\lim_{k \to \infty} \mathbb{E}_{t,p,z,h} \left[ \int_{t}^{t_\kappa^\omega} A^u g_u^* (s, P^u_s, Z_s, H_s) \, ds \right] = 0.
\]

Further,
\[
0 = \mathbb{E}_{t,p,z,h} \left[ \int_{t}^{t_\kappa^\omega} A^u g_u^* (s, P^u_s, Z_s, H_s) \, ds \right] = \mathbb{E}_{t,p,z,h} \left[ A^u g_u^* (t, p, z, h) \right] = A^u g_u^* (t, p, z, h),
\]

whereby the second equality is justified by Lebesgue’s differentiation theorem (cf. Rudin (1987), Chapter 7) and since \((t, p, z, h)\) has been arbitrarily chosen, (2.4.4) is established.

Let \( \tilde{u}_\kappa^u \) be an allocation rule that is equal to \( u(t) \equiv u \in U \) (a constant) on the interval \([t, \tau_\kappa^u]\) and equal to the equilibrium \( u^* \) outside that interval, that is
\[
\tilde{u}_\kappa^u (s) = u \mathbbm{1}_{[t, \tau_\kappa^u]} (s) + u^* (s) \mathbbm{1}_{[\tau_\kappa^u, T]} (s) \quad (2.4.5)
\]
2.4. SUFFICIENCY AND NECESSITY

\[ (u \mathbb{1}_{[t,\tau_k^u]}(s) + u^*(s) \mathbb{1}_{[\sigma_k^u,T]}(s)) \mathbb{1}_{A_k^u} + u_1^*(s) \mathbb{1}_{[\sigma_k^u,T]}(s) \mathbb{1}_{A_k^u} \]

(2.4.6)

\[ = (u \mathbb{1}_{[t,\sigma_k^u]}(s) + u^*(s) \mathbb{1}_{[\sigma_k^u,T]}(s)) \mathbb{1}_{A_k^u} + u_{t+a_k} \mathbb{1}_{A_k^u} \]

(2.4.7)

where (2.4.6) follows from the definition of \( \tau_k^u \), cf. (2.4.3). Moreover, as the right-hand bracket of (2.4.6) is for each fixed \( k \in \mathbb{N} \) easily seen to be a function of feedback type as \( u_{t+c} \) in Definition 2.3.5, we can equate it to (2.4.7).

**Lemma 2.4.8.** Consider an equilibrium control \( u^* \), the control \( \tilde{u}_{\tau_k} \) given by (2.4.5) and the function \( F_u \) defined by (2.3.1). Then we have

\[ \lim_{k \to \infty} \frac{F_u^*(t,p,z,h) - F_{\tilde{u}_{\tau_k}}(t,p,z,h)}{a_k} = -A^u F_u^*(t,p,z,h). \]

(2.4.8)

**Proof.** According to Dynkin’s formula,

\[ \mathbb{E}_{t,p,z,h}[F_u^*(\tau_k^u, P_{\tau_k}^u, Z_{\tau_k}^u, H_{\tau_k}^u)] = F_u^*(t,p,z,h) + \int_t^{\tau_k^u} A^u \mathbb{E}_{s,P_{\tau_k}^u, Z_s, H_s} F_u^*(s, P_s^u, Z_s, H_s) \, ds , \]

and we observe that

- the integral limits in the previous equation are \( t \) and \( \tau_k^u \), therefore we can denote \( P_s^u \) by \( P_s^u \) and \( A^u \mathbb{E}_{\tau_k} \) by \( A^u \) on the random interval \( (t, \tau_k^u) \).
- as the starting time point is \( \tau_k^u \), it holds that

\[ F_u^*(\tau_k^u, P_{\tau_k}^u, Z_{\tau_k}^u, H_{\tau_k}^u) = \mathbb{E}_{\tau_k, P_{\tau_k}^u, Z_{\tau_k}^u, H_{\tau_k}^u} \left[ P_T^u - \frac{\gamma}{2} (P_T^u)^2 + \gamma P_T^u D(H_T) - D(H_T) - \frac{\gamma}{2} D(H_T)^2 \right] \]

\[ = \mathbb{E}_{\tau_k, P_{\tau_k}^u, Z_{\tau_k}^u, H_{\tau_k}^u} \left[ P_T^u - \frac{\gamma}{2} (P_T^u)^2 + \gamma P_T^u D(H_T) - D(H_T) - \frac{\gamma}{2} D(H_T)^2 \right] . \]
Using these two observations, we rewrite

\[ F^*_u(t, p, z, h) + \mathbb{E}_{t, p, z, h} \left[ \int_t^{\tau^u_k} A^u F^*_{k^u}(s, P^u_{s}, Z_s, H_s) \, ds \right] \]

\[ = \mathbb{E}_{t, p, z, h} \left[ F^*_u(\tau^u_k, P^u_{\tau^u_k}, Z^u_{\tau^u_k}, H^u_{\tau^u_k}) \right] \]

\[ = \mathbb{E}_{t, p, z, h} \left[ \mathbb{E}_{\tau^u_k, P^u_{\tau^u_k}, Z^u_{\tau^u_k}, H^u_{\tau^u_k}} \left[ P^u_{\tau^u_k} - \frac{\gamma}{2} (P^u_{\tau^u_k})^2 + \gamma P^u_{\tau^u_k} D(H_T) \right. \right. \]

\[ - \left. D(H_T) - \frac{\gamma}{2} D(H_T)^2 \right] \]

\[ = \mathbb{E}_{t, p, z, h} \left[ P^u_{\tau^u_k} - \frac{\gamma}{2} (P^u_{\tau^u_k})^2 + \gamma P^u_{\tau^u_k} D(H_T) - D(H_T) - \frac{\gamma}{2} D(H_T)^2 \right] \]

\[ = F^*_u(t, p, z, h). \]

Finally, we use dominated convergence and Lebesgue’s differentiation theorem similarly as in the proof of Lemma 2.4.7 to deduce that

\[ \lim_{k \to \infty} F^*_u(t, p, z, h) - F^*_{\tilde{u}^u_k}(t, p, z, h) \]

\[ = \lim_{k \to \infty} \mathbb{E}_{t, p, z, h} \left[ \mathbb{1}_{A_k}(\omega) \int_{t}^{t + a_k} A^u F^*_{k^u}(s, P^u_{s}, Z_s, H_s) \, ds \right] \]

\[ = \mathbb{E}_{t, p, z, h} \left[ \lim_{k \to \infty} \mathbb{1}_{A_k}(\omega) \int_{t}^{t + a_k} A^u F^*_{k^u}(s, P^u_{s}, Z_s, H_s) \, ds \right] \]

\[ = -A^u_{t, p, z, h} F^*_u(t, p, z, h), \]

which is what we have set out to prove. \( \square \)

**Lemma 2.4.9.** Consider an equilibrium control \( u^* \), the control \( \tilde{u}^u_k \) given by (2.4.5) and the function \( g_u \) defined by (2.3.1). Then we have

\[ \lim_{k \to \infty} g^*_u(t, p, z, h)^2 - g^*_{\tilde{u}^u_k}(t, p, z, h)^2 \]

\[ = -2 g^*_u(t, p, z, h) A^u g^*_u(t, p, z, h). \]

(2.4.9)
Proof. Using similar techniques as before, the following calculation yields

\[
\begin{align*}
&\lim_{k \to \infty} \frac{g_{u^*}(t, p, z, h)^2 - g_{\tilde{u}_{k}}(t, p, z, h)^2}{a_k} \\
&= - \lim_{k \to \infty} \frac{g_{\tilde{u}_{k}}(t, p, z, h)^2 - g_{u^*}(t, p, z, h)^2}{a_k} \\
&= - \lim_{k \to \infty} \frac{\left(\mathbb{E}_{t, p, z, h} \left[ P_{T_k}^{\tilde{u}_{k}} - D(H_T) \right] \right)^2 - g_{u^*}(t, p, z, h)^2}{a_k} \\
&= - \lim_{k \to \infty} \frac{\left(\mathbb{E}_{t, p, z, h} \left[ \mathbb{E}_{\tau_k} \tilde{u}_{k}^{\tilde{u}_{k}} \left[ P_{T_k}^{u^*} - D(H_T) \right] \right] \right)^2 - g_{u^*}(t, p, z, h)^2}{a_k} \\
&= - \lim_{k \to \infty} \frac{\left(\mathbb{E}_{t, p, z, h} \left[ g_{u^*}(\tau_k, P_{\tau_k}^{u^*}, Z_{\tau_k}, H_{\tau_k}) \right] \right)^2 - g_{u^*}(t, p, z, h)^2}{a_k} \\
&= - \lim_{k \to \infty} \frac{1}{a_k} \left( g_{u^*}(t, p, z, h) + \mathbb{E}_{t, p, z, h} \left[ \int_{t}^{\tau_k} A^u g_{u^*}(s, P_{s}^{u^*}, Z_{s}, H_{s}) \, ds \right] \right)^2 \\
&= - 2 g_{u^*}(t, p, z, h) A^u g_{u^*}(t, p, z, h),
\end{align*}
\]

where the abbreviation T.P. stands for tower property.

Lemma 2.4.10. Consider an equilibrium control \( u^* \), the control \( \tilde{u}_{k} \) given by \([2.4.5]\) and the value function \( J \) specified in Definition \([2.3.4]\). Then it holds that

\[
- \lim_{k \to \infty} \frac{J(t, p, z, h, u^*) - J(t, p, z, h, \tilde{u}_{k})}{a_k} = A^u V(t, p, z, h) + G^u g_{u^*}(t, p, z, h).
\]
Proof of Theorem 2.4.4. The proof is conducted in four steps:

**Step 1:** We show the boundary conditions. The boundary conditions $V(T, p, z, h) = p - D(h)$ and $g_{u^*}(T, p, z, h) = p - D(h)$ are met by the equilibrium control law $u^*$, which follows from Definition 2.3.2 and Definition 2.3.5.

**Step 2:** Observe that $\mathcal{A}u^*g_{u^*}(t, p, z, h) = 0$ is stated by Lemma 2.4.7.

**Step 3:** We show that $\mathcal{A}u^*V(t, p, z, h) + \mathcal{G}u^*g_{u^*}(t, p, z, h) = 0$. Recall from Definition 2.3.4 and Definition 2.3.5 that $V(t, p, z, h) = F_{u^*}(t, p, z, h) + \frac{\gamma}{2} g_{u^*}^2(t, p, z, h)$. Following a similar line of reasoning as in the proof of Lemma 2.4.7, one can show that $\mathcal{A}u^*F_{u^*}(t, p, z, h) =$...
2.4. SUFFICIENCY AND NECESSITY

So we have

\[ A^u V(t, p, z, h) + G^u g_u^*(t, p, z, h) = \frac{\gamma}{2} A^u g_u^2(t, p, z, h) + \gamma g_u^*(t, p, z, h) L^u g_u^*(t, p, z, h) - \frac{\gamma}{2} L^u g_u^2(t, p, z, h) \]

\[ = \gamma g_u^*(t, p, z, h) A^u g_u^*(t, p, z, h) = 0, \]

whereby the last equality follows from Step 2.

So far, we have shown that the regular equilibrium \((u^*, V(t, p, z, h), F_u^*(t, p, z, h), g_u^*(t, p, z, h))\) is a prospective solution of the extended HJB system (2.3.5). Therefore we are left showing that \(u^*\) is indeed maximal in the first row of (2.3.5).

**Step 4:** We show that \(0 \geq A^u V(t, p, z, h) + G^u g_u^*(t, p, z, h)\).

In the following calculation, the first inequality follows by Definition of the equilibrium control \(u^*\), cf. Definition 2.3.5. Observe that

\[ 0 \geq -\liminf_{c \to 0} \frac{J(t, p, z, h, u^*) - J(t, p, z, h, u_{t+c})}{c} \]

\[ = -\liminf_{c \to 0} \left( \frac{J(t, p, z, h, u^*) - J(t, p, z, h, u_{t+c})}{c} A_k(\omega) \right) \]

\[ + \frac{J(t, p, z, h, u^*) - J(t, p, z, h, u_{t+c})}{c} A_k(\omega) \]

\[ = -\liminf_{k \to \infty} \frac{J(t, p, z, h, u^*) - J(t, p, z, h, u_{t+a_k})}{c_k} A_k(\omega) \]

\[ -\liminf_{k \to \infty} \frac{J(t, p, z, h, u^*) - J(t, p, z, h, u_{t+a_k})}{c_k} A_k(\omega). \]

Note that Lemma 2.4.6 implies that

\[ \liminf_{k \to \infty} \frac{J(t, p, z, h, u^*) - J(t, p, z, h, u_{t+c_k})}{c_k} A_k(\omega) = 0. \]
Consequently, we deduce from (2.4.7) that

\[
- \liminf_{k \to \infty} \frac{J(t, p, z, h, u^*) - J(t, p, z, h, u_{t+a_k})}{a_k} \mathbb{1}_{A_k}^*(\omega) = - \liminf_{k \to \infty} \frac{J(t, p, z, h, u^*) - J(t, p, z, h, \tilde{u}_{t+a_k})}{a_k},
\]

and Lemma 2.4.10 concludes the proof.

\[\square\]

### 2.5 Explicit Solution

In the sequel, let \( D \equiv 0 \), i.e., we consider an investor aiming at receiving a high expected payoff while keeping its variance low. In this special case the extended HJB system (2.3.5) admits explicit closed-form solutions. Some notational definitions are in order:

- \( \sigma_S := (\sigma_{ij})_{1 \leq i \leq m, 1 \leq j \leq d} \), i.e., \( \sigma_S \in \mathbb{R}^{m \times d} \).
- \( \tilde{\sigma}_S := \sigma_S \sigma_S^\top \), i.e., \( \tilde{\sigma}_S \in \mathbb{R}^{m \times m} \). Note that \( \tilde{\sigma}_S \) is a symmetric matrix.
- \( \tilde{\sigma}_S_i := (\tilde{\sigma}_{s_{i1}}, \ldots, \tilde{\sigma}_{s_{im}})^\top \), i.e., \( \tilde{\sigma}_S_i \in \mathbb{R}^m \).
- \( \sigma_S_i := (\sigma_{i1}, \ldots, \sigma_{id})^\top \), i.e., \( \sigma_S^k \in \mathbb{R}^d \) for every \( i \in \{1, \ldots, m\} \).
- \( \rho_S := (\rho_{ij})_{1 \leq i \leq m, 1 \leq j \leq k} \), i.e., \( \rho_S \in \mathbb{R}^{m \times k} \).
- \( \tilde{\rho}_S := \rho_S \rho_S^\top \).
- \( \rho_{S_i} := (\rho_{i1}, \ldots, \rho_{ik})^\top \), i.e., \( \rho_{S_i} \in \mathbb{R}^k \).
- \( \mu := (\mu_1, \ldots, \mu_n)^\top \), i.e., \( \mu \in \mathbb{R}^n \).
- \( \tilde{\mu} := \mu - 1r \).
- \( \Delta P_t^u(x, \bar{x}) := u_S^\top(t)\rho_S x + u_Y(t)\eta_L(t, \lambda_{t-}, Y_{t-}, \bar{x}) \), i.e., \( \Delta P_t^u(x, \bar{x}) \in \mathbb{R} \).
- \( \Delta Z_t(x, \bar{x}) := \text{Diag}(S_{t-})\rho_S x, \tilde{\sigma}(t, \lambda_{t-}, \bar{x}), Y_{t-}\eta_L(t, \lambda_{t-}, Y_{t-}) \), i.e., \( \Delta Z_t(x, \bar{x}) \in \mathbb{R}^{m+2} \).
2.5. EXPLICIT SOLUTION

\[ \mu_i(t, Z_t) := \begin{cases} 
\mu_i S_i^t, & i = 1, \ldots, m, \\
\mu_\lambda(t, \lambda_t), & i = m + 1, \\
(r + \nu_L(t, \lambda_t, Y_t))Y_t, & i = m + 2,
\end{cases} \]
i.e., \( \mu(t, Z_t) = (\mu_1(t, Z_t), \ldots, \mu_{m+2}(t, Z_t)) \in \mathbb{R}^{m+2}. \)

\[ \sigma_{ij}(t, Z_t) := \begin{cases} 
\sigma_{ij} S_i^t, & 1 \leq i \leq m, 1 \leq j \leq d, \\
\sigma_\lambda(t, \lambda_t), & i = m + 1, j = d + 1, \\
\sigma_L(t, \lambda_t, Y_t)Y_t, & i = m + 2, j = d + 1, \\
0, & \text{else},
\end{cases} \]
i.e., \( \sigma(t, Z_t) \in \mathbb{R}^{(m+2) \times (d+1)}. \)

\[ Q^u_i(t, Z_t) := \begin{cases} 
S^t_i u^u S_i(t, \lambda_t) \sigma_\lambda(t, \lambda_t) \sigma_L(t, \lambda_t, Y_t), & i = m + 1, \\
u_Y(t) \sigma_\lambda(t, \lambda_t) \sigma_L(t, \lambda_t, Y_t), & i = m + 2,
\end{cases} \]
i.e., \( Q^u(t, Z_t) = (Q^u_1(t, Z_t), \ldots, Q^u_{m+2}(t, Z_t)) \in \mathbb{R}^{m+2}. \)

Inspired by Basak and Chabakauri (2010) and Björk and Murgoci (2010), we make the following Ansatz:

\[ V(t, p, z) = A(t)p + B(t, z), \]
\[ g(t, p, z) = a(t)p + b(t, z). \] (2.5.1)

The goal is finding the functions \( A, a, B, b \) as well as the equilibrium control laws of feedback type. Clearly, the functions \( A, a, B, b \) are assumed to satisfy the necessary regularity conditions and the limits induced by applying the operators \( A, L \) and \( G \) are assumed to exist accordingly. Consider the first line in the system (2.3.5) and define

\[ \Xi(u_S(t), u_Y(t)) := \mathcal{L}^u V(t, p, z) + \mathcal{G}^u g(t, p, z). \]

Omitting details at this stage, an application of Itô’s formula for jump-
diffusions yields

\[
\Xi(u_S(t), u_Y(t)) = \sum_{i=1}^{m} \left\{ \frac{\partial u_S^i(t)\tilde{\sigma}_Su_S(t)}{\partial u_S(t)} + \frac{1}{2} \sum_{j=1}^{m} \tilde{\sigma}^i_{Sj}u_S^j(t) \right\}^2 \quad \text{with} \quad \sum_{i=1}^{m} \tilde{\sigma}^i_{Sj} = \tilde{\sigma}^i_S.
\]

For arbitrary \( i \in \{1, \ldots, m\} \), we consider the following FOC. Observe that the interchange of differentiation and integration is justified by our assumptions.
2.5. EXPLICIT SOLUTION

\[ \frac{\partial \Xi}{\partial u_S(t)} = A(t)\hat{\mu}_i - \gamma a(t)^2 \hat{\sigma}_S^T u_S(t) - \gamma a(t) \sigma_S^T \sum_{j=1}^{m} b_{z_j}(t, Z_t) S^T_{ij} \sigma_S \\
- \gamma a(t) \int_{R_{k+1}\{0\}} (a(t) \Delta P_\ell^u(x, \bar{x}) \\
+ b(t, Z_t + \Delta Z_t(x, \bar{x})) - b(t, Z_t)) \rho_S^T x \vartheta_{\mathcal{X}}(dx, d\bar{x}) \\
= A(t)\hat{\mu}_i - \gamma a(t)^2 \hat{\sigma}_S^T u_S(t) - \gamma a(t) \sigma_S^T \sum_{j=1}^{m} b_{z_j}(t, z) S^T_{ij} \sigma_S \\
- \gamma a(t) \int_{R_{k}\{0\}} (a(t) u_{1i}^T(t) \rho_S x \\
+ b(t, Z_t + \Delta Z_t(x, 0)) - b(t, Z_t)) \rho_S^T x \vartheta_{\mathcal{X}}(dx) \equiv 0 \\
\Leftrightarrow A(t)\hat{\mu}_i - \gamma a(t) \sigma_S^T \sum_{j=1}^{m} b_{z_j}(t, z) S^T_{ij} \sigma_S \\
- \gamma a(t) \int_{R_{k}\{0\}} (b(t, Z_t + \Delta Z_t(x, 0)) - b(t, Z_t)) x^T \vartheta_{\mathcal{X}}(dx) \rho_S \\
= \gamma a(t)^2 (\hat{\sigma}_S^T + \xi \rho_S^T, \rho_S^T) u_S(t). \]

We use the following abbreviations in the sequel:

\[ \xi := \int_{R_{k}\{0\}} xx^T \vartheta_{\mathcal{X}}(dx), \]
\[ \hat{\eta}_L(t, \lambda_t, Y_t) := \int_{R_{k}\{0\}} \eta_L(t, \lambda_t, Y_t, \bar{x})^2 \vartheta_{\mathcal{X}}(d\bar{x}), \]
\[ b_1(t, Z_t) := \int_{R_{k}\{0\}} (b(t, Z_t + \Delta Z_t(x, 0)) - b(t, Z_t)) x \vartheta_{\mathcal{X}}(dx), \quad (2.5.3) \]
\[ b_2(t, Z_t) := \int_{R_{k}\{0\}} (b(t, Z_t + \Delta Z_t(0, \bar{x})) - b(t, Z_t)) \eta_L(t, \lambda_t, Y_t, \bar{x}) \vartheta_{\mathcal{X}}(d\bar{x}). \quad (2.5.4) \]

Note that in the optimum an equality of type \* needs to hold for every \( u_S(t) \),
so using the just defined functions and matrix-vector notation, we see that the vector \( u_S^*(t) \) has to satisfy

\[
\begin{aligned}
\quad \quad u_S^*(t) \\
= \frac{(\tilde{\sigma} S + \tilde{\rho} S \xi)^{-1}}{\gamma a(t)^2} \left( A(t) \tilde{\mu} - \gamma a(t) \sigma S \sum_{j=1}^{m} b_{z_j}(t, Z_i) S_j \sigma S_j - \gamma a(t) \rho S b_1(t, Z_i) \right),
\end{aligned}
\] (2.5.5)

where the symbol \( \star \) indicates the optimality of the strategy. Next we compute

\[
\frac{\partial \Xi}{\partial u_Y(t)} = \begin{aligned}
A(t) &\nu_L(t, \lambda_t, Y_t) - \gamma a(t)^2 \sigma_L^2(t, \lambda_t, Y_t) u_Y(t) \\
- \gamma a(t) &\left( b_{z_{m+1}}(t, Z_i) \sigma_L(t, \lambda_t, Y_t) + b_{z_{m+2}}(t, Z_i) \sigma_L^2(t, \lambda_t, Y_t) Y_t \right) \\
- \gamma a(t) &\int_{R \setminus \{0\}} \left( a(t) u_Y(t) \gamma a(t, \lambda_t, Y_t, \bar{x}) \\
&+ b(t, Z_t + \Delta Z_i(0, \bar{x})) - b(t, Z_t) \right) \eta_L(t, \lambda_t, Y_t, \bar{x}) \vartheta_X(d\bar{x})
\end{aligned}
\]

\[
= \begin{aligned}
A(t) &\nu_L(t, \lambda_t, Y_t) - \gamma a(t)^2 \sigma_L^2(t, \lambda_t, Y_t) u_Y(t) \\
- \gamma a(t) &\left( b_{z_{m+1}}(t, Z_i) \sigma_L(t, \lambda_t, Y_t) + b_{z_{m+2}}(t, Z_i) \sigma_L^2(t, \lambda_t, Y_t) Y_t \right) \\
- \gamma a(t) &\gamma a(t, \lambda_t, Y_t, \bar{x}) u_Y(t) - \gamma a(t) b_2(t, Z_t) \equiv 0 \\
&
\Leftrightarrow u_Y^*(t) = \left( \gamma a(t)^2 \sigma_L^2(t, \lambda_t, Y_t) + \gamma a(t, \lambda_t, Y_t) \right)^{-1} \cdot \left( A(t) \nu_L(t, \lambda_t, Y_t) \\
- \gamma a(t) &\left( b_{z_{m+1}}(t, Z_i) \sigma_L(t, \lambda_t, Y_t) + b_{z_{m+2}}(t, Z_i) \sigma_L^2(t, \lambda_t, Y_t) Y_t \right) \\
- \gamma a(t) b_2(t, Z_t) \right).
\end{aligned}
\]

Observe that the optimal control does not depend on \( p \). We next plug \( u^* \) into (2.5.2). Then we can apply separation of variables to the first line of (2.3.5). This leads to an ordinary differential equation (ODE) for \( A \) and a PIDE for \( B \). The ODE for \( A \) is given by

\[
\dot{A}(t) + A(t)r = 0, \\
A(T) = 1,
\]
2.5. EXPLICIT SOLUTION

and the solution is easily seen to be \( A(t) = e^{(T-t)} \). The PIDE for \( B \) is given by

\[
\dot{B}(t, Z_t) + A(t) (\bar{\mu}^T u_\bar{S}^*(t) + u_\bar{V}^*(t)\nu_L(t, \lambda_t, Y_t)) + \nabla Z B(t, Z_t)^\top \mu(t, Z_t)
\]

\[
+ \frac{1}{2} \text{Tr} \left( \sigma(t, Z_t)^\top H_Z(B(t, Z_t))\sigma(t, Z_t) \right)
\]

\[
- \frac{\gamma}{2} \left( u_\bar{S}^*(t)^\top \bar{\sigma} S u_\bar{S}^*(t) + u_\bar{V}^*(t)^2 \sigma^2(t, \lambda_t, Y_t) \right) a(t)^2
\]

\[
- \frac{\gamma}{2} \text{Tr} \left( \sigma(t, Z_t)^\top \nabla_Z b(t, Z_t)^\top \nabla_Z b(t, Z_t) \sigma(t, Z_t) \right) - \gamma a(t) \nabla_Z b(t, Z_t)^\top Q u^*(t, Z_t)
\]

\[
+ \int_{\mathbb{R}^{k+1} \setminus \{0\}} \left( B(t, Z_t + \Delta Z(t, \bar{x})) - B(t, Z_t) \right)
\]

\[
- (\Delta Z(t, \bar{x}))^\top \nabla_Z B(t, Z_t) \right) \partial_{X, \bar{X}}(dx, d\bar{x})
\]

\[
\quad - \frac{\gamma}{2} \int_{\mathbb{R}^{k+1} \setminus \{0\}} \left( a(t) \Delta P^u_{t}[u^*(x, \bar{x}) + b(t, Z_t + \Delta Z(t, \bar{x}))]
\right.
\]

\[
\left. - b(t, Z_t) \right)^2 \partial_{X, \bar{X}}(dx, d\bar{x}) = 0.
\]

(2.5.6)

Note that \( \Delta P^u_{t}[u^*(x, \bar{x}) \right) \) in the last line of (2.5.6) means the jump of the portfolio process where the investor is allocating optimally. For solving the latter PIDE, we need to find the functions \( a \) and \( b \). To do so, we use the third equation of the system (2.3.5) (for the special case \( D \equiv 0 \)), namely \( A^u g^u_{t}(t, p, z) = 0 \). Following the Ansatz \( g(t, p, z) = a(t)p + b(t, z) \), we obtain

\[
\dot{a}(t)p + \dot{b}(t, Z_t) + a(t)(pr + \bar{\mu}^T u_\bar{S}^*(t) + \nu_L(t, \lambda_t, Y_t) u_\bar{V}^*(t)) + \nabla Z b(t, Z_t)^\top \mu(t, Z_t)
\]

\[
+ \frac{1}{2} \text{Tr} \left( \sigma(t, Z_t)^\top H_Z(b(t, Z_t))\sigma(t, Z_t) \right)
\]

\[
+ \int_{\mathbb{R}^{k+1} \setminus \{0\}} \left( b(t, Z_t + \Delta Z(t, \bar{x})) - b(t, Z_t) \right)
\]

\[
- (\Delta Z(t, \bar{x}))^\top \nabla_Z b(t, Z_t) \right) \partial_{X, \bar{X}}(dx, d\bar{x}) = 0.
\]

(2.5.7)

with suitable boundary conditions for \( a \) and \( b \). Using separation of variables
again, we find the ODE

\[
\dot{a}(t) + a(t) r = 0, \\
a(T) = 1,
\]

leading to \( a(t) = e^{(T-t)} \). Observe that \( A(t) = a(t) \), so we can cancel some terms in the optimal strategies. Several further definitions are in order:

- \( \Theta_1^i := \tilde{\mu}^T (\tilde{\sigma}_S + \tilde{\rho}_S \xi)^{-1} \), i.e., \( \Theta_1^i \in \mathbb{R}^m \),
- \( \Theta_2(t, \lambda_t, Y_t) := \frac{\nu_L(t, \lambda_t, Y_t)}{\sigma^2_L(t, \lambda_t, Y_t)} \),
- \( C(t, \lambda_t, Y_t, x, \bar{x}) := \Theta_1 \rho S x + \Theta_2(t, \lambda_t, Y_t) \eta_L(t, \lambda_t, Y_t, \bar{x}) \),
- \( \phi_1(t, Z_t) := \begin{cases} \Theta_1 \sigma S \sigma S_i S_i^t, & i = 1, \ldots, m, \\ \Theta_2(t, \lambda_t, Y_t) \sigma \sigma_L(t, \lambda_t, Y_t), & i = m + 1, \\ \Theta_2(t, \lambda_t, Y_t) \sigma^2_L(t, \lambda_t, Y_t)^t Y_t, & i = m + 2, \end{cases} \) i.e., \( \phi_1(t, Z_t) = (\phi_1(t, Z_t), \ldots, \phi_{m+2}(t, Z_t)) \in \mathbb{R}^{m+2} \).

Inserting \( u_S^i \) and \( u_Y^i \) into (2.5.7) and manipulating terms, we find that

\[
\begin{align*}
\dot{b}(t, Z_t) + \frac{\Theta_1}{\gamma} \bar{\mu} - \Theta_1 \sigma S & \sum_{j=1}^m b_{z_j}(t, Z_t) S_i^t \sigma S_i - \Theta_1 \rho S b_1(t, Z_t) \\
+ \frac{\Theta_2(t, \lambda_t, Y_t) \nu_L(t, \lambda_t, Y_t)}{\gamma} & - \Theta_2(t, \lambda_t, Y_t) \left( b_{z_{m+1}}(t, Z_t) \sigma_\lambda(t, \lambda_t) \sigma_L(t, \lambda_t, Y_t) \right) \\
+ b_{z_{m+2}}(t, Z_t) \sigma^2_L(t, \lambda_t, Y_t) & - \Theta_2(t, \lambda_t, Y_t) b_2(t, Z_t) + \nabla_Z b(t, Z_t)^T \mu(t, Z_t) \\
+ \frac{1}{2} \text{Tr} & (\sigma(t, Z_t)^T H_Z (b(t, Z_t)) \sigma(t, Z_t)) \\
+ \int_{\mathbb{R}^{k+1}\{0\}} & \left( b(t, Z_t + \Delta Z_t(x, \bar{x})) - b(t, Z_t) - (\Delta Z_t(x, \bar{x}))^T \nabla_Z b(t, Z_t) \right) \mathcal{L}_X \lambda (dx, d\bar{x}) = 0
\end{align*}
\]

\[
\Rightarrow \dot{b}(t, Z_t) + \frac{\Theta_1}{\gamma} \bar{\mu} + \frac{\Theta_2(t, \lambda_t, Y_t) \nu_L(t, \lambda_t, Y_t)}{\gamma} - \Theta_1 \sigma S \sum_{j=1}^m b_{z_j}(t, Z_t) S_i^t \sigma S_i \\
- \Theta_2(t, \lambda_t, Y_t) \left( b_{z_{m+1}}(t, Z_t) \sigma_\lambda(t, \lambda_t) \sigma_L(t, \lambda_t, Y_t) + b_{z_{m+2}}(t, Z_t) \sigma^2_L(t, \lambda_t, Y_t) \right)
\]
2.5. EXPLICIT SOLUTION

Observe that (2.5.8) is a linear PIDE and therefore solvable. In the sequel, we present a Feynman-Kac solution [Kromer et al. (2015)]:

\[ b(t, z) = E^p_{t,z} \left[ \int_t^T \Theta_1 \tilde{\mu} + \Theta_2 (s, \lambda^*_s, Y^*_s) \nu_L(s, \lambda^*_s, Y^*_s) \, ds \right] = \frac{\Theta_1 \tilde{\mu}}{\gamma} (T - t) + E^p_{t,z} \left[ \int_t^T \Theta_2 (s, \lambda^*_s, Y^*_s) \nu_L(s, \lambda^*_s, Y^*_s) \, ds \right]. \]  

Note that this form of the function \( b \) implies that the last two terms inside
the brackets in formula for \( u^*_S(t) \) given by (2.5.5) are equal to zero. The dynamics of \( Z^* \) under the measure \( \mathbb{P}^* \) read

\[
dZ^*_t = \left( \mu(t, Z^*_t) - \phi^{(1)}(t, Z^*_t) - \int_{\mathbb{R}^{k+1}\setminus\{0\}} \Delta Z^*_t(x, \bar{x}) C(t, \lambda^*_t, Y^*_t, x, \bar{x}) \vartheta_{X,\bar{X}}(dx, d\bar{x}) \right) dt
+ \sigma(t, Z^*_t) \, dW^*_t + \int_{\mathbb{R}^{k+1}\setminus\{0\}} \Delta Z^*_t(x, \bar{x}) \, J^*_t(x, \bar{x}) \, \tilde{J}_{X,\bar{X}}(dt, dx, d\bar{x}),
\]

with \( Z_0 = (S_0, \lambda_0, Y_0)^T \in \mathbb{R}^{m+2} \). The density process

\[
\Phi_t = \frac{d\mathbb{P}}{d\mathbb{P}^*} \bigg|_{\mathcal{F}_t}
\]

solves the SDE

\[
d\Phi_t = \Phi_t \psi^{(1)}(t, Z_t) \, d\tilde{W}_t + \Phi_t - \int_{\mathbb{R}^{k+1}\setminus\{0\}} \psi^{(2)}(t, Z_t, x, \bar{x}) \, \tilde{J}_{X,\bar{X}}(dt, dx, d\bar{x}),
\]

with \( \Phi_0 = 1 \). Moreover, \( \psi^{(1)}(t, z) \) is given as solution of the system of equations

\[
\sigma(t, z) \cdot \psi^{(1)}(t, z) = -\phi^{(1)}(t, z).
\]

Note that there exists at least one solution to this system because \( m \leq d \) (cf. the explanations preceding (2.2.2)) and \( \lambda \) and \( Y \) are driven by the same Brownian motion \( \tilde{W} \). In particular, it holds that \( \psi^{(1)}_{d+1}(t, Z_t) = \frac{-\phi^{(1)}_{m+1}(t, Z_t)}{\sigma(t, \lambda_t)} = \frac{-\phi^{(1)}_{m+1}(t, Z_t)}{\sigma(t, \lambda_t, Y_t)} \). Provided that \( C(t, \lambda_t, Y_t, x, \bar{x}) < 1 \) for Lebesgue-almost all \( t \in [0, T] \) and \( \vartheta_{X,\bar{X}}(dx, d\bar{x}) \)-a.s., it holds that

\[
\psi_2(t, Z_t, x, \bar{x}) = -C(t, \lambda_t, Y_t, x, \bar{x}).
\]
If \( \Phi \) is a positive martingale (see e.g. [Kazamaki (1979)]), then, according to the Girsanov theorem, \( \mathbb{P}^{*} \) is equivalent to \( \mathbb{P} \) and

\[
d\hat{W}_t = -\psi_1(t, Z_t) \, dt + dW_t^*,
\]

\[
\vartheta^*_{X,\tilde{X}}(dx, d\bar{x}) = (1 - C(t, \lambda_t, Y_t, x, \bar{x})) \, \vartheta_{X,\tilde{X}}(dx, d\bar{x}).
\]

Finally, we can represent the solution to the PIDE (2.5.6) as

\[
B(t, z) := \mathbb{E}_{t,z} \left[ \int_t^T \left( e^{r(T-s)} \left( \bar{\mu}^\top u_S^*(s) + u_Y^*(s) \nu_L(s, \lambda_s, Y_s) \right)
\right.
\]

\[
- \frac{\gamma}{2} \left( u_S^*(s)^\top \bar{\sigma} S u_S^*(s) + u_Y^*(s)^2 \sigma_L^2(s, \lambda_s, Y_s) \right) e^{2r(T-s)}
\]

\[
- \gamma e^{r(T-s)} \nabla b(s, z)^\top \tilde{Q} u^*(s, Z_s)
\]

\[
- \frac{\gamma}{2} \int_{\mathbb{R}^{k+1}\setminus\{0\}} \left( e^{r(T-s)} \Delta P_s^u(x, \bar{x}) + b(s, Z_{s-} + \Delta Z_s(x, \bar{x}))
\right.
\]

\[
- b(s, Z_{s-})^2 \vartheta_{X,\tilde{X}}(dx, d\bar{x}) \bigg] ds.
\]

Note that the expectation in (2.5.11) is calculated under the physical measure \( \mathbb{P} \). We summarize the most important part of the previous discussion in the following theorem:

**Theorem 2.5.1.** Consider the extended HJB system (2.3.5) for the case \( D \equiv 0 \). For any \( t \in [0, T) \) the optimal amounts to be invested in the stocks and the longevity asset are given by

\[
u_S^*(t) = \frac{\tilde{\mu}(\tilde{\sigma}_S + \tilde{\rho}_S \xi)^{-1}}{\gamma e^{r(T-t)}},
\]

(2.5.12)
CHAPTER 2. MV HEDGING OF UNIT LINKED LIFE INSURANCE

\[ u^*_Y(t) = (\gamma e^{(T-t)}(\sigma^2_L(t, \lambda_t, Y_t) + \eta_L(t, \lambda_t, Y_t)))^{-1} \cdot \left( \nu_L(t, \lambda_t, Y_t) 
- \gamma (b_{zm+1}(t, Z_t) \sigma_\lambda(t, \lambda_t) \sigma_L(t, \lambda_t, Y_t) + b_{zm+2}(t, Z_t) \sigma^2_L(t, \lambda_t, Y_t) Y_t) 
- \gamma b_2(t, Z_t) \right), \]

(2.5.13)

with the function \( b \) given by (2.5.9), while \( b_2 \) is defined by (2.5.4). In addition, the equilibrium value function \( V \) decomposes into

\[ V(t, p, z) = A(t)p + B(t, z), \]

with \( A(t) = e^{r(T-t)} \) and \( B(t, z) \) given by (2.5.11). The expected optimal terminal wealth \( g^{u*} \) decomposes into

\[ g^{u*}(t, p, z) = a(t)p + b(t, z), \]

with \( a(t) = e^{r(T-t)} \) and \( b(t, z) \) given by (2.5.9).

2.6 Numerical Results

In this part we exemplify Theorem 2.5.1. The pricing of the longevity asset is done under some pricing measure \( Q \) while the optimization is performed under the objective measure \( \mathbb{P} \). Hence, we need to know the dynamics of \( \lambda \) and \( Y \) under both measures. Further, we need to choose a process modeling the force of mortality \( \lambda \) that is nonnegative a.s. The Cox-Ingersoll-Ross (CIR) process supplemented by positive jumps (we refer to it as JCIR process in the sequel) is a good candidate for several reasons. The CIR process cannot become negative and belongs to the class of affine models allowing for a closed-form formula of the zero-bond price and the JCIR process preserves these properties (Brigo and Mercurio (2006)). Hence, we assume in this section that our previously outlined theory still holds when modeling the force of mortality by a JCIR process. The positive jumps are modeled by a compound Poisson process. Thereby the number of jumps is counted by the homogeneous Poisson process \( \tilde{N} = (\tilde{N}_t)_{t \in [0, T]} \) with constant intensity \( g_\lambda > 0 \)
and the jump sizes are independent and follow an exponential distribution with mean \( \zeta > 0 \). Thus, the Lévy measure is given by \( \vartheta_X(d\bar{x}) = \varrho \lambda f(d\bar{x}) \), with \( f \) denoting the probability density function of an exponentially distributed random variable. Let \( \psi_1(t,Z_t) = \kappa \sqrt{\lambda_t}, \kappa > 0 \), be the market price of Brownian risk in the longevity market and denote by \( \psi_2 > -1 \) the market price of jump risk. For parameters \( \beta, \sigma, \theta > 0 \) and \( \bar{\theta} := \theta + \frac{(1+\psi_2)\varrho \lambda}{\beta} \), consider the dynamics of \( \lambda \) given by

\[
\frac{d\lambda_t}{\lambda_t} = \beta(\bar{\theta} - (\beta + \kappa \sigma \lambda) \lambda_t - \psi_2 \varrho \lambda) \, dt + \sigma \sqrt{\lambda_t} \, d\bar{W}_t + \int_{\mathbb{R} \setminus \{0\}} \bar{x} \, \tilde{J}_X(dt, d\bar{x}), \tag{2.6.1}
\]

with \( \lambda_0 > 0 \). Defining

\[
\frac{d\tilde{W}_t}{\tilde{W}_t} = dW_t^Q + \kappa \sqrt{\lambda_t} \, dt,
\]

\[
\frac{d\vartheta^Q_X(d\bar{x})}{\vartheta^Q_X(d\bar{x})} = (1 + \psi_2) \, \vartheta_X(d\bar{x}), \tag{2.6.2}
\]

a straightforward calculation shows that the \( Q \)-dynamics of \( \lambda \) reads

\[
\frac{d\lambda_t}{\lambda_t} = \beta(\bar{\theta} - \lambda_t) \, dt + \sigma \sqrt{\lambda_t} \, dW_t^Q + \int_{\mathbb{R} \setminus \{0\}} \bar{x} \, J_X(dt, d\bar{x}), \tag{2.6.3}
\]

\( \lambda_0 > 0 \), which is the classical JCIR model under \( Q \). From this we can easily deduce that an appropriate choice of parameters and the starting value \( \lambda_0 \) preserves nonnegativity. Note that we have specified the market prices of risk such that the model is tractable under both measures. In particular, the representation \( 2.6.1 \) is consistent with the general setup in \( 2.2.3 \), and \( 2.6.3 \) is the starting point for pricing. For \( r \geq 0 \) and \( t \in [0,T) \), we are interested in the price of the zero-coupon longevity bond

\[
L_\lambda(t,T) = \mathbb{E}_Q \left[ e^{-\int_t^T (\lambda_s + r) \, ds} \bigg| \mathcal{F}_t \right].
\]

Defining the auxiliary process

\[
\tilde{L}_\lambda(t,T) := e^{r(T-t)} L_\lambda(t,T) = \mathbb{E}_Q \left[ e^{-\int_t^T \lambda_s \, ds} \bigg| \mathcal{F}_t \right],
\]
the affine structure of (2.6.3) yields that

\[ \tilde{L}_\lambda(t, T) = A_\lambda(t, T) \alpha_\lambda(t, T) e^{-B_\lambda(t, T) \lambda_t}, \]  

(2.6.4)

for deterministic functions \( A_\lambda, \alpha_\lambda, \) and \( B_\lambda \) to be found in Chapter 22 of Brigo and Mercurio (2006). Recall that the dollar value at time \( t \) of an investment in \( L_\lambda \) at time \( t = 0 \) is given by

\[ Y_t = e^{-\int_0^t \lambda_s \, ds} L_\lambda(t, T). \]

The next proposition characterizes the dynamics of \( Y \) assuming that \( \lambda \) is modeled by a JCIR process.

**Proposition 2.6.1.** Consider the process \( \lambda \) given by (2.6.1) and the specification (2.6.2) of the pricing measure \( Q \). The dollar value process \( Y \) of an investment in the longevity asset \( L_\lambda \) at time \( t = 0 \) is given by

\[ \frac{dY_t}{Y_t} = \left( r + B_\lambda(t, T) \sigma_\lambda \kappa \lambda_t - \psi_2 \int_{R \setminus \{0\}} (e^{-B_\lambda(t, T) x} - 1) d\tilde{X}(d\bar{x}) \right) dt 
- B_\lambda(t, T) \sigma_\lambda \sqrt{\lambda_t} \, d\tilde{W}_t + \int_{R \setminus \{0\}} (e^{-B_\lambda(t, T) x} - 1) \tilde{J}_\lambda(dt, d\bar{x}), \]  

(2.6.5)

with \( Y_0 = L_\lambda(0, T) \).

**Proof.** Using (2.6.4), we define \( f(t, \lambda_t) = \tilde{L}_\lambda(t, T) \). Then a standard calculation shows that

\[ df(t, \lambda_t) = \lambda_t f(t, \lambda_t) dt + f'(t, \lambda_t) \sigma_\lambda \sqrt{\lambda_t} \, dW_t^Q 
+ \int_{R \setminus \{0\}} (f(t, \lambda_t + \bar{x}) - f(t, \lambda_t -)) \tilde{J}_\lambda^Q(dt, d\bar{x}). \]  

(2.6.6)

We easily see from (2.6.4) that

\[ f'(t, \lambda_t) = -B_\lambda(t, \lambda_t) f(t, \lambda_t), \]

\[ f(t, \lambda_t + \bar{x}) - f(t, \lambda_t -) = f(t, \lambda_t -) \left( e^{-B_\lambda(t, T) \bar{x}} - 1 \right). \]

Plugging this back into (2.6.6) and translating \( df \) to the measure \( P \), we obtain

\[ df(t, \lambda_t) = f(t, \lambda_t -) \left( \lambda_t + B_\lambda(t, T) \sigma_\lambda \kappa \lambda_t \right) \]
2.6. NUMERICAL RESULTS

\[-\psi_2 \int_{\mathbb{R}\setminus\{0\}} \left( e^{-B_\lambda(t,T)x} - 1 \right) \vartheta_X (dx) \, dt \]

\[-B_\lambda(t,T)\sigma_\lambda \sqrt{\lambda_t} \, dW_t + \int_{\mathbb{R}\setminus\{0\}} \left( e^{-B_\lambda(t,T)x} - 1 \right) \tilde{J}_X (dt, dx) \].

Integration by parts then yields (2.6.5). We conclude the proof by remarking that the function $B_\lambda$ is suitably integrable.

We further deduce from (2.6.5) (cf. (2.2.5)) that

$$
\nu_L(t, \lambda_t, Y_t) = \nu_L(t, \lambda_t) = B_\lambda(t, T) \sigma_\lambda \lambda_t - \psi_2 \int_{\mathbb{R}\setminus\{0\}} \left( e^{-B_\lambda(t,T)x} - 1 \right) \vartheta_X (dx),
$$

$$
\sigma_L(t, \lambda_t, Y_t) = \sigma_L(t, \lambda_t) = -B_\lambda(t, T) \sigma_\lambda \sqrt{\lambda_t},
$$

$$
\eta_L(t, \lambda_t, Y_t, \bar{x}) = \eta_L(t, \bar{x}) = e^{-B_\lambda(t,T)x} - 1,
$$

$$
\tilde{\eta}_L(t, \lambda_t, Y_t) = \tilde{\eta}_L(t) = \int_{\mathbb{R}\setminus\{0\}} \eta_L(t, \bar{x})^2 \vartheta_X (dx).
$$

In order to implement the strategies $u^*_S$ and $u^*_Y$, we need to specify the function $b(t, z) = b(t, \lambda)$ from (2.5.9). One can calculate that the density process (2.5.10) is given as solution of the SDE

$$
\frac{d\Phi_t}{\Phi_{t-}} = -\frac{\nu_L(t, \lambda_t)}{\sigma_L^2(t, \lambda_t) + \tilde{\eta}_L(t)} \sigma_L(t, \lambda_t) \, dW_t - \frac{\nu_L(t, \lambda_t)}{\sigma_L^2(t, \lambda_t) + \tilde{\eta}_L(t)} \int_{\mathbb{R}\setminus\{0\}} \eta_L(t, \bar{x}) \, \tilde{J}_X (dt, dx),
$$

(2.6.7)

$\Phi_0 = 1$, i.e., $\Phi_t$ is the stochastic exponential of the integrated right-hand side of (2.6.7).

For simplicity we restrict to the case $d = k = 1$ for the simulation, that is, there is only one stock traded on the market. We model the jumps of the stock price process by a homogeneous Poisson process $N = (N_t)_{t \in [0,T]}$ with intensity $\varrho_S > 0$. Denote the compensated version by $\tilde{N} = (\tilde{N}_t)_{t \in [0,T]}$. The dynamics of the stock price then reads

$$
\frac{dS_t}{S_{t-}} = \mu \, dt + \sigma \, dW_t + \rho \, d\tilde{N}_t,
$$

(2.6.8)
Table 2.1: Parameter values

<table>
<thead>
<tr>
<th>$P_0$</th>
<th>$T$</th>
<th>$r$</th>
<th>$\gamma$</th>
<th>$S_0$</th>
<th>$\mu$</th>
<th>$\sigma$</th>
<th>$\rho$</th>
<th>$\phi_S$</th>
<th>$\beta$</th>
<th>$\sigma_{\lambda}$</th>
<th>$\theta$</th>
<th>$\kappa$</th>
<th>$\psi_2$</th>
<th>$\phi_{\lambda}$</th>
<th>$\lambda_0$</th>
<th>$\varsigma$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>10</td>
<td>.02</td>
<td>2</td>
<td>1</td>
<td>.06</td>
<td>.1</td>
<td>.1</td>
<td>.4</td>
<td>.3</td>
<td>.1</td>
<td>.2</td>
<td>-.2</td>
<td>.5</td>
<td>.05</td>
<td>.001</td>
<td></td>
</tr>
</tbody>
</table>

Table 2.2: Expectation and Variance

\[
\begin{array}{c|c|c|c|c|}
\text{E}[S_T] & \text{Var}[S_T] & \text{E}[\lambda_T] & \text{Var}[\lambda_T] \\
1.822 & 1.633 & 0.096 & 0.0102 \\
\end{array}
\]

Table 2.3: Expectation and Variance with different horizons

with $S_0 > 0$. Finally, we see that

\[
\Theta_1 = \frac{\tilde{\mu}}{\sigma + \rho \xi};
\]

\[
\Theta_2(t, \lambda_t, Y_t) = \Theta_2(t, \lambda_t) = \frac{\nu_L(t, \lambda_t)}{\sigma_L^2(t, \lambda_t) + \eta_L(t)}.
\]

Therefore the function $b$ is given by

\[
b(t, \lambda_t) = \frac{\Theta_1 \tilde{\mu}}{\gamma} (T - t) + \mathbb{E}_{t, \lambda_t} \left[ \frac{\Phi_T}{\Phi_t} \int_t^T \frac{\Theta_2(s, \lambda_s) \nu_L(s, \lambda_s)}{\gamma} ds \right],
\]

and Theorem 2.5.1 implies that the optimal strategies in this market setup read

\[
u_S^*(t) = \frac{\tilde{\mu}}{(\sigma^2 + \rho^2 \phi_S) \gamma e^{r(T-t)}};
\]

\[
u_Y^*(t) = \left( (\sigma_L^2(t, \lambda_t) + \eta_L(t)) \gamma e^{r(T-t)} \right)^{-1} \left( \nu_L(t, \lambda_t) - \gamma b_L(t, \lambda_t) \sigma_{\lambda} \sqrt{\lambda_t} \sigma_L(t, \lambda_t) \right.
\]

\[ - \gamma \int_{\mathbb{R} \setminus \{0\}} \left( b(t, \lambda_t + \bar{x}) - b(t, \lambda_t) \right) \eta_L(t, \bar{x}) \vartheta_{\bar{x}}(d\bar{x}) \right). \]
2.6. NUMERICAL RESULTS

\[
\begin{array}{ccccccc}
\text{(A)} & \text{(B)} & \text{(C)} & \text{(D)} & \text{(E)} & \text{(F)} \\
\mathbb{E}[P_{T_l}^*] & 1.4244 & 1.4377 & 1.4399 & 1.4365 & 1.4379 & 1.4373 \\
\text{ROER} & 3.54\% & 3.63\% & 3.65\% & 3.62\% & 3.63\% & 3.63\% \\
\text{Var}[P_{T_l}] & 0.1009 & 0.1128 & 0.1077 & 0.1079 & 0.109 & 0.1083 \\
\end{array}
\]

Table 2.4: Expectation and Variance: (A): without longevity asset; (B): ignoring jumps; (C): Brownian risk only; (D): std. normally distributed jump sizes of \( S \); (E): \( T_L = 15 \); (F): \( T_L = 25 \).

In Table 2.1 the assigned parameter values are summarized. We chose values that are typical in the literature. This leads to the expectations and variances of respectively the stock price and the force of mortality given in Table 2.2. These quantities have been calculated using the formulas given by Lemma 2.4.1. In order to compare the expected payoffs under different scenarios, we consider the rate of expected return (ROER) given by \[ \text{ROER} = \frac{\ln(\mathbb{E}[P_{T_l}^*])}{T}. \] Following the optimal strategies (2.6.9) and (2.6.10) yields the expectation, ROER and variance of the optimal terminal wealth displayed in the left panel of Table 2.3. Any modification made in the sequel will be compared against these values. The other two panels in Table 2.3 show the expected value, ROER and variance of the terminal wealth when increasing the horizon to respectively 15 and 25 years. We see that the expected terminal payoffs and ROERs lie significantly above the final payoffs one would receive from investing in the riskless asset solely. The overall variance is of course slightly increasing when the investment horizon is prolonged, for that reason the insurance company is induced to invest less in the risky asset causing a minor decrease in the ROER over time.

In Table 2.4 several further scenarios are analyzed. In Panel (A) the expectation, ROER and variance of the optimal terminal wealth without investing in the longevity asset are displayed. From (2.6.9) we see that the amount invested in the stock is the same as in the previously described case while the money allocated to the longevity asset before is now invested in the riskless asset. Comparing the values in Panel (A) of Table 2.4 to the left panel of Table 2.3 we see a slight decrease in the expected terminal wealth and in the ROER and a slight increase in the variance. Hence, the overall
effect of not investing in the longevity asset is relatively small. Next, in Panel (B), we investigate the change in mean, ROER and variance of the optimal terminal wealth when keeping \( \lambda \) and \( S \) as in (2.6.1) and (2.6.8) respectively, but erroneously assuming that \( \lambda \) and \( S \) do not exhibit jumps. This corresponds to the scenario that an investor observes the variance of the stock and the force of mortality, but naively ascribes it to the Brownian components. Note that in this case the values of \( u^\star S \) do not change. We observe that the expected value (1.4377) and the ROER (3.63\%) are not significantly different compared to the benchmark, while the variance increased to 0.1128, which corresponds to a rise of 7.4\%. This slight increase in the variance stems from the fact that the hedging strategy does no longer account for the presence of jumps. However, the effect is relatively small, which is in line with our previous findings in Panel (A) of Table 2.4 because the impact of the investment in the longevity asset under the allocation rule \( u^\star Y \) on the mean, ROER and variance of the optimal terminal wealth is relatively low in general. Hence, buying the longevity asset on average yields a higher payoff than the riskless investment and leads to some diversification. When hedging a terminal condition linked to the mortality rate in the insurance pool as discussed in Example 2.3.1, the effect observed is likely to be stronger. In addition, our findings indicate that erroneously ignoring the jumps the force of mortality exhibits would then also lead to a significantly higher variance.

In Panel (C) of Table 2.4 we investigate the effect of setting the jump intensity of \( S \) and \( \lambda \) to zero, so all uncertainty is stemming from Brownian risk. Thereby the expected values and variances of \( S_T \) and \( \lambda_T \) have been kept stable at the values depicted in Table 2.2 by adjusting the volatility parameters of the respective Brownian parts. However, the mean of the optimal terminal wealth just changed to 1.4399, the ROER marginally increased to 3.65\%, while the variance of the optimal terminal wealth slightly changed to 0.1077. The same phenomenon is shown in Panel (D): we replaced the constant jump of the Poisson process \( N \) in the dynamics of \( S \) by standard normally distributed jump sizes while keeping mean and variance of \( S \) the same: we obtained a mean of 1.4365, an ROER of 3.62\% and a variance of 0.1079. Thus, we tentatively conclude that as long as we know mean and
variance of the stock and the force of mortality, the expected optimal terminal wealth, the ROER and the variance are robust.

So far we assumed that the time to maturity of the longevity asset coincides with the insurance horizon. We investigated the change of mean, ROER and variance of the optimal terminal wealth when the time to maturity of the longevity asset, say $T_L$, is longer than the insurance horizon, which has been kept fixed at $T = 10$. The results are displayed in Panel (E) and Panel (F) of Table 2.4. We see that neither the expected value, the ROER nor the variance of the terminal optimal wealth change significantly. This result is highly important from a practical point of view because one cannot expect to find longevity assets whose times to maturity coincide with the insurance horizon, there are too few of them offered and traded. Our results show that picking an asset with a longer time to maturity does in particular not add to the variance of the terminal wealth.

Finally, Figure 2.1 displays a path of the optimal portfolio process. The jumps of the underlying processes are clearly visible and the path shows a positive trend. In Figure 2.2, paths of the optimal dollar amounts to be invested are plotted. From (2.6.5) we can see that the excess rate of return can become negative inducing the return of the longevity asset to be below the riskless rate. Thus, taking a short position in $Y$ at several points in time is optimal and this can be clearly seen in Panel (B). Furthermore, as maturity is approaching, the optimal amount invested in the longevity asset is increasing because the function $B_{\lambda}(\cdot, T)$ in (2.6.5) is tending to zero as $t$ approaches $T$. The economic meaning for this investment behavior is that close to maturity, the price of the longevity bond converges to 1 (cf. (2.2.4)) and does not exhibit much variation anymore. Thus, the investment in the longevity bond becomes less risky.

We conclude by remarking that a numerical analysis incorporating the hedging of a terminal condition should be conducted. In this case, closed-form solutions to the extended HJB system (2.3.5) cannot be obtained anymore, so one needs to resort to more involved numerical methods. We leave this direction for future research.
Figure 2.1: Optimal Portfolio Process

(a) Stock

(b) Longevity Asset

Figure 2.2: Optimal Dollar Amounts
2.A Appendix

We provide the formulas to calculate the moments of $\lambda_T$ displayed in Table 2.2.

**Lemma 2.A.1.** Consider the JCIR process given by (2.6.1), i.e.,

$$d\lambda_t = [\beta \tilde{\theta} - (\beta + \kappa \sigma_\lambda)\lambda_t - \psi_2 \varrho \lambda_\varsigma] \, dt + \sigma_\lambda \sqrt{\lambda_t} \, d\bar{W}_t + \int_{\mathbb{R} \setminus \{0\}} \bar{x} \, \tilde{J}_X(dt, d\bar{x}),$$

with $\lambda_0 > 0$. We have that

$$E[\lambda_t] = \frac{\beta \theta}{\beta + \kappa \sigma_\lambda} \left(1 - e^{-(\beta + \kappa \sigma_\lambda)T}\right) + \lambda_0 e^{-(\beta + \kappa \sigma_\lambda)T} + \frac{\varrho \lambda_\varsigma}{\beta + \kappa \sigma_\lambda} \left(1 - e^{-(\beta + \kappa \sigma_\lambda)T}\right),$$

$$\text{Var}[\lambda_t] = \frac{\beta \sigma_\lambda^2 \theta}{2(\beta + \kappa \sigma_\lambda)^2} \left(1 - e^{-(\beta + \kappa \sigma_\lambda)T}\right) + \frac{\lambda_0 \sigma_\lambda^2}{\beta + \kappa \sigma_\lambda} e^{-(\beta + \kappa \sigma_\lambda)T} \left(1 - e^{-(\beta + \kappa \sigma_\lambda)T}\right)$$

$$+ \frac{\varrho \lambda_\varsigma^2}{\beta + \kappa \sigma_\lambda} \left(1 - e^{-2(\beta + \kappa \sigma_\lambda)T}\right) + \frac{\sigma_\lambda^2 \varrho \lambda_\varsigma}{\beta + \kappa \sigma_\lambda} \left(1 - e^{-(\beta + \kappa \sigma_\lambda)T}\right)$$

$$+ \frac{\sigma_\lambda^2 \varrho \lambda_\varsigma}{2(\beta + \kappa \sigma_\lambda)^2} \left(e^{-2(\beta + \kappa \sigma_\lambda)T} - 1\right).$$

(2.A.1)

**Proof.** Denote the imaginary unit by $i$. For notational simplicity, we define the following:

$$a := \beta + \kappa \sigma_\lambda,$$

$$b := \frac{\beta \theta}{a},$$

$$g(t, y) := 1 + y \tilde{g}(t),$$

$$\tilde{g}(t) := -\frac{\sigma_\lambda^2}{2a} \left(1 - e^{-aT}\right),$$

$$\psi(t, y) := \frac{ye^{-at}}{g(t, y)},$$

$$f_1(t, y) := g(t, y) \frac{2ab}{\pi^2},$$

$$f_2(t, y) := g(t, y) \frac{2ab}{\pi^2}.$$
\[ f_2(t, y) := \exp \left( \int_0^t \int_0^\infty \left( e^{\bar{x}\psi(s, y)} - 1 \right) \vartheta_X(d\bar{x})ds \right). \]

According to Jin et al. (2016), the characteristic function of \( \lambda_t \) reads
\[ \mathbb{E}[e^{y\lambda_t}] = f_1(t, y) e^{\lambda_0\psi(t, y)} f_2(t, y). \]

So it holds that
\[
\frac{\partial}{\partial y} \mathbb{E}[e^{y\lambda_t}] = f_1(t, y) e^{\lambda_0\psi(t, y)} f_2(t, y) \left[ \frac{2abg'(t, y)}{\sigma_s^2 g(t, y)} + \frac{\lambda_0 e^{-at}}{g(t, y)^2} \right. \\
\left. + \int_0^t \int_0^\infty \frac{e^{\bar{x}\psi(s, y) - as\bar{x}}}{g(s, y)^2} \vartheta_X(d\bar{x})ds \right] \\
= \mathbb{E}[e^{y\lambda_t}] \cdot \left[ \frac{b(1 - e^{-at})}{g(t, y)} + \frac{\lambda_0 e^{-at}}{g(t, y)^2} + \int_0^t \int_0^\infty \frac{e^{\bar{x}\psi(s, y) - as\bar{x}}}{g(s, y)^2} \vartheta_X(d\bar{x})ds \right], \tag{2.A.2}
\]

and from this we easily deduce the first moment of \( \lambda_t \):
\[
\mathbb{E}[\lambda_t] = (-i)^1 \cdot \frac{\partial}{\partial y} \mathbb{E}[e^{y\lambda_t}] \bigg|_{y=0} = (-i)^1 \cdot 1 \cdot \left[ b(1 - e^{-at}) + \lambda_0 e^{-at} + \frac{\theta_\lambda S}{a} (1 - e^{-at}) \right] \\
= b(1 - e^{-at}) + \lambda_0 e^{-at} + \frac{\theta_\lambda S}{a} (1 - e^{-at}).
\]

Re-substitution of the abbreviations \( a, b \) yields \( \mathbb{E}[\lambda_t] \) given by (2.A.1). To calculate the second moment of \( \lambda_t \), we need to determine the second derivative of the characteristic function w.r.t. \( y \). The structure of the first derivative given by (2.A.2) will turn out useful. For notational simplicity, we abbreviate the term in square brackets in (2.A.2) by \([...]\) in the sequel. We find
\[
\frac{\partial^2}{\partial y^2} \mathbb{E}[e^{y\lambda_t}] = \frac{\partial}{\partial y} \left( \mathbb{E}[e^{y\lambda_t}] \cdot [...]\right) \]
2.A. APPENDIX

\[
\frac{\partial}{\partial y} \mathbb{E}[e^{y\lambda t}] \cdot [...] + \mathbb{E}[e^{y\lambda t}] \cdot \left( \frac{\partial}{\partial y} [...] \right)
\]

\[
= \mathbb{E}[e^{y\lambda t}] \cdot [...] + \mathbb{E}[e^{y\lambda t}] \cdot \left( \frac{\partial}{\partial y} [...] \right)
\]

\[
= \mathbb{E}[e^{y\lambda t}] \left( [...] + \frac{\partial}{\partial y} [...] \right)
\]

We need to calculate that

\[
\frac{\partial [...]}{\partial y} = b(1 - e^{-at})\tilde{g}(t) + \frac{2\lambda_0 e^{-at}\tilde{g}(t)}{g(t, y)^2}
\]

\[
= \int_0^t \int_0^\infty \left( x^2 \frac{e^{x\psi(s,y)-2as}}{g(s, y)^2} - \frac{2 e^{x\psi(s,y)-as}g(s)}{g(s, y)^3} \right) \vartheta(x) (d\bar{x}) ds,
\]

in order to argue that the second moment is given by

\[
\mathbb{E}[(\lambda_t)^2]
\]

\[
= (-1)^2 \cdot \frac{\partial^2}{\partial y^2} \mathbb{E}[e^{y\lambda t}] \bigg|_{y=0}
\]

\[
= \left( - \mathbb{E}[\lambda_t]^2 + b(1 - e^{-at})\tilde{g}(t) + 2\lambda_0 e^{-at}\tilde{g}(t) + \frac{\xi^2 \vartheta_\lambda}{a} (e^{-2at} - 1) + 2 \frac{\vartheta_\lambda}{a} \tilde{g}(t)
\]

\[
- \frac{\sigma^2 \vartheta_\lambda}{2a^2} (e^{-2at} - 1) \right)
\]

\[
= \mathbb{E}[\lambda_t]^2 - b(1 - e^{-at})\tilde{g}(t) - 2\lambda_0 e^{-at}\tilde{g}(t) - \frac{\xi^2 \vartheta_\lambda}{a} (e^{-2at} - 1) - 2 \frac{\vartheta_\lambda}{a} \tilde{g}(t)
\]

\[
+ \frac{\sigma^2 \vartheta_\lambda}{2a^2} (e^{-2at} - 1).
\]

Finally, subtraction of the square of the first moment from the previous expression and re-substitution of the abbreviations \(a, b, \tilde{g}(t)\) gives the formula for the variance of \(\lambda_t\) provided by (2.A.1). \(\square\)
Chapter 3

Extensions of Mean-Variance Portfolio Selection

3.1 Introduction

It is commonly appreciated when the identification of risk is intuitive, easily understandable and mathematically tractable. These are the major reasons that in theory and practice the risk of an investment portfolio is frequently measured in terms of the variance of its terminal value. Additionally, in a mean-variance problem, the tradeoff between risk and return is explicit in the sense that the interplay between gains in the expected payoff at the expense of increases in the variance is directly observable. Due to the symmetry of the second central moment, however, the variance punishes potential gains and losses equally. Clearly, unexpected gains are highly acknowledged and it is the downside risk that investors are concerned about. In this paper we present an alternative optimization criterion where the variance is substituted by the (accumulated) local volatility of the terminal payoff. This approach combines two important aspects: First of all, it is a non-symmetric criterion reflecting investors’ preferences typically better. Second, we entirely preserve the intuitive notion of the mean-variance problem by using widely known and easily understandable risk measures; furthermore, in the mean-local volatility setting the tradeoff between risk and return is still explicit in the
There are various portfolio optimization techniques circumventing a symmetric punishment of gains and losses. For instance, given some utility function, it is common practice to optimize the expected utility of the portfolio’s terminal payoff, see e.g. Merton (1969), Cox and Huang (1989), Mongin (1997), Goll and Kallsen (2000), Cohen and Natoli (2003), Pang (2006), Kraft et al. (2013); for detailed comparisons with the mean-variance approach we recommend Kroll et al. (1984). Utility functions are defined so that they properly mimic investors’ preferences, however, it is rather difficult to elicit a particular utility function from an investor. Further, the tradeoff between risk and return is implicit making this approach less intuitive. While commenting on the Capital Asset Pricing Model and the Intertemporal Capital Asset Pricing Model, Cochrane (2005) explains that bad news about future stock returns increases the marginal utility of wealth (and of consumption). Hence, following this line of thought, stocks and portfolios performing well on such news are appreciated (see Section 9.3 in Cochrane (2005)). The local volatility measures the correlation with the financial market. In fact, it does so ex-ante, i.e., as soon as one plugs in the terminal payoff. By allowing for a wide class of penalty functions in this paper, we can punish local movements of the terminal payoff in the same direction as the financial market more severely than payoffs being inversely related. This way portfolios reacting in an investor’s interest to negative news about the stock market return can be constructed easily. Obviously, the local-volatility approach bridges the gap between keeping an explicit risk-return tradeoff and formulating a criterion accounting for a higher marginal utility in recessions without the need of knowing an explicit utility function. Obtaining this property when employing other measures of risk is a highly involved procedure requiring many restrictive assumptions. For instance, the Value at Risk and Tail-Risk are also well known and extensively used asymmetric risk measures (see e.g. Jorion (2007) for details), but they are law-invariant through only depending on the distribution of the final payoff. Thus, they do not account for the correlation with the underlying financial market.

The mean-variance criterion is immovably linked to the name Markowitz
and his highly celebrated work Markowitz (1952), in which such a problem in one period is solved. Due to the quadratic non-linearity and non-recursiveness of the variance part, this criterion is a time-inconsistent one, see for instance the explanations and examples in Cui et al. (2012) for a discrete time and in Karnam et al. (2017) for a continuous time case. Loosely speaking, time-inconsistency causes a strategy that is optimal today to be potentially suboptimal in the future. Investors ignoring this suboptimality are said to pre-commit, see Strotz (1955) for a comprehensive elaboration on this term. According to Björk and Murgoci (2010), when facing a time-inconsistent problem one can either ignore it and follow the initially optimal policy or reformulate the problem. Extensions of Markowitz’s seminal work to the pre-commitment case in the multi-period setting are done by Smith (1967) in discrete time and Li and Zhou (2000) in continuous time, further generalizations are to be found in Li and Ng (2000), Li et al. (2002), Lim and Zhou (2002), Yin and Zhou (2003), Lim (2004), Bielecki et al. (2005), Xiong and Zhou (2007) and references in these papers. A reformulation in game-theoretic terms in order to circumvent time-inconsistency issues and obtain a dynamically optimal policy can for instance be found in Björk and Murgoci (2010), Basak and Chabakauri (2010), Czichowsky (2013), Björk et al. (2014) and Bosserhoff and Stadje (2019a). A time-consistent version of the mean-local-volatility problem is solved in Pistorius and Stadje (2017) under a short-selling constraint. Originally, the mean-local-volatility problem we consider in this paper belongs to the class of time-inconsistent stochastic control problems. Since Bellman’s dynamic programming approach (see e.g. Pham (2009) for details) fails on this class, different solution techniques have to be used. For instance Li and Zhou (2000) and Hu et al. (2012) follow a more general approach on stochastic linear-quadratic problems into which the mean-variance one can be embedded. Clearly, such a strategy does not work in our case because we punish asymmetrically. We instead discretize the mean-local-volatility problem. The finiteness of the arising filtration enables optimizing over the whole space, thus, we do not need a particular solution technique requiring time-consistency. We allow for arbitrary many trading time points per unit interval. Note that our solution approach is
highly useful from a practical and numerical point of view. Subsequently, we suggest a way to establish the convergence to the continuous time optimal strategy. Since our way of solving the problem is rather technical involving a lot of long-winded formulas, we restrict to Binomial and Brownian markets, respectively.

This paper is structured as follows: Section 2 provides the basic model and elaborates on the local volatility process of some terminal payoff. Moreover, the optimization problem is introduced and existence issues of the optimal solution are illustrated. Subsequently, in Section 3 the discretized version of the aforementioned mean-local-volatility optimization problem is considered. First, the local volatility process in discrete time is derived. Second, the problem is solved. Third, the solution is exemplified along a time-dependent version of the mean-variance problem. Section 4 contains conditions under which the optimal discrete time policies converge to their continuous time counterpart. Section 5 concludes and suggests future research directions.

**Notation.** Denote by \( \mathbb{R}_+ \) the positive real numbers. On some probability space \((\Omega, \mathcal{F}, \mathbb{P})\), equalities and inequalities between random variables are understood to hold \( \mathbb{P} \)-almost surely (a.s.). Any two random variables \( X \) and \( Y \) are identified if they are equal a.s.; if they are equal in distribution, we write \( X \overset{D}{=} Y \). We use the notation \( L^p(\Omega) \) for the space of \( \mathbb{R} \)-valued \( \mathcal{F} \)-measurable random variables \( X \) such that \( (\mathbb{E}[|X|^p])^{1/p} < \infty \), for \( p \in [1, \infty) \). Accordingly, for some \( T > 0 \), denote by \( L^p(\Omega \times [0, T]) \) the space of \( \mathbb{R} \)-valued processes \( X = (X_t)_{t \in [0, T]} \) such that \( (\mathbb{E}[\int_0^T X^p_t \, ds])^{1/p} < \infty \), for \( p \in [1, \infty) \). Define \( \sum_{i=m}^n x_i = 0 \) if \( m > n \) and call it empty sum. We write \( f(x) \in \mathcal{O}(g(x)) \) if and only if there exists some \( C > 0 \) and a real number \( x_0 \) such that \( |f(x)| \leq Cg(x) \) for all \( x \geq x_0 \). We denote \( x^+ = \max(x, 0) \) and \( x^- = \min(x, 0) \).

### 3.2 Model Setup

Let \( T > 0 \) be the planning horizon. Consider the probability space \((\Omega, \mathcal{F}, \mathbb{P})\) that is equipped with a standard Brownian motion \( W = (W_t)_{t \in [0, T]} \). Let
(\mathcal{F}) = (\mathcal{F}_t)_{t \in [0,T]}$ be the right-continuous completion of the natural filtration generated by $W$. The financial market under consideration is assumed to be frictionless and consists of two assets. The riskless asset is denoted by $B$ and satisfies

$$
 dB_t = r B_t \, dt,
$$

with $B_0 = 1$ and some constant interest rate $r \geq 0$. The risky asset we denote by $S$ and it is given as solution to the SDE

$$
 \frac{dS_t}{S_t} = \mu \, dt + \sigma \, dW_t,
$$

with $S_0 > 0, \mu \in \mathbb{R}$ and $\sigma > 0$. Note that this financial market corresponds to the standard Black-Scholes model and it is complete. Next we specify the set of admissible trading strategies.

**Definition 3.2.1.** We denote the set of admissible trading strategies on $[0, T]$ by $\mathcal{Y}$ and define

$$
 \mathcal{Y} := \{ y | y \text{ is } S\text{-integrable, predictable and } y \in L^2(\Omega \times [0,T]) \}.
$$

At any time $t$, the quantity $y_t$ indicates the absolute number of stocks an investor holds. The wealth process induced by the strategy $y = (y_t)_{t \in [0,T]}$ is denoted by $V^y = (V^y_t)_{t \in [0,T]}$, which is defined as solution to the SDE

$$
 dV^y_t = y_t \, dS_t + (V^y_t - y_t S_t) r \, dt,
$$

with initial wealth $V^y_0 = v_0 > 0$. The dynamics of $V^y$ given by (3.2.3) show that money is exchanged between $B$ and $S$ in a self-financing manner. Observe that the square integrability of $y$ ensures the existence and pathwise uniqueness of a solution to (3.2.3). Moreover, $\mathbb{E}[(V^y_t)^2] < \infty$ for all $t \in [0,T]$. This implies the existence of an $\mathbb{R}$-valued stochastic process $H^V = (H^V_t)_{t \in [0,T]}$ such that

$$
 V^y_T = \mathbb{E}[V^y_T] + \int_0^T H^V_t \, dW_t,
$$

(3.2.4)
where $H^V$ is predictable w.r.t. $(\mathcal{F})$, and it holds that $\int_0^T (H^V_t)^2 \, dt < \infty$ a.s. and $H^V \in L^2(\Omega \times [0, T])$ (cf. Jeanblanc et al. (2009), Theorem 11.2.8.1). In the literature the possibility to express $V^y_T$ as in (3.2.4) is known as predictable representation property. The Clark-Ocone formula characterizes $H^V$ as the Malliavin derivative of $V^y_T$, see Theorem 6.35 in Nunno et al. (2008). Informally speaking, the random variable $H^V_t$ reflects the local impact of a movement of $W$ in the infinitesimal time interval $[t, t + dt]$ on the terminal payoff. To see this, note that

$$E[V^y_T|\mathcal{F}_t] = E[V^y_T] + \int_0^t H^V_s \, dW_s,$$

and therefore

$$E[V^y_T|\mathcal{F}_{t+dt}] - E[V^y_T|\mathcal{F}_t] = H^V_t(W_{t+dt} - W_t). \quad (3.2.5)$$

Consequently, we name the process $H^V$ the local volatility of $V^y_T$. From (3.2.5) it becomes obvious that positive values of $H^V$ imply a positive correlation of the portfolio’s terminal wealth w.r.t. the driving Brownian motion $W$ and vice versa. Thus, for an investor identifying risk with the terminal wealth’s correlation with the stochastic drivers, using $H^V$ to quantify and control the risk is appropriate. The way in which risk is to be controlled is determined by the choice of the penalty function, which we generally define as follows:

**Definition 3.2.2.** A function $g : [0, T] \times \mathbb{R} \rightarrow [0, \infty)$ is called penalty function if it is strictly convex and differentiable in the second component.

A natural target is the achievement of a high expected terminal payoff while keeping control of the accumulated local volatility. To this end, consider an arbitrary penalty function $g$ and define the target functional, say $J$, by

$$J(y) := E\left(V^y_T - \gamma \int_0^T g(t, H^V_t) \, dt\right), \quad (3.2.6)$$

for some risk-aversion parameter $\gamma > 0$. 

Example 3.2.3. As outlined in the introduction to this paper, investors appreciate portfolios whose terminal payoffs are inversely related to the movements of the financial market. Identifying these with the increments of the driving Brownian motion, positive values of $H^V$ should be punished more strongly than negative ones. Thus, a natural choice for the penalty function is

$$g(\cdot, x) = a(x^+)^p + b(x^-)^q,$$

for $a > 0$, $b \geq 0$, $p, q \geq 1$ and $p \geq q$. We recommend Pistorius and Stadje (2017) for further examples for the choice of $g$.

The optimization problem over the time horizon $[0, T]$ boils down to the maximization of $J(y)$ over $y \in \mathcal{Y}$, in symbols:

$$\Psi_0 := \sup_{y \in \mathcal{Y}} J(y).$$  \hspace{1cm} (II)

A control $y^* \in \mathcal{Y}$ is consequently called optimal for Problem (II) if $J(y^*) = \Psi_0$. Denote the set of admissible controls over $[t, T]$ by $\mathcal{Y}_{[t, T]}$ and the corresponding target functional by $J_t$. The dynamic counterpart to Problem (II) is given by

$$\Psi_t := \text{ess sup}_{y \in \mathcal{Y}_{[t, T]}} J_t(y).$$  \hspace{1cm} (3.2.7)

Labeling the optimal control in (3.2.7) over $[t, T]$ by $y_t^{\star}$ and assuming for the moment it does exist, we categorize the optimization problem as follows.

**Definition 3.2.4.** We say that the optimal control problem in (3.2.7) is time-consistent if

$$y_{s}^{t, \star} = y_{s}^{\star}, \ 0 \leq t \leq s \leq T.$$  \hspace{1cm} (3.2.8)

If (3.2.8) fails to hold, the optimal control problem in (3.2.7) is said to be time-inconsistent.

From the Itō isometry it is easy to see that the maximization of (3.2.6) for the choice $g(\cdot, x) = x^2$ is the mean-variance portfolio selection problem, so it is a special case of Problem (II). Thus, we see that Problem (II) undoubtedly belongs to the class of time-inconsistent control problems (see for
instance Yong and Zhou (1999) or Pham (2009) for illustrations on the time-
inconsistency of the mean-variance criterion).

A natural way to approach the existence of $y^\star$ is to show that $-J$ is a
convex, lower-semicontinuous and coercive functional on the reflexive Banach
space $\mathcal{Y}$. Then, according to Corollary 3.23 in Brezis (2011), the existence of
a solution would follow. However, we have the following example.

Example 3.2.5. Consider a trivial probability space and suppose $X$ is living
on this space; then $X$ is a constant. Since

$$\text{Var}[X] - E[X] = -X,$$

for large $X$ the negative of the mean-variance target functional cannot be
bounded from below by the square of $X$.

Recalling that $V^y_T \in L^2(\Omega)$, we conclude from Example 3.2.5 that $-J$ is
not coercive in the special case that $g(\cdot, x) = x^2$. Needless to say, it cannot
be coercive for general penalty functions $g$ either. Thus, the aforementioned
method cannot be used to show that Problem (II) has an optimal solution
if it exists at all. Indeed, Karnam et al. (2017) explain that it is oftentimes
not known whether time-inconsistent problems actually possess a solution,
therefore, an assumption on the existence is typically found in the literature.

Time-inconsistency also causes the dynamic programming principle to fail
(Yong and Zhou (1999)). An intuitive starting point to solve Problem (II)
is finding a more explicit representation of $H^V$ to infer further character-
istics, but this is only possible in a few special cases, see Øksendal (2010)
for concrete examples; in particular, the current setting is significantly more
complicated than the examples in the just mentioned source. For these rea-
sons we approach the discretized version of Problem (II) first and then loose
some words about convergence.

### 3.3 Discrete Time Problem

Assume w.l.o.g. $T \equiv 1$ in this section. We still work with the continuously
compounded interest rate $r \geq 0$. A discrete time model for the stock price
is built by choosing a natural number \( n \in \mathbb{N} \) and constructing a binomial tree that takes \( n \) steps per unit time. Let \( t_0 = 0 \) and \( t_i = i/n \) for \( i \in \{1, \ldots, n\} \), so subsequent trading time points are equidistantly spaced, a generalization to non-equidistant time points is of course feasible. Let \( X_i^{(n)} \) be independent and identically distributed (i.i.d.) Bernoulli random variables such that \( \mathbb{P}(X_i^{(n)} = \pm 1) = 1/2 \) for every \( i \in \{1, \ldots, n\} \). Consider the scaled random walk \( R_i^{(n)} = (R_i^{(n)})_{t \in [0,1]} \) that is constant on each of the intervals \([t_i, t_{i+1})\) with components at time \( t_i \) given by

\[
R_i^{(n)} = \sum_{l=1}^{i} \frac{1}{\sqrt{n}} X_l^{(n)}, \quad i = 1, 2, \ldots, n. \tag{3.3.1}
\]

For fixed \( n \) denote the filtration generated by the random walk by \( (\mathcal{F}_t^{(n)})_{t \in [0,1]} \). Define the up-factor \( u^{(n)} \) and the down-factor \( d^{(n)} \) of the stock by

\[
u^{(n)} := \exp\left(\frac{\sigma}{\sqrt{n}} + \frac{\mu - \sigma^2}{2}/n\right), \quad d^{(n)} := \exp\left(-\frac{\sigma}{\sqrt{n}} + \frac{\mu - \sigma^2}{2}/n\right).
\]

The stock price \( S_{t_i}^{(n)} \) at time \( t_i \) then results from \( i \) movements of the random walk plus the drift multiplied by \( t_i \), i.e.,

\[
S_{t_i}^{(n)} = \exp\left(\frac{i}{n} X_i^{(n)} + (\mu - \sigma^2/2)t_i\right), \tag{3.3.2}
\]

and we set \( S_{t_0} = S_0 = 1 \) for convenience. For each \( i \in \{0, 1, \ldots, n-1\} \), denote by \( y_{t_i}^{(n)} \) the absolute amount of shares of \( S^{(n)} \) bought at time \( t_i \). We allow it to be random through dependence on \( S_{t_i}^{(n)} \), which is indicated by writing \( y_{t_i}^{(n)} (S_{t_i}^{(n)}) \). Such controls are known as feedback controls in the literature. Forcing the portfolio process \( V^{(n)} \) to be self-financing, it necessarily satisfies the recursion

\[
V_{t_{i+1}}^{(n)} = y_{t_i}^{(n)} (S_{t_i}^{(n)}) S_{t_{i+1}}^{(n)} + e^{r/n} \left( (V_{t_i}^{(n)} - y_{t_i}^{(n)} (S_{t_i}^{(n)}) S_{t_i}^{(n)}) \right), \quad i = 0, 1, 2, \ldots, n-1, \tag{3.3.3}
\]
CHAPTER 3. EXTENSIONS OF MV PORTFOLIO SELECTION

with initial value $V_0^{(n)} = v_0$. For the $\mathcal{F}_1^{(n)}$-measurable random variable $V_1^{(n)}$ we apply the predictable representation theorem for Bernoulli random walks stating the existence of a bounded $(\mathcal{F}^{(n)})$-predictable process $H^{(n)}$ such that

$$V_1^{(n)} = E[V_1^{(n)}] + \sum_{j=1}^n H_{t_j}^{(n)} X_j^{(n)}/\sqrt{n}. \quad (3.3.4)$$

Since $(\mathcal{F}^{(n)})$ is finite, we can calculate the process $H^{(n)}$ explicitly. To this end, we use the following algorithm that can be deduced from a close analysis of the proof of Theorem 9.43 in [Klenke (2014)].

**Algorithm 3.3.1.** For any stochastic process $\phi^{(n)} = (\phi_{t_i}^{(n)})_{i \in \{1,\ldots,n\}}$ adapted to $(\mathcal{F}^{(n)})$, set

$$\phi_{t_i}^{(n)} = \begin{cases} \phi_{t_i}^{(n,+)}, & \text{if } X_i^{(n)} = 1, \\ \phi_{t_i}^{(n,-)}, & \text{if } X_i^{(n)} = -1. \end{cases}$$

Define $\phi_1^{(n)} := V_1^{(n)}$. Recursively calculate

$$H_{t_i}^{(n)} := \frac{\phi_{t_i}^{(n,+)} - \phi_{t_i}^{(n,-)}}{2/\sqrt{n}},$$

$$\phi_{t_i}^{(n)} = \phi_{t_i}^{(n,-)} + H_{t_i}^{(n)}/\sqrt{n}.$$  

In the sequel the following abbreviations are repeatedly used:

$$\alpha^{(n)} := \frac{u^{(n)} - d^{(n)}}{2},$$

$$\tilde{\alpha}^{(n)} := \frac{u^{(n)} + d^{(n)}}{2},$$

$$\beta^{(n)} := \tilde{\alpha}^{(n)} - e^{r/n},$$

$$\theta^{(n)} := \frac{\beta^{(n)}}{\alpha^{(n)}} = \frac{e^{(\mu - \sigma^2)/n} \cosh \left( \frac{\sigma}{\sqrt{n}} \right) - e^{r/n}}{e^{(\mu - \sigma^2)/n} \sinh \left( \frac{\sigma}{\sqrt{n}} \right)}, \quad (3.3.6)$$

$$y_{t_j}^{(n)} (S_{t_j}^{(n)}) := y_{t_j}^{(n)} (S_{t_j}^{(n)}) S_{t_j}^{(n)},$$

$$P_j^{(n)} := \sum_{i=0}^j \binom{j}{i} y_{t_j}^{(n)} (u^{(n)})^{j-i} (d^{(n)})^i. \quad (3.3.7)$$
3.3. DISCRETE TIME PROBLEM

Note that \( \tilde{y}^{(n)}(\cdot) \) in (3.3.7) refers to the total monetary amount invested at time \( t_j \). As the process \( H^{(n)} \) is calculated by backward recursion, we state the following theorem also in this style by counting backwards from the terminal time 1.

**Theorem 3.3.2.** For some fixed \( n \in \mathbb{N} \), consider the terminal payoff \( V^{(n)}_1 \) and its predictable representation (3.3.4). For any \( k \in \{0, 1, \ldots, n - 1\} \) it holds that

\[
H^{(n)}_{t_{n-k}} = \sqrt{n} \sum_{i=0}^{k-1} \left( \frac{\beta^{(n)} e^{ir/n}}{2^{k-i}} \sum_{j=0}^{k-i} \frac{k-i-2j}{k-i} \tilde{y}^{(n)}_{t_{n-i-1}} \left(S^{(n)}_{t_{n-i-1}} u^{(n,k-j)} d^{(n)}_j \right) \right)
+ \sqrt{n} \tilde{y}^{(n)}_{t_{n-k-1}} (S_{t_{n-k-1}}) e^{kr/n} \alpha^{(n)}.
\]

(3.3.8)

**Proof.** Throughout the proof, we use the following abbreviations:

\[
A^{(n)}_{t_i} := (v_0 - y^{(n)}_0) e^{ir/n},
\]

\[
f^{(n)}(i) := \sum_{j=0}^{n-i-1} \left( y^{(n)}_{t_{j+1}} (S^{(n)}_{t_{j+1}}) - y^{(n)}_i (S^{(n)}_{t_{j}}) \right) S^{(n)}_{t_{j+1+1}} e^{(n-i-j)r/n},
\]

i.e., \( A^{(n)}_{t_i} \) is the total amount of money the initial investment in the riskless asset has grown to up to time \( t_i \) and \( f^{(n)}(i) \) is the trading gains until time \( t_i \). It is straightforward to verify that

\[
f^{(n)}(i) = f^{(n)}(i+1) + \left( y^{(n)}_{t_{n-i-1}} (S^{(n)}_{t_{n-i-1}}) - y^{(n)}_{t_{n-i-1}} (S^{(n)}_{t_{n-i-1}}) \right) S^{(n)}_{t_{n-i-1}} e^{r/n}.
\]

(3.3.9)

Thus, the recursive representation of \( V^{(n)} \) in (3.3.3) can be solved to

\[
V^{(n)}_{t_{i+1}} = \tilde{y}^{(n)}_{t_i} (S^{(n)}_{t_{i+1}}) S^{(n)}_{t_{i+1}} + A^{(n)}_{t_{i+1}} - f^{(n)}(n-i).
\]

(3.3.10)

The remainder of the proof follows Algorithm 3.3.1 and is split into two steps:
**CHAPTER 3. EXTENSIONS OF MV PORTFOLIO SELECTION**

Step 1 is to prove that the process $\phi^{(n)} = (\phi_{t_{n-k}}^{(n)})_{k \in \{0, \ldots, n-1\}}$ is of the following form:

$$
\phi_{t_{n-k}}^{(n)} = \sum_{i=0}^{k-2} \frac{j(n) e^{ir/n}}{2k-1-i} \sum_{j=0}^{k-1-i} \left( \frac{k-1-i}{j} \right) y_{t_{n-i}}^{(n)} (S_{t_{n-k}}^{(n)} u^{(n)} + \alpha^{(n)}) + f^{(n)}(k).
$$

We give a proof by induction on $k$.

$k = 1$:

According to Algorithm [3.3.1] we have

$$
\phi_{t_1}^{(n)} = \phi_{t_n}^{(n)} = y_{t_{n-1}}^{(n)} (S_{t_{n-1}}^{(n)} u^{(n)} + \alpha^{(n)}) + A_{t_n}^{(n)} - f^{(n)}(1),
$$

as well as

$$
\phi_{t_{n-1}}^{(n, +)} = y_{t_{n-1}}^{(n)} (S_{t_{n-1}}^{(n)} u^{(n)} + A_{t_n}^{(n)} - f^{(n)}(1)),
\phi_{t_{n-1}}^{(n, -)} = y_{t_{n-1}}^{(n)} (S_{t_{n-1}}^{(n)} d^{(n)} + A_{t_n}^{(n)} - f^{(n)}(1)).
$$

Consequently, we find

$$
H_{t_n}^{(n)} = \frac{\phi_{t_{n-1}}^{(n, +)} - \phi_{t_{n-1}}^{(n, -)}}{2/\sqrt{n}} = \sqrt{n} y_{t_{n-1}}^{(n)} (S_{t_{n-1}}^{(n)} \alpha^{(n)}),
\phi_{t_{n-1}}^{(n)} = y_{t_{n-1}}^{(n)} (S_{t_{n-1}}^{(n)} \tilde{\alpha}^{(n)} + A_{t_n}^{(n)} - f^{(n)}(1),
$$

which concludes the base case.

**Induction step:**

Suppose the claim (3.3.11) holds true for arbitrary $k \in \{1, \ldots, n - 1\}$. We prove that it is true then for $\bar{k} := k + 1$ by following Algorithm [3.3.1]:

$$
\phi_{t_{\bar{k}-k}}^{(n, +)} = \sum_{i=0}^{k-2} \frac{j(n) e^{ir/n}}{2k-1-i} \sum_{j=0}^{k-1-i} \left( \frac{k-1-i}{j} \right) y_{t_{n-i}}^{(n)} (S_{t_{n-k}}^{(n)} u^{(n)} + \alpha^{(n)}) + f^{(n)}(k).
$$
Inserting and summarizing yields

\[
- \left( y_{t_{n-k}}^{(n)} (S_{t_{n-k}}^{(n)} u^{(n)}) - y_{t_{n-k}}^{(n)} (S_{t_{n-k}}^{(n)} u^{(n)} e^{kr/n}) \right) S_{t_{n-k}}^{(n)} u^{(n)} e^{kr/n}
= \sum_{i=0}^{k-2} \beta^{(n)} e^{ir/n} \sum_{j=0}^{k-1-i} \left( k - 1 - i \right) y_{t_{n-i-1}}^{(n)} (S_{t_{n-k}}^{(n)} u^{(n)k-1-i-j} d^{(n)^{j+1}}) + \sum_{i=0}^{k-2} \beta^{(n)} e^{ir/n} \sum_{j=0}^{k-1-i} \left( k - 1 - i \right) y_{t_{n-i-1}}^{(n)} (S_{t_{n-k}}^{(n)} u^{(n)k-1-i-j} d^{(n)^{j+1}}).
\]

\[
\phi_{t_{n-k}}^{(n,-)} = \sum_{i=0}^{k-2} \beta^{(n)} e^{ir/n} \sum_{j=0}^{k-1-i} \left( k - 1 - i \right) y_{t_{n-i-1}}^{(n)} (S_{t_{n-k}}^{(n)} u^{(n)k-1-i-j} d^{(n)^{j+1}}) + \sum_{i=0}^{k-2} \beta^{(n)} e^{ir/n} \sum_{j=0}^{k-1-i} \left( k - 1 - i \right) y_{t_{n-i-1}}^{(n)} (S_{t_{n-k}}^{(n)} u^{(n)k-1-i-j} d^{(n)^{j+1}}).
\]

Next, we need to calculate \( H_{t_{n-k}}^{(n)} = \frac{\phi_{t_{n-k}}^{(n,+) - \phi_{t_{n-k}}^{(n,-)}}}{2\sqrt{n}} \). For notational convenience we shorten the terms and introduce the following abbreviations:

\[
U^{(n)}(i, j, k) := y_{t_{n-i-1}}^{(n)} (S_{t_{n-k}}^{(n)} u^{(n)k-1-i-j} d^{(n)^{j+1}}) S_{t_{n-k}}^{(n)} u^{(n)k-1-i-j} d^{(n)^{j+1}},
\]

\[
D^{(n)}(i, j, k) := y_{t_{n-i-1}}^{(n)} (S_{t_{n-k}}^{(n)} u^{(n)k-1-i-j} d^{(n)^{j+1}}) S_{t_{n-k}}^{(n)} u^{(n)k-1-i-j} d^{(n)^{j+1}}.
\]

Inserting and summarizing yields

\[
\sum_{i=0}^{k-2} \beta^{(n)} e^{ir/n} \sum_{j=0}^{k-1-i} \left( k - 1 - i \right) y_{t_{n-i-1}}^{(n)} (S_{t_{n-k}}^{(n)} u^{(n)k-1-i-j} d^{(n)^{j+1}}) + \sum_{i=0}^{k-2} \beta^{(n)} e^{ir/n} \sum_{j=0}^{k-1-i} \left( k - 1 - i \right) y_{t_{n-i-1}}^{(n)} (S_{t_{n-k}}^{(n)} u^{(n)k-1-i-j} d^{(n)^{j+1}}).
\]

\[
= \sum_{i=0}^{k-2} \beta^{(n)} e^{ir/n} \sum_{j=0}^{k-1-i} \left( k - 1 - i \right) \left( U^{(n)}(i, j, \bar{k}) - D^{(n)}(i, j, k) \right).
\]
Using that \( U^{(n)}(i, j + 1, \bar{k}) = D^{(n)}(i, j, k) \) and that \( A^{(n)}_n - f^{(n)}(\bar{k}) = 0 \), we find

\[
H_{i,n-k}^{(n)} / \sqrt{n} = \sum_{i=0}^{\bar{k}-2} \frac{\beta^{(n)} e^{ir/n}}{2^{k-1-i}} \sum_{j=0}^{k-1-i} \left( k - 1 - i \right) (U^{(n)}(i, j, \bar{k}) - D^{(n)}(i, j, k)) \\
+ \bar{y}^{(n)}_{i,n-k} (S^{(n)}_{i,n-k}) \alpha^{(n)} e^{(\bar{k}-1)r/n}. 
\]

(3.3.13)

Hence, we can calculate

\[
\begin{align*}
\phi^{(n)}_{i,n-k} &= \sum_{i=0}^{\bar{k}-2} \frac{\beta^{(n)} e^{ir/n}}{2^{k-1-i}} \sum_{j=0}^{k-1-i} \left( k - 1 - i \right) D^{(n)}(i, j, k) + \bar{y}^{(n)}_{i,n-k} (S^{(n)}_{i,n-k}) d^{(n)} e^{(\bar{k}-1)r/n} - f^{(n)}(\bar{k}) \\
&= \sum_{i=0}^{\bar{k}-2} \frac{\beta^{(n)} e^{ir/n}}{2^{k-1-i}} \sum_{j=0}^{k-1-i} \left( k - 1 - i \right) (U^{(n)}(i, j, \bar{k}) + D^{(n)}(i, j, k)) \\
&+ \bar{y}^{(n)}_{i,n-k} (S^{(n)}_{i,n-k}) \alpha^{(n)} e^{(\bar{k}-1)r/n} + A^{(n)}_n - f^{(n)}(\bar{k}).
\end{align*}
\]

In order to conclude the induction step, the following observation is crucial:

Fix \( i \in \{0, \ldots, \bar{k} - 2\} \) and consider the simplification

\[
\begin{align*}
&\sum_{j=0}^{k-1-i} \left( k - 1 - i \right) (U^{(n)}(i, j, \bar{k}) + D^{(n)}(i, j, k)) \\
=&U^{(n)}(i, 0, \bar{k}) + \sum_{j=1}^{k-1-i} \left( k - 1 - i \right) U^{(n)}(i, j, \bar{k}) \\
+ \sum_{j=0}^{k-1-i-1} \left( k - 1 - i \right) D^{(n)}(i, j, k) + D^{(n)}(i, k - 1 - i, k) \\
=&U^{(n)}(i, 0, \bar{k}) + \sum_{j=0}^{k-1-i-1} \left( k - 1 - i \right) U^{(n)}(i, j + 1, \bar{k})
\end{align*}
\]
3.3. DISCRETE TIME PROBLEM

\[
\begin{align*}
\sum_{j=0}^{k-1-i-1} \binom{k-1-i}{j} D^{(n)}(i,j,k) + D^{(n)}(i,k-1-i,k) \\
= U^{(n)}(i,0,\bar{k}) + \sum_{j=0}^{k-1-i-1} \binom{\bar{k}-1-i}{j+1} U^{(n)}(i,j+1,\bar{k}) + D^{(n)}(i,k-1-i,\bar{k}) \\
= U^{(n)}(i,0,\bar{k}) + \sum_{j=1}^{k-1-i} \binom{\bar{k}-1-i}{j} U^{(n)}(i,j,\bar{k}) + U^{(n)}(i,\bar{k}-1-i,\bar{k}) \\
= \sum_{j=0}^{k-1-i} \binom{\bar{k}-1-i}{j} U^{(n)}(i,j,\bar{k}),
\end{align*}
\]

where the third equality uses the recurrence relation of binomial coefficients.

Thus, we find that

\[
\phi^{(n)}_{t_n-k} = \sum_{i=0}^{\bar{k}-2} \frac{\beta^{(n)} e^{ir/n}}{2^{k-1-i}} \sum_{j=0}^{\bar{k}-1-i} \binom{\bar{k}-1-i}{j} U^{(n)}(i,j,\bar{k}) + \tilde{y}^{(n)}_{t_n-k}(S^{(n)}_{t_n-k}) e^{(k-1)\bar{r}/n} \tilde{\alpha}^{(n)} + A^{(n)}_{t_n} - f^{(n)}(\bar{k})
\]

and as \( k \) has been chosen arbitrarily, the principle of mathematical induction implies that (3.3.11) is true for all \( k \in \{1, \ldots, n-1\} \).

**STEP 2** is the verification of (3.3.8).

We can see the form of \( H^{(n)} \) already from (3.3.13). The following observation leads to a considerable simplification. Fixing \( i \in \{0, 1, \ldots, k-1\} \) and using the recurrence relation of binomial coefficients in the third equality below, we calculate that
\[
\sum_{j=0}^{k-1-i} \binom{k-1-i}{j} \left( U^{(n)}(i, j, k + 1) - D^{(n)}(i, j, k) \right)
\]
\[
= \sum_{j=1}^{k-1-i} \binom{k-1-i}{j} U^{(n)}(i, j, k + 1) - \sum_{j=0}^{k-1-i-1} \binom{k-1-i}{j} D^{(n)}(i, j, k)
\]
\[
+ U^{(n)}(i, 0, k + 1) - D^{(n)}(i, k - 1 - i, k)
\]
\[
= \sum_{j=0}^{k-1-i-1} \frac{k-i-2(j+1)}{k-i} \binom{k-i}{j+1} U^{(n)}(i, j, k + 1) + U^{(n)}(i, 0, k + 1)
\]
\[
- D^{(n)}(i, k - 1 - i, k)
\]
\[
= \sum_{j=1}^{k-1-i} \frac{k-i-2j}{k-i} \binom{k-i}{j} U^{(n)}(i, j, k + 1) + U^{(n)}(i, 0, k + 1)
\]
\[
- D^{(n)}(i, k - 1 - i, k)
\]
\[
= \sum_{j=0}^{k-i} \frac{k-i-2j}{k-i} \binom{k-i}{j} U^{(n)}(i, j, k + 1).
\]

Consequently, we obtain the following more condensed representation of \(H^{(n)}_{t_{n-k}}\):

\[
\frac{H^{(n)}_{t_{n-k}}}{\sqrt{n}} = \sum_{i=0}^{k-1} \beta^{(n)} e^{ir/n} \sum_{j=0}^{k-i} \frac{k-i-2j}{k-i} \binom{k-i}{j} U^{(n)}(i, j, k + 1)
\]
\[
+ \gamma^{(n)}_{t_{n-k-1}} (S^{(n)}_{t_{n-k-1}}) e^{kr/n} \alpha^{(n)},
\]

and inserting the definition of \(U^{(n)}\) from (3.3.12) finally yields (3.3.8).

We exemplify the lengthy formula for \(H^{(n)}\) in the following.

**Example 3.3.3.** Let \(n = 2\), then \(t_0 = 0\), \(t_1 = 1/2\), \(t_2 = 1\). Suppose the stock price increases twice until the terminal time 1. Denote the corresponding
portfolio value by $V_1^{(2,++)}$. According to (3.3.4), it must hold that

$$V_1^{(2,++)} = E[V_1^{(2)}] + \frac{H_{t_1}}{\sqrt{2}} + \frac{H_{t_2}^{(2,+)}}{\sqrt{2}},$$

where $H_{t_2}^{(2,+)}$ indicates that the stock price moved up in the first period. Writing $\tilde{y}_{t_0}(1) = \tilde{y}_{t_0}^{(2)}$, we note that

$$E[V_1^{(2)}] + \frac{H_{t_2}^{(2)}}{\sqrt{2}} + \frac{H_{t_2}^{(2,+)}}{\sqrt{2}}$$

$$= \beta^{(2)} \left( e^{r/2} \tilde{y}_{t_0}^{(2)} + \frac{1}{2} \left( \tilde{y}_{t_1}^{(2)} (u^{(2)}) + \tilde{y}_{t_1}^{(2)} (d^{(2)}) \right) \right) + \frac{\beta^{(2)}}{2} \left( \tilde{y}_{t_1}^{(2)} (u^{(2)}) - \tilde{y}_{t_2}^{(2)} (d^{(2)}) \right)$$

$$+ \tilde{y}_{t_0}^{(2)} e^{r/2} \alpha^{(2)} + \tilde{y}_{t_1}^{(2)} (u^{(2)}) \alpha^{(2)} + ve^r$$

$$= \tilde{y}_{t_0}^{(2)} e^{r/2} (u^{(2)} - e^{r/2}) + \tilde{y}_{t_1}^{(2)} (u^{(2)}) (u^{(2)} - e^{r/2}) + ve^r = V_1^{(2,++)}.$$

and the final identity is readily verified from respectively (3.3.3) or (3.3.10).

After establishing the form of $H^{(n)}$, we introduce the discrete time analogue of Problem (Π) and naturally call it Problem (Π(n)). The target functional is obviously given by

$$J(y^{(n)}) := E \left( V_1^{(n)} - \frac{1}{n} \sum_{k=0}^{n-1} g(t_{n-k}, H_{t_{n-k}}^{(n)}) \right),$$

and the corresponding optimization problem reads

$$\Psi^{(n)} := \max_{y^{(n)}} J(y^{(n)}).$$

Note that the function $g$ is by definition strictly convex and differentiable in the second component, therefore the first derivative of $g$ is injective. In order to simplify the subsequent theorems containing the optimal strategies, we make the following assumption:
Assumption 3.3.4. The function $g$ decomposes multiplicatively into $g(t, x) = f(t) \cdot z(x)$ for some strictly convex and differentiable function $z$ such that $z'$ is surjective and a continuous function $f : [0, T] \to \mathbb{R}_+$. Consequently, $g'(t, x) = f(t) \cdot z'(x)$ and the inverse of $z'$ is denoted by $h$, in symbols: $h := (z')^{-1}$.

The next theorem gives the solution to Problem $(\Pi^{(n)})$. We indicate the optimality of the strategy and the dollar amount by writing $\tilde{y}^{(n,*)}$ and $\tilde{\eta}^{(n,*)}$, respectively; further, $m^*$ denotes the realized number of downward movements of the stock. The following abbreviation turns out useful in the sequel.

$$C^{(n)}(k, m) := \sum_{i=0}^{k-1} \frac{\beta^{(n)} e^{r/n}}{2^{k-i}} \sum_{j=0}^{k-i} \frac{k-i-2j}{k-i} \left( \frac{k-i}{j} \right) \tilde{y}^{(n,*)}_{t_{n-i-1}} \left( u^{(n)} n^{-1-(m+j)} d^{(n) m+j} \right).$$

(3.3.14)

Theorem 3.3.5. Consider the optimization problem $(\Pi^{(n)})$ and suppose Assumption 3.3.4 is satisfied. For $k = 0$ and $m^* \in \{0, 1, \ldots, n-1\}$, the optimal dollar amount is given by

$$\tilde{y}^{(n,*)}_{t_{n-1}} (u^{(n)} n^{-1-m^*} d^{(n) m^*}) = h \left( \frac{\theta^{(n)} (1 + \theta^{(n)})^{m^*} (1 - \theta^{(n)})^{n-1-m^*}}{\alpha^{(n)} \sqrt{n}} \right),$$

(3.3.15)

and for any $k \in \{1, \ldots, n-1\}$ and $m^* \in \{0, 1, \ldots, n-k-1\}$ the optimal dollar amount is recursively given by

$$\tilde{y}^{(n,*)}_{t_{n-k-1}} (u^{(n)} n^{-k-1-m^*} d^{(n) m^*}) = h \left( \frac{\theta^{(n)} (1 + \theta^{(n)})^{m^*} (1 - \theta^{(n)})^{n-k-1-m^*}}{e^{k r/n} \alpha^{(n)} \sqrt{n}} \right) - \sqrt{n} C^{(n)}(k, m^*),$$

(3.3.16)

with $C^{(n)}(\cdot, \cdot)$ defined by (3.3.14).
3.3. DISCRETE TIME PROBLEM

Proof. Using (3.3.10), one can easily calculate that

\[
E[V_t^{(n)}] = \beta(n) \sum_{j=0}^{n-1} \frac{e^{(n-1-j)r/n}}{2^j} P_j^{(n)} + ve^r,
\]

\[
E[g(t_{n-k}, H_t^{(n)})] = \sum_{m=0}^{n-k-1} \frac{1}{2^{n-k-1}} \binom{n-k-1}{m} g\left(t_{n-k}, \sqrt{n} C^{(n)}(k, m) + \sqrt{n} y_{t_{n-k-1}}^{(n)}(u^{(n)}(n-k-1-m), d^{(n)}(m)) e^{kr/n} \alpha^{(n)} \right).
\]

(3.3.17)

Suppose \( k = n-1 \) and we want to find the optimal initial investment. Then solving the corresponding first order condition (FOC) yields

\[
\frac{\partial J(y^{(n)})}{\partial y_{t_0}} = \beta(n)e^{(n-1)r/n}
\]

\[
- \frac{\gamma}{t_0} g'\left(t_0, \sqrt{n} C^{(n)}(n-1, 0) + \sqrt{n} y_t^{(n)} e^{(n-1)r/n} \alpha^{(n)} \right) - \frac{\sqrt{n} \gamma g'(t_0)}{e^{(n-1)r/n} \alpha^{(n)} \sqrt{n}} \equiv 0
\]

\[
\Rightarrow y_{t_0}^{(n, *)} = \frac{h\left(\frac{\sqrt{n} \gamma g'(f(t_0))}{e^{(n-1)r/n} \alpha^{(n)} \sqrt{n}} \right)}{e^{(n-1)r/n} \alpha^{(n)} \sqrt{n}}.
\]

(3.3.18)

The rest of the proof follows an inductive argument: fix \( k^{*} \in \{1, \ldots, n-1\} \) and suppose the optimal investments \( (\tilde{y}_{t_{n-j}}^{(n, *)})_{j \in \{1, \ldots, k^{*}\}} \) are given by (3.3.16), then we can calculate \( \tilde{y}_{t_{n-k^{*}-1}}^{(n)} \). For ease of exposition we conduct the proof w.l.o.g. for the case \( m^{*} = 0 \) solely, the generalization is technical, but immediate.

\( k = n-2; \)

For finding \( \tilde{y}_{t_1}^{(n)}(u^{(n)}) \), we calculate

\[
\frac{\partial J(y^{(n)})}{\partial y_{t_1}^{(n)}(u^{(n)})} = \beta(n)e^{(n-2)r/n}
\]

\[
- \frac{\gamma}{2n} g'\left(t_1, \sqrt{n} C^{(n)}(n-2, 0) + \sqrt{n} \tilde{y}_{t_1}^{(n)}(u^{(n)}) e^{(n-2)r/n} \alpha^{(n)} \right) \sqrt{n} e^{(n-2)r/n} \alpha^{(n)}
\]
− \frac{\gamma}{n} g\left(t_0, \sqrt{n} C(n) (n - 1, 0) + \tilde{y}_{t_0}^{(n)} e^{(n-1)r/n} \alpha(n) \sqrt{n} \right) \sqrt{n} \frac{\partial C(n)(n - 1, 0)}{\partial \tilde{y}_{t_0}^{(n)} (u(n))} \\
\equiv 0.

Now it is straightforward to deduce from (3.3.14) that
\[
\frac{\partial C(n)(n - 1, 0)}{\partial \tilde{y}_{t_1}^{(n)} (u(n))} = \frac{\beta(n) e^{(n-2)r/n}}{2},
\]
and (3.3.18) yields
\[
\sqrt{n} C(n)(n - 1, 0) + \tilde{y}_{t_0}^{(n)} e^{(n-1)r/n} \alpha(n) \sqrt{n} = h \left( \sqrt{n} g(t_0) \theta(n) \right).
\]
Inserting these findings, we immediately obtain
\[
\frac{\partial J(y(n))}{\partial \tilde{y}_{t_1}^{(n)} (u(n))} \equiv 0 \\
\Rightarrow \theta(n) - \frac{\gamma}{\sqrt{n}} g\left(t_1, \sqrt{n} C(n)(n - 2, 0) + \sqrt{n} \tilde{y}_{t_1}^{(n)} (u(n)) e^{(n-2)r/n} \alpha(n) \right) \\
- \theta(n) \frac{\gamma}{\sqrt{n}} g\left(t_0, h \left( \frac{\sqrt{n}}{g(t_0)} \theta(n) \right) \right) = 0 \\
\Rightarrow \tilde{y}_{t_1}^{(n, \star)} (u(n)) = \frac{h \left( \frac{\sqrt{n}}{g(t_1)} \theta(n) (1 - \theta(n)) \right) - \sqrt{n} C(n)(n - 2, 0)}{e^{(n-2)r/n} \alpha(n) \sqrt{n}},
\]
which concludes the base case.

\textbf{Induction step:}

Still considering \( m^\star = 0 \), assume that for any \( k^\star \) in \( \{0, 1, \ldots, n - 2\} \) it holds that
\[
\tilde{y}_{t_{n-k^\star-2}}^{(n, \star)} (u(n)^{n-k^\star-2}) = \frac{h \left( \frac{\sqrt{n}}{g(t_{n-k^\star-2})} \theta(n) (1 - \theta(n))^{n-k^\star-2} - \sqrt{n} C(n)(k^\star + 1, 0) \right)}{e^{(k^\star+1)r/n} \alpha(n) \sqrt{n}}.
\]

We calculate \( \tilde{y}_{t_{n-k^\star-1}}^{(n)} (u(n)^{n-k^\star-1}) \):
3.3. DISCRETE TIME PROBLEM

Hence, we can draw the following chain of conclusions:

\[
\frac{\partial J(y^{(n)})}{\partial y_{n-k^*-1}^{(n)}(u^{(n)}|_{n-k^*-1})} = \frac{\beta^{(n)} e^{k^* r/n}}{2n-1-k^*} + \frac{\gamma \sqrt{n} e^{k^* r/n} \alpha^{(n)}}{2n-1-k^*} \left( t_{n-k^*-1}, \sqrt{n} C^{(n)}(k^*, 0) \right) + \sqrt{n} y_{n-k^*-1}^{(n)}(u^{(n)}|_{n-k^*-1}) e^{k^* r/n} \alpha^{(n)}
\]

\[
- \sum_{k=k^*+1}^{n-1} \left( \frac{1}{2n-k-1} g' \left( t_{n-k-1}, C^{(n)}(k, 0) \sqrt{n} + \sqrt{n} y_{n-k-1}^{(n)}(u^{(n)}|_{n-k-1}) e^{k r/n} \alpha^{(n)} \right) \right) \cdot \sqrt{n} \frac{\partial C^{(n)}(k, 0)}{\partial y_{n-k^*-1}^{(n)}(u^{(n)}|_{n-k^*-1})} \frac{\gamma}{n} \equiv 0.
\]

Observe that for any \( k \in \{k^* + 1, \ldots, n - 1\} \), (3.3.14) implies that

\[
\frac{\partial C^{(n)}(k, 0)}{\partial y_{n-k^*-1}^{(n)}(u^{(n)}|_{n-k^*-1})} = \frac{\beta^{(n)} e^{k^* r/n}}{2k^*-1}.
\]

Moreover, (3.3.19) yields that for any \( k \in \{k^* + 1, \ldots, n - 1\} \) we have

\[
C^{(n)}(k, 0) \sqrt{n} + \sqrt{n} y_{n-k-1}^{(n)}(u^{(n)}|_{n-k-1}) e^{k r/n} \alpha^{(n)} = h \left( \frac{\sqrt{n}}{\gamma} \int (t_{n-k-1}) \theta^{(n)}(1 - \theta^{(n)})^{n-1-k} \right).
\]

Hence, we can draw the following chain of conclusions:

\[
\frac{\partial J(y^{(n)})}{\partial y_{n-k^*-1}^{(n)}(u^{(n)}|_{n-k^*-1})} \equiv 0
\]

\[
\Rightarrow \theta^{(n)} - \frac{\gamma}{\sqrt{n}} g' \left( t_{n-k^*-1}, \sqrt{n} C^{(n)}(k^*, 0) + \sqrt{n} y_{n-k^*-1}^{(n)}(u^{(n)}|_{n-k^*-1}) e^{k^* r/n} \alpha^{(n)} \right)
\]

\[
- \theta^{(n)} \frac{\gamma}{\sqrt{n}} \sum_{k=k^*+1}^{n-1} \frac{\sqrt{n} \theta^{(n)} (1 - \theta^{(n)})^{n-1-k}}{\gamma} = 0
\]

\[
\Rightarrow \theta^{(n)} - \frac{\gamma}{\sqrt{n}} g' \left( t_{n-k^*-1}, \sqrt{n} C^{(n)}(k^*, 0) + \sqrt{n} y_{n-k^*-1}^{(n)}(u^{(n)}|_{n-k^*-1}) e^{k^* r/n} \alpha^{(n)} \right)
\]
\[
\begin{align*}
&\quad - \theta^{(n)^2} \sum_{k=0}^{n-2-k^*} (1 - \theta^{(n)})^k = 0 \\
&\quad \Rightarrow \theta^{(n)} - \frac{\gamma}{\sqrt{n}} g' \left( t_{n-k^*-1}, \sqrt{n} C^{(n)}(k^*, 0) + \sqrt{n} \widetilde{y}_{i_{n-k^*-1}}^{(n)} (u^{(n)-k^*-1}) e^{k^* r/n \alpha^{(n)}} \right) \\
&\quad \quad - \theta^{(n)^2} (1 - \theta^{(n)})^{n-2-k^*} - (1 - \theta^{(n)})^{n-2-k^*} + 1 = 0 \\
&\quad \Rightarrow \theta^{(n)}(1 - \theta^{(n)})^{n-1-k^*} \\
&\quad = \frac{\gamma}{\sqrt{n}} g' \left( t_{n-k^*-1}, \sqrt{n} C^{(n)}(k^*, 0) + \sqrt{n} \widetilde{y}_{i_{n-k^*-1}}^{(n)} (u^{(n)-k^*-1}) e^{k^* r/n \alpha^{(n)}} \right) \\
&\quad \Rightarrow y_{i_{n-k^*-1}}^{(n)}(u^{(n)-k^*-1}) \\
&\quad = \frac{h \left( \frac{\sqrt{n}}{\gamma J^{(n)}} \theta^{(n)} (1 - \theta^{(n)})^{n-1-k^*} \right) - \sqrt{n} C^{(n)}(k^*, 0)}{\sqrt{n} e^{k^* r/n \alpha^{(n)}}}.
\end{align*}
\]

Note that we used the geometric sum in the third step above. Since \( k^* \) has been arbitrarily chosen, the principle of mathematical induction implies that \((3.3.15)\) and \((3.3.16)\) are candidates for local maximums. Since \( E[V^{(n)}] \) is linear and \( g \) is convex in the second component by definition, \( J(y^{(n)}) \) is concave in \( y^{(n)} \) and the stationary points found are global maximums. \qed

\textbf{Remark 3.3.6.} Note that the existence of a solution to Problem \((\Pi^{(n)})\) is established in Theorem 3.3.5 under the relatively mild Assumption 3.3.4, while arguing for the existence of a solution to the continuous time version is significantly more complicated (if it exists at all).

Figuratively speaking, in Theorem 3.3.5 the strategies are specified for the stock price having reached a certain node in the binomial tree. Clearly, a version with the random variable \( S_{t_{n-k-1}}^{(n)} \) being the argument of the strategy instead of a particular realization is desirable. Since the formulation in Theorem 3.3.5 utilizes the number of downward movements of the stock price, the following observation is useful: at time point \( t_{n-k-1} \), the random walk has moved \( n - k - 1 \) times. The number of downward movements is easily seen to be given by

\[
\# \text{ downward movements} = \frac{n - k - 1 - \sum_{i=1}^{n-k-1} X_i^{(n)}}{2}.
\]
Clearly, one could write down the optimal solution also for the number of upward movements at a certain point in time with the final result obviously being the same. To be consistent with Theorem 3.3.5, we formulate the following corollary for the downward movements:

**Corollary 3.3.7.** Suppose the conditions of Theorem 3.3.5 are satisfied. For \( k = 0 \), the optimal dollar amount is given by

\[
\hat{y}_{n-1}(S^{(n)}_{t_{n-1}}) = \frac{h}{\alpha^{(n)} \sqrt{n}} \left( \frac{\sqrt{n}}{\gamma f(t_{n-1})} \theta^{(n)} (1 - \theta^{(n)})^{\frac{1}{2}(n-1)} \left( \frac{1 - \theta^{(n)}}{1 + \theta^{(n)}} \right)^{\frac{1}{2} \sum_{l=1}^{n-1} X^{(n)}_l} \right),
\]

(3.3.20)

and for any \( k \in \{1, \ldots, n-1\} \) the optimal dollar amount is recursively found by

\[
\hat{y}_{n-k-1}(S^{(n)}_{t_{n-k-1}}) = \frac{1}{e^{kr/n} \alpha^{(n)} \sqrt{n}} \left( h \left( \frac{\sqrt{n}}{\gamma f(t_{n-k-1})} \theta^{(n)} (1 - \theta^{(n)})^{\frac{1}{2}(n-k-1)} \left( \frac{1 - \theta^{(n)}}{1 + \theta^{(n)}} \right)^{\frac{1}{2} \sum_{l=1}^{n-k-1} X^{(n)}_l} \right) - \sqrt{n} \ C^{(n)} \left( k, \frac{n - k - 1 - \sum_{l=s+1}^{n-k-1} X^{(n)}_l}{2} \right) \right),
\]

(3.3.21)

with \( C^{(n)}(\cdot, \cdot) \) defined by (3.3.14).

**Remark 3.3.8.** Fix \( s \in \{0, 1, \ldots, n - 1\} \) and consider the problem \( \Psi^{(n)}_{t_s} = \max_{y^{(n)}} J_s(y^{(n)}) \), i.e., the version of Problem \((\Pi^{(n)})\) evaluated from time \( t_s \) onwards. For any \( k \in \{1, \ldots, n - 1 - s\} \), the corresponding optimal dollar amount \( \hat{y}_{n-k-1}^{(s,n)}(S^{(n)}_{t_{n-k-1}}) \) is straightforwardly seen to be given by

\[
\hat{y}_{n-k-1}^{(s,n)}(S^{(n)}_{t_{n-k-1}}) = \frac{1}{e^{kr/n} \alpha^{(n)} \sqrt{n}} \left( h \left( \frac{\sqrt{n}}{\gamma f(t_{n-k-1})} \theta^{(n)} (1 - \theta^{(n)})^{\frac{1}{2}(n-k-1)} \left( \frac{1 - \theta^{(n)}}{1 + \theta^{(n)}} \right)^{\frac{1}{2} \sum_{l=s+1}^{n-k-1} X^{(n)}_l} \right) - \sqrt{n} \ C^{(n)} \left( k, \frac{n - k - 1 - \sum_{l=s+1}^{n-k-1} X^{(n)}_l}{2} \right) \right),
\]
and the case \( k = 0 \) is immediate from (3.3.20). Hence, we see that
\[
\tilde{y}_{t_{n-k-1}}^{(s,n,\star)}(S_{t_{n-k-1}}^{(n)}) \neq \tilde{y}_{t_{n-k-1}}^{(n,\star)}(S_{t_{n-k-1}}^{(n)}),
\]
that is, Problem \((\Pi^{(n)})\) is a time-inconsistent optimal control problem (cf. Definition 3.2.4).

Next we state a corollary describing the solution to the familiar mean-variance portfolio selection problem in discrete time under the incorporation of time-dependence. We denote the optimal dollar amount to be invested by \( \tilde{y}_{t_{n-k-1}}^{(n,\star,\text{MV})} \). The linearity of the payoff function \( h \) in this special case leads to a significant simplification of the strategies, for it holds that \( h(x) = \frac{x}{2} \). This particular form allows us to describe the error term explicitly and make some inference about its behaviour for large \( n \).

**Corollary 3.3.9.** In addition to the conditions of Theorem 3.3.5, suppose that \( h(x) = x/2 \), i.e., Problem \((\Pi^{(n)})\) is the time-dependent mean-variance portfolio selection problem. For \( k = 0 \) it holds that the optimal amount of money to be invested in the stock is given by
\[
\tilde{y}_{t_{n-1}}^{(n,\star,\text{MV})}(S_{t_{n-1}}^{(n)}) = \frac{\theta^{(n)}(1 - \theta^{(n)})^2 \frac{1}{2} (n-1) \left( \frac{1 - \theta^{(n)}}{1 + \theta^{(n)}} \right)^{\frac{1}{2}} \sum_{l=1}^{n-1} X_l^{(n)}}{2 \gamma \alpha^{(n)} f(t_{n-1})}. \tag{3.3.22}
\]

For any \( k \in \{1, \ldots, n - 1\} \), it holds that
\[
\tilde{y}_{t_{n-k-1}}^{(n,\star,\text{MV})}(S_{t_{n-k-1}}^{(n)}) = \frac{1}{2 \gamma e^{kr/n} \alpha^{(n)} f(t_{n-k-1})} \left( \theta^{(n)} (1 - \theta^{(n)})^2 \frac{1}{2} (n-k-1) \left( \frac{1 - \theta^{(n)}}{1 + \theta^{(n)}} \right)^{\frac{1}{2}} \sum_{l=1}^{n-k-1} X_l^{(n)} \right)
\cdot \left( 1 + \sum_{a=1}^{k} M_{k,a}^{(n)} (\theta^{(n)})^{2a} \right), \tag{3.3.23}
\]
3.3. DISCRETE TIME PROBLEM

with constants $M_k^{(n)}$ that are recursively defined by

$$M_{k,1}^{(n)} = \sum_{i=0}^{k-1} f(t_{n-k-1}) f(t_{n-i-1}),$$

$$M_{k,j+1}^{(n)} = \sum_{i=j}^{k-1} M_{i,j}^{(n)} f(t_{n-k-1}), \quad j = 1, \ldots, k.$$

Proof. Note that (3.3.22) follows immediately from (3.3.20). The rest of the proof is based on an inductive argument: assuming we know the form of $\tilde{y}_{k}^{(n, \star, MV)}(S_{t_n-1})$, we can deduce $\tilde{y}_{k}^{(n, \star, MV)}$. $k = 1$:

First, a straightforward calculation shows that

$$C^{(n)} \left( 1, \frac{1}{2} \left( n - 2 - \sum_{l=1}^{n-2} X_l^{(n)} \right) \right) = \frac{\theta^{(n)^2} (1 - \theta^{(n)^2})^{\frac{1}{2}(n-2)} \left( \frac{1 - \theta^{(n)}}{1 + \theta^{(n)}} \right)^{\frac{1}{2} \sum_{l=1}^{n-2} X_l^{(n)}}}{2 \gamma f(t_{n-1})}.$$

Then we easily see from (3.3.21) that

$$\tilde{y}_{n-2}^{(n, \star, MV)}(S_{t_n-2}) = \frac{\theta^{(n)} (1 - \theta^{(n)^2})^{\frac{1}{2}(n-2)} \left( \frac{1 - \theta^{(n)}}{1 + \theta^{(n)}} \right)^{\frac{1}{2} \sum_{l=1}^{n-2} X_l^{(n)}}}{2 \gamma e^{r/n} \alpha^{(n)} f(t_{n-2})} \left( 1 + M_{1,1}^{(n)} \theta^{(n)^2} \right),$$

with $M_{1,1}^{(n)} = \frac{f(t_{n-2})}{f(t_{n-1})}$.

Induction step:

Suppose for any $k \in \{0, 1, n-2\}$ that the optimal dollar amount is given by

$$\tilde{y}_{n-k-1}^{(n, \star, MV)}(S_{t_{n-k-1}}) = \frac{\theta^{(n)} (1 - \theta^{(n)^2})^{\frac{1}{2}(n-k-1)} \left( \frac{1 - \theta^{(n)}}{1 + \theta^{(n)}} \right)^{\frac{1}{2} \sum_{l=1}^{n-k-1} X_l^{(n)}}}{2 \gamma e^{r/n} \alpha^{(n)} f(t_{n-k-1})} \left( 1 + \sum_{a=1}^{k} M_{k,a}^{(n)} (\theta^{(n)})^{2a} \right).$$
Define $\tilde{k} := k + 1$ and

$$
\tilde{M}_{k,1} := \frac{1}{f(t_{n-i-1})},
$$

(3.3.24)

$$
\tilde{M}_{k,j+1} := \sum_{i=j}^{k} M_{i,j}^{(n)}, \quad j \in \{1, \ldots, k\}.
$$

Then we find that

$$
C^{(n)} \left( \tilde{k}, \frac{1}{2} \left( n - \tilde{k} - 1 - \sum_{i=1}^{n-\tilde{k}-1} X_i^{(n)} \right) \right)
$$

$$
= \sum_{i=0}^{\tilde{k}-1} \frac{\beta^{(n)} e^{ir/n}}{2^{k-i}} \sum_{j=0}^{\tilde{k}-i} \frac{\tilde{k} - i - 2j}{\tilde{k} - i} \left( \tilde{k} - i \right)
$$

$$
\cdot g_{t_{i+1}}^{(n,*,MV)} \left( u(n)^{n-i-1-\left( \sum_{i=1}^{n-\tilde{k}-1} X_i^{(n)} \right) + j} \right)
$$

$$
\cdot \left( (1 + \theta^{(n)})^{\frac{1}{2} \left( n - \tilde{k} - 1 - \sum_{i=1}^{n-\tilde{k}-1} X_i^{(n)} \right) + j} \cdot (1 - \theta^{(n)})^{n-i-1-\frac{1}{2} \left( n - \tilde{k} - 1 - \sum_{i=1}^{n-\tilde{k}-1} X_i^{(n)} \right) - j} \right)
$$

$$
= \sum_{i=0}^{\tilde{k}-1} \frac{\theta^{(n)}^2}{2^{k-i+1} f(t_{n-i-1})} \left( 1 + \sum_{a=1}^{i} M_{i,a}^{(n)} (\theta^{(n)})^{2a} \right)
$$

$$
\cdot \left( \sum_{j=0}^{\tilde{k}-i} \frac{\tilde{k} - i - 2j}{\tilde{k} - i} \left( \tilde{k} - i \right) \left( (1 + \theta^{(n)})^{\frac{1}{2} \left( n - \tilde{k} - 1 - \sum_{i=1}^{n-\tilde{k}-1} X_i^{(n)} \right) + j} \right)
$$

$$
\cdot (1 - \theta^{(n)})^{n-i-1-\frac{1}{2} \left( n - \tilde{k} - 1 - \sum_{i=1}^{n-\tilde{k}-1} X_i^{(n)} \right) - j} \right)
$$

$$
= \theta^{(n)}^2 \left( 1 + \theta^{(n)} \right)^{\frac{1}{2} (n-\tilde{k}-1)} (1 - \theta^{(n)})^{\frac{1}{2} (n-\tilde{k}-1)} \left( \frac{1 - \theta^{(n)}}{1 + \theta^{(n)}} \right) \frac{1}{2} \sum_{i=1}^{n-\tilde{k}-1} X_i^{(n)}
$$

$$
\cdot \sum_{i=0}^{\tilde{k}-1} \left( \frac{1 - \theta^{(n)}}{2^{k-i+1} f(t_{n-i-1})} \left( 1 + \sum_{a=1}^{i} M_{i,a}^{(n)} (\theta^{(n)})^{2a} \right) \right)
$$
\[ \sum_{j=0}^{\bar{k}-i} \frac{\bar{k}-i-2j}{\bar{k}-i} \binom{\bar{k}-i}{j} \left( 1 + \theta^{(\bar{k})} \right)^j \left( 1 - \theta^{(\bar{k})} \right)^j \]

\[ = \frac{\theta^{(n)} 2^k \sum_{i=1}^{\bar{k}-1} \left( \frac{1}{2 \gamma f(t_{n-i-1})} + \sum_{a=1}^{i} M_{i,a}^{(n)} (\theta^{(n)})^{2a} \right)}{1 + \sum_{a=1}^{i} M_{i,a}^{(n)} (\theta^{(n)})^{2a}} \]

where the last equality holds because (3.3.24) implies that

\[ \sum_{i=0}^{\bar{k}-1} \frac{\theta^{(n)} 2^k \sum_{i=1}^{\bar{k}-1} \left( \frac{1}{2 \gamma f(t_{n-i-1})} + \sum_{a=1}^{i} M_{i,a}^{(n)} (\theta^{(n)})^{2a} \right)}{1 + \sum_{a=1}^{i} M_{i,a}^{(n)} (\theta^{(n)})^{2a}} = \sum_{a=1}^{k} \hat{M}_{k,a}^{(n)} (\theta^{(n)})^{2a}. \]

Set \( M_{k,a}^{(n)} := \hat{M}_{k,a}^{(n)} \cdot f(t_{n-k-1}) \) for all \( a \in \{1, \ldots, \bar{k}\} \). Using (3.3.21), we then conclude that

\[ \gamma^{(n, \gamma, MV)} (S_{n-k-1}) \]

\[ = \frac{1}{\epsilon^{kr/n} \alpha(n) \sqrt{n}} \left( \frac{\sqrt{n}}{2 \gamma f(t_{n-k-1})} + \sum_{l=1}^{n-k-1} X_l^{(n)} \right) \left( \frac{1}{1 - \theta(n)^2} \right)^{1/2} \left( \frac{1}{n \gamma f(t_{n-k-1})} \right) \]

\[ \times \left( 1 - \theta(n)^{2} \right) \left( \frac{1}{2 \gamma f(t_{n-k-1})} + \sum_{a=1}^{i} M_{i,a}^{(n)} (\theta^{(n)})^{2a} \right) \]

\[ = \frac{\theta^{(n)} \left( 1 - \theta(n)^{2} \right) \left( \frac{1}{2 \gamma f(t_{n-k-1})} \right)}{2 \gamma^{kr/n} \sqrt{n} \alpha(n) \sqrt{n}} \left( \frac{1}{1 - \theta(n)^{2}} \right)^{1/2} \]
i.e., the principle of mathematical induction implies that (3.3.23) is true for any \( k \in \{1, \ldots, n - 1\} \).

We remark that (3.3.24) gives a straightforward way to calculate the constants in the error term in (3.3.23). Moreover, it holds that \( \theta(n)^2 \in \mathcal{O}(1/n) \) (cf. (3.A.2)), so the error term is vanishing for large \( n \).

### 3.4 Convergence

In order to consider the convergence of the optimal policies, we need to define a certain notion for the convergence of stochastic processes first. Let \( \mathbb{D}[0, T] \) denote the space of \( \mathbb{R} \)-valued sample paths that are right continuous with left limits (abbreviated by the French acronym càdlàg in the sequel). The space \( \mathbb{D}[0, T] \) is usually equipped with the Skorohod topology, that is formally defined as follows (see Billingsley (1999), Section 12).

**Definition 3.4.1.** A sequence \((x^{(n)})_{n \in \mathbb{N}}\) converges to \( x \) in \( \mathbb{D}[0, T] \) in the Skorohod topology if and only if there is a sequence \( \lambda_n : [0, T] \to [0, T] \) of strictly increasing continuous functions, so called time changes, such that for every \( t_0 \in [0, T] \), \( \sup_{t \leq t_0} |\lambda_n(t) - t| \to 0 \) and \( \sup_{t \leq t_0} |x^{(n)}(\lambda_n(t)) - x(t)| \to 0 \). If the limit process \( x \) is continuous, convergence in the Skorohod topology is equivalent to the convergence in the uniform metric topology.

Using the convergence in the Skorohod topology, we next define the weak convergence of stochastic processes.

**Definition 3.4.2.** A sequence of càdlàg stochastic processes \((Z^{(n)})_{n \in \mathbb{N}}\) (possibly defined on different probability spaces) is said to converge weakly to a càdlàg process \( Z \), denoted by \( Z^{(n)} \overset{w}{\rightarrow} Z \), if \( \lim_{n \to \infty} \mathbb{E}[f(Z^{(n)})] = \mathbb{E}[f(Z)] \) for every bounded continuous function \( f : \mathbb{D}[0, T] \to \mathbb{R} \).

In the following we provide conditions under which the sequence of discrete time optimal trading strategies converges weakly to its continuous time counterpart. Likewise, the convergence of the optimal value functions is obtained.
Assumption 3.4.3. Suppose the continuous time optimal strategy $y^*$ solving Problem (Π) exists. We assume that

(i) for some continuous function $h_1 : [0,T] \times \mathbb{R}_+ \to \mathbb{R}$, it holds that $y^*_t = h_1(t, S_t)$ almost everywhere.

(ii) the approximating sequence of discrete time strategies $(\bar{y}^{(n)}_t)_{n \in \mathbb{N}}$ with $\bar{y}^{(n)}_t := h_1(t, S^{(n)}_t)$ induces a sequence of terminal values $(V^{(n)}_T)_{n \in \mathbb{N}}$ satisfying $\sup_{n \in \mathbb{N}} \mathbb{E}[(V^{(n)}_T)^{2+\epsilon}] < \infty$ for some $\epsilon > 0$.

(iii) the penalty function $g$ satisfies $g(x) \leq K(1 + |x|^2)$ for some constant $K > 0$ for every $x \in \mathbb{R}$.

Theorem 3.4.4. Suppose that Assumption 3.4.3 is satisfied. Then it holds that

$$\liminf_{n \to \infty} J(y^{(n,*)}) \geq J(y^*).$$

(3.4.1)

Proof.

Step 1 is to show that the family of random variables $(\int_0^T (H^{(n)}_s)^2 \, ds)_{n \in \mathbb{N}}$ is uniformly integrable.

For ease of exposition take w.l.o.g. $T \equiv 1$ in this part of the proof. For fixed $n \in \mathbb{N}$, the process $H^{(n)} = (H^{(n)}_s)_{s \in [0,1]}$ is constant on each of the intervals $[t_i, t_{i+1})$, $i = 0, 1, \ldots, n-1$ (cf. (3.3.4)). Using (3.3.1), we define the discrete time martingale $Z^{(n)}$ by

$$Z^{(n)}_{t_i} := \sum_{j=1}^i H^{(n)}_{t_j} X^{(n)}_j \frac{1}{\sqrt{n}} = \int_0^{t_i} H^{(n)}_s \, dR^{(n)}_s, \quad i = 0, 1, \ldots, n,$$

with the last integral being a stochastic integral in discrete time. For $\epsilon > 0$, we set $p := 2 + \epsilon$. Denote the quadratic variation of $Z^{(n)}$ by $[Z^{(n)}, Z^{(n)}]$. Then it holds that

$$\sup_{n \in \mathbb{N}} \mathbb{E} \left[ \left( \int_0^1 (H^{(n)}_s)^2 \, ds \right)^{p/2} \right] = \sup_{n \in \mathbb{N}} \mathbb{E} \left[ ([Z^{(n)}, Z^{(n)}])^{p/2} \right]$$
for some constant $c_p$; the first equality above follows from the definition of the quadratic variation, the first inequality is implied by the Burkholder-Davis-Gundy inequality (Protter (2005), Chapter IV, Theorem 48), the second inequality follows from Doob’s maximal quadratic inequality. The last inequality holds due to the second part of Assumption 3.4.3 and the representation (3.3.4).

Step 2 is to establish the claim of the theorem.
Consider the scaled random walk $R^{(n)}$ given by (3.3.1) extended in the obvious way to the interval $[0, T]$ and the driving Brownian motion $W$ in (3.2.2). According to Duffie (1988), Section 2.2, it holds that

$$\lim_{n \to \infty} R^{(n)} \overset{w}{\to} W.$$ 

Since $W$ is continuous, we know from Definition 3.4.1 that this convergence even holds in the supremum norm. Since the function $h_1$ is continuous by Assumption 3.4.3 and as the discrete and continuous time financial markets considered in Example 6.1 in Duffie and Protter (1992) coincide with our general setting, we deduce from their results (in particular their Theorem 4.4 and Lemma 5.1; cf. also Definition 4.1) that the following convergences hold:

$$\lim_{n \to \infty} S^{(n)} \overset{w}{\to} S,$$

$$\lim_{n \to \infty} \left( y^{(n)}, S^{(n)} \right) \overset{w}{\to} \left( y^*, S \right),$$

$$\lim_{n \to \infty} \left( \widetilde{y}^{(n)}, R^{(n)}, S^{(n)}, \int_0^1 y^{(n)} \ dS^{(n)} \right) \overset{w}{\to} \left( y^*, W, S, \int_0^1 y^* \ dS \right). \quad (3.4.2)$$
Observe that $V_T^{(n, \bar{g}(n))} = v_0 + \int_0^T \bar{y}_t^{(n)} dS_t^{(n)}$ and $V_T^{y^*} = v_0 + \int_0^T y_t^* dS_t$. There exists some probability space $(\bar{\Omega}, \bar{\mathcal{F}}, \bar{\mathbb{P}})$ such that the convergence in (3.4.2) holds a.s. In particular, there are random variables $(V_T^{(n, \bar{g}(n))})_{n \in \mathbb{N}}$ and $V_T^{y^*}$ as well as processes $\bar{R}^{(n)}$ and $\bar{W}$ that are defined on this space such that $(V_T^{(n, \bar{g}(n))}, \bar{R}^{(n)}) \overset{D}{=} (V_T^{y^*}, \bar{W})$ for each $n \in \mathbb{N}$ and $(V_T^{y^*}, W) \overset{D}{=} (V_T^{y^*}, \bar{W})$ satisfying

$$\lim_{n \to \infty} (V_T^{(n, \bar{g}(n))}, \bar{R}^{(n)}) = (V_T^{y^*}, \bar{W}), \quad \bar{\mathbb{P}}\text{-a.s.} \quad (3.4.3)$$

Since $\sup_{n \in \mathbb{N}} \mathbb{E}[(V_T^{(n, \bar{g}(n))})^{2+\epsilon}] < \infty$, the convergence in (3.4.3) also holds in $L^2(\bar{\Omega})$. Additionally, since $\lim_{n \to \infty} \sup_{0 \leq t \leq T} |\bar{R}_t^{(n)} - \bar{W}_t| = 0$, $\bar{\mathbb{P}}$-a.s., Corollary 3.2 in [Briand et al. (2001)] implies that for every $\delta > 0$ it holds that

$$\lim_{n \to \infty} \mathbb{P}_\delta \left( \int_0^T \left( \bar{H}_t^{(n, \bar{g}(n))} - \bar{H}_t^{y^*} \right)^2 dt \geq \delta \right) = 0. \quad (3.4.4)$$

According to STEP 1, it holds that

$$\sup_{n \in \mathbb{N}} \mathbb{E}_\bar{\mathbb{P}} \left[ \left( \int_0^T \left( \bar{H}_t^{(n, \bar{g}(n))} \right)^2 dt \right)^{1+\epsilon/2} \right] < \infty.$$

Thus, the *Vitali convergence theorem* implies that the convergence established in (3.4.4) extends to hold in $L^2(\bar{\Omega} \times [0, T])$. Hence, since $g$ grows at most quadratically due to the third part of Assumption 3.4.3, it follows that

$$\lim_{n \to \infty} \mathbb{E}_\bar{\mathbb{P}} \left( \int_0^T g(\bar{H}_t^{(n, \bar{g}(n))}) dt \right) = \mathbb{E}_\bar{\mathbb{P}} \left( \int_0^T g(\bar{H}_t^{y^*}) dt \right).$$

Denoting the corresponding version of $J$ given by (3.2.6) on the space $(\bar{\Omega}, \bar{\mathcal{F}}, \bar{\mathbb{P}})$ by $\bar{J}$, we showed that

$$\lim_{n \to \infty} \bar{J}(\bar{g}^{(n)}) = \bar{J}(y^*).$$

Finally, since $\bar{J}(\bar{g}^{(n)}) = J(\bar{g}^{(n)})$ for each $n \in \mathbb{N}$ and $\bar{J}(y^*) = J(y^*)$, we conclude that

$$\liminf_{n \to \infty} J(y^{(n, \ast)}) \geq \liminf_{n \to \infty} J(\bar{g}^{(n)}) = \lim_{n \to \infty} J(\bar{g}^{(n)}) = J(y^*),$$

where the previous inequality holds due to the optimality of $y^{(n, \ast)}$. \(\square\)
Note that the main idea to establish (3.4.1) is to approximate the discrete time optimal strategy reasoning from the continuous time solution. In order to obtain the reverse direction - and thereby the desired convergence - the continuous time optimal policy needs to be approximated utilizing the discrete time optimal strategies. Thus, the next assumption ensures a certain functional structure of the discrete time optimal policies such that a continuous time strategy can be approximated. We allow it to depend on the number of trading time points \( n \in \mathbb{N} \).

**Assumption 3.4.5.** Consider the sequence of discrete time optimal strategies \((y^{(n, \star)})_{n \in \mathbb{N}}\). We assume that

(i) for some sequence of functions \((h^{(n)})_{n \in \mathbb{N}}\) and a function \(h_2: [0, T] \times \mathbb{R}_+ \to \mathbb{R}\) being continuous for each \( n \in \mathbb{N} \) satisfying that for every \( \delta > 0 \) there exists \( N \in \mathbb{N} \) such that for all \( n \geq N \) and any \((t, s) \in [0, T] \times \mathbb{R}_+\) with \(|h_{2}^{(n)}(t, s) - h_2(t, s)| < \delta\), it holds that \(y^{(n, \star)}_t = h_{2}^{(n)}(t, S_t^{(n)})\). The approximated continuous time optimal strategy \(z\) with \(z_t = h_2(t, S_t)\) is admissible, i.e., \(z \in \mathcal{Y}\).

(ii) the sequence of optimal terminal values \((V_T^{(n, y^{(n, \star)})})_{n \in \mathbb{N}}\) satisfies
\[
\sup_{n \in \mathbb{N}} \mathbb{E}[|V_T^{(n, y^{(n, \star)})}|^{2+\epsilon}] < \infty \quad \text{for some} \quad \epsilon > 0.
\]

**Theorem 3.4.6.** Suppose Assumption 3.4.3 and Assumption 3.4.5 are satisfied and the optimal solution to Problem (I) is unique. Then it holds that
\[
\lim_{n \to \infty} J(y^{(n, \star)}) = J(y^*),
\]
\[
\lim_{n \to \infty} y^{(n, \star)} \xrightarrow{w} y^*.
\]

**Proof.** Following analogous arguments as in the proof of Theorem 3.4.4 we can immediately see that
\[
\limsup_{n \to \infty} J(y^{(n, \star)}) = \lim_{n \to \infty} J(y^{(n, \star)}) = J(z) \leq J(y^*), \quad (3.4.5)
\]
where the last inequality holds due to the optimality of \(y^*\). Combining (3.4.5)
with Theorem 3.4.4, the statement

\[ \lim_{n \to \infty} J(y^{(n,*)}) = J(y^*) \]

is established. Since the maximum \( y^* \) is unique by assumption, it follows that \( z = y^* \). Hence, we see that

\[ \lim_{n \to \infty} y^{(n,*)} \xrightarrow{w} y^*, \]

which establishes the second statement. \( \square \)

3.5 Conclusion

In this paper we suggest a mean-local-volatility portfolio selection criterion that mimics the preferences of an investor usually better than the mean-variance one. Measuring risk using the local volatility of the terminal payoff accounts for its correlation with the underlying financial market. Since this criterion gives rise to a time-inconsistent problem, Bellman’s dynamic programming principle is not applicable. We discretize the problem and by utilizing the finiteness of the arising filtration, the problem can be solved under relatively mild assumptions. The solution is given in closed-form. Subsequently, we provide conditions under which the discrete time optimal strategies converge to their continuous time counterpart. The next step is searching for weaker conditions under which convergence holds. We leave this as suggestion for future research.

3.A Appendix

Lemma 3.A.1. Let \( N \in \mathbb{N} \) and \( x \neq 1 \). Then it holds that

\[ \sum_{j=0}^{N} \frac{N - 2j}{N} \binom{N}{j} \left( \frac{1 + x}{1 - x} \right)^j = -2^N (1 - x)^{-N} x. \] (3.A.1)
CHAPTER 3. EXTENSIONS OF MV PORTFOLIO SELECTION

Proof. We give a proof by induction on $N$:

$N = 1$: 
Observe that

$$
\sum_{j=0}^{1} \frac{1 - 2j}{1} \binom{1}{j} \left(\frac{1 + x}{1 - x}\right)^j = 1 - \frac{1 + x}{1 - x} = -2^1(1 - x)^{-1}x,
$$

so (3.A.1) holds true for $N = 1$.

Induction Step: 
Suppose (3.A.1) holds true for some fixed $N \in \mathbb{N}$. We show that it is also valid for $N + 1$. A straightforward calculation yields that

$$
\sum_{j=0}^{N+1} \frac{N + 1 - 2j}{N + 1} \binom{N + 1}{j} \left(\frac{1 + x}{1 - x}\right)^j = \sum_{j=0}^{N} \frac{N + 1 - 2j}{N + 1} \binom{N + 1}{j} \left(\frac{1 + x}{1 - x}\right)^j - \left(\frac{1 + x}{1 - x}\right)^{N+1}
$$

$$
= \sum_{j=0}^{N} \binom{N}{j} \left(\frac{1 + x}{1 - x}\right)^j - \left(\frac{1 + x}{1 - x}\right)^{N+1}
$$

$$
= \left(1 + \frac{1 + x}{1 - x}\right)^N - \left(1 + \frac{1 + x}{1 - x}\right)^{N+1} = -2^{N+1}(1 - x)^{-(N+1)x}.
$$

Since $N$ has been arbitrarily chosen, the principle of mathematical induction implies that (3.A.1) is true for all $N \in \mathbb{N}$ and $x \neq 1$.

Lemma 3.A.2. Consider $\theta^{(n)}$ defined by (3.3.6). It holds that

$$
\theta^{(n)} \in O\left(\frac{1}{\sqrt{n}}\right).
$$

Proof. Define $f(x) := \frac{\cosh(\sqrt{x})^{-1}}{\sinh(\sqrt{x})}$. The series expansion of $f$ around 0 is easily seen
to be given by
\[ f(x) = \frac{\sqrt{x}}{2} - \frac{x^{3/2}}{24} + \mathcal{O}(x^{5/2}). \]
Since \( \lim_{n \to \infty} e^{(\mu - x^2/n)/n} = \lim_{n \to \infty} e^{r/n} = 1 \), we can neglect respectively normalize these terms to 1 when evaluating the behavior of \( \theta^{(n)} \) for large \( n \). But then we see that \( \theta^{(n)} \) behaves for large \( n \) as \( f(x) \) does around 0, so we conclude that \( \theta^{(n)} \in \mathcal{O}\left(\frac{1}{\sqrt{n}}\right) \).
\[ \square \]
Chapter 4

Robustness of Delta Hedging in a Jump-Diffusion Model

4.1 Introduction

This chapter is based on [Bosserhoff and Stadje (2019b)]. In the scientific literature on pricing and hedging of contingent claims as well as in practice, one oftentimes presumes against better knowledge that some stochastic model mimics the developments at the stock market appropriately. However, even if an investor is aware of the model type (for example Markovian diffusion, jump-diffusion, infinite activity pure jump process, etc.), functions such as the drift, the volatility as well as the jump sensitivity are to be specified. Determining an allocation rule based on misspecified model parameters and trading in the true stock may result in a severe violation of the hedging objective. As Delta strategies yield a perfect hedge in complete markets, applying them in incomplete markets is tempting as well and often done in practice. This gives rise to studying their general robustness properties. In this paper we closely look at the performance of Delta strategies in jump-diffusion financial markets and provide sufficient conditions under which a Delta hedge yields a superhedge of some European contingent claim \( h(S(T)) \) in expectation.

In addition to assuming a deterministic interest rate and that the mis-
CHAPTER 4. ROBUSTNESS OF DELTA HEDGING

specified stock price process is Markov, the fundamental assumption enabling the variety of results obtained in this paper is the convexity of $h$, which allows to prove that the option price function is convex in the current stock level. For a diffusion setting, El Karoui et al. (1998) show this property employing the theory of stochastic flows; Hobson (1998) provides a simplification using coupling techniques and Ekström et al. (2005) investigate convexity properties when the claim depends on several underlying assets.

For example, the convexity of the option value function enables us to deduce an ordering result. To be more precise, we show that if the model volatility and jump sensitivity systematically overestimate the true ones, the model option price dominates the corresponding market option price. A general comparison result for the solution of one-dimensional stochastic differential equations (SDEs) can be found in Peng and Zhu (2006). Ordering results by deriving sufficient conditions for the convexity of Euler schemes, generalizations to multi-dimensional special semimartingales and path-dependent options are to be found in Bergenthum and Rüschendorf (2006). Extensions to more general discretization schemes and applications to Bermudan option prices are discussed in Pagès (2016). Predictable representation results are used in Arnaudon et al. (2008). Hobson (2010) employs coupling arguments to draw comparisons among option prices in various stochastic volatility models, see also Criens (2019). For convex ordering results with pathwise Itô calculus see Köpfer and Rüschendorf (2019). A general overview of the impact of model uncertainty on pricing of contingent claims is provided in Cont (2006).

However, contrary to most of the above literature, we are mainly interested in the robustness of Delta hedging strategies. We show that the terminal value of the self-financing Delta hedging portfolio dominates the true claim on average as soon as the misspecified volatility and jump sensitivity dominate the true ones. Hence, we call a Delta strategy robust if the hedging portfolio yields a superhedge in expectation. In local volatility models the superhedge coincides with a perfect hedge; consequently, El Karoui et al. (1998) call a Delta strategy robust if the physical Delta strategy is an almost sure (a.s.) superhedge for the claim as soon as the model volatility systemat-
4.1. INTRODUCTION

ically overestimates the market volatility. Schied and Stadje (2007) establish the robustness of the Delta hedging strategy for general path-dependent options in local volatility models. They show that a sufficient condition for the robustness of the Delta strategy in every local volatility model is the directional convexity of the payoff function. Gapeev et al. (2011) investigate the robust hedging problem when log-returns of the stock price are Gaussian and self-similar and the investor is not sure whether the market is efficient.

Further, there are many works in the literature addressing the robustness of superhedging with dominating or non-dominating measures and its link to optimal transport problems, see for instance Neufeld and Nutz (2013), Possamaï et al. (2013), Dolinsky and Soner (2014, 2015) or Nutz (2015). In these works, the investor typically tries to find a hedge which for a whole class of models constitutes a superhedge. Hence, as long as the true model belongs to the set which the investor chooses for her calculation, a superhedge will be obtained. For results on robust utility maximization see for instance Matoussi et al. (2015), Herrmann et al. (2017), and the references therein. The above approach is fundamentally different from ours. First of all, in our analysis the true model does not enter the calculations of the investor. We instead analyze under which conditions for convex payoff functions the resulting hedge is robust in the sense that on average it does overestimate than underestimate the payoff if we use as input the physical prices instead of the model price under which the hedge has been calculated. Second, instead of superhedging we consider Delta hedging which allows for a negative hedging error, but is much less expensive. In fact, superhedging in many cases leads to a trivial solution and due to its expenses is often not used in practice. Instead, accepting a certain risk for reducing hedging expenses is not uncommon and for instance also done in quantile hedging (Föllmer and Leukert (1999)), time-consistent and time-inconsistent mean-variance hedging (Lim (2004), Černý and Kallsen (2007)) or utility indifference pricing (see Carmona (2009) and the references therein as well as Laeven and Stadje (2014)).

Throughout this paper, we consider an investor modeling the stock price based on Lévy jump processes, but also allowing for a Brownian component.
CHAPTER 4. ROBUSTNESS OF DELTA HEDGING

Such models are known to be more appropriate for mimicking the true stock price development as observed at the stock exchange than pure Brownian models, see for instance Cont (2001) as well as Cont and Tankov (2012) and references therein. A good overview of the theory of Lévy processes can be found in Applebaum (2009), Bertoin (2009) and Sato (2013), while Barndorff-Nielsen et al. (2012) and Kyprianou et al. (2006) focus on its applications in finance.

This paper is structured as follows. In Section 2 the basic model is provided. In Section 3 the main result is Theorem 4.3.2 stating the convexity of the contingent claim’s price. The major part of Section 3 is devoted to its proof. Section 4 shows that the Delta hedge leads to a super-replication of $h(S(T))$ in expectation for a wide class of models if the true volatility and jump sensitivity are dominated. It is also shown that the misspecified option’s price is larger than the true one. Section 5 starts by introducing the robust pricing operator and discussing several examples. After proving a non-smooth version of Itô’s lemma for the case of finite jump activity, it is established that in each model the hedging error induced by the Delta strategy is a submartingale, which allows deducing robustness properties. In Section 6, we consider a.s. superhedges and argue that under Markovian assumptions the payoff is not necessarily linear only if the jump sensitivity is independent of the jump size and the stock price process is only driven by either the jump process or the diffusion component. For the latter case the robustness of the replicating strategy is exemplified in the pure Poisson case. We remark that all our robustness results are stated for the case of a systematic overestimation of the true volatility and jump sensitivity, but they immediately generalize to an underestimation and lead to corresponding subhedges in expectation respectively a.s in the pure Poisson setting.

Notation. Denote by $\mathbb{R}_+$ the positive real numbers. For any $T > 0$ and a function $f : [0, T] \times \mathbb{R} \to \mathbb{R}$, we write $f \in C^{i,j}$ if $f$ is $i$ (resp. $j$) times continuously differentiable w.r.t. the $i$th (resp. $j$th) variable. Further, for $t \in (0, T)$ and $f : [0, T] \times \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ such that $f \in C^{1,2,\cdot}$, we define $\hat{f}(t, \cdot, \cdot) := \frac{\partial}{\partial t} f(t, \cdot, \cdot)$ and for any $s \in \mathbb{R}$ we denote $f'(\cdot, s, \cdot) := \frac{\partial}{\partial s} f(\cdot, s, \cdot)$. For any
4.2. MODEL SETUP

function $f : \mathbb{R} \to \mathbb{R}$, we may write $f^\prime_+(x) = \lim_{y \downarrow x} \frac{f(x)-f(y)}{x-y}$ for the right-hand derivative and $f^\prime_-(x) = \lim_{y \uparrow x} \frac{f(x)-f(y)}{x-y}$ for the left-hand derivative of $f$ at $x$ provided they exist. Denote the Borel $\sigma$-algebra of the set $\mathcal{X}$ by $\mathcal{B}(\mathcal{X})$.

On some probability space $(\Omega, \mathcal{F}, \mathbb{P})$, equalities and inequalities between random variables are understood to hold $\mathbb{P}$-a.s.; two random variables are identified if they are equal a.s. We write $L^p$ for the space of $\mathbb{R}$-valued $\mathcal{F}$-measurable random variables $X$ such that $\|X\|_{L^p} := \left( \mathbb{E}[|X|^p] \right)^{1/p} < \infty$, for $p \in [1, \infty)$.

The equivalence between any two probability measures $\mathbb{P}$ and $\mathbb{Q}$ is denoted by $\mathbb{P} \sim \mathbb{Q}$. For some semimartingale $Y$, we write $\mathcal{E}(Y)$ for its stochastic exponential. The minimum of two numbers $a, b \in \mathbb{R}$ is denoted by $a \wedge b$. We write $\text{sgn}$ for the sign function.

4.2 Model Setup

Consider a continuous-time setting with a horizon $T > 0$. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space that is equipped with a standard one-dimensional Brownian motion $W = (W(t))_{t \in [0, T]}$ and a Poisson random measure $J(dt, dz)$ on $[0, T] \times \mathbb{R} \setminus \{0\}$, being independent of $W$, with respective intensity measure $\nu(dz)dt$. Denote its compensated version by $\tilde{J}(dt, dz) = J(dt, dz) - \nu(dz)dt$.

Let $(\mathcal{F}_t)_{t \in [0, T]}$ be the right-continuous completion of the filtration generated by $W$ and $J$. Throughout the paper we assume that two assets are continuously traded in a frictionless financial market. One of them is the money market whose price at any time $t \in [0, T]$ we denote by $M(t)$ with

$$M(t) = e^{\int_0^t r(u) du},$$

(4.2.1)

for some deterministic interest rate process $r$ satisfying $\int_0^T |r(u)| \, du < \infty$.

The other asset is denoted by $S$ and henceforth regarded as the true stock price process whose realization is displayed at the stock exchange. Its price process $S = (S(t))_{t \in [0, T]}$ satisfies the SDE

$$\frac{dS(t)}{S(t-)} = r(t) \, dt + \sigma(t) \, dW(t) + \int_{\mathbb{R} \setminus \{0\}} \eta(t, z) \, \tilde{J}(dt, dz),$$

(4.2.2)
CHAPTER 4. ROBUSTNESS OF DELTA HEDGING

with $S(0-) = S(0) > 0$. The processes $\sigma$ and $\eta$ are assumed to be $(\mathcal{F}_t)_{t \in [0,T]}$-predictable, $\sigma$ is non-negative and satisfies $\int_0^T \sigma(t)^2 \, dt < \infty$ a.s., while $\eta$ is strictly larger than $-1$ and satisfies $\int_0^T \int_{\mathbb{R}\setminus\{0\}} \eta(t,z)^2 \vartheta(dz) \, dt < \infty$. We denote the jump of $S$ at time $t$ by $\Delta S(t,z) = S(t^-)\eta(t,z)$. In (4.2.2) the mean rate of return is equal to the interest rate $r(t)$, therefore under $\mathbb{P}$ the discounted version of $S$ is a local martingale. As we are mainly interested in the calculation of prices of contingent claims in this work, we restrict to risk-neutral modeling and do not consider the statistical probability measure of $S$. We note, however, that under the physical measure only the drift in (4.2.2) and (4.2.5) below would change which does not have any consequences on hedging and pricing regardless of possible misspecification. The measure $\mathbb{P}$, under which $S$ in (4.2.2) is specified, can be thought of as reference risk-neutral measure. We impose the following assumption:

**Assumption 4.2.1.** We assume that the local martingale $\tilde{S} := S/M$ is a square-integrable martingale, that is, $\tilde{S} = (\tilde{S}(t))_{t \in [0,T]}$ is a martingale and it holds that

$$\mathbb{E}[\tilde{S}(t)^2] < \infty, \quad t \in [0,T].$$

In order to consider options written on $S$, we define payoff functions as follows:

**Definition 4.2.2.** A payoff function is a convex function $h : \mathbb{R}_+ \to \mathbb{R}$ having bounded one-sided derivatives, that is

$$|h'_\pm(x)| \leq L, \quad x > 0,$$

for some positive constant $L$.

In the sequel, $h$ refers to a non-further specified arbitrary payoff function. A European contingent claim is a non-path-dependent contract paying $h(S(T))$ at time $T$. As the financial market defined by (4.2.1) - (4.2.2) is generically incomplete, $h(S(T))$ is not necessarily perfectly replicable. This gives rise to the study of possible hedging strategies. We call a bounded predictable process $y = (y(t))_{t \in [0,T]}$ a self-financing trading strategy, and the induced
portfolio process $P^y = (P^y(t))_{t \in [0,T]}$ described by the SDE

$$dP^y(t) = P^y(t)r(t) \, dt + y(t)[dS(t) - r(t)S(t) \, dt], \quad P^y(0) > 0, \quad (4.2.3)$$

whose solution is actually given by

$$P^y(t) = M(t) \left[ P^y(0) + \int_0^t y(u-) \, d\tilde{S}(u) \right], \quad t \in [0,T], \quad (4.2.4)$$

is the hedging portfolio. Changes in the value of the portfolio process defined by (4.2.3) are caused only by movements in the assets’ price processes and trading gains. In particular, no money is inserted or withdrawn. Due to Assumption 4.2.1 the process $P^y/M$ is a martingale.

Suppose an investor seeking to hedge $h(S(T))$ knows that the dynamics of $S$ is driven by a Brownian motion supplemented by jumps, but in her pricing and hedging model incorrectly specifies the volatility process and the jump sensitivity. Define the misspecified stock price process $S^x_m = (S^x_m(t))_{t \in [0,T]}$ as solution of the SDE

$$\frac{dS^x_m(t)}{S^x_m(t-)} = r(t) \, dt + \gamma(t, S^x_m(t)) \, dW(t) + \int_{\mathbb{R}\setminus\{0\}} \tilde{\gamma}(t, S^x_m(t-), z) \, \tilde{J}(dt, dz). \quad (4.2.5)$$

The dependence on the initial price $S^x_m(0) = x > 0$ is expressed by the superscript $x$. When referring to the realized price at some time $t > 0$, we are going to add it to the superscript. Denote the jump of $S^x_m$ at time $t$ by $\Delta S^x_m(t, z) = S^x_m(t-)\tilde{\gamma}(t, S^x_m(t-), z)$. The functions $\gamma$ and $\tilde{\gamma}$ are respectively the misspecified volatility and jump sensitivity. These functions are presumed to be random only through their dependence on the stock price $S^x_m$. The subscript $m$ indicates the generation of the stock price by the misspecification of the volatility and the jump sensitivity. This paper investigates the impact of a systematic overestimation of the latter on Delta hedging strategies of $h(S(T))$. To be more precise, we assume that

$$\sigma(t) \leq \gamma(t, S(t)) \quad \text{and} \quad \text{sgn}(\tilde{\gamma}(t, \tilde{S}^x_m(t), z) - \eta(t, z)) = \text{sgn}(\eta(t, z)), \quad (4.2.6)$$
CHAPTER 4. ROBUSTNESS OF DELTA HEDGING

d\mathbb{P} \times dt and d\mathbb{P} \times dt \times \vartheta(dz)\text{-a.s.} The second part of the previous condition obviously means that a positive jump sensitivity is always systematically overestimated and a negative one underestimated, i.e.,

\[ \hat{\gamma}(t, S(t), z) \geq \eta(t, z), \text{if } \eta(t, z) \geq 0, \]
\[ \hat{\gamma}(t, S(t), z) \leq \eta(t, z), \text{if } \eta(t, z) < 0. \]

Under (4.2.6) robustness properties are established in this paper. In order to enable this we need the following assumption:

**Assumption 4.2.3.** Consider the process \( S^x_m \) defined by (4.2.5).

(i) Assume that \( \gamma : [0, T] \times \mathbb{R}_+ \to \mathbb{R} \) is continuous and bounded from above. Defining \( \rho(t, s) := s\gamma(t, s) \), suppose that \( \rho'(t, s) \) is continuous in \((t, s)\) and locally Lipschitz continuous and bounded in \( s \in \mathbb{R}_+ \), uniformly in \( t \in [0, T] \).

(ii) Assume that \( \tilde{\gamma} : [0, T] \times \mathbb{R}_+ \times \mathbb{R}\{0\} \to \mathbb{R} \) is continuous, bounded from above and \( \tilde{\gamma}(t, s, z) > -1 \). Defining \( \tilde{\rho}(t, s, z) := s\tilde{\gamma}(t, s, z) \), suppose that \( \tilde{\rho}'(t, s, z) \) is continuous in \((t, s, z)\), locally Lipschitz continuous and bounded in \( s \in \mathbb{R}_+ \), uniformly in \( t \in [0, T] \), \( \tilde{\rho}'(t, s, z) > -1 + \epsilon \), for some \( \epsilon > 0 \), and that there exists a constant \( L > 0 \) such that

\[ \int_{\mathbb{R}\{0\}} (\tilde{\rho}(t, s_1, z) - \tilde{\rho}(t, s_2, z))^2 \vartheta(dz) \leq L \cdot |s_1 - s_2|^2, \]
\[ \int_{\mathbb{R}\{0\}} \tilde{\rho}'(t, s, z)^2 \vartheta(dz) \leq L. \]

Denote the discounted version of \( S^x_m \) by \( \tilde{S}^x_m := S^x_m / M \). Define

\[ \mathcal{Q}_{em} := \{ Q \sim P | \tilde{S}^x_m \text{ is a martingale w.r.t. } Q \}, \]

i.e., \( \mathcal{Q}_{em} \) is the set of all equivalent martingale measures (EMMs) of \( \tilde{S}^x_m \) (see Lemma 4.A.1 for a detailed characterization). For some subset \( \mathcal{M} \subseteq \mathcal{Q}_{em} \), we call a stochastic process \( X \) an \( \mathcal{M} \)-(sub-/super-)martingale if it is a (sub-/super-)martingale w.r.t. all measures \( Q \in \mathcal{M} \).
4.3 Convexity of European Contingent Claim Value

The convexity of the European contingent claim value is the main tool in the proofs in subsequent sections. We formally define it under the reference measure $\mathbb{P}$ as follows:

**Definition 4.3.1.** The misspecified value at time $t$ of the European contingent claim with payoff function $h$ is

$$v_m(t, x) := \mathbb{E} \left[ h(S^t_m(T)) e^{-\int_t^T r(u) du} \right], \quad t \in [0, T], \ x > 0.$$ 

The next theorem is formulated for the time-zero misspecified price only. The generalization to arbitrary $(t, s) \in [0, T] \times \mathbb{R}_+$ is immediate.

**Theorem 4.3.2.** Consider the process $S^x_m$ described by (4.2.5), suppose that Assumption 4.2.1 and Assumption 4.2.3 are satisfied. Then the European contingent claim value $v_m(x) := v_m(0, x)$ is convex in $x$.

**Proof.** The proof is conducted in six steps:

**STEP 1** is to prove: if $0 < x < y$, then $S^x_m(T) \leq S^y_m(T)$.

Define $\tau_1 := \inf \{ t \geq 0 : S^x_m(t) > S^y_m(t) \}$ and by contradiction suppose $\tau_1 \land T < T$. Consequently, it holds that $S^x_m(\tau_1) < S^y_m(\tau_1)$ and in addition

$$S^x_m(\tau_1) + \Delta S^x_m(\tau_1, z) > S^y_m(\tau_1) + \Delta S^y_m(\tau_1, z)$$

$$\Rightarrow S^x_m(\tau_1) - S^y_m(\tau_1) > \Delta S^y_m(\tau_1, z) - \Delta S^x_m(\tau_1, z)$$

$$= \tilde{\rho}(\tau_1, S^y_m(\tau_1), z) - \tilde{\rho}(\tau_1, S^x_m(\tau_1), z).$$

Observe that

$$\tilde{\rho}(\tau_1, S^y_m(\tau_1), z) - \tilde{\rho}(\tau_1, S^x_m(\tau_1), z) = \int_{S^y_m(\tau_1)}^{S^y_m(\tau_1)} \tilde{\rho}'(\tau_1, s, z) \ ds$$

$$> (1 + \varepsilon)(S^y_m(\tau_1) - S^x_m(\tau_1)).$$
so in total we obtain

\[ 0 > \epsilon (S^y_m(\tau_1^-) - S^x_m(\tau_1^-)), \]

which is obviously a contradiction. Thus, we conclude that \( \tau_1 \land T = T \). This yields that \( \tau_2 := \inf\{t \geq 0 : S^x_m(t) \geq S^y_m(t)\} \land T = \inf\{t \geq 0 : S^x_m(t) = S^y_m(t)\} \land T \). If \( \tau_2 \land T = \tau_2 \), then strong uniqueness for (4.2.5) (cf. Øksendal and Sulem [2005], Theorem 1.19) implies that \( S^x_m(t) = S^y_m(t) \) for all \( t \in [\tau_2, T] \).

To sum up, we see that \( S^x_m(T) \leq S^y_m(T) \).

**Step 2** is to note that if \( 0 < x < y \), then by convexity of \( h \) and Step 1 it holds that

\[
\begin{align*}
  h(S^y_m(T)) - h(S^x_m(T)) &\leq h'_+(S^y_m(T)) \cdot (S^y_m(T) - S^x_m(T)), \\
  h(S^y_m(T)) - h(S^x_m(T)) &\geq h'_-(S^x_m(T)) \cdot (S^y_m(T) - S^x_m(T)).
\end{align*}
\]

**Step 3** is to prove: if \( x, y > 0, x \neq y \), then \( \phi(t) := E\left[\left(\frac{S^y_m(t) - S^x_m(t)}{M(t)}\right)^2\right] \leq 3(y - x)^2e^{6L^2t}, t \in [0, T] \) and \( L > 0 \).

We find that

\[
\begin{align*}
  \phi(t) &\leq 3(y - x)^2 + 3E \left( \int_0^t \left( \frac{\rho(u, S^y_m(u)) - \rho(u, S^x_m(u))}{M(u)} \right)^2 du \right) \\
  &\quad + 3E \left( \int_0^t \int_{\mathbb{R}\setminus\{0\}} \left( \frac{\tilde{\rho}(u, S^y_m(u), z) - \tilde{\rho}(u, S^x_m(u), z)}{M(u)} \right)^2 \vartheta(dz) du \right) \\
  &\leq 3(y - x)^2 + 6L^2 \int_0^t \phi(u) \, du,
\end{align*}
\]

whereby the first inequality follows from an elementary inequality and Itô’s isometry while the second one is justified by Lipschitz continuity and Tonelli’s theorem. Next, Grönwall’s inequality implies

\[
\phi(t) \leq 3(y - x)^2e^{6L^2t}, t \in [0, T],
\]

and observing that the right-hand side of (4.3.3) is increasing in \( t \) gives the
4.3. CONVEXITY OF EUROPEAN CONTINGENT CLAIM VALUE

Step 4 is to prove: if \( x, y > 0, x \neq y \), then

\[
v'_{m,+}(x) = \lim_{y \downarrow x} \frac{v_m(y) - v_m(x)}{y - x} = \mathbb{E} \left[ h'_+ (S^x_m(T)) \xi^x_m(T) \right], 
\]

(4.3.4)

\[
v'_{m,-}(x) = \lim_{y \uparrow x} \frac{v_m(y) - v_m(x)}{y - x} = \mathbb{E} \left[ h'_- (S^x_m(T)) \xi^x_m(T) \right], 
\]

(4.3.5)

with

\[
\xi^x_m(T) = \mathcal{E} \left( \int_0^T \rho'(u, S^x_m(u^-)) \, dW(u) + \int_0^T \int_{\mathbb{R} \setminus \{0\}} \tilde{\rho}'(u, S^x_m(u^-), z) \, \tilde{J}(du, dz) \right). 
\]

Observe that our assumptions allow us to employ the theory of stochastic flows for general semimartingales (we use Protter (2005), Chapter V.7, Theorem 39 in the sequel; see also Kunita (2004) and the references therein): for almost all \( \omega \in \Omega \) the function \( x \mapsto S^x_m(t) \) is continuously differentiable. Phrased differently, there exists \( N_1 \) with \( \mathbb{P}(N_1) = 0 \) such that for all \( \omega \in \Omega \setminus N_1 \) the function \( D^x_m(t) := (\partial / \partial x)S^x_m(t) \) is defined. We only consider such \( \omega \) in the sequel. Then \( D^x_m(t) \) solves the SDE given by

\[
\frac{dD^x_m(t)}{D^x_m(t-)} = \left[ r(t) \, dt + \rho'(t, S^x_m(t)) \, dW(t) + \int_{\mathbb{R} \setminus \{0\}} \tilde{\rho}'(t, S^x_m(t), z) \, \tilde{J}(dt, dz) \right],
\]

with \( D^x_m(0) = 1 \). An application of Itô’s formula then yields

\[
D^x_m(t) = \xi^x_m(t) M(t), 
\]

(4.3.6)

\[
\xi^x_m(t) = \mathcal{E} \left( \int_0^t \rho'(u, S^x_m(u^-)) \, dW(u) + \int_0^t \int_{\mathbb{R} \setminus \{0\}} \tilde{\rho}'(u, S^x_m(u^-), z) \, \tilde{J}(du, dz) \right). 
\]

(4.3.7)

We only argue for (4.3.4) since (4.3.5) is established analogously. Observe that

\[
\limsup_{y \downarrow x} \frac{v_m(y) - v_m(x)}{y - x} = \mathbb{E} \left[ \limsup_{y \downarrow x} \left( h'_+ (S^y_m(T)) \frac{S^y_m(T) - S^x_m(T)}{M(T)(y - x)} \right) \right] \leq \mathbb{E} \left[ \limsup_{y \downarrow x} h'_+ (S^y_m(T)) \frac{S^y_m(T) - S^x_m(T)}{M(T)(y - x)} \right]. 
\]
= \mathbb{E}[h'_+(S^x_m(T))\xi^x_m(T)],

whereby the uniform integrability used in the first equality is implied by Step 3 and the last equality holds since \( h \) has bounded one-sided right-continuous derivatives and because of (4.3.6). Conversely, it follows similarly from (4.3.2) that

\[
\liminf_{y \downarrow x} \frac{v_m(y) - v_m(x)}{y - x} = \mathbb{E}[h'_+(S^x_m(T))\xi^x_m(T)].
\]

We conclude that (4.3.4) holds.

**Step 5** is to prove: \( \xi^x_m (t) = (\xi^x_m(t))_{t \in [0,T]} \) given by (4.3.7) is a positive martingale.

Since \( \rho' \) and \( \tilde{\rho}' \) are bounded in \( s \) (uniformly in \( t \) respectively in \( (t,z) \)), the process

\[
\left( \int_0^t \rho'(u, S^x_m(u-)) \, dW(u) + \int_0^t \int_{R \setminus \{0\}} \tilde{\rho}'(u, S^x_m(u-), z) \, J(du, dz) \right)_{t \in [0,T]}
\]

is a martingale of bounded mean oscillation under \( \mathbb{P} \), also referred to as \( \text{BMO}(\mathbb{P})\)-martingale\(^1\). Moreover, as \( \tilde{\rho}' > -1 + \epsilon \), Kazamaki’s criterion (see Kazamaki (1979)) yields the claim.

**Step 6** is to prove: \( \nu'_m, \pm \) is non-decreasing.

Define a new probability measure \( \mathbb{P}^x \) on \( (\Omega, \mathcal{F}) \) by \( d\mathbb{P}^x / d\mathbb{P} = \xi^x_m(T) \). According to Step 5, we can apply Girsanov’s theorem (cf. Lemma [I.A.1]) to deduce that

\[
W^x(t) = W(t) - \int_0^t \rho'(u, S^x_m(u)) \, du
\]

\(^1\)For a martingale \( X \) with càdlàg paths, denote the quadratic variation by \([X, X]\) and define the \( H^p \) norm of \( X \) for any \( p \geq 1 \) by \(|X||_{H^p} := \mathbb{E} \left[ [X, X]^p_T / 2 \right]^{1/p} \). A martingale \( X \) is said to be of \( \text{BMO} \) if it is in \( H^2 \) and there exists a constant \( c \) such that for any stopping time \( \tau \leq T \) it holds that \( \mathbb{E} \left[ (X_T - X_\tau)^2 | F_\tau \right] \leq c^2 \) a.s. (see Protter (2005), Chapter IV.4 for further details).
4.3. CONVEXITY OF EUROPEAN CONTINGENT CLAIM VALUE

is a $\mathbb{P}^x$- Brownian motion and

$$\tilde{J}^x(dt, dz) = \tilde{J}(dt, dz) - \tilde{\rho}'(t, S^x_m(t), z) \vartheta(dz)dt$$

is a $\mathbb{P}^x$- compensated Poisson random measure. In particular, the $\mathbb{P}^x$-compensator $\vartheta^x(dz)dt$ of $J(dt, dz)$ is given by

$$\vartheta^x(dz)dt := (\tilde{\rho}'(t, S^x_m(t), z) + 1) \vartheta(dz)dt.$$

Consequently, we can rewrite (4.2.5) as

$$dS^x_m(t) = S^x_m(t) r(t) dt + \rho(t, S^x_m(t)) \rho'(t, S^x_m(t)) dt + \rho(t, S^x_m(t)) dW^x(t)$$

$$+ \int_{\mathbb{R}\setminus\{0\}} \tilde{\rho}(t, S^x_m(t-), z) \tilde{J}^x(dt, dz)$$

$$+ \int_{\mathbb{R}\setminus\{0\}} \tilde{\rho}(t, S^x_m(t-), z) \tilde{\rho}'(t, S^x_m(t-), z) \vartheta(dz)dt,$$

with $S^x_m(0) = x$, and note that uniqueness in law holds for solutions to this SDE (Applebaum (2009), Chapter 6). Define $\tilde{S}^x_m$ as solution to

$$d\tilde{S}^x_m(t) = \tilde{S}^x_m(t) r(t) dt + \rho(t, \tilde{S}^x_m(t)) \rho'(t, \tilde{S}^x_m(t)) dt + \rho(t, \tilde{S}^x_m(t)) dW(t)$$

$$+ \int_{\mathbb{R}\setminus\{0\}} \tilde{\rho}(t, \tilde{S}^x_m(t-), z) \tilde{J}(dt, dz)$$

$$+ \int_{\mathbb{R}\setminus\{0\}} \tilde{\rho}(t, \tilde{S}^x_m(t-), z) \tilde{\rho}'(t, \tilde{S}^x_m(t-), z) \vartheta(dz)dt,$$

with $\tilde{S}^x_m(0) = x$. Observe that the process $\tilde{S}^x_m$ has the same distribution under $\mathbb{P}$ as the process $S^x_m$ under $\mathbb{P}^x$. Calculating $v_{m,\pm}'$ under $\mathbb{P}^x$, we conclude from STEP 4 that $v_{m,+}'(x) = \mathbb{E}^x[h_+(S^x_m(T))]$ and $v_{m,-}'(x) = \mathbb{E}^x[h_-(S^x_m(T))]$ and therefore

$$v_{m,\pm}'(x) = \mathbb{E}[h_{\pm}'(\tilde{S}^x_m(T))], \ x > 0. \quad (4.3.8)$$

For $0 < x < y$, following the same line of reasoning as in STEP 1, one can show that $\tilde{S}^x_m(T) \leq \tilde{S}^y_m(T)$. Since $h_{\pm}'$ is non-decreasing, monotonicity of the expected value implies that $v_{m,\pm}'$ is also non-decreasing, which is equivalent to $v_m(x)$ being convex w.r.t. $x$. \qed
Note that Theorem 4.3.2 generalizes Theorem 5.2 in El Karoui et al. (1998) to jump-diffusions. We conclude from Theorem 4.3.2 that the Delta strategy for the misspecified model always exists and is bounded. The following example shows that the condition \( \rho'(t, s, z) > -1 + \epsilon, \epsilon > 0 \), enforced in the second part of Assumption 4.2.3 is not only necessary in Step 5 of the proof of Theorem 4.3.2 but that it is also inevitable to establish the monotonicity of the mapping \( x \mapsto S^x_m(t) \) in Step 1, without which convexity would not hold.

**Example 4.3.3.** Suppose \( r \equiv 0, \rho'(t, s, z) = -2 \) and the stock price is driven by a compensated homogeneous Poisson process \( \tilde{N} = (\tilde{N}(t))_{t \in [0, T]} \) with intensity \( \lambda > 0 \). Assuming the constant of integration is equal to 4, the stock price dynamics is given by

\[
dS^x_m(t) = (-2S^x_m(t-) + 4) \, d\tilde{N}(t).
\]

Denote the first jump time of the Poisson process by \( \tau := \inf\{t > 0 : N(t) = 1\} \). Then we obviously have

\[
S^x_m(\tau) = x + (-2x + 4) \cdot (1 - \lambda\tau).
\]

Choosing \( \lambda = 0.1 \) and defining \( B := \{\omega \in \Omega : \tau(\omega) < 5\} \), we see that

\[
S^1_m(\tau)1_B = (3 - 0.02\tau)1_B > 2 = S^2_m(\tau)1_B,
\]

i.e., the monotonicity property no longer holds a.s. because \( \mathbb{P}(B) > 0 \). Obviously, the parameters can be chosen such that \( B \) has probability arbitrarily close to one.

### 4.4 Robustness of the Delta Hedge

In this section we consider an investor intending to approximate \( h(S(T)) \) by means of its Delta strategy. It is assumed that the investor bases her computation of the Delta strategy on the misspecified model price (4.2.5) under
the reference measure $\mathbb{P}$. We first characterize the induced hedging error and subsequently deduce from its characteristics certain robustness properties. Throughout this and the next section we need the following:

**Definition 4.4.1.** We say a measure $Q \in Q_{em}$ satisfies Condition (I) if it holds for every $(t, s) \in [0, T] \times \mathbb{R}_+$ that there exists some constant $L > 0$ so that

\[
\int_{\mathbb{R}\{0\}} \tilde{\rho}(t, s, z)^2 \vartheta_Q(dz) \leq L \cdot (1 + |s|^2), \text{ and}
\]

\[
S^x_m \in L^2(dQ \times dt), \quad x > 0.
\]

We denote

\[
Q := \{ Q \in Q_{em} \mid Q \text{ satisfies Condition (I)} \}.
\]

Consider for all $x \in \mathbb{R}_+$ and $t \in [0, T]$ the partial integro-differential equation (PIDE) given by

\[
0 = \dot{g}(t, x) + r(t)xg'(t, x) + \frac{1}{2}x^2 \gamma^2(t, x)g''(t, x) - r(t)g(t, x)
\]

\[
+ \int_{\mathbb{R}\{0\}} (g(t, x + x\tilde{\gamma}(t, x, z)) - g(t, x) - x\tilde{\gamma}(t, x, z)g'(t, x)) \vartheta(dz),
\]

\[
g(x, T) = h(x).
\]

**(4.4.1)**

**Assumption 4.4.2.** We assume the existence of a classical solution $g : [0, T] \times \mathbb{R}_+ \rightarrow \mathbb{R}$ to the PIDE (4.4.1) whose derivatives in the second variable are bounded by a polynomial function of $x$, uniformly in $t \in [0, T]$.

For results on the existence of classical solutions to (4.4.1) see for instance [Bensoussan and Lions (1982)] or [Cont and Voltchkova (2005)] (for the purely Brownian case conditions are discussed in [Friedman (1964)]). The Feynman-Kac theorem (cf. [Kromer et al. (2015) and the references therein]) then implies that the solution $g$ to the PIDE (4.4.1) coincides with the misspecified value of the contingent claim from Definition 4.3.1, i.e.,

\[
g(t, x) = v_m(t, x) = \mathbb{E} \left[ h(S^x_m(T)) e^{-\int^T_t r(u) \, du} \right], \quad t \in [0, T], \quad x > 0.
\]

**(4.4.2)**
We immediately obtain the following corollary that is necessary to extract trading strategies and uniform bounds from price functions:

**Corollary 4.4.3.** Suppose the conditions of Theorem 4.3.2 are satisfied. Then the convex European contingent claim value function \( v_m \) given by (4.4.2) satisfies

\[
|v_{m,\pm}'(x)| \leq ||h_\pm'||_{L^\infty} = \sup_{y \in \mathbb{R}^+} |h_\pm'(y)|, \; x > 0.
\]

**Proof.** The claim is immediate from Definition 4.2.2 and (4.3.8). \( \square \)

Note that \( v_m' \) is an admissible trading strategy. Suppose now the investor follows the Delta strategy \( v_m' = (v_m'(t, \cdot))_{t \in [0,T]} \). The trading is done in the physical stock \( S \), whose price is described by (4.2.2). Then the corresponding self-financing hedging portfolio \( P^{v_m'} = (P^{v_m'}(t))_{t \in [0,T]} \) solves the SDE

\[
dP^{v_m'}(t) = P^{v_m'}(t)r(t) \, dt + v_m'(t, S(t-))[dS(t) - r(t)S(t) \, dt], \tag{4.4.3}
\]

and the initial capital to set up the hedging portfolio coincides with the initial misspecified price of the claim, i.e., \( P^{v_m'}(0) = v_m(0, x) \). Observe that \( P^{v_m'} \) is typically not Markov. We formally define the hedging error (in the misspecified model) \( e_m = (e_m(t))_{t \in [0,T]} \) by

\[
e_m(t) := P^{v_m'}(t) - v_m(t, S(t)). \tag{4.4.4}
\]

This hedging error at time \( t \) obviously displays the difference between the value of the hedging portfolio and the misspecified claim price upon observing \( S(t) \) quoted at the stock exchange. We are particularly interested in the hedging error’s value at time \( T \) because \( v_m(T, S(T)) = h(S(T)) \). So \( e(T) \) indicates how far off an investor following \( P^{v_m'} \) is from the claim’s payoff due to both misspecification and market incompleteness. The next proposition gives an explicit formula for the discounted hedging error \( e_m/M = (e_m(t)/M(t))_{t \in [0,T]} \).

**Proposition 4.4.4.** Suppose Assumption 4.2.1, Assumption 4.2.3 and Assumption 4.4.2 are satisfied. Consider the hedging error \( e_m \) defined by (4.4.4).
4.4. ROBUSTNESS OF THE DELTA HEDGE

The discounted hedging error reads

\[
\frac{e_m(t)}{M(t)} = \frac{1}{2} \int_0^t \frac{1}{M(u)} v''_m(u, S(u)) S(u)^2 [\gamma(u, S(u))^2 - \sigma(u)^2] \, du \\
+ \int_0^t \int_{\mathbb{R}\setminus\{0\}} \frac{1}{M(u)} \left( v_m(u, S(u-)) + S(u-\tilde{\gamma}(u, S(u-), z)) \\
- v_m(u, S(u-)) + S(u-)\eta(u, z)) \\
+ v'_m(u, S(u)) S(u-) (\eta(u, z) - \tilde{\gamma}(u, S(u-), z)) \right) \, \vartheta(dz) du \\
- \int_0^t \int_{\mathbb{R}\setminus\{0\}} \frac{1}{M(u)} \left( v_m(u, S(u-)) + S(u-\eta(u, z)) - v_m(u, S(u-)) \\
- S(u-)\eta(u, z)v'_m(u, S(u)) \right) \, \tilde{J}(du, dz).
\]

(4.4.5)

**Proof.** Observe that

\[
dv_m(t, S(t)) = \dot{v}_m(t, S(t)) \, dt + v'_m(t, S(t)) \, dS(t) + \frac{1}{2} v''_m(t, S(t)) S(t)^2 \sigma(t)^2 \, dt \\
+ \int_{\mathbb{R}\setminus\{0\}} \left( v_m(t, S(t-)) + S(t-)\eta(t, z)) - v_m(t, S(t-)) \\
- v'_m(t, S(t-)) S(t-)\eta(t, z) \right) \, J(dt, dz) \\
= r(t) v_m(t, S(t)) \, dt + v'_m(t, S(t)) [dS(t) - r(t) S(t) \, dt] \\
+ \frac{1}{2} v''_m(t, S(t)) S(t)^2 [\sigma(t)^2 - \gamma(t, S(t))^2] \, dt \\
+ \int_{\mathbb{R}\setminus\{0\}} \left( v_m(t, S(t-)) + S(t-)\eta(t, z)) - v_m(t, S(t-)) \\
- S(t-)\eta(t, z)v'_m(t, S(t)) \right) \, \tilde{J}(dt, dz) \\
+ \int_{\mathbb{R}\setminus\{0\}} \left( v_m(t, S(t-)) + S(t-)\eta(t, z)) - v_m(t, S(t-)) + S(t-)\tilde{\gamma}(t, S(t-), z)) \right)
\]
\[ + v'_m(t, S(t)) S(t-) (\tilde{\gamma}(t, S(t-), z) - \eta(t, z)) \] \[ \vartheta(dz) dt, \]

where the first equality is an application of Itô's lemma and the second one follows from (4.4.1). Next we calculate the difference \( de_m(t) = d(P v'_m(t) - v_m(t, S(t))) \). This yields

\[ de_m(t) = (P v'_m(t) - v_m(t, S(t))) r(t) dt \]

\[ + \frac{1}{2} v''_m(t, S(t)) S(t)^2 [\gamma(t, S(t))^2 - \sigma(t)^2] dt \]

\[ + \int_{\mathbb{R} \setminus \{0\}} \left( v_m(t, S(t-) + S(t-) \tilde{\gamma}(t, S(t-), z) - v_m(t, S(t-) + S(t-) \eta(t, z)) \right. \]

\[ \left. - v'_m(t, S(t-)) S(t-) (\tilde{\gamma}(t, S(t-), z) - \eta(t, z)) \right) \vartheta(dz) dt \]

\[ - \int_{\mathbb{R} \setminus \{0\}} \left( v_m(t, S(t-) + S(t-) \eta(t, z)) - v_m(t, S(t-)) \right. \]

\[ \left. - S(t-) \eta(t, z) v'_m(t, S(t-)) \right) \tilde{J}(dt, dz). \]

An application of the integration by parts formula to \( d(e_m(t)/M(t)) \) then establishes (4.4.5).

The following lemma provides among others valuable insight into monotonicity properties of the (expected) discounted hedging error, but it is formulated in slightly more general terms and used repeatedly throughout the remainder of this paper.

**Lemma 4.4.5.** Consider the process \( S \) given by (4.2.2) and let \( g \) be some function that is convex in the second component. Assume further that

\[ \sigma(t) \leq \gamma(t, S(t)) \text{ and } \text{sgn}(\tilde{\gamma}(t, S(t), z) - \eta(t, z)) = \text{sgn}(\eta(t, z)), \quad (4.4.6) \]

\( d\mathbb{P} \times dt \) and \( d\mathbb{P} \times dt \times \vartheta(dz) \)-a.s. Then the process \( \Pi = (\Pi(t))_{t \in [0,T]} \) defined
4.4. ROBUSTNESS OF THE DELTA HEDGE

by

\[ \Pi(t) := \frac{1}{2} \int_0^t g''(u, S(u))S(u)^2[\gamma(u, S(u))^2 - \sigma(u)^2] \, du \]

\[ + \int_0^t \int_{\mathbb{R}\setminus\{0\}} \left( g(u, S(u) + S(u-\tilde{\gamma}(u, S(u), z)) - g(u, S(u)+S(u-\eta(u, z))) + g'(u, S(u-))S(u-)(\eta(u, z) - \tilde{\gamma}(u, S(u), z)) \right) \vartheta(dz) du, \]

(4.4.7)

is non-decreasing a.s.

Proof. The first integral in (4.4.7) is easily seen to be non-decreasing because \( g''(\cdot, \cdot) > 0 \) by convexity of \( g \) and the left part of the condition (4.4.6). Regarding the integrand of the double integral, observe that

\[ g(t, S(t-)+S(t-\tilde{\gamma}(t, S(t-), z)) - g(t, S(t-)+S(t-\eta(t, z))) + g'(t, S(t-))S(t-)(\eta(t, z) - \tilde{\gamma}(t, S(t-), z)) \]

\[ \geq S(t-) (\tilde{\gamma}(t, S(t-), z) - \eta(t, z)) (g'(t, S(t-)+S(t-\eta(t, z))) - g'(t, S(t-))) \geq 0, \]

where the first inequality follows from the convexity of \( g \) and the second inequality holds because the right part of the condition (4.4.6) and the fact that \( g' \) is monotonically non-decreasing imply that the two brackets always have the same sign inducing a non-negative product. We conclude that the double integral is non-decreasing as well and so is the process \( \Pi \). 

We see that Lemma 4.4.5 implies that the first two summands of the discounted hedging error (4.4.5) are non-decreasing. However, the third term cannot be positive because it is a martingale. The fact that the hedging error does not exhibit any a.s. monotonicity properties despite a systematic overestimation has a straightforward economic interpretation: due to the market incompleteness, the Delta strategy does neither provide a perfect hedge nor an a.s. sub- or superhedge. Nonetheless, it is possible to deduce robustness properties of the expected discounted hedging error. To this end,
consider the set of EMMs $Q$ and its characterization in Lemma 4.1. Define

$$Q_0 := \{ Q \in Q \mid \vartheta_Q(A) \leq \vartheta(A) \text{ for any } A \in \mathcal{B}(\mathbb{R}\setminus\{0\}) \}. \quad (4.4.8)$$

Note that the set of Lévy measures satisfying the inequality in the definition of $Q_0$ is convex. Let $\text{cl}(\text{conv}(Q_0))$ denote the closure of the convex hull of $Q_0$. The next theorem states that the discounted hedging error is a $\text{cl}(\text{conv}(Q_0))$-submartingale meaning that it is a submartingale w.r.t. every measure $Q \in \text{cl}(\text{conv}(Q_0))$.

**Theorem 4.4.6.** Suppose the conditions of Proposition 4.4.4 are satisfied. If

$$\sigma(t) \leq \gamma(t, S(t)) \text{ and } \text{sgn}(\tilde{\gamma}(t, S(t), z) - \eta(t, z)) = \text{sgn}(\eta(t, z)),$$

$d\mathbb{P} \times dt$ and $d\mathbb{P} \times dt \times \vartheta(dz)$-a.s., then the induced discounted hedging error $e_m/M$ specified by (4.4.5) is a $\text{cl}(\text{conv}(Q_0))$-submartingale. In particular,

$$\inf_{Q \in \text{cl}(\text{conv}(Q_0))} \mathbb{E}_Q \left[ \frac{e_m(t)}{M(t)} \right] \geq 0 \text{ for all } t \in [0, T].$$

**Proof.**

**Step 1** is to observe that for every $Q \in Q$ and some function $f : \mathbb{R}_+ \to \mathbb{R}$ such that $f'$ is bounded by some constant $K > 0$, it holds that

$$\int_{\mathbb{R}\setminus\{0\}} f'(S(t-))^2S(t)^2 \vartheta_Q(dz) \leq K^2 \int_{\mathbb{R}\setminus\{0\}} S(t)^2 \tilde{\gamma}(t, S(t), z)^2 \vartheta_Q(dz) \leq K^2 L(1 + |S(t)|^2). \quad (4.4.9)$$

**Step 2** is to establish the claim of the theorem.

Let $Q \in Q_0$ arbitrary. Performing a change of measure from $\mathbb{P}$ to $Q$ in (4.4.5), the resulting compensated jump integral is consequently defined under $Q$, see (4.A.2) for the specification of $\tilde{J}_Q(\cdot, \cdot)$. From Corollary 4.4.3 and (4.4.9) we see that the process

$$\left( \int_0^t \int_{\mathbb{R}\setminus\{0\}} \left( v_m(u, S(u-) + S(u-)\eta(u, z)) - v_m(u, S(u-)) \right) \, d\tilde{J}_Q(u, z) \right)$$
4.4. ROBUSTNESS OF THE DELTA HEDGE

$$- S(u-)\eta(u, z)v'_m(u, S(u-)) \int_Q(du, dz)_{t \in [0, T]}$$

is a true martingale under each $Q$. Thus, fixing $0 \leq w < t \leq T$, we can calculate that

$$\mathbb{E}_Q \left[ \frac{e_m(t)}{M(t)} | \mathcal{F}_w \right] = \frac{e_m(w)}{M(w)} + \mathbb{E}_Q \left[ \frac{1}{2} \int_w^t \frac{1}{M(u)} v''_m(u, S(u)) S(u)^2 [\gamma(u, S(u))^2 - \sigma(u)^2] \, du \right.$$

$$+ \int_w^t \int_{\mathbb{R} \setminus \{0\}} \frac{1}{M(u)} \left( v_m(u, S(u-) + S(u-)\tilde{\gamma}(u, S(u-), z)) - v_m(u, S(u-)) + v'_m(u, S(u))S(u-) \eta(u, z) \right.$$  

$$+ \theta(u, z)(v_m(u, S(u-) + S(u-)\eta(u, z)) - v_m(u, S(u-)))$$  

$$- S(u-)\eta(u, z)v'_m(u, S(u))) \right) \psi(dz)du \bigg| \mathcal{F}_w \bigg] \geq \frac{e_m(w)}{M(w)},$$

(4.4.10)

where the last inequality follows from combining Theorem 4.3.2 with Lemma 4.4.5 and because $\theta(t, z) \geq 0$, $d\mathbb{P} \times dt \times \psi(dz)$-a.s., whenever $Q \in \mathcal{Q}_0$. This implies that $e_m/M$ is a $\mathcal{Q}_0$-submartingale. Moreover, we easily conclude from (4.4.10) that the hedging error has a non-negative value at any time $t \in [0, T]$ under each measure $Q \in \mathcal{Q}_0$ since choosing $w = 0$ yields

$$\mathbb{E}_Q \left[ \frac{e_m(t)}{M(t)} | \mathcal{F}_0 \right] \geq \frac{e_m(0)}{M(0)} = 0.$$  

(4.4.11)

If (4.4.11) holds for every $Q \in \mathcal{Q}_0$, it clearly also holds for the closure of its convex hull. □

Compared to the reference measure $\mathbb{P}$, two properties of the measures
Q ∈ Q_0 are noteworthy. First, under the measures Q ∈ Q_0, the process S^x_m is not necessarily Markov. It is only assumed to be Markov under the measure the Delta hedging strategy is calculated. Second, the measures contained in Q_0 have less probability mass on the jump part. From an investor’s point of view this goes along with a risk decrease in jump uncertainty.

As the discounted hedging error starts at zero and is a Q_0-submartingale, it is increasing on average. Thus, (4.4.4) implies that the hedging portfolio \( \tilde{P}^{v_m} \) tends to result in a super-replication of \( h(S(T)) \).

Using our results on the hedging error, we close this section with a corollary stating that systematic overestimations of the volatility and the jump sensitivity lead to a domination of the true contingent claim price by the misspecified price. The reason is that the overestimations make the option more valuable benefiting from the additional volatility.

**Corollary 4.4.7.** Suppose the conditions of Theorem 4.4.6 are satisfied. Then the process \( (v_m(t, S(t))/M(t))_{t \in [0,T]} \) with \( v_m \) given by (4.4.2) is a supermartingale. In particular, \( v_m(0, S(0)) \geq E[h(S(T))/M(T)] \). So the misspecified model also gives a higher (more conservative) claim price.

**Proof.** Since \( e_m/M \) is a submartingale under \( P \) and Assumption 4.2.1 implies that \( P^{v_m}/M \) given by (4.4.3) is a martingale, we conclude from (4.4.4) that \( (v_m(t, S(t))/M(t))_{t \in [0,T]} \) is a supermartingale. Consequently, it holds that

\[
v_m(0, S(0)) = \frac{v_m(0, S(0))}{M(0)} \geq E\left[\frac{v_m(T, S(T))}{M(T)}\right] = E[h(S(T))/M(T)].
\]

\( \square \)

### 4.5 Robust Pricing

Previously we obtained structural properties of the hedging error that hold true for a certain subset of \( Q \). In this section we analyze general subsets \( M \subseteq Q \) such that the hedging error exhibits robustness properties. Normalizing the interest rate to zero for convenience for the remainder of the paper, we
define the pricing operator at time $t$ of the European claim $h(S(T))$ in the
misspecified model by

$$C_m^\mathcal{M}(t) := \text{ess sup}_{Q \in \mathcal{M}} \mathbb{E}_Q \left[ h(S_{m}^{x}(T)) \right| \mathcal{F}_t].$$ (4.5.1)

In the sequel we omit the superscript $\mathcal{M}$ unless there is ambiguity. A non-
exhaustive list of examples for the choice of $\mathcal{M}$ leading to a well known robust
pricing operator $C_m$ is the following:

i) **COMPLETE MARKET**: The set $\mathcal{M}$ can be a singleton set. Then its
only element is the reference risk-neutral measure $\mathbb{P}$. Thus, the market
is complete and $C_m$ provides a perfect hedge. For robustness properties
in this case, we refer to the references mentioned in the introduction or
Corollary 4.6.4 below.

ii) **GOOD-DEAL BOUND PRICING**: Pricing contingent claims by only re-
quiring the absence of arbitrage leads to a relatively wide range of
prices; Cochrane and Saá-Requejo (2000) outline that prices beyond
a certain benchmark correspond to unreasonably good deals. To over-
come this, they suggest to narrow the no-arbitrage bounds by imposing
bounds on the Sharpe ratio in a Brownian setting. In Björk and Slinko
(2006) these results are extended to Lévy processes, see Jaschke and
Küchler (2001) for the connection between good-deal bounds and co-
herent risk measures, and Staum (2004) for a fundamental theorem
of asset pricing for good-deal bounds. Using the notation of Lemma
4.A.1 one specifies

$$\mathcal{M} := \left\{ Q \in \mathcal{Q} \left| \psi(t)^2 + \int_{\mathbb{R}\setminus\{0\}} \theta(t,z)^2 \vartheta(dz) \leq B, \right. \right.$$ 

$$\text{for Lebesgue-almost all } t \in [0,T] \right\},$$

for some constant $B > 0$. The inequality in the specification of $\mathcal{M}$
is equivalent to bounding the Sharpe ratio by $B$. The process $C_m$ in
this case is called the **upper good deal price bound process**, the lower
one is constructed accordingly. Consequently, all prices in-between are
considered to be good deals. Arai and Fukasawa (2014) describe the good-deal bounds through convex risk measures. Another possibility of narrowing down the pricing interval is based on the notion of acceptability in terms of a collection of test measures and associated floors, see Carr et al. (2001) for details.

iii) Ball scenarios: A measure $P$ qualifies to be a reference risk-neutral measure if it is plausible from an investor’s point of view. To obtain an equally reasonable price range of the claim $h(S(T))$ in (4.5.1), one can consider market models $Q \in \mathcal{Q}$ in a small ball around the reference model $P$. Noting that the choice $\psi = \theta = 0$ in (4.5.2) leads to $P$, one specifies

$$\mathcal{M} := \{Q \in \mathcal{Q} \mid |\psi(t)| \leq B_1, |\theta(t,z)| \leq B_2, \text{ for Lebesgue-almost all } t \in [0,T] \text{ and } \vartheta(dz)\text{-a.s.}\},$$

for some constants $B_1, B_2 > 0$.

iv) Coherent risk measures: Consider for a terminal payment obligation $X$ the minimization problem

$$\tilde{\rho}(X) := \inf_{H \in \mathcal{V}_T} \rho(X - H),$$

where $\rho$ is a coherent risk measure, see for instance Artzner et al. (1999) for a precise definition, and $\mathcal{V}_T$ is the set of hedging instruments. It is well known that there exists a set $\mathcal{R}$ of probability measures such that $\rho(X) = \sup_{Q \in \mathcal{R}} \mathbb{E}_Q[X]$ and it follows from Corollary 3.9 in Barrieu and El Karoui (2009) that

$$\tilde{\rho}(X) = \sup_{Q \in \mathcal{R} \cap \mathcal{Q}_{em}} \mathbb{E}_Q[X].$$

In particular, $\tilde{\rho}$ is of the form (4.5.1).

Denote by $\mathcal{P}$ the predictable $\sigma$-algebra on $[0,T] \times \Omega$ w.r.t. $(\mathcal{F}_t)$. 

Remark 4.5.1. If the set $\mathcal{M}$ is of the form

$$\mathcal{M} = \{Q \in \mathcal{Q}_{em} | (\psi(t), \theta(t, \cdot)) \in H(t, \omega)\},$$

for a compact $\mathcal{P} \times \mathcal{B}(\mathbb{R}) \times \mathcal{B}(L^2(\vartheta(\text{d}z)))$-measurable set $H$, there exists $\widehat{Q} \in \mathcal{M}$ such that for all $t \in [0, T]$: $C_m(t, x) = \mathbb{E}_{\widehat{Q}}[h(S^x_m(T)) | \mathcal{F}_t]$.

An assumption ensuring the Markovian structure of the robust price under the misspecified model is needed. To this end, we define the set of probability measures $\mathcal{A}$ leading to a Markov process $S^x_m$ as

$$\mathcal{A} := \{Q | Q \text{ is a probability measure on } (\Omega, \mathcal{F}) \text{ and } S^x_m \text{ is Markov w.r.t. } Q\}.$$ (4.5.2)

Assumption 4.5.2. We assume that we can restrict ourselves to Markov processes in (4.5.1), i.e., $C_m(t) = \text{ess sup}_{Q \in \mathcal{M} \cap \mathcal{A}} \mathbb{E}_Q[h(S^x_m(T)) | \mathcal{F}_t]$.

Assumption 4.5.2 is typically satisfied in examples because non-trivial prices usually give rise to a measure $Q$ such that $S^x_m$ is Markov. Due to Assumption 4.5.2, we can express the essential supremum in (4.5.1) as deterministic function of time and $S^x_m(t)$, i.e.,

$$C_m(t) = C_m(t, S^x_m(t)), \quad (4.5.3)$$

for $C_m : [0, T] \times \mathbb{R}_+ \to \mathbb{R}$. We deduce another corollary from Theorem 4.3.2.

Corollary 4.5.3. Suppose the conditions of Theorem 4.3.2 and Assumption 4.5.2 are satisfied. Then the function $C_m$ is convex and has bounded one-sided derivatives in the second variable.

Observe that the price $C_m$ at time $T$ is equal to the value of the claim itself because

$$C_m(T, S^T_m(T)) = \text{ess sup}_{Q \in \mathcal{M} \cap \mathcal{A}} \mathbb{E}_Q[h(S^T_m(S^m(T)(T))) = h(S_m(T)). \quad (4.5.4)$$

The following regularity assumption is needed for later applications of Itō’s lemma:
CHAPTER 4. ROBUSTNESS OF DELTA HEDGING

Assumption 4.5.4. We assume that either \(C \in C^{1,2}\) or, if \(\gamma \equiv 0\) and \(\int_{\mathbb{R} \setminus \{0\}} |z| \vartheta(dz) < \infty\), that \(C \in C^{1,1}\) or, if \(\gamma \equiv 0\) and \(\vartheta(\mathbb{R} \setminus \{0\}) < \infty\), that \(C\) is locally Lipschitz in \(t\).

Remark 4.5.5. If \(\gamma \equiv 0\) and \(\int_{\mathbb{R} \setminus \{0\}} |z| \vartheta(dz) < \infty\), then for Itô’s formula (and therefore all results in the sequel) to hold, it is sufficient that \(C \in C^{1,1}\).

The next lemma shows that the conditions enforced in the third part of Assumption 4.5.4 indeed yield Itô’s formula. Recall that \(C'_{m,+}\) denotes the right-hand derivative of \(C_m\).

Proposition 4.5.6. Suppose that \(\vartheta(\mathbb{R} \setminus \{0\}) < \infty\) and \(\gamma \equiv 0\). If \(C_m\) is continuously differentiable in time, then Itô’s formula holds with \(\dot{C}_m = C'_{m,+}\).

Proof. Fix \(t \in [0,T]\) and define

\[
\lambda(t) := -\int_{\mathbb{R} \setminus \{0\}} \tilde{\gamma}(t, S_m^x(t), z) \vartheta(dz),
\]

\[
S^J(t) := \int_0^t \int_{\mathbb{R} \setminus \{0\}} \ln(1 + \tilde{\gamma}(u, S_m^x(u), z)) J(du, dz).
\]

Assumption 4.2.3 and the condition \(\vartheta(\mathbb{R} \setminus \{0\}) < \infty\) ensure that \(\lambda(t)\) is bounded for every \(t \in [0,T]\). Suppressing as usual the dependence on \(\omega\), we can define some function \(f\) by \(f(t, s^J) = C_m(t, \exp(\int_0^t \lambda(u) \, du + s^J))\).

Observe that \(f\) is locally Lipschitz continuous in \(t\). Thus, it is almost everywhere (a.e.) differentiable in \(t\) by the Lebesgue theorem (as it is absolutely continuous) and equal to the integral of its derivative. Corollary 4.5.3 implies that \(C_m\) is a.e. differentiable in its second component and since its derivative is equal to \(C'_{m,+}\), the derivative of \(f\) w.r.t. \(t\) is a.e. equal to

\[
\dot{f}(t, S^J(t)) = \dot{C}_m(t, S_m^x(t)) + S_m^x(t) \lambda(t) C'_{m,+}(t, S_m^x(t)).
\] (4.5.5)

As \(\vartheta(\mathbb{R} \setminus \{0\}) < \infty\) and \(\gamma \equiv 0\), \(J(dt, dz)\) corresponds to a compound Poisson process whose jump times we denote by \((\tau_n)_{n \geq 1}\) with \(\tau_0 = 0\). It is
\[ C_m(t, S^x_m(t)) = f(t, S^J(t)) = \sum_{n \geq 1, \tau_n \leq t} \left( \int_{\tau_{n-1}}^{\tau_n} \dot{f}(u, S^J(u)) \, du + f(\tau_n, S^J(\tau_n)) - f(\tau_{n-1}, S^J(\tau_{n-1})) \right) \]

\[ = \int_0^t \dot{f}(u, S^J(u)) \, du + \sum_{0 \leq u \leq t} (f(u, S^J(u)) - f(u, S^J(u))) \]

\[ = \int_0^t \dot{C}_m(t, S^x_m(t)) \, du + \int_0^t C'_m(u, S^x_m(u)) \, dS^x_m(u) \]

\[ + \int_0^t \int_{\mathbb{R} \setminus \{0\}} \left( C_m(u, S^x_m(u)) - C_m(u, S^x_m(u)) \right) J(du, dz), \]

where the first equation holds because \( S^J \) is constant between the jumps and by the Lebesgue theorem explained before. This gives Itô’s formula. \( \square \)

**Remark 4.5.7.** In the sequel we will for the ease of exposition with a slight abuse of notation denote \( C'_m := C'_{m,+} \) and \( C''_m \cdot \gamma^2 = C'_m \cdot \sigma^2 := 0 \) in case \( \gamma = \sigma = 0 \) even if \( C''_m \) formally does not exist.

Suppose some investor chooses to follow the trading strategy \( C'_m \)

\[ = (C'_m(t, \cdot))_{t \in [0, T]} \] that is computed with respect to the misspecified model and trades in the physical stock \( S \). Then the corresponding self-financing hedging portfolio is given by

\[ P^{C'_m}(t, S(t)) = C_m(0, x) + \int_0^t C'_m(u, S(u)) \, dS(u). \] (4.5.6)

We denote by \( E^{C'_m} = (E^{C'_m}(t))_{t \in [0, T]} \) the hedging error induced, and formally define it as

\[ E^{C'_m}(t) := P^{C'_m}(t, S(t)) - C_m(t, S(t)). \] (4.5.7)

Hence, \( E^{C'_m}(T) \) gives the difference how far off the terminal value of the hedging portfolio \( P^{C'_m} \) is from the payoff \( h(S(T)) \). The following integrability
CHAPTER 4. ROBUSTNESS OF DELTA HEDGING

A condition is needed for the proof of Theorem 4.5.9 below.

**Assumption 4.5.8.** In addition to Assumption 4.2.3, suppose that for all \((t, s) \in [0, T] \times \mathbb{R}_+\) there exists some \(B > 0\) such that

\[
\int_{\mathbb{R} \setminus \{0\}} \ln(1 + \tilde{\gamma}(t, s, z))^2 \vartheta(dz) \leq B,
\]

\[
\int_{\mathbb{R} \setminus \{0\}} \tilde{\gamma}(t, s, z) - \ln(1 + \tilde{\gamma}(t, s, z)) \vartheta(dz) \leq B.
\]

The following theorem analyzes the hedging error \(E_{C_m}^\prime\) and states that it is a submartingale w.r.t. each measure \(Q \in \mathcal{M}\) if the volatility and the jump sensitivity are systematically overestimated.

**Theorem 4.5.9.** Suppose Assumption 4.2.1, Assumption 4.5.2, Assumption 4.5.4 and Assumption 4.5.8 are satisfied. Consider the hedging portfolio \(P_{C_m}^\prime\) given by (4.5.6) and the corresponding hedging error \(E_{C_m}^\prime\) specified by (4.5.7). If

\[
\sigma(t) \leq \gamma(t, S(t)) \quad \text{and} \quad \text{sgn}(\tilde{\gamma}(t, S(t), z) - \eta(t, z)) = \text{sgn}(\eta(t, z)),
\]

\(d\mathbb{P} \times dt\) and \(d\mathbb{P} \times dt \times \vartheta(dz)\)-a.s., then the hedging error \(E_{C_m}^\prime\) is an \(\mathcal{M}\)-submartingale. In particular, \(\inf_{Q \in \mathcal{M}} \mathbb{E}_Q[E_{C_m}^\prime(t)] \geq 0\) for all \(t \in [0, T]\).

**Proof.** The proof is divided into two steps. Recall that the monotonicity of the mapping \(x \mapsto S^x_{m}(t)\) for any \(t\) established in Step 1 of the proof of Theorem 4.3.2 holds up to some set \(\mathcal{N}_1\) of measure zero.

**Step 1** is to show that for fixed \((t, \omega) \in [0, T] \times \Omega \setminus \mathcal{N}_1\) it holds that

\[
\lim_{x \downarrow 0} S^x_{m}(t) = 0 \quad \text{and} \quad \lim_{x \uparrow \infty} S^x_{m}(t) = \infty.
\]

Regarding the first statement, let \((x_k)_{k \in \mathbb{N}}\) be a strictly positive sequence satisfying \(\lim_{k \to \infty} x_k = 0\). Observe that

\[
\lim_{k \to \infty} \mathbb{E}[^{S^x_{m}(t)}] = \lim_{k \to \infty} \mathbb{E}[^{S^x_k(t)}] = \lim_{k \to \infty} S^x_{m}(0) = \lim_{k \to \infty} x_k = 0,
\]

that is, \(S^x_{m}(t)\) converges to 0 in \(L^1\). Hence, there exists a subsequence \((x'_{k})_{k \in \mathbb{N}}\) along which \(S^x_{m}(t)\) converges to 0 a.s. As the mapping \(x \mapsto S^x_{m}(t)\)
is monotonously increasing, the aforementioned subsequence must coincide with \((x_k)_{k \in \mathbb{N}}\).

Regarding the second statement, let \((x_k)_{k \in \mathbb{N}}\) be an arbitrary sequence such that \(\lim_{k \to \infty} x_k = \infty\). By contradiction assume that

\[
\mathbb{P} \left[ \sup_{x > 0} S_{m}^x (t) < \infty \right] = \mathbb{P} \left[ \lim_{k \to \infty} S_{m}^{x_k} (t) < \infty \right] > 0. \tag{4.5.8}
\]

Note that the previous equality holds by monotonicity. Define

\[
B := \left\{ \omega \in \Omega \setminus \mathcal{N}_1 \middle| \lim_{k \to \infty} \left( \int_0^t \gamma(u, S_{m}^{x_k} (u)) \ dW(u) 
+ \int_0^t \int_{\mathbb{R}\setminus\{0\}} \ln(1 + \tilde{\gamma}(u, S_{m}^{x_k} (u) \cdot z)) \tilde{J}(du, dz) \right) = -\infty \right\}.
\]

Then we see that

\[
\mathbb{P} \left[ \sup_{x > 0} S_{m}^x (t) < \infty \right] = \mathbb{P} \left[ \lim_{k \to \infty} S_{m}^{x_k} (t) < \infty \right]
= \mathbb{P} \left[ \lim_{k \to \infty} x_k \exp \left( \int_0^t \gamma(u, S_{m}^{x_k} (u)) \ dW(u) - \frac{1}{2} \int_0^t \gamma(u, S_{m}^{x_k} (u))^2 \ du 
+ \int_0^t \int_{\mathbb{R}\setminus\{0\}} \ln(1 + \tilde{\gamma}(u, S_{m}^{x_k} (u) \cdot z)) \tilde{J}(du, dz)
- \int_0^t \int_{\mathbb{R}\setminus\{0\}} (\tilde{\gamma}(u, S_{m}^{x_k} (u) \cdot z) - \ln(1 + \tilde{\gamma}(u, S_{m}^{x_k} (u) \cdot z))) \vartheta(dz)du \right) < \infty \right]
\leq \mathbb{P} \left[ \lim_{k \to \infty} \left( \int_0^t \gamma(u, S_{m}^{x_k} (u)) \ dW(u) - \frac{1}{2} \int_0^t \gamma(u, S_{m}^{x_k} (u))^2 \ du 
+ \int_0^t \int_{\mathbb{R}\setminus\{0\}} \ln(1 + \tilde{\gamma}(u, S_{m}^{x_k} (u) \cdot z)) \tilde{J}(du, dz)
- \int_0^t \int_{\mathbb{R}\setminus\{0\}} (\tilde{\gamma}(u, S_{m}^{x_k} (u) \cdot z) - \ln(1 + \tilde{\gamma}(u, S_{m}^{x_k} (u) \cdot z))) \vartheta(dz)du \right) = -\infty \right]
= \mathbb{P}(B).
\]

The last equality is justified by Assumption 4.5.8. Observe that
0 \leq X_k := \int_0^t \gamma(u, S_m^x(u)) \, dW(u) \\
+ \int_0^t \int_{\mathbb{R}\setminus\{0\}} \ln(1 + \tilde{\gamma}(u, S_m^x(u)-, z)) \, \tilde{J}(du, dz) \cdot 1_B \\
\leq \int_0^t \gamma(u, S_m^x(u)) \, dW(u) \\
+ \int_0^t \int_{\mathbb{R}\setminus\{0\}} \ln(1 + \tilde{\gamma}(u, S_m^x(u)-, z)) \, \tilde{J}(du, dz) \in L^1.

Note that again Assumption [4.5.8] ensures that the previous term is in $L^1$. Then, by the definition of $B$, it holds that

$$\lim_{k \to \infty} X_k = \begin{cases} 
\infty, & \text{on } B, \\
0, & \text{else.}
\end{cases}$$

It follows from (4.5.8) that $\mathbb{P}(B) > 0$, so we clearly see that

$$\lim_{k \to \infty} \mathbb{E}[X_k] = \infty. \quad (4.5.9)$$

However, as

$$\mathbb{E}[X_k^2] \\
= \mathbb{E} \left[ \left| \int_0^t \gamma(u, S_m^x(u)) \, dW(u) \\
+ \int_0^t \int_{\mathbb{R}\setminus\{0\}} \ln(1 + \tilde{\gamma}(u, S_m^x(u)-, z)) \, \tilde{J}(du, dz) \right|^2 \cdot 1_B \right] \\
\leq \mathbb{E} \left[ \left( \int_0^t \gamma(u, S_m^x(u)) \, dW(u) \\
+ \int_0^t \int_{\mathbb{R}\setminus\{0\}} \ln(1 + \tilde{\gamma}(u, S_m^x(u)-, z)) \, \tilde{J}(du, dz) \right)^2 \right] \\
\leq 2 \mathbb{E} \left[ \left( \int_0^t \gamma(u, S_m^x(u)) \, dW(u) \right)^2 \right]$$
4.5. ROBUST PRICING

\[ + 2 \mathbb{E} \left[ \left( \int_0^t \int_{\mathbb{R}\setminus\{0\}} \ln(1 + \tilde{\gamma}(u, S_m^x(u)), z) \, \tilde{J}(du, dz) \right)^2 \right] \]

\[ = 2 \mathbb{E} \left[ \int_0^t \gamma(u, S_m^x(u))^2 \, du \right] \]

\[ + 2 \mathbb{E} \left[ \int_0^t \int_{\mathbb{R}\setminus\{0\}} \ln(1 + \tilde{\gamma}(u, S_m^x(u)), z))^2 \, \vartheta(dz)du \right] \leq 2T(B_1^2 + B_2^2), \]

for constants \( B_1, B_2 > 0 \), the Cauchy-Schwartz inequality implies that

\[ \mathbb{E}[X_k] \leq \sqrt{\mathbb{E}[X_k^2]} \leq \sqrt{2T(B_1^2 + B_2^2)}, \]

which contradicts (4.5.9). Hence, we conclude that

\[ \mathbb{P} \left[ \sup_{x > 0} S_m^x(t) < \infty \right] = 0 \Rightarrow \mathbb{P} \left[ \sup_{x > 0} S_m^x(t) = \infty \right] = 1. \]

**Step 2** is to establish the claim of the theorem.

An application of Itô’s lemma, compensating the jump-integral and (4.4.9) yield

\[ C_m(t, S_m^x(t)) \]

\[ = C_m(0, x) + \int_0^t \dot{C}_m(u, S_m^x(u)) \, du + \int_0^t C''_m(u, S_m^x(u-)) \, dS_m^x(u) \]

\[ + \frac{1}{2} \int_0^t C''_m(u, S_m^x(u)) \, S_m^x(u)^2 \gamma(u, S_m^x(u))^2 \, du \]

\[ + \int_0^t \int_{\mathbb{R}\setminus\{0\}} \left( C_m(u, S_m^x(u-)) + \Delta S_m^x(u, z) - C_m(u, S_m^x(u-)) \right. \]

\[ - \Delta S_m^x(u, z)C'_m(u, S_m^x(u-)) \left. \right) \tilde{J}_Q(du, dz) \]

\[ + \int_0^t \int_{\mathbb{R}\setminus\{0\}} \left( C_m(u, S_m^x(u-)) + \Delta S_m^x(u, z) - C_m(u, S_m^x(u-)) \right. \]

\[ - \Delta S_m^x(u, z)C'_m(u, S_m^x(u-)) \left. \right) (1 - \theta(u, z)) \, \vartheta(dz)du. \]

Note that the compensated jump integral in (4.5.10) gives rise to a true martingale under each measure \( \mathbb{Q} \in \mathcal{M} \) due to Corollary 4.5.3 and (4.4.9).
Now it follows directly from (4.5.1) that $C_m$ is a supermartingale. Hence, the predictable process $A = (A(t))_{t \in [0,T]}$ given by

$$A(t) = -\int_0^t \dot{C}_m(u, S_m^x(u)) \, du - \frac{1}{2} \int_0^t C''_m(u, S_m^x(u)) \, S_m^x(u)^2 \, \gamma(u, S_m^x(u))^2 \, du$$

$$- \int_0^t \int_{\mathbb{R}\{0\}} \left(C_m(u, S_m^x(u-)) + \Delta S_m^x(u, z) - C_m(u, S_m^x(u-))ight)(1 - \theta(u, z)) \, \vartheta(dz) \, du$$

must be increasing. Let $0 \leq w < t \leq T$. We conclude from the Lebesgue differentiation theorem (cf. Rudin (1987), Chapter 7) that the following inequality holds Lebesgue-a.s.:

$$0 \leq \lim_{w \to t} \frac{A(t) - A(w)}{t - w} = \lim_{w \to t} \frac{1}{t - w} \left(-\int_w^t \dot{C}_m(u, S_m^x(u)) \, duight)$$

$$- \frac{1}{2} \int_w^t C''_m(u, S_m^x(u)) \, S_m^x(u)^2 \, \gamma(u, S_m^x(u))^2 \, du$$

$$- \int_w^t \int_{\mathbb{R}\{0\}} \left(C_m(u, S_m^x(u)) + S_m^x(u) \tilde{\gamma}(u, S_m^x(u), z) - C_m(u, S_m^x(u))ight)$$

$$- S_m^x(u) \tilde{\gamma}(u, S_m^x(u), z) C'_m(u, S_m^x(u))) \right)(1 - \theta(u, z)) \, \vartheta(dz) \, du$$

$$= - \left(\dot{C}_m(t, S_m^x(t)) + \frac{1}{2} C''_m(t, S_m^x(t)) \, S_m^x(t)^2 \, \gamma(t, S_m^x(t))^2 + \int_{\mathbb{R}\{0\}} \left(C_m(t, S_m^x(t)) + S_m^x(t) \tilde{\gamma}(t, S_m^x(t), z) - C_m(t, S_m^x(t))ight)$$

$$- S_m^x(t) \tilde{\gamma}(t, S_m^x(t), z) C'_m(t, S_m^x(t))) \right)(1 - \theta(t, z)) \, \vartheta(dz) \right) \right).$$

(4.5.11)

Observe that (4.5.11) must hold for every $S_m^x(t)$. Since we proved in Step 1 that the image of the mapping $x \mapsto S_m^x(t)$ is the entire positive reals, we see that (4.5.11) holds true when replacing $S_m^x(t)$ by some arbitrary $s \in \mathbb{R}_+$. 
This implies that we can substitute $S_{m}^z(t)$ in particular by the true stock price $S(t)$. Using the definition of $E^{C_m'}$ in (4.5.7), we obtain from (4.5.10) that

$$E^{C_m'}(t) = P^{C_m'}(t, S(t)) - C_{m}(t, S(t))$$

$$= - \int_{0}^{t} \dot{C}_{m}(u, S(u)) \, du - \frac{1}{2} \int_{0}^{t} C_{m}''(u, S(u)) \, S(u)^2 \, \sigma(u)^2 \, du$$

$$- \int_{0}^{t} \int_{\mathbb{R}\setminus\{0\}} \left( C_{m}(u, S(u-)) + \Delta S(u, z) \right) - C_{m}(u, S(u-))$$

$$- \Delta S(u, z) C_{m}''(u, S(u-)) \left( 1 - \theta(u, z) \right) \vartheta(dz) \, du$$

Fix again $0 \leq w < t \leq T$. Since $C_{m}''$ is uniformly bounded according to Corollary 4.5.3, we can deduce from (4.4.9) that

$$\mathbb{E}_{Q}[E^{C_{m}'}(t)|\mathcal{F}_{w}] \geq E^{C_{m}'}(w) - \mathbb{E}_{Q}\left[ \int_{w}^{t} \dot{C}_{m}(u, S(u)) \, du \right]$$

$$+ \frac{1}{2} \int_{w}^{t} C_{m}''(u, S(u)) \, S(u)^2 \, \gamma(u, S(u))^2 \, du$$

$$+ \int_{w}^{t} \int_{\mathbb{R}\setminus\{0\}} \left( C_{m}(u, S(u-)) + S(u-\gamma(u, S(u-), z)) - C_{m}(u, S(u-)) \right)$$

$$- S(u-\gamma(u, S(u-), z)) C_{m}''(u, S(u-)) \left( 1 - \theta(u, z) \right) \vartheta(dz) \, du \left| \mathcal{F}_{w} \right]$$

$$\geq E^{C_{m}'}(w),$$

whereby the first inequality follows from Corollary 4.5.3 combined with Lemma 4.4.5 and the second one from (4.5.11). Hence, $E^{C_{m}'}$ is a submartingale under each measure $Q \in \mathcal{M}$. \qed
CHAPTER 4: ROBUSTNESS OF DELTA HEDGING

Remark 4.5.10. In contrast to Kunita (2004), who uses Kolmogorov’s criterion, we show the bijectivity of stochastic flows under the milder assumptions that the derivatives of the coefficients of $S^x_m$ are locally Lipschitz (instead of globally Lipschitz) and that the mapping $x \mapsto x + \hat{\rho}(t,x,z)$ is homeomorphic only on $\mathbb{R}_+$ (instead of being homeomorphic on $\mathbb{R}$) for fixed $(t,z) \in [0,T] \times \mathbb{R}\{0\}$. Further results on stochastic flows for non-Lipschitz coefficients for multi-dimensional SDEs in a Brownian and a Brownian-Poisson filtration can be found in Zhang (2005) and Qiao and Zhang (2008).

By definition the hedging error’s initial value $E^{C'_{\text{m}}}(0)$ is equal to zero and since the process is a submartingale w.r.t. any measure $Q \in \mathfrak{M}$, it is in expectation non-negative. By (4.5.7), this implies in turn that the hedging portfolio $P^{C'_{\text{m}}}$ super-replicates $h(S(T))$ on average. For that reason Theorem 4.5.9 comes down to the statement that trading according to $C'_{\text{m}}$ is robust with regard to overestimation of the volatility and the jump sensitivity in the sense that the hedging error has a positive price at time 0 under any market model $Q \in \mathfrak{M}$, i.e., $\inf_{Q \in \mathfrak{M}} \mathbb{E}^Q [E^{C'_{\text{m}}}(t)] \geq 0$ for any $t \in [0,T]$.

4.6 Robust Superhedging

In this section we turn to the computation and properties of the superhedging strategy. We first characterize the superhedging price function by comparing it to the theory outlined in the previous section. Subsequently we discuss several special cases in Theorem 4.6.2. We define the cost of super-replication at time $t \in [0,T)$, say $\bar{C}_{\text{m}}(t)$, by

$$\bar{C}_{\text{m}}(t) := \text{ess inf} \left\{ c(t) \in L^2(\mathcal{F}_t) : \text{there exists a self-financing trading} \right\}$$

strategy $y$ such that $P \left( c(t) + \int_t^T y(u-) dS^t_{m,x}(u) \geq h(S^t_{m,x}(T)) \right) = 1$,

\begin{equation}
(4.6.1)
\end{equation}

i.e., it is the least amount of money needed to setup a superhedge. As $h'$ is uniformly bounded, the superhedging price process $\bar{C}_{\text{m}}$ is well-defined. In
4.6. ROBUST SUPERHEDGING

Kramkov (1996) it is shown that

$$\tilde{C}_m(t) = \text{ess sup}_{Q \in \mathcal{Q}_{em}} \mathbb{E}_Q \left[ h(S_{m}^{t,x}(T)) \middle| \mathcal{F}_{t} \right].$$  \hspace{1cm} (4.6.2)

The equality of (4.6.1) and (4.6.2) is known as superhedging duality. Recall the definition of the set $\mathcal{A}$ in (4.5.2) and put $\mathcal{S} := \mathcal{Q}_{em} \cap \mathcal{A}$. The following assumption has an important impact on robustness properties of the superhedge.

Assumption 4.6.1. In (4.6.2) we assume that

$$\tilde{C}_m(t) = \text{ess sup}_{Q \in \mathcal{S}} \mathbb{E}_Q \left[ h(S_{m}^{t,x}(T)) \middle| \mathcal{F}_{t} \right].$$

Assumption 4.6.1 enables us to write

$$\tilde{C}_m(t) = C^S_m(t, S_{m}^{t,x}(t))$$ \hspace{1cm} (4.6.3)

for some deterministic function $C^S_m : [0, T] \times \mathbb{R}_+ \rightarrow \mathbb{R}$. Corollary 4.5.3 implies that $C^S_m$ is convex and has a bounded one-sided derivative in the second component.

We remark that the superhedging price process is not a special case induced by the robust pricing operator $C_m$ in (4.5.1) because

$$\mathcal{M} \cap \mathcal{A} \subseteq \mathcal{Q} \cap \mathcal{A} \subseteq \mathcal{Q}_{em} \cap \mathcal{A} = \mathcal{S},$$

and the set $\mathcal{S}$ in (4.6.3) is in general larger than the other ones by not being restricted to those EMMs satisfying Condition (I) (cf. Definition 4.4.1).

The following theorem is a particular consequence of the assumption of the Markov property on the superhedging price.

Theorem 4.6.2. Consider the process $S_{m}^{x}$ given by (4.2.5), suppose Assumption 4.2.1, Assumption 4.5.8, Assumption 4.6.1 and Assumption 4.5.4 (with $C_m$ replaced by $C^S_m$) are satisfied. Let $\tilde{\gamma}(t, s, z) \neq 0$ for all $(t, s, z) \in [0, T] \times \mathbb{R}_+ \times \mathbb{R} \setminus \{0\}$.

(i) If $\gamma(t, s) > 0$ for all $(t, s) \in [0, T] \times \mathbb{R}_+$, then the superhedging price
function is linear in the second variable, i.e.,

\[ C_m^S(t, S^x_m(t)) = a(t) + b(t)S^x_m(t), \tag{4.6.4} \]

for some deterministic functions of time \(a\) and \(b\).

(ii) Suppose \(\gamma \equiv 0\). If for every \((t, s) \in [0, T] \times \mathbb{R}_+\) there exist \(z_1, z_2 \in \mathbb{R} \setminus \{0\}, z_1 \neq z_2\), in the support of \(\vartheta\) such that \(\tilde{\gamma}(t, s, z_1) \neq \tilde{\gamma}(t, s, z_2)\), then the superhedging price function is linear in the second variable taking the same form as in (4.6.4).

Proof. The proof is conducted in three steps:

**Step 1** is to show that if \(g : \mathbb{R}_+ \to \mathbb{R}\) is convex and differentiable, \(\tilde{\gamma}(t, s, z) \neq 0\) for every \((t, s, z) \in [0, T] \times \mathbb{R}_+ \times \mathbb{R} \setminus \{0\}\), and for fixed \((t, z) \in [0, T] \times \mathbb{R} \setminus \{0\}\) it holds for every \(s \in \mathbb{R}_+\) that

\[ g(s + s\tilde{\gamma}(t, s, z)) = g(s) + s\tilde{\gamma}(t, s, z)g'(s), \]

then \(g\) is linear.

The previous equation states that the tangent in the point \(s\) is revisited by \(g\) in \(s + s\tilde{\gamma}(t, s, z)\). Since \(g\) is convex, its graph lies above all of its tangents. Thus, we conclude that \(g\) is equal to its tangent in \(s\) on the interval \([s, s + s\tilde{\gamma}(t, s, z)]\). Since \(s\) has been arbitrarily chosen, we see that \(g\) is linear on each such interval. We are left arguing that \(g\) is linear on its entire domain, but this is immediate because \(\tilde{\gamma}(t, s, z) \neq 0\) for all \((t, s, z) \in [0, T] \times \mathbb{R}_+ \times \mathbb{R} \setminus \{0\}\) and \(s + s\tilde{\gamma}(\cdot, s, \cdot) \to 0\) for \(s \to 0\) and \(s + s\tilde{\gamma}(\cdot, s, \cdot) \to \infty\) for \(s \to \infty\) as \(\tilde{\gamma}\) is bounded and strictly larger than \(-1\).

**Step 2** is to show that if \(g : \mathbb{R}_+ \to \mathbb{R}\) is convex and differentiable, \(\gamma(t, s, z) \neq 0\) for all \((t, s, z) \in [0, T] \times \mathbb{R}_+ \times \mathbb{R} \setminus \{0\}\), and for every \(z_1, z_2 \in \mathbb{R} \setminus \{0\}\) and \((t, s) \in [0, T] \times \mathbb{R}_+\) such that \(\tilde{\gamma}(t, s, z_1) \neq \tilde{\gamma}(t, s, z_2)\) it holds that

\[ \frac{g(s + s\tilde{\gamma}(t, s, z_1)) - g(s)}{s\tilde{\gamma}(t, s, z_1)} = \frac{g(s + s\tilde{\gamma}(t, s, z_2)) - g(s)}{s\tilde{\gamma}(t, s, z_2)}, \tag{4.6.5} \]
4.6. ROBUST SUPERHEDGING

then $g$ is linear.

Assume w.l.o.g. $0 < \gamma(t, s, z_1) < \gamma(t, s, z_2)$. An easy calculation allows us to deduce from (4.6.5) that

$$
\frac{g(s + s\tilde{\gamma}(t, s, z_1)) - g(s)}{s\tilde{\gamma}(t, s, z_1)} = \frac{g(s + s\tilde{\gamma}(t, s, z_2)) - g(s + s\tilde{\gamma}(t, s, z_1))}{s(\tilde{\gamma}(t, s, z_2) - \tilde{\gamma}(t, s, z_1))}.
$$

(4.6.6)

Observe that

$$
g(s + s\tilde{\gamma}(t, s, z_1)) - g(s) = \int_s^{s + s\tilde{\gamma}(t, s, z_1)} g'(u) \, du
\leq g'(s + s\tilde{\gamma}(t, s, z_1))
\leq \int_{s + s\tilde{\gamma}(t, s, z_1)}^{s + s\tilde{\gamma}(t, s, z_2)} g'(u) \, du
= \frac{g(s + s\tilde{\gamma}(t, s, z_2)) - g(s + s\tilde{\gamma}(t, s, z_1))}{s(\tilde{\gamma}(t, s, z_2) - \tilde{\gamma}(t, s, z_1))},
$$

where the two inequalities follow because $g$ is convex. As (4.6.6) states that all the terms in the previous chain of (in-)equalities are equal, we conclude that

$$
g'(u) = g'(s + s\tilde{\gamma}(t, s, z_1)) \text{ for all } u \in [s, s + s\tilde{\gamma}(t, s, z_1)],
g'(u) = g'(s + s\tilde{\gamma}(t, s, z_1)) \text{ for all } u \in [s + s\tilde{\gamma}(t, s, z_1), s + s\tilde{\gamma}(t, s, z_2)],
$$

i.e., $g$ is linear on the interval (the upper and the lower bound of the interval might switch depending on the sign of $\tilde{\gamma}$). Referring to STEP 1, one can show by analogous arguments that $g$ is linear on its entire domain.

**STEP 3** is to establish the claim of the theorem.

An application of Itô’s lemma to $C^2_m$ yields
CHAPTER 4. ROBUSTNESS OF DELTA HEDGING

\[ C^S_m(t, S^x_m(t)) = C^S_m(0, x) + \int_0^t C^S_m(u, S^x_m(u)) \, du + \int_0^t C^S_m(u, S^x_m(u)) \, dS^x_m(u) \]

\[ + \frac{1}{2} \int_0^t C^S_m(u, S^x_m(u)) S^x_m(u)^2 \gamma(u, S^x_m(u))^2 \, du \]

\[ + \int_0^t \int_{\mathbb{R} \setminus \{0\}} \left( C^S_m(u, S^x_m(u)) - \Delta S^x_m(u, z) \right) \, J(du, dz) \]

\[ + \int_0^t \int_{\mathbb{R} \setminus \{0\}} \left( C^S_m(u, S^x_m(u)) - \Delta S^x_m(u, z) C^S_m(u, S^x_m(u)) \right) \, \psi(dz) \, du. \]

(4.6.7)

The optional decomposition theorem (cf. Theorem 1 in Föllmer and Kabanov (1997)) implies the existence of some predictable \( S^x_m \)-integrable process \( \pi(t) = (\pi(t))_{t \in [0,T]} \) and an increasing adapted process \( A = (A(t))_{t \in [0,T]} \) with \( A(0) = 0 \) such that for every \( t \in [0, T] \) we have

\[ C^S_m(t, S^x_m(t)) = C^S_m(0, x) + \int_0^t \pi(u^-) \, dS^x_m(u) - A(t). \]

(4.6.8)

Equating (4.6.7) and (4.6.8) and using that the process \( A \) is of finite variation because it is increasing, the calculation of the quadratic variation yields the following implication:

\[ 0 = \int_0^t \left( C^S_m(u, S^x_m(u)) - \pi(u^-) \right)^2 S^x_m(u)^2 \gamma(u, S^x_m(u))^2 \, du \]

\[ + \int_0^t \int_{\mathbb{R} \setminus \{0\}} \left( \left( C^S_m(u, S^x_m(u)) - \pi(u^-) \right) S^x_m(u) \gamma(u, S^x_m(u), z) \right) \]

\[ + C^S_m(u, S^x_m(u)) + \Delta S^x_m(u, z) \]
Recall from Step 4 of the proof of Theorem 4.3.2 that the function \( x \mapsto S_m^x(t) \) is differentiable in \( x \) up to a set of measure zero \( \mathcal{N}_1 \). The term in brackets in the last two lines of (4.6.9) is equal to zero for any fixed \( \omega \in \Omega \) up to some set of measure zero \( \mathcal{N}_2 \). Define \( \mathcal{N} := \mathcal{N}_1 \cup \mathcal{N}_2 \). In the sequel, we restrict to \( \omega \in \Omega \backslash \mathcal{N} \). Define \( \mathcal{T}(\omega) := \{ t \in [0, T] : \Delta S_m^x(t, z, \omega) \neq 0 \} \). Step 1 of the proof of Theorem 4.5.9 states that the image of the mapping \( x \mapsto S_m^x(t) \) is \( \mathbb{R}_+ \). Thus, fixing \( (t, s, z) \in [0, T] \times \mathbb{R}_+ \times \mathbb{R} \backslash \{ 0 \} \), it must hold that

\[
C_m^S(t, s + s\gamma(t, s, z)) - C_m^S(t, s) - s\dot{\gamma}(t, s, z)\pi(t, s) = 0. \tag{4.6.10}
\]

We turn to the discussion of the cases (i) and (ii):

(i) If \( \gamma(t, s) \neq 0 \) for all \( (t, s) \in [0, T] \times \mathbb{R}_+ \), we learn from the first line of (4.6.9) that (4.6.10) holds with \( \pi(t, s) = S_m^S(t, t) \). Consequently, Step 1 implies that \( C_m^S(t, \cdot) \) is linear in the second variable for every \( t \in \mathcal{T} \). Since the set of jump times \( \mathcal{T}(\omega) \) is dense in \( [0, T] \) for every fixed \( \omega \in \Omega \backslash \mathcal{N} \) (see Cont and Tankov [2012], p.84 for a proof) and as \( C_m^S \) is continuous by Assumption 4.5.4, we conclude that \( C_m^S(t, \cdot) \) is linear in the second argument for every \( t \in [0, T] \).

(ii) Suppose \( \gamma \equiv 0 \). Let \( z_1, z_2 \in \mathbb{R} \backslash \{ 0 \} \), such that \( \tilde{\gamma}(t, s, z_1) \neq \tilde{\gamma}(t, s, z_2) \). By continuity of \( \tilde{\gamma} \) and (4.6.10) it must hold that

\[
\frac{C_m^S(t, s + s\tilde{\gamma}(t, s, z_1)) - C_m^S(t, s)}{s\tilde{\gamma}(t, s, z_1)} = \frac{C_m^S(t, s + s\tilde{\gamma}(t, s, z_2)) - C_m^S(t, s)}{s\tilde{\gamma}(t, s, z_2)}.
\]
We immediately see that Step 2 implies that $C^S_m(t, \cdot)$ must be linear in the second component for every $t \in T$. By the same argument as in (i) we conclude linearity for every $t \in [0, T]$.

**Remark 4.6.3.** We remark that Theorem 4.6.2 is a robustness result in incomplete markets upon assuming that the superhedging price process is Markov. A perfect hedge is possible due to the induced linearity of the payoff function. Thus, robustness is trivial.

Finally, we broach the issue of robustness in a complete market. As El Karoui et al. (1998) established the robustness of the perfect hedging strategy w.r.t. volatility misspecification in a diffusion setting, we focus on pure jump processes. The following corollary treats the case $\gamma \equiv 0$ and $\vartheta(dz) = \lambda \mathds{1}_{\{\alpha\}}(dz)$ for some $\lambda > 0$ and $\alpha \in \mathbb{R}\setminus\{0\}$. Hence, the jump component corresponds to a homogeneous Poisson process. In this case the risk-neutral measure is unique and consequently the set $Q_{em}$ is a singleton containing $\mathbb{P}$ solely. To make this explicit, we denote the contingent claim price at time $t$ by $C^p_m(t, S^x_m(t))$.

**Corollary 4.6.4.** Suppose all assumptions of Theorem 4.6.2 are satisfied, $\gamma \equiv 0$ and $\vartheta(dz) = \lambda \mathds{1}_{\{\alpha\}}(dz)$ for some $\lambda > 0$ and $\alpha \in \mathbb{R}\setminus\{0\}$. The replicating Delta-strategy in the misspecified model is given by $\pi = (\pi(t, S^x_m(t)))_{t \in [0, T]}$ with

$$
\pi(t, S^x_m(t)) = \frac{C^p_m(t, S^x_m(t-)) + \Delta S^x_m(t, \alpha)) - C^p_m(t, S^x_m(t-))}{\Delta S^x_m(t, \alpha)}.
$$

If in addition

$$
\text{sgn}(\tilde{\gamma}(t, S^x_m(t), \alpha) - \eta(t, \alpha)) = \text{sgn}(\eta(t, \alpha)),
$$

$d\mathbb{P} \times dt$-a.s., then following this strategy and trading in the real stock $S$ yields an a.s. super-replication of $h(S(T))$, i.e.,

$$
C^p_m(0, x) + \int_0^T \pi(u, S(u-)) dS(u) \geq h(S(T)).
$$
Proof. We start with (4.6.7) and use (4.6.8) to prove the corollary. Since the market is complete, the claim \( h(S_m^x(T)) \) is perfectly replicable and the process \( A \) in (4.6.8) (allowing for the interpretation of the cash amount that can be withdrawn at time \( t \)) is then (cf. (4.6.1)) equal to zero, i.e.,

\[
A(t) = \int_0^t \dot{C}_m^p(u, S_m^x(u)) \, du \\
+ \lambda \int_0^t \left( C_m^p(u, S_m^x(u-)) + \Delta S_m^x(u, \alpha) \right) \, du = 0,
\]

for Lebesgue-almost all \( t \in [0, T] \). Then the Lebesgue differentiation theorem implies that (4.5.11) holds with equality:

\[
0 = \dot{C}_m^p(t, S_m^x(t)) + \lambda \left( C_m^p(t, S_m^x(t)) + S_m^x(t) \tilde{\gamma}(t, S_m^x(t), \alpha) \right) - C_m^p(t, S_m^x(t)) - S_m^x(t) \tilde{\gamma}(t, S_m^x(t), \alpha) C'_m(t, S_m^x(t)) \right).
\]

(4.6.11)

We can replace \( S_m^x(t) \) in (4.6.11) by the the true stock price \( S(t) \) because (4.6.11) holds for all \( s \in \mathbb{R}_+ \) (see Step 2 of the proof of Theorem 4.5.9 for detailed arguments). Since \( C_m^p(T, S(T)) = h(S(T)) \), trading in the physical stock \( S \) yields

\[
C_m^p(0, x) + \int_0^T \left( C_m^p(u, S(u)) + \Delta S(u, \alpha) - C_m^p(u, S(u-)) \right) \, dS(u) \\
= C_m^p(0, x) + \int_0^T C_m^{p'}(u, S(u-)) \, dS(u) \\
+ \int_0^T \left( C_m^p(u, S(u)) + \Delta S(u, \alpha) - C_m^p(u, S(u-)) \right) \tilde{J}(du, dz) \\
= h(S(T)) - \left( \int_0^T \dot{C}_m^p(u, S(u)) \, du \right).
\]
\[ + \lambda \int_0^T \left( C^p_m(u, S(u-)) - C^p_m(u, S(u)) \right) \, du \]
\[ \geq h(S(T)) - \left( \int_0^T \dot{C}^p_m(u, S(u)) \, du \right. \]
\[ + \lambda \int_0^T \left( C^p_m(u, S(u-)) + S(u-) \hat{\gamma}(u, S(u-), \alpha) \right) \]
\[ \left. - S(u-) \hat{\gamma}(u, S(u-), \alpha) C^p_m(u, S(u-)) \right) \, du \]
\[ = h(S(T)), \]

where the second equality is justified by Itô’s formula and the last inequality follows from the convexity of \( C^p_m \) in the second component (cf. Corollary 4.5.3) combined with Lemma 4.4.5. The final equality is immediate from (4.6.11). \( \square \)

### 4.A Appendix

**Lemma 4.A.1.** Consider the process \( S^x_m \) given by \((4.2.5)\) w.r.t. the reference measure \( \mathbb{P} \). Then it holds that \( \mathbb{Q} \in \mathcal{Q}_{em} \) if and only if there exist predictable processes \( \psi : [0, T] \to \mathbb{R} \) and \( \theta : [0, T] \times \mathbb{R} \setminus \{0\} \to \mathbb{R} \) with \( \theta(t, z) < 1 \), and

\[ \gamma(t, \tilde{S}^x_m(t)) \psi(t) + \int_{\mathbb{R} \setminus \{0\}} \hat{\gamma}(t, \tilde{S}^x_m(t), z) \theta(t, z) \, \theta(dz) = 0, \quad (4.A.1) \]

\( d\mathbb{P} \times dt \text{-a.s.} \), such that the process \( \xi = (\xi(t))_{t \in [0, T]} \) with

\[ \xi(t) := \left( - \int_0^t \psi(u) \, dW(u) - \frac{1}{2} \int_0^t \psi(u)^2 \, du \right. \]
\[ + \int_0^t \int_{\mathbb{R} \setminus \{0\}} \ln(1 - \theta(u, z)) \, \tilde{J}(du, dz) \]
\[ + \int_0^t \int_{\mathbb{R}\setminus\{0\}} (\ln(1 - \theta(u, z)) + \theta(u, z)) \, \vartheta(dz)\, du \]

is well-defined and satisfies \( \mathbb{E}_\mathbb{P}[\xi(T)] = 1 \). Define

\[
\begin{align*}
    dW_Q(t) &= \psi(t) \, dt + dW(t), \\
    \tilde{J}_Q(dt, dz) &= \theta(t, z) \, \vartheta(dz)\, dt + \tilde{J}(dt, dz). 
\end{align*}
\] (4.4.2)

Then \( W_Q \) is a standard Brownian motion w.r.t \( Q \) and \( \tilde{J}_Q(dt, dz) \) is the \( Q \)-compensated version of the Poisson random measure \( J(dt, dz) \).

**Proof.** See for instance Öksendal and Sulem (2005), Chapter 1.4.
Bibliography


Short CV

Personal Details

Name      Frank Bosserhoff
Born in   Bad Soden am Taunus, 1989

Education

01/2016 – 01/2020  Doctoral Candidate at Ulm University
10/2013 – 10/2015  Studies of Finance (with major in Mathematics) at Ulm University
                  Degree: Master of Science
                  Degree: Bachelor of Science
09/2012 – 12/2012  Exchange semester at the Chinese University of Hong Kong
                  Degree: Abitur
Work and Professional Experience

Since 01/2016  Research associate
Ulm University, Institute of Insurance Science

04/2018 – 09/2019  Study Guide ("Studienlotse")
Ulm University

11/2014 – 03/2015  Corrector
Ulm University, Institute of Mathematical Finance

Scholarships

Since 11/2013  Beta Gamma Sigma
Membership in International Honor Society

01/2016 – 12/2017  Stiftung WiMa Ulm
Ph.D. Grant

11/2011  Maastricht University -"Top 3% award"
Reimbursement of tuition fees due to outstanding study achievements

02/2011 – 10/2015  Studienstiftung des Deutschen Volkes
Stipendiary of scholarship foundation

Contributed Talks

09/2017  4th Fudan-Ulm Symposium on Insurance & Finance
Fudan University

03/2018  13th German Probability and Statistics Days
Freiburg University

03/2019  2nd Uni Ulm Actuarial Day
Ulm University

07/2019  23rd International Congress on Insurance: Mathematics and Economics (IME)
Technical University of Munich
Publications and Preprints


Ehrenwörtliche Erklärung

Ich versichere hiermit, dass ich die Arbeit selbstständig angefertigt habe und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt sowie die wörtlich oder inhaltlich übernommenen Stellen als solche kenntlich gemacht und die zur Zeit gültige Satzung der Universität Ulm zur Sicherung guter wissenschaftlicher Praxis beachtet habe.

Ich habe bisher weder an der Universität Ulm oder an einer anderen deutschen oder ausländischen Hochschule mit Promotionsrecht eine Zulassung zur Promotion beantragt noch habe ich den Doktorgrad Dr.rer.nat. bereits erworben oder bin bisher in einem früheren Promotionsverfahren für den Doktorgrad Dr.rer.nat. oder dieselbe Dissertation an der Universität Ulm oder einer anderen deutschen oder ausländischen Hochschule mit Promotionsrecht endgültig gescheitert.

Ulm, November 2019

Frank Bosserhoff