Hedging in a Financial Market with Regime-Switching

by

Adam Daniel Gomes

A thesis
presented to the University of Waterloo
in fulfillment of the
requirement for the degree of
Doctor of Philosophy
in
Electrical and Computer Engineering

Waterloo, Ontario, Canada, 2021

Examining Committee Membership

The following served on the Examining Committee for this thesis. The decision of the Examining Committee is by majority vote.

Supervisor: Andrew Heunis
Professor, Dept. of Electrical and Computer Engineering,
University of Waterloo

Internal Member: Ravi Mazumdar
Professor, Dept. of Electrical and Computer Engineering,
University of Waterloo

Internal Member: Patrick Mitran
Professor, Dept. of Electrical and Computer Engineering,
University of Waterloo

Internal-External Member: Mario Ghossoub
Professor, Dept. of Statistics and Actuarial Science,
University of Waterloo

External Member: George Yin
Professor, Dept. of Mathematics,
University of Connecticut

Author's Declaration

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Abstract

It is well-known that in the complete standard financial market model driven solely by Brownian motion, one can always hedge a given contingent claim starting from an appropriate initial wealth. In other words, there always exists an initial investment amount and a trading strategy from which one can produce enough wealth to pay off a given contingent claim. Furthermore, one can always produce just the right amount of wealth to settle the claim. That is, one exactly hedges, rather than super-hedges (produces excess wealth), a contingent claim in a standard financial market. In more general market models, such as those where there are constraints on the amount one can invest or when there are other processes driving the randomness in the market, it is not immediately obvious whether one can exactly hedge, or even merely super-hedge, a given contingent claim.

We consider hedging problems in a generalization of the well-studied standard Brownian motion market model, namely the regime-switching market model. A standard Brownian motion market model is not very robust as it can only handle small-scale persistent changes in market behaviour. The regime-switching market model, on the other hand, is able to handle large-scale occasional changes in market behaviour, along with small-scale persistent changes, by using a continuous-time Markov chain in addition to a multi-dimensional Brownian motion to drive the randomness in the market. This generalization comes at a cost, however; adding the additional source of randomness renders the financial market incomplete. El-Karoui and Quenez [16] and Cvitanic and Karatzas [10] solved hedging problems in an incomplete Brownian motion market model by introducing the cumulative consumption process and the space of dual processes. These tools allowed them to show that one could in fact promise to super-hedge a given contingent claim in their incomplete market. We use these same tools to handle the incompleteness of the regime-switching market, along with more advanced stochastic analysis, namely the study of discontinuous local martingales, to handle the discontinuity of the paths of the Markov chain. We show that under a certain integrability condition, one can always promise to super-hedge a given contingent claim in a financial market with regime-switching. Furthermore, we characterize both the minimum initial wealth, called the price of the contingent claim, and the trading strategy needed to hedge the contingent claim.

We further generalize the problem of hedging in a regime-switching market model by including convex portfolio constraints, introduced by Cvitanic and Karatzas [10], and margin requirements, of the kind introduced by Cuoco and Liu [7]. These additions allow us to model, for example, markets where there are restrictions on investments and interest rates that are higher for borrowing than for lending. Once again we show that in such a market, under a certain integrability condition, one can always promise to super-hedge a

given contingent claim. Furthermore, we characterize the minimum initial wealth and the trading strategy needed to hedge the contingent claim. We then show that under specific optimality conditions, one can exactly hedge a given contingent claim without producing an excess amount of wealth at the end of trade. In other words, we provide conditions that allow one to almost surely hedge a contingent claim without requiring them to consume wealth through a cumulative consumption process.

Lastly, we address the problem of approximate hedging in a regime-switching market model, where one tries to hedge a given contingent claim with initial wealth less than the price of the claim. Since the price of the contingent claim is the minimal initial wealth one needs to almost-surely hedge the claim, if one were to begin trading with an initial wealth lower than this price, there is a non-zero probability of them failing to settle the claim. In this case, an investor should trade in an optimal way so that their expected loss from hedging is minimized.We use this approach to solve the approximate hedging problem in a regime-switching market model with portfolio constraints and margin requirements. Using convex duality and tools of non-smooth convex analysis, we show that there does exist an optimal trading strategy that minimizes a specific cost criterion when starting from a lower initial wealth than the price of the given contingent claim.

Acknowledgements

I wish to express my gratitude to my supervisor, Dr. Andrew Heunis. I began my doctoral program with very little knowledge in the field and Andrew had the patience to help me through many of the challenges I faced. There is no way I would have been able to complete this degree without Andrew's mentorship and guidance. I also particularly enjoyed our weekly meetings, where we not only discussed research, but had many other insightful conversations. I would like to thank my committee members Dr. Ravi Mazumdar, Dr. Patrick Mitran, Dr. Mario Ghossoub, and Dr. George Yin, for taking the time to read my thesis and providing invaluable comments. I would also like to thank Dr. Daniel Miller for our many conversations and providing me with resources that greatly aided my work. Lastly, I would like to thank my master's degree supervisor Dr. David Wang, who, although I was no longer supervised by him, was always available to provide guidance on both academics and life in general.

I would like to thank the members of UWHaptics lab, especially Madeleine Wang, Yousef Sawires, Keegan Fernandes, Tina Wu, Elaine Huang, and Monisha Naik, where together we had great discussions and solved some challenging problems.

I would like to thank my parents Rupert and Sue, my sister Christine, my brother-in-law Melroy, and my brother Aaron for their support throughout my life. Lastly, I would like thank my friends who helped me throughout this degree, in particular, Aman Muthreja, Jessica Chen, Jamie Waugh, Lilly Zheng, Hummer Li, Lux Jeeva, Aaron Leung, Takin Tadayon, Stanislav Bochkarev, Francis Li, Ben Chang, and the members of the UWHaptics Lab.

Dedication

This is dedicated to my mom, dad, and brother and best friend.

Table of Contents

Chapter 1

Introduction

In finance, hedging is a method of limiting one's risk in investment. Simply put, hedging is the practice of taking a position in a certain investment in order to counteract the risk associated with an opposing investment. For example, an investor can take a position in renewable energy sources, but also invest in fossil fuels to limit his or her risk of loss in renewable energy. In mathematical finance, hedging is the act of finding both an initial investment amount and a strategy of saving or trading so that one can guarantee they will be able to cover some monetary obligation at a future point in time. This obligation is often called a *contingent claim* or a *derivative security*, and its value can be reproduced by a combination of stocks available for investment in the stock market. The future point in time when the obligation must be settled is called the *expiration date*. One way an investor can hedge a contingent claim is to invest the obligation amount (if known) into a bank account and withdraw the amount on the expiration date. However, if the obligation amount is a random variable whose value is unknown until the expiration date, this method of hedging cannot work since one could not possibly know the future value of the contingent claim with 100% certainty. Another approach to hedging is for the investor to begin with as little money as possible such that they can find a strategy of trading in the stock market which guarantees the obligation is settled on the expiration date. Such a method of hedging is called *almost sure hedging* and solving the *almost sure hedging problem* requires one to find both the an initial investment and a strategy for investing into each stock, called a *portfolio process*, such that the given contingent claim can be payed off on the expiration date. If the investor's portfolio process successfully hedges the contingent claim, this strategy is called a *hedging strategy* and consists of the portfolio process called a *hedging portfolio* and the initial investment. The least initial investment for which a hedging strategy exists is called the *price* of the the contingent claim, and together with associated hedging portfolio, is called the *optimal hedging strategy*. Solving the hedging problem, which amounts to finding the optimal hedging strategy under certain conditions that will be discussed shortly, is the primary goal of this thesis.

In order to solve the hedging problem, it is necessary to construct a model of the financial market. This market model seeks to mimic the behaviour of the real-life stock market. Additionally, the model must be amenable to mathematical analysis so that investors can use it to find solutions to investment problems. Market models have been studied since the early 1900's, starting with the work of Louis Bachelier [2], and have now become quite sophisticated. One such market model, which the hedging problem of this thesis is based in, is called a *regime-switching* market model. Regime-switching models allow for the market to undergo "shocks" at random times. At any point in time, the market is assumed to be in some *regime*, for example a bull market where stock prices are generally rising. When a shock occurs, the market's behaviour can fundamentally change. This is called a *regime-switch* and represents the sudden change of the market, such as from a bull market to a bear market where stock prices are now generally falling. Famous examples of a regime-switch would be the Wall Street crash of 1929 or even the pandemic of 2020. There are many types of regime-switching market models which exist in the literature. In this thesis we choose a very general regime-switching market model as it offers significant advantages over other less general regime-switching market models. For example, our model allows the use of stochastic volatility which, for instance, Markov modulated regime-switching models do not. Simply put, the regime-switching market model used in this work allows for a more realistic model of the stock market.

The study of hedging in mathematical finance begins with the *portfolio selection* work of Markowitz [35] and the *pricing* work of Black, Scholes and Merton [4] [36], in what is called the *standard financial market model* (no regime-switching). In very simple terms, portfolio selection describes the method of *how to invest* wealth in the stock market, and pricing describes *how much* to initially invest into the stock market. The almost sure hedging problem can be thought of as a combination of both problems. Since the papers of Markowitz, Black and Scholes, and Merton, both portfolio selection problems and pricing problems have grown in complexity to include items such as *portfolio constraints* and *margin requirements*. Based on the type of industry an investor works on behalf of, limitations on the amount an investor can invest may be imposed. For example, there may be a no *short-selling* restriction on an investor working on behalf of a pension fund, meaning that the investor can never have negative holdings on a stock, as often times pension funds may face regulatory requirements to never short-sell. These types of requirements on investments act as constraints on the hedging portfolio, and as a result, are called *portfolio constraints*. Other considerations, such as differing interest rates for borrowing and lending, also make these problems more complex. These payments can be collectively grouped into what are called *margin requirements*. As we will see in this thesis, the presence of portfolio constraints and margin requirements make problems in mathematical finance more challenging to solve.

The almost sure hedging problem in a standard financial market model without market constraints was originally solved in the aforementioned pricing work of Black, Scholes and Merton [4] [36] using the properties of *continuous martingales*. When the almost sure hedging problem is instead posed in a regime-switching market model, the random changes in regimes requires the analysis of more general *discontinuous local martingales*. Furthermore, the concept of *consumption of wealth* needs to be introduced in the regime-switching market model to make the almost sure hedging problem tractable. In the standard financial market model, finding the optimal hedging strategy only requires the investor to solve for a hedging portfolio and price of the contingent claim. However, once portfolio constraints are imposed on the investor or additional sources of randomness assumed in the market, the market model is rendered *incomplete*, and the investor needs to strategically spend a certain amount of his or her money elsewhere to exactly hedge the contingent claim. This strategy of spending is called the investor's *cumulative consumption process*, and in addition to the hedging portfolio and price of the contingent claim, forms the investor's optimal hedging strategy. The addition of regime-switching alone (that is, without market constraints) to the financial market renders the market model incomplete; thus, requiring the use of a cumulative consumption as part of the investor's hedging strategy. As will be shown in this thesis, finding this optimal consumption process is far from trivial and the difficulty is further increased when portfolio constraints and margin requirements are added to the regime-switching model. This leads us to solve three separate, but related, cases of the almost sure hedging problems in this thesis. The first is a classical problem within a standard financial market model without regime-switching, and thus without consumption, which has already been established in the literature. However, purely for background purposes and to enhance readability, we treat this problem in some detail because it will then be used as a bootstrap to the hedging problems of genuine interest, namely within a regime-switching model. In Chapter 4 we move on to the *unconstrained hedging problem* in a regime-switching market model, and again use the results to develop a solution to our general problem of interest, the *constrained hedging problem* in a regime-switching market model, in Chapter 5.

The problems discussed above are all very conservative forms of hedging. Their solutions promise the investor can cover his or her obligation with 100% certainty. As a result, the investor will usually be required to begin trading with a very large initial wealth to successfully pay off the contingent claim. Often times this initial wealth could be unacceptably high, which leads to the idea of *approximate hedging*. Here one gives up trying to hedge exactly from an initial wealth and merely begins with some initial wealth that the agent can "afford" but which may be well below the price of the given contingent claim. The goal of the agent is now to trade in such a way as to *minimize* a specific risk criterion which imposes a penalty when the agent falls short of being able to fund the contingent claim. This is essentially a problem of stochastic optimal control with a convex but very non-smooth cost criterion and has been solved in a standard financial market with portfolio constraints, but without regime-switching or margin requirements. In Chapter 6 of this thesis we describe and solve the approximate hedging problem in a regime-switching market model with portfolio constraints and margin requirements.

1.1 Outline of Thesis

In Chapter 2, we provide the historical development of the hedging problem and methods of solution from the literature. We give reasons for the specific choice of the regime-switching model, including reasons for portfolio constraints and margin requirements, and describe both the almost sure hedging problem and the approximate hedging problem in very non-technical terms.

In Chapter 3, we define the standard financial market model and the almost sure hedging problem in such a market in precise mathematical terms. We review the solution to this well-studied problem and get the reader familiar to the methods and tools which will be used for the remainder of the thesis. We also discuss the decorated *Black-Scholes* formula as an important example of the hedging problem.

In Chapter 4, we define the regime-switching market model in mathematical terms and introduce the necessary mathematical spaces and processes needed for working with such a market model. We introduce the cumulative consumption process and the space of dual process that allow us to analyze the unconstrained hedging problem, and determine conditions for which there exists a solution to the hedging problem. Using the tools of stochastic calculus, specifically a general martingale representation theorem and the Doob-Meyer decomposition, we provide a characterization for the price and associated optimal hedging strategy for a given contingent claim.

In Chapter 5, we introduce portfolio constraints and margin requirements to the hedging problem in a regime-switching market model. By applying many of the methods introduced in Chapter 4, along with tools of convex analysis, we find a solution to the almost sure hedging problem. That is, we characterize the price and optimal hedging strategy of a given contingent claim. We then show that under a set of equivalent conditions, one can hedge a given contingent claim without being required to consume wealth.

In Chapter 6, we turn our attention to approximate hedging in a constrained regime-switching market model. We prove the existence of a hedging strategy that minimizes a specific risk criterion. We proceed very much in the same way as Cvitanic [8] as most of the results carry over directly from the approximate hedging problem in a constrained standard financial market model. Many of the results in this chapter require the theory of convex analysis, with a major result relying on the special use of the Komlós theorem.

In Chapter 7, we give the conclusion and outline briefly some future areas of investigation related to the work in this thesis.

Supplementary results for Chapters 4 and 5 are given in Appendix A while lengthy proofs of key results are given in Appendix B. Appendix C and Appendix D review standard definitions and results from stochastic processes and convex analysis theory that may be referred to throughout the thesis document.

Chapter 2

Background

In this thesis we the address the problems of *almost sure hedging* and *approximate hedging* in a regime-switching market model. In both problems, the goal of an investor is to find trading strategies that guarantee certain conditions are met. However, to define both problems in detail, and thus state such conditions, we need to introduce essential concepts of mathematical finance. In this chapter, we will introduce such concepts in informal terms. We begin by reviewing the market models of mathematical finance, including the regime-switching market model in non-technical terms. We will give reasons for why the regime-switching market model is an improvement over prior market models. Once the regime-switching model is constructed, the hedging problems can be precisely defined. We will discuss the historical development of hedging problems in mathematical finance and then give brief, non-mathematical descriptions of the almost sure hedging problem and the approximate hedging problem. Full mathematical treatment of each hedging problem is given in Chapter 3 through Chapter 6 of this thesis.

2.1 Financial Market Models

Stock market investors want to model the stock market to have a better understanding of how to invest their money and how secure their investments are. For example, they may want to know if they should invest all their money in one stock, or in many stocks, and if so, in how many stocks? The main question that arises is: how can one capture the behaviour of a stock market in a mathematical model? The most prominent model is based on the idea that the stock market moves randomly. This randomness is classically described by a *Brownian motion*, a phenomenon that can be observed physically in nature. A botanist named Robert Brown noticed that pollen grains jittered in a random fashion when suspended in water. It was subsequently discovered that this jittery random movement can be modeled in terms of a random process which is usually known as Brownian motion. Historically, the first truly mathematical description of Brownian motion occurs in connection with mathematical finance, and is due to the French mathematician Louis Bachelier [2], who noticed that stock market prices appear to jitter in a random fashion. Bachelier derived the first mathematical model of Brownian motion in order to model stock prices. At the time of Bachelier's discovery in 1900, his work was not given very much attention and it was only about fifty years later that Bachelier's paper began to be recognized as a ground-breaking piece of work. At this time the economist Robert Merton published two additional significant papers based on the Brownian motion process, as we summarize in the next paragraph. It may not be out of place to note here that Einstein [14][15], in 1905, also introduced a mathematical description of Brownian motion, essentially identical to that of Bachelier, but from a completely different point of view, namely as a model in the kinetic theory of ideal gases. The mathematical models of Brownian motion due to Bachelier and Einstein preceded the axiomatization of probability based on measure theory. They were largely based on physical intuition, and did not really conform to rigorous standards of formulation and proof. That final stage in the development of a mathematically rigorous theory of Brownian motion based on measure theory was accomplished by Norbert Wiener [49] during the 1920's.

Robert Merton [36] used the mathematical model of Brownian motion established by Wiener to develop the continuous-time market model of stock prices. Specifically, Merton described stock prices as *geometric Brownian motion*, which is a continuous-time stochastic processes. Having stock prices modeled by a geometric Brownian motion implies that they are lognormally distributed, which makes sense as stock prices cannot assume negative values. Expressed mathematically, a stock has price $S(t)$ at time t that satisfies the stochastic differential equation

$$\frac{dS(t)}{S(t)} = \mu dt + \sigma dW(t) \tag{2.1}$$

where $\{W(t)\}$ is a Brownian motion and μ and σ are constants called the *market coefficients*. μ is called the *expected rate of return* and σ is called the *volatility* of the stock. The stock price process $\{S(t) : t \geq 0\}$ is a stochastic process which is a set of random variables indexed by time. In a complementary paper [37], Merton modeled stock prices with market coefficients which are no longer constants, but depend on the stock price and time to further generalize the model. This model is given by the following stochastic differential equation:

$$\frac{dS(t)}{S(t)} = \mu(t, S(t))dt + \sigma(t, S(t))dW(t). \tag{2.2}$$

This model allows for greater latitude when fitting actual market data to the model. For example, there is empirical evidence that suggests that a low stock price increases the stock price volatility more than a high stock price (see Black [3]). In that case, $\sigma(t, S(t))$ would increase more when $S(t)$ is small than when $S(t)$ is large. The model with constant coefficients (2.1) would not be able to reproduce such behaviour.

Harrison and Kreps [20] and Harrison and Pliska [21] developed mathematical models of the stock market where the stock price processes are general stochastic processes instead of just geometric Brownian motion. This allows the market coefficients, σ and μ, to be general stochastic processes as well, such as in the equation

$$\frac{dS(t,\omega)}{S(t,\omega)} = \mu(t,\omega)dt + \sigma(t,\omega)dW(t,\omega). \tag{2.3}$$

The market coefficients in this model are specified processes $\{\mu(t)\}$ and $\{\sigma(t)\}$, and as such are functions of both time and the underlying probability space through ω. Modeling the market coefficients as random processes increases the generality of the model past that of (2.2). For example, it allows for the use of stochastic volatility models which provide a better fit of actual market data to the model as shown by Hull [25]. From a technical standpoint the main requirement on the processes μ and σ in (2.2) is that these be adapted to *some filtration* with respect to which W is a Brownian motion (for otherwise the stochastic integral in (2.3) is undefined). There is nevertheless considerable leeway in satisfying this requirement, which greatly increases the generality of the model (2.3). Of particular importance is the special case in which the processes μ and σ are adapted to the filtration which is *generated* by the Brownian motion W. Such models (and multidimensional extensions thereof) are typically known as *standard financial market models* or *Brownian motion market models*. It is perhaps worth pointing out that the special case of standard market models provides a very rich setting in which to address numerous challenging problems in mathematical finance. For example, the definitive work of Karatzas and Shreve [31] is devoted almost entirely to (multidimensional) standard market models. In such models one can regard the Brownian motion W as "driving" or "determining" the processes μ and σ (indeed the Doob measurability theorem makes this mathematically explicit), and therefore the Brownian motion W is the *sole source* of randomness in standard market models; the Brownian motion W "drives" the persistent, short-duration, and small-scale microeconomic changes which are a characteristic feature of the market parameters. The scope of (2.3) however goes well beyond standard market models, important though these are. Thus, one can specify that processes μ and σ are adapted to a filtration which is *strictly larger* than the filtration generated by W, but with respect to which W nevertheless remains a Brownian motion. For example, this is the case when one stipulates some additional process, such as a Markov process $\{\alpha(t)\}$, which is *independent of* W, and then requires that the processes μ and σ in (2.3) be adapted to the *joint filtration* of α and W. This joint filtration is of course larger than the filtration of W, but W is nevertheless still a Brownian motion with respect to the joint filtration in view of the independence of α and W. There are now two sources of randomness in the market model, namely the Brownian motion W and the Markov process α, and the additional "degree of freedom" arising from availability of the Markov process provides scope for constructing more realistic and accurate market models than would be obtained from standard market models in which the Brownian motion is the only source of randomness. The so-called *regime-switching* market models, with which this thesis is much concerned, are in fact precisely of this kind, and we now turn to a discussion of these models.

2.2 The Regime-Switching Market Model

Financial traders have long been familiar with distinct "regimes" or "states" of a market. For example "bullish" and "bearish" have clear significance as possible regimes of the market. Other examples of regimes which influence the market are of course also possible, such as consumer confidence (e.g. "strong", "moderate" or "weak"). Whatever the significance of the regime states, when accounting for regimes in stock price models of the form (2.1)-(2.3) above, one must clearly allow dependence of the expected rate of return μ and the volatility σ on the regime state. The transition or switch from one regime to another (such bullish to bearish or the opposite) is generally caused by non-deterministic large-scale but typically rarely occurring events, for example the sudden insolvency of a major bank, the election of an eccentric politician, an unexpected viral pandemic, or the discovery of a major deposit of some resource (such as oil). Hamilton [19] introduced the use of a finite-state Markov chain to model switches from one regime to another, it being assumed that there are finitely many regimes associated with the market, with the Markov states corresponding to the regime states.

If one denotes this finite-state Markov chain by $\{\alpha(t)\}$ then a simple price model which includes regime-switching is of the form

$$\frac{dS(t)}{S(t)} = \mu(\alpha(t))dt + \sigma(\alpha(t))dW(t) \tag{2.4}$$

in which μ and σ are real-valued functions defined on the set of regimes (i.e. Markov states), and it is assumed that the Markov chain α and Brownian motion W are independent. This independence is essential for technical reasons, for it ensures that W is a Brownian motion with respect to the joint filtration of α and W; since the processes $\{\mu(\alpha(t))\}$ and $\{\sigma(\alpha(t))\}$ are of course adapted to this joint filtration one sees that (2.4) makes sense as a stochastic differential equation. There are, in addition, clear economic reasons for stipulating that α and W be independent, and we shall indicate these shortly. At this point we note that (2.4), in which the market parameters are said to be *Markov modulated* (in the sense that the market parameters at any instant t are completely determined by the regime state $\alpha(t)$ at the same instant), is a clear improvement on the model (2.1), in which the expected rate of return μ and volatility σ are just constant. Because of its simplicity and tractability the Markov modulated model (2.4) is used in the great majority of works on regime-switching, for both hedging and portfolio optimization, including for example Buffington and Elliott [5], Di Masi *et-al* [12], Jobert and Rogers [27], Sotomayor and Cadenillas [48], Yao *et-al* [51], Yin *et-al* [52] Zhang [53], and Zhou and Yin [54]. Nevertheless, despite this wide use, the model (2.4) is not without problems. For example the volatility $\sigma(\alpha(t))$ is constant within each regime state, thus one cannot model stochastic volatility in a fixed regime state, and more generally one cannot treat the market parameters as stochastic processes in their own right within a given regime state. For this reason, in the present thesis, we shall instead use the more general market model (2.3), in which it is stipulated that the market parameter processes μ and σ are adapted to the joint filtration of the Brownian motion W and a finite-state Markov chain $\{\alpha(t)\}$ which models regime-switching, with α and W being independent. Of course this independence is necessary for W to be a Brownian motion with respect to the joint filtration, so that (2.3) makes sense as a stochastic differential equation. In view of the Doob measurability theorem one sees that the market parameters μ and σ are non-anticipatively determined by the processes W and α, which are the sources of randomness for the model. The Brownian motion W drives short-duration and small-scale but persistent microeconomic changes in the market parameters (exactly as noted above for standard market models), while the finite-state Markov chain α on the other hand drives occasional long-duration, large-scale macroeconomic changes. The independence of the Brownian motion W and the Markov chain α amounts to the reasonable assumption that the micro- and macroeconomic effects are independent.

There is one final point concerning regime-switching models which must be emphasized, namely the occasional misconception that the goal of regime switching is to introduce jump-discontinuities into the price model. This is most certainly not the case. No matter how dire (or promising) the situation, one does not get a jump discontinuity in stock prices in the transition from a bull to a bear (or a bear to a bull) market, although prices could change rapidly. The goal of regime switching is to model the consequences of remaining locked in a particular regime state (such as a bull or a bear market) for a random but usually extended period of time, resulting from the dependence of the market parameters μ and σ on the regime state. This is not to say that price models with jump discontinuities are of no interest. For example, Merton [37] introduces a price model given by the jump diffusion:

$$\frac{dS(t)}{S(t)} = \mu dt + \sigma dW(t) + dq(t),$$

which has a form similar to (2.1) (i.e. constant market parameters μ and σ) but includes a compensated Poisson process $\{q(t)\}$ as an additional term on the right to contribute jumps in the price. This is not a regime switching model because the market parameters are not influenced by any regime state. Indeed, such a model could never account for the persistently rising/falling price between regime-switches which are the hallmark of a bull/bear market; all one would get is a discontinuous change in price, followed by normal market behaviour until another discontinuous change in price, followed by normal market behaviour until another discontinuous change in price, and so on. In general, jump discontinuities in stock prices are not a relevant concern in regime-switching market models.

2.3 The Almost Sure Hedging Problem

The motivation for having a mathematical model of the financial market is to address complex problems involving an investor in the market. These problems can be of pricing, such as finding the selling price of a call option, or of portfolio selection, essentially determining how one should "best" distribute their wealth among a given set of assets. To solve a portfolio selection problem, an investor is required to find a *portfolio process* that allows him or her to attain some financial goal at a fixed, future point in time. A portfolio process represents how much wealth an investor has invested in each asset in the stock market at every instant in time between the beginning and end of trade. For example, consider an investor who has \$10,000 available to invest in the stock market. The investor would like to invest this money so that, in thirty years time, the value of his or her investment has grown to \$100,000. In this case, the investor must find the portfolio process which allows him or her to do this, and as a result, solve their portfolio selection problem. In making investment choices, the investor can operate in two possible trading environments, namely: (i) be free to distribute their current wealth at every instant among the assets without any constraints on this distribution; or (ii) be compelled by various externally-imposed trading regulations to respect some *portfolio constraints* when allocating their wealth among the assets. An example of (ii) would be the prohibition on going *short* (i.e. having negative dollars invested in an asset) on some or all stocks. Pension funds often place this kind of restriction on investments. The choice of portfolio can also be effected by certain regulations such as differing interest rates in the market. These regulations can collectively be grouped into *margin requirements* on the portfolio. Typically, investment problems with portfolio constraints and margin requirements are more challenging than those in unconstrained markets. In Chapters 3 and 4 of this thesis, we will assume that the portfolios we deal with are *unconstrained* and without margin requirements. In Chapter 5 we generalize the methods used in Chapters 3 and 4 concerning unconstrained market models to deal with portfolio selection in a market with both constrained portfolios and margin requirements.

The birth of portfolio theory, the theory of solving portfolio selection problems, is generally attributed to Markowitz [35], who published “Portfolio Selection” in 1952. He considered a problem of selecting a portfolio of investments which minimizes the risk of return for a given level of expected return on investments. He realized that investors should consider the combined risk and return of a portfolio of assets, rather than just the risk and return of individual assets. Since Markowtiz’s celebrated paper, portfolio selection problems have grown in complexity. These newer problems may take place in more sophisticated markets, such as the previously described regime-switching market, or they may utilize more complex optimization criteria. Examples of optimization criteria include minimizing risk (e.g. *mean-variance portfolio selection*), maximizing utility, or just simply obtaining a specific dollar amount at the end of trade. Portfolio selection problems in regime-switching market models, with varying risk criteria, have received a fair amount of attention. Two works which address portfolio selection in the setting of regime-switching are those of Zhou and Yin [54], who apply stochastic LQ-control to the problem of mean-variance portfolio selection, and Sotomayor and Cadenillas [48], who use dynamic programming to maximize the *discounted expected utility from consumption* (utility maximization). In both Zhou and Yin [54] and Sotomayor and Cadenillas [48] the regime-switching model is incorporated in a fairly simple way. The market coefficients were assumed to be Markov-modulated as in (2.4), where the market parameters are *completely determined* by the state of the regime-switching Markov chain at the same instant of time. The work of Heunis [23] solves a utility maximization portfolio selection problem in a very general regime-switching market model. The regime-switching market model in Heunis [23] takes the form of (2.3), where the market coefficients are adapted to the *joint filtration generated by* the combined effect of the Brownian motion and the regime-switching Markov chain. The optimality results in Heunis [23] will motivate certain choices and constructions which are made in this thesis. Some of these optimality results were also developed by Donnelly and Heunis [13] who solved a mean-variance portfolio selection problem in the same general regime-switching market. Chapters 3 through 5 of this thesis is concerned with the selection of a portfolio that allows an investor to almost surely meet an obligation at the end of trade in a regime-switching market model, while Chapter 6 will be concerned with a risk minimizing portfolio selection problem in the same financial market model.

Shifting our focus to pricing problems, a pricing problem generally involves answering the following question: what is the "fair price" one can charge for a given financial instrument? This financial instrument often times happens to be a *derivative security*, which is also known as a *contingent claim*. A derivative security, is a financial contract whose value is *derived* from the value of underlying (and more "basic") instruments such as stocks or bonds. Some common derivative securities include put options, call options and future contracts. Derivative securities, just like stocks themselves, are usually traded on exchanges, but they can also be created as private contracts between financial institutions and their clients. In fact, the world-wide market for derivative securities is so large that it is valued in the trillions of dollars! The price for which a publicly traded derivative is sold is usually determined by the law of supply and demand; however, privately sold derivatives tend be be sold at a price which both parties agree to be "fair". Even for publicly traded derivatives, the fact that they are described in terms of underlying stocks or bonds, whose price history is known, suggests there should be a rational way of pricing them through the value of the underlying assets. Finding this price is the solution to the pricing problem of a derivative security. If the price is not made fair, an arbitrage opportunity (an opportunity to exploit the price difference between two identical financial instruments) presents itself, and some investor will "lose" their position. Solving this particular pricing problem, the pricing of derivative securities, is one of the main goals of *financial hedging*.

The hedging of a derivative security is a problem generally faced by financial institutions that engage in selling some sort of derivative product designed to reduce their client's risk. By selling such a product, the financial institution has taken on the client's possible risk and will need to invest in other financial instruments so as to minimize its own exposure to the risk. This method of investment, to minimize incurred risk, is called *financial hedging*. If the institution correctly priced the derivative security which it sold to the client, they should conceivably be able to duplicate the value of the derivative by trading in the stock market. That is an investor employed by the institution can find some portfolio process, whose assets are the underlying stocks and bonds of the derivative, such that the value of the portfolio is the same as the value of the derivative security at each point in time. The portfolio process that achieves this duplication is called the *hedging portfolio*. In the case of the financial institution described above, this hedging portfolio is used to remove the risk incurred by the sale of the security to its client. This is the essence of the most common form of financial hedging, called *almost sure hedging*, which will now be described in more detail.

The almost sure hedging problem involves an investor or agent who must determine some *least initial wealth*, such that by appropriately trading from this initial wealth in a market of stocks and bonds, the agent earns sufficient wealth to pay off some obligation, called a contingent claim, at a fixed future time T (almost surely!). Finding this least initial wealth can be classified as a *pricing problem*, since the investor wants to find the smallest dollar amount in which to invest so that he or she can meet a terminal wealth goal. As described earlier, this initial wealth is the "fair" price for which the contingent claim is sold. It goes without saying that the appropriate strategy for investing the least initial wealth among the stocks and bonds in order to pay off the contingent claim must also be determined, or at least characterized. This strategy is called the *hedging portfolio*, and determination of this portfolio falls within the scope of *portfolio selection.* In summary, the agent invests the least initial wealth in accordance with the hedging portfolio, and is assured that the value of the hedging portfolio almost surely exceeds the value of the contingent claim at close of trade T. The mathematics of almost sure hedging in a standard Brownian motion market model, without regime-switching, is well understood and comprehensively treated in many texts on mathematical finance, such as Karatzas and Shreve [31]. Despite this, we summarize the essence of this problem in Chapter 3 of the thesis, since it is essential background for the almost sure hedging problem in a regime-switching market model. This latter problem is not currently well understood, and is the main goal of the research in this thesis. Chapters 4 and 5 which follows are devoted to solving the almost sure hedging problem in a regime-switching market model. In Chapter 4 we address the almost sure hedging problem in a regime-switching market model and in Chapter 5 we build upon the results of Chapter 4 to solve the almost sure hedging problem in a regime-switching market model with both portfolio constraints and margin requirements.

A specific example of an almost sure hedging problem is the pricing of a *European call option.* A European call option is a type of derivative security where the buyer of the option has the right, but not the obligation, to buy the underlying stock at a pre-specified price (called the *strike price*) on a specific date (called the *expiration date*) regardless of the market value of the stock. If the market value of the stock is greater than the strike price at the expiration date, the buyer will exercise his/her right to buy the stock. If the market price is below the strike price on the expiration date, the buyer will decline buying the stock and will have lost only his or her initial investment (the price of the option). The seller of the call option needs to determine a price for the option that is fair for both the buyer and the seller. Specifically, the seller would like to find a selling price for the option such that when that amount is invested through a hedging portfolio, the seller will be able to make back the strike price less the sell price of the option with probability one. In fact, this price of the option needs to be the lowest possible dollar amount such that the seller can hedge their losses, to avoid an arbitrage opportunity. Generally, when a market has the form (2.3), the solution to the almost sure hedging problem of a European call option is not easily computable. However, if the interest rate and volatility of each stock are both constant, as in (2.1), there is a formulaic method for determining both the selling price of the European call option and the associated hedging portfolio. This method comprises the celebrated *Black and Scholes formula* for the price of a European call option. By way of general background, and because of its interest and importance, the Black-Scholes formula is briefly discussed in Example 3.3.13.

The modern theory of pricing contingent claims began with the work of Samuelson [44], Samuelson and Merton [45], Black and Scholes [4], and Merton [36]. These works dealt with a market model in which the market coefficients are kept constant, as in (2.1), and led to the development of the Black-Scholes option pricing model which began a revolution in finance with the widespread creation of numerous derivative securities. The Black-Scholes formula gives a computable version of the price of the option, which was useful in a time where computational power was not as advanced as it is today. Again, it is important to note that the Black-Scholes method is crucially dependent on accurate knowledge of the volatility parameter σ (see (2.1)), and the success of the formula is tied to how well the volatility parameter matches its real-world counterpart. Empirical evidence has shown that the true volatility is usually not constant, and this led in turn to procedures for estimating volatility changes, and the use of general stochastic volatilities. More contemporary methods for options pricing, such as the model developed by Heston [22], allows volatilities to be modelled as stochastic processes. In this thesis we shall not be concerned with such matters, and will adopt the point of view introduced by Harrison and Kreps [20] and Harrison and Pliska [21], namely that the market coefficients are stochastic processes given as part of the market model, and the stock prices are modeled by multi-dimensional generalizations of the basic relation (2.3). In this model the filtration with respect to which the market coefficient processes μ and σ are adapted is an essential element, since this filtration is the mathematical expression of the underlying randomness in the market model. The almost sure hedging problem is much-studied and well understood in the case where the filtration is generated only by the Brownian motions in the stock-price models (i.e. the process W in (2.3)), and, by way of providing adequate background for this thesis, we summarize the essential aspects of this problem in Chapter 3. Our main research goal involves the case in which this filtration is generated *jointly by the Brownian motion together with a finite-state Markov chain* (which models the regime-switching discussed above). There is a considerable literature on hedging in the very simple and special case for which the market coefficients are just Markov-modulated, as at (2.4) (we mention only two works of this kind, namely Jobert and Rogers [27] and Yao *et-al* [51]). Despite the attention it has received, this special case of somewhat limited interest, for the reasons discussed above (following (2.4)), and the solution methods involved are typically highly problem-specific and rely crucially on the simplicity of the Markov-modulated coefficients. Considerably more challenging is the case where the market coefficients are non-anticipatively determined by both the Brownian motion and the Markov chain together. We devote Chapter 4 to the hedging problem formulated at this level of generality.

Our description of the almost sure hedging problem has so far dealt with a market in which there are no portfolio constraints or margin requirements, that is at every instant the investor can *freely distribute* wealth among all the assets in the market with no margin payments. As noted earlier, portfolio constraints and margin requirements make problems in mathematical finance much more challenging to solve. Karatzas, Lehoczky, Shreve and Xu [29] were the first to solve the problem of constrained utility maximization in what they describe to be an *incomplete market*. This market is a Brownian motion market model where the prices of the n risky assets are driven by a d-dimensional Brownian motion. The *incompleteness* arises since n is assumed to be strictly smaller than d. This makes it impossible for an agent to "hedge without risk" every random contingent claim. That is, there exists contingent claims for which the probability of an investor defaulting is non-zero. Their method of solution relies upon fictitiously completing the incomplete market. Fictitious stocks are carefully chosen so that the optimal portfolio will not invest in them, and as a result, the optimal portfolio in the fictitious market is a potential solution to the incomplete market. By creating optimal portfolios in many fictitious markets, the optimal portfolio in the original incomplete market is the portfolio that minimizes a given expected utility function.

El-Karoui and Quenez [16][17] were the first to produce results for the almost sure hedging problem in an incomplete market - the same market posed by Karatzas, Lehoczky, Shreve and Xu [29]. El-Karoui and Quenez showed that in such a market one could not promise to exactly hedge a given contingent claim. However, they did show that one could always *super-hedge* a given contingent claim. That is, an investor can always find a least initial wealth and portfolio process that almost surely produces more wealth than a given contingent claim. Using the same fictitious-completion approach as Karatzas et al. these authors were able to derive formulas for the least initial wealth and super-hedging portfolio of a contingent claim. Additionally, El-Karoui and Quenez introduced the notion of a *cumulative consumption process* which allows an investor to *consume* wealth during the trading interval. They showed that with the inclusion of a specifically chosen consumption process, the least initial wealth and super-hedging portfolio does, in fact, exactly hedge a given contingent claim. This notion of a cumulative consumption process is described in

Chapter 4 and will be essential for the results developed in Chapters 4 and 5 of this thesis.

Cvitanic and Karatzas [10] extended the approach of El-Karoui and Quenez [16][17] to the case of *general convex constraints on portfolio proportions*. These convex portfolio constraints are far more general than the constraints imposed on the market by Karatzas *et-al* [29]. Cvitanic and Karatzas developed representations of the least initial wealth and super-hedging portfolio in such a constrained market by use of the *Doob-Meyer decomposition* and *martingale representation theorem*. Their more probabilistic treatment of the hedging problem is based upon very deep methods developed for stochastic control problems. Cvitanic and Karatzas also show how to modify their approach for hedging to include very simple margin requirements, specifically *higher interest rates for borrowing than for investing*; however, in the absence of portfolio constraints. Follmer and Kramkov [18] have extended the approach of El-Karoui and Quenez [16][17] and Cvitanic and Karatzas [10] beyond the Brownian motion case to very general market models, with the asset prices being represented by abstract semimartingales. Needless to say, with this general model, one can establish only *existence* of a constrained super-hedging strategy, without the rather concrete characterizations of the super-hedging strategy that one finds in El-Karoui and Quenez [16][17] and Cvitanic and Karatzas [10] for the Brownian motion case. As a result, in Chapter 5 of this thesis, we extend many of the methods of Cvitanic and Karatzas to develop a characterization for the optimal hedging strategy in the almost sure hedging problem in a regime-switching market model with both convex portfolio constraints and general margin requirements.

Often times the least initial wealth required to super-hedge a contingent claim is so high that it is unrealistic for an investor to pay such a price. For example, if an agent sells a call option on one share of stock, and cannot borrow money, he must sell the option for the price of the stock to come out even. However, nobody would pay this price for an option if they can just buy the stock itself. Cvitanic [8] studied the problem of *approximate hedging*, that is, the problem of minimizing the expected loss of hedging with an initial investment less than the price of a given contingent claim. In a Brownian motion market model with general convex constraints, Cvitanic shows that starting from some initial wealth, there always exists a hedging strategy that minimizes a specific convex loss function. What makes this problem difficult is the fact that the loss function is not smooth, and therefore, one cannot simply resort to tools of smooth convex analysis. Instead, Cvitanic uses methods such as the Komlós theorem and non-smooth convex analysis to prove the existence of a risk-minimizing hedging strategy. In the last portion of this thesis we will study the approximate hedging problem in a constrained regime-switching market model. Specifically, in Chapter 6 of this thesis, we will show the existence of a risk-minimizing hedging strategy in a regime-switching market with both convex portfolio constraints and general margin requirements using a procedure very similar to that of Cvitanic.

Chapter 3

The Unconstrained Hedging Problem in a Brownian Motion Market Model

The subject of this chapter is the almost sure hedging of contingent claims in a standard financial market model. Specifically, if the seller of a contingent claim is required to pay off some strictly positive amount at some final time only, the least initial wealth and the associated portfolio process that almost surely promises the agent can pay off the claim will be found. This chapter provides a summary of the main elements of unconstrained hedging, as found in §5.2, §5.3 and §5.6 in Karatzas and Shreve [31]. All of the results in this chapter are well known, and do not constitute new research. In particular the regime-switching in the market model, which is the main goal of our research, is not addressed here. However, the results of this chapter are essential prior background for posing the problem we intend to address and for establishing a framework which we shall extend to include regime-switching in Chapters 4 and 5.

3.1 Market Model

The model that will be used in this chapter is a continuous-time market model over a time horizon $[0, T]$ for a constant $T \in (0, \infty)$. The market model comprises the basic elements stipulated in Conditions 3.1.1 and 3.1.3.

Condition 3.1.1. *A standard N-dimensional Brownian motion $\{\mathbf{W}(t); t \in [0, T]\}$ with scalar entries $W_n(t)$ where $n = 1, \ldots, N$, is given on the common complete probability space $(\Omega, \mathcal{F}, P)$ and is the source of randomness in the model. With $\mathcal{N}(P)$ the collection of* *all P-null events in $\mathcal{F}$, that is $\mathcal{N}(P) \triangleq \{A \in \mathcal{F} : P(A) = 0\}$, define the usual augmented filtration $\{\mathcal{F}_t\}$ as*

$$\{\mathcal{F}_t\} \triangleq \sigma\{\mathbf{W}(s), s \in [0,t]\} \vee \mathcal{N}(P) \tag{3.1}$$

for all $t \in [0,T]$. In particular, $\{\mathcal{F}_t\}$ represents the information available to investors.

Notation 3.1.2. The notation $\mathcal{F}^*$ is used to denote the $\{\mathcal{F}_t\}$-progressively measurable σ-algebra on $[0,T] \times \Omega$, and $\eta \in \mathcal{F}^*$ is used to indicate that the process $\eta : [0,T] \times \Omega \to \mathbb{R}^N$ is $\{\mathcal{F}_t\}$-progressively measurable. The qualifier "a.s." always refers to the probability P on $(\Omega, \mathcal{F})$, while the qualifier "a.e." refers to the product measure $\lambda \otimes P$ on $\mathcal{B}([0,T]) \times \mathcal{F}$, where λ denotes the Lebesgue measure on the Borel σ-algebra of $[0,T]$, denoted $\mathcal{B}([0,T])$.

Condition 3.1.3. *The market comprises a single **risk-free asset** with price $\{S_0(t); t \in [0,T]\}$ and several **risky assets** with prices $\{S_n(t); t \in [0,T]\}$, $n = 1, \ldots, N$, modeled by the stochastic differential equations*

$$dS_0(t) = r(t)S_0(t)dt, \qquad dS_n(t) = S_n(t)(b_n(t)dt + \sum_{m=1}^{N} \sigma_{nm}(t)dW_m(t)), \tag{3.2}$$

*with $S_0(0) \triangleq 1$, and $S_n(0)$, $n = 1, \ldots, N$, being given, strictly positive constants. The **risk-free interest rate** $\{r(t)\}$, the entries $\{b_n(t)\}$ of the $\mathbb{R}^N$-valued **rate of return** $\{\boldsymbol{b}(t)\}$, and the entries $\{\sigma_{nm}(t)\}$ of the $N \times N$ matrix-valued **volatility process** $\{\boldsymbol{\sigma}(t)\}$ are given uniformly bounded and $\{\mathcal{F}_t\}$-progressively measurable scalar processes on $[0,T] \times \Omega$. Using $||\boldsymbol{z}||$ for the Euclidean norm and $\boldsymbol{z}^\top$ for the transpose of a vector $\boldsymbol{z} \in \mathbb{R}^N$, we shall suppose that there exists a constant $\kappa \in (0,\infty)$ such that $\boldsymbol{z}^\top \boldsymbol{\sigma}(t,\omega)\boldsymbol{\sigma}(t,\omega)^\top \boldsymbol{z} \geq \kappa ||\boldsymbol{z}||^2$ for all $(\boldsymbol{z}, t, \omega) \in \mathbb{R}^N \times [0,T] \times \Omega$.*

Remark 3.1.4. The *risk-free interest rate process* $\{r(t)\}$, the *mean rate of return process* $\{\boldsymbol{b}(t)\}$, and the *volatility process* $\{\boldsymbol{\sigma}(t)\}$, stipulated at Condition 3.1.3 are assumed known, and are together called the *market coefficient processes* of the market model.

Remark 3.1.5. From now on Conditions 3.1.1 and 3.1.3 will be assumed without explicit mention for all theorems, lemmas and propositions in this chapter.

Remark 3.1.6. In view of the constant $\kappa \in (0,\infty)$ postulated in Condition 3.1.3, we get from (2.4) and (2.5) on p. 90 of Xu and Shreve [50], the existence of a constant $\kappa_1 \in (1,\infty)$ such that

$$\begin{aligned} &\frac{1}{\kappa_1} \max\{||(\boldsymbol{\sigma}(t,\omega))^{-1}\boldsymbol{z}||, ||(\boldsymbol{\sigma}(t,\omega)^\top)^{-1}\boldsymbol{z}||\} \\ &\leq ||\boldsymbol{z}|| \leq \kappa_1 \min\{(||(\boldsymbol{\sigma}(t,\omega))^{-1}\boldsymbol{z}||, ||(\boldsymbol{\sigma}(t,\omega)^\top)^{-1}\boldsymbol{z}||\} \end{aligned} \tag{3.3}$$

for all $(\boldsymbol{z}, t, \omega) \in \mathbb{R}^N \times [0,T] \times \Omega$.

Notation 3.1.7. Define the $\mathbb{R}^N$- valued *market price of risk* $\{\boldsymbol{\theta}(t),\ t \in [0,T]\}$ by

$$\boldsymbol{\theta}(t) \triangleq \boldsymbol{\sigma}^{-1}(t)(\boldsymbol{b}(t) - r(t)\mathbf{1}) \tag{3.4}$$

for all $t \in [0,T]$ where $\mathbf{1}$ is the N-dimensional vector of unit entries. From Condition 3.1.3 and Remark 3.1.6, the process $\boldsymbol{\theta}$ is $\{\mathcal{F}_t\}$-progressively measurable and uniformly bounded, i.e., $\kappa_{\boldsymbol{\theta}} \triangleq \sup_{(t,\omega)} ||\boldsymbol{\theta}(t,\omega)|| < \infty$.

Before defining the problem of unconstrained hedging we make some preliminary comments in the following remarks and text on how an agent trades in the market just defined:

Remark 3.1.8. It is assumed throughout that, at each and every instant t in the trading interval $[0,T]$, an agent (or investor) allocates his or her *total wealth* among the risk-free asset (with price $S_0(t)$) and the risky assets (with prices $S_n(t)$, $n = 1,2,\ldots,N$). To model this investment we introduce the process $\{\pi_0(t), t \in [0,T\}$, which denotes the *monetary (or dollar)* amount allocated by the agent to the risk-free asset. We shall always assume that this process is $\{\mathcal{F}_t\}$-progressively measurable, that is $\pi_0 \in \mathcal{F}^*$. To model investment in the N-risky assets we define the following space Π of *portfolio processes*:

$$\Pi \triangleq \left\{ \boldsymbol{\pi} : [0,T] \times \Omega \to \mathbb{R}^N \middle| \boldsymbol{\pi} \in \mathcal{F}^*,\ \int_0^T ||\boldsymbol{\pi}(t)||^2 dt < \infty \text{ a.s.} \right\}. \tag{3.5}$$

The interpretation of elements $\boldsymbol{\pi}$ of Π is as follows: Each scalar component $\pi_n(t)$, for $n = 1 \ldots N$, of the $\mathbb{R}^N$-vector $\boldsymbol{\pi}(t)$, for $t \in [0,T]$, indicates the monetary (or dollar) amount allocated by the agent to the risky asset with price $S_n(t)$ at the instant $t \in [0,T]$. It then follows that the *total wealth* of the agent is the sum of the amounts in the risk-free and risky assets, that is the total wealth of the agent is given by the process $\{X(t), t \in [0,T]\}$ defined as follows:

$$X(t) \triangleq \sum_{n=0}^{N} \pi_n(t) = \boldsymbol{\pi}(t)^\top \mathbf{1} + \pi_0(t), \qquad t \in [0,T], \tag{3.6}$$

in which $\mathbf{1}$ is the N-dimensional vector with unit entries.

Denote by x the *initial wealth* of the agent at start of trade $t = 0$, that is

$$x \triangleq \sum_{n=0}^{N} \pi_n(0). \tag{3.7}$$

From now on it will be assumed that the agent begins trading with *non-negative* initial wealth, that is

$$x \in [0,\infty). \tag{3.8}$$

Remark 3.1.9. A basic assumption that we shall adopt in the present chapter, and which is common throughout much of mathematical finance, is that the agent follows a *self funded trading strategy* over the entire trading interval $t \in [0, T]$. This means that apart from the initial wealth x assigned to the agent at start of trade $t = 0$, over the remaining trading interval $t \in (0, T]$, there is never any infusion of cash to or removal of cash from the investor's wealth by an external source (e.g. a wealthy relative). The total investor wealth at every $t \in [0, T]$ is distributed *entirely* among the risk-less and risky assets (so that (3.6) holds). In particular there are never any payments out of the investor's wealth for consumption.

With the stipulation of self-funded trading, as in Remark 3.1.9, it is a standard result of mathematical finance (see §1.2 - §1.3 in Methods of Mathematical Finance by Karatzas and Shreve [31]) that the wealth process X of the agent necessarily satisfies the following stochastic integral equation:

$$\begin{aligned} X(t) = x + \int_0^t \Big\{ r(s)X(s) + \boldsymbol{\pi}(s)^\top \boldsymbol{\sigma}(s)\boldsymbol{\theta}(s) \Big\} ds \\ + \int_0^t \boldsymbol{\pi}(s)^\top \boldsymbol{\sigma}(s) d\mathbf{W}(s), \quad t \in [0, T], \end{aligned} \tag{3.9}$$

in which of course x is the initial wealth of the agent and $\boldsymbol{\pi} \in \Pi$ is the portfolio process according to which the wealth of the agent is allocated among the N *risky* assets (see Remark 3.1.8). The remarkable thing about (3.9) is that the process $\{\pi_0(t)\}$, giving the allocation of investor wealth to the riskless asset with price $\{S_0(t)\}$ (see Remark 3.1.8) does not appear anywhere, and only the portfolio process $\boldsymbol{\pi} \in \Pi$ appears in the equation. The reason for this, as a detailed derivation of (3.9) (which we do not give here) would make clear, is that the assumption of a *self-funded* trading strategy renders $\pi_0(t)$ a redundant variable which can be eliminated using the relation (3.6).

Remark 3.1.10. From the stochastic integral equation (3.9) it is clear that the wealth process X is *completely determined* by the initial wealth x (which is effectively the "initial value") and the portfolio process $\boldsymbol{\pi} \in \Pi$ (which is effectively a "control input" decided upon by the agent). From now on we shall denote the wealth process of the agent not by X but instead by $X^{(x,\boldsymbol{\pi})}$, to make clear the dependence of the wealth process on the initial wealth x and portfolio process $\boldsymbol{\pi} \in \Pi$, so that (3.9) will be written as follows:

$$\begin{aligned} X^{(x,\boldsymbol{\pi})}(t) = x + \int_0^t \Big\{ r(s)X^{(x,\boldsymbol{\pi})}(s) + \boldsymbol{\pi}(s)^\top \boldsymbol{\sigma}(s)\boldsymbol{\theta}(s) \Big\} ds \\ + \int_0^t \boldsymbol{\pi}(s)^\top \boldsymbol{\sigma}(s) d\mathbf{W}(s), \quad t \in [0, T], \end{aligned} \tag{3.10}$$

or, in the more usual notation of stochastic differential equations, as

$$dX^{(x,\boldsymbol{\pi})}(t) = \{r(t)X^{(x,\boldsymbol{\pi})}(t) + \boldsymbol{\pi}(t)^\top \boldsymbol{\sigma}(t)\boldsymbol{\theta}(t)\}dt + \boldsymbol{\pi}(t)^\top \boldsymbol{\sigma}(t)d\mathbf{W}(t), \quad t \in [0,T], \tag{3.11}$$

with the initial condition

$$X^{(x,\boldsymbol{\pi})}(0) = x. \tag{3.12}$$

The SDE (3.11) - (3.12) is simple enough that we can explicitly solve for the wealth $X^{(x,\boldsymbol{\pi})}(t)$ in terms of x and $\boldsymbol{\pi}$. In fact, from Condition 3.1.3 we see that the risk-free asset has the price

$$S_0(t) = \exp\left\{\int_0^t r(\tau)d\tau\right\}, \qquad t \in [0,T]. \tag{3.13}$$

It then follows easily from Ito's formula that the wealth process is given explicitly in terms of $x \in [0,\infty)$ and $\boldsymbol{\pi} \in \Pi$ by the expression

$$\begin{aligned} X^{(x,\boldsymbol{\pi})}(t) &= S_0(t)x + S_0(t)\left\{\int_0^t \frac{1}{S_0(\tau)}\boldsymbol{\pi}(\tau)^\top \boldsymbol{\sigma}(\tau)\boldsymbol{\theta}(\tau)d\tau\right\} \\ &\quad + S_0(t)\left\{\int_0^t \frac{1}{S_0(\tau)}\boldsymbol{\pi}(\tau)^\top \boldsymbol{\sigma}(\tau)d\mathbf{W}(\tau)\right\}, \ \ t \in [0,T]. \end{aligned} \tag{3.14}$$

3.2 Definition of the Hedging Problem

In the present section we shall define the problem of unconstrained hedging with reference to the preliminaries outlined in the previous section. We begin by defining a *contingent claim*. A *contingent claim* B is a specified strictly positive $\mathcal{F}_T$-measurable random variable. That is

$$B : (\Omega, \mathcal{F}, P) \mapsto (0,\infty), \qquad B \in \mathcal{F}_T. \tag{3.15}$$

Since B is $\mathcal{F}_T$-measurable it incorporates all of the "randomness" in the market over the entire trading interval $t \in [0,T]$, that is one sees a realization $B(\omega)$ of this random variable only at the close of trade $t = T$. The random variable B defines an obligation on the part of the agent in the following sense: starting from some initial wealth $x \in [0,\infty)$, the agent must come up with some portfolio process $\boldsymbol{\pi} \in \Pi$ such that

$$X^{(x,\boldsymbol{\pi})}(T) = B \ \text{ a.s.} \tag{3.16}$$

In other words, the agent starts with an initial wealth $x \in [0,\infty)$ and must invest in the risky assets in such a way that sufficient wealth is generated at close of trade $t = T$ for the agent to be able to pay off the contingent claim B. For both technical and financial reasons we restrict attention to portfolio processes $\boldsymbol{\pi} \in \Pi$ for which the corresponding wealth process is *non-negative over the entire trading interval.* That is, for a given initial wealth $x \in [0, \infty)$, the agent is restricted to portfolio processes $\boldsymbol{\pi}$ in the set $\mathcal{A}(x)$ defined as follows:

$$\mathcal{A}(x) \triangleq \left\{ \boldsymbol{\pi} \in \Pi \middle| X^{(x,\boldsymbol{\pi})}(t) \geq 0 \text{ a.s. for all } t \in [0,T] \right\}. \tag{3.17}$$

Remark 3.2.1. Observe that $\mathcal{A}(x) \neq \emptyset$ for each $x \in [0, \infty)$. To see this take $\boldsymbol{\pi} = 0$. In accordance with Remark 3.1.8 this amounts to investing *zero wealth* in the risky assets (so that the entire wealth is invested in the risk-free asset with price S_0.) With this $\boldsymbol{\pi}$ one sees that the last two terms on the right side of (3.14) are zero so that

$$X^{(x,\boldsymbol{\pi})}(t) = S_0(t)x \geq 0, \qquad t \in [0,T], \tag{3.18}$$

and therefore $\boldsymbol{\pi} \in \mathcal{A}(x)$. Thus $\mathcal{A}(x) \neq \emptyset$ for each $x \in [0, \infty)$.

Remark 3.2.2. An agent who begins with initial wealth $x \in [0, \infty)$ and determines some $\pi \in \mathcal{A}(x)$ such that (3.16) holds is said to *hedge the contingent claim* B from the initial wealth x, and the corresponding portfolio process $\boldsymbol{\pi} \in \mathcal{A}(x)$ is called an *unconstrained hedging strategy* from the initial wealth x.

Remark 3.2.3. In light of the preceding discussion, from now on we restrict to pairs $(x, \boldsymbol{\pi})$ for which $x \in [0, \infty)$ and $\boldsymbol{\pi} \in \mathcal{A}(x)$, so that the wealth process $X^{(x,\boldsymbol{\pi})}$ is non-negative over the trading interval $t \in [0,T]$. From a financial point of view this is a reasonable requirement, for it merely insists that the agent never go into debt during the trading interval (negative wealth amounts to debt). From a mathematical viewpoint, removal of the condition of non-negative wealth (essentially by requiring $\boldsymbol{\pi} \in \Pi$ instead of $\boldsymbol{\pi} \in \mathcal{A}(x)$), which allows the agent to temporarily take on debt during the trading interval, leads to a rather intractable hedging problem, which is in fact not well understood to this day. We shall take care to point out the essential role played by non-negativity of the wealth whenever this condition is used.

In light of the previous discussion of the hedging problem it is clear that the initial wealth $x \in [0, \infty)$ is a factor of great importance. Indeed, from (3.14) one sees that if the initial wealth x is too small then there may not even exist any $\boldsymbol{\pi} \in \mathcal{A}(x)$ such that (3.16) holds, that is there fails to exist an unconstrained hedging strategy from the initial wealth x. Therefore, one of our principal goals is to determine the *least possible* initial wealth for which there does exist a hedging strategy. To make this precise we define the set Λ comprising all the initial x for which a hedging strategy does exist, that is

$$\Lambda \triangleq \left\{ x \in [0, \infty) \middle| \; \exists \boldsymbol{\pi} \in \mathcal{A}(x) \quad \text{s.t. } X^{(x,\boldsymbol{\pi})}(T) = B \text{ a.s.} \right\}. \tag{3.19}$$

In view of (3.19) and (3.17) we can express Λ in expanded form as follows:

$$\Lambda = \left\{ x \in [0,\infty) \,\middle|\, X^{(x,\boldsymbol{\pi})}(T) = B \text{ a.s. for some } \boldsymbol{\pi} \in \Pi \right. \\ \left. \text{such that } X^{(x,\boldsymbol{\pi})}(t) \geq 0 \text{ a.s. for all } t \in [0,T] \right\}. \tag{3.20}$$

Remark 3.2.4. There is of course no guarantee that Λ is non-empty. Indeed, as we shall see later on (at Remark 3.3.7), if the contingent claim random variable B is stipulated to be "unreasonably large" then Λ will be empty. However, assuming that Λ is indeed non-empty and furthermore attains its infimum at some $\hat{u} \in [0,\infty)$ (that is $\hat{u} \in \Lambda$ and $\hat{u} \leq x$ for all $x \in \Lambda$), it is evident that $x = \hat{u}$ is the least initial wealth from which one can hedge the contingent claim B. This least initial wealth is usually called *the price of the contingent claim* B. The most important goals of this chapter can now be stated as follows:

(a) Determine conditions on the contingent claim random variable B which ensure that Λ is non-empty.

(b) With the conditions in (a) in force establish that there exists some $\hat{u} \in \Lambda$ such that $\hat{u} \leq x$ for all $x \in \Lambda$. It then follows from (3.19) that $x = \hat{u}$ is the *least initial wealth* from which one can hedge the contingent claim B.

(c) With $x = \hat{u}$ characterize some $\boldsymbol{\pi} \in \mathcal{A}(x)$ such that (3.16) holds. This portfolio process hedges the contingent claim B from the least initial wealth $\hat{u}$.

3.3 Solution to the Hedging Problem

Having outlined the problem of unconstrained hedging in Remark 3.2.4 we shall now address this problem. To this end an essential role will be played by the so-called *state price density process*.

Definition 3.3.1. The *state price density process* $\{H_0(t),\ t \in [0,T]\}$ for the market model defined by Conditions 3.1.1 and 3.1.3 is the process with values in $(0,\infty)$ defined as follows:

$$H_0(t) \triangleq \exp\left\{ -\int_0^t r(s)ds \right\} \mathcal{E}(-\boldsymbol{\theta} \bullet \mathbf{W})(t), \qquad t \in [0,T], \tag{3.21}$$

in which we use the following abbreviated notation for stochastic integrals

$$(\boldsymbol{\theta} \bullet \mathbf{W})(t) \triangleq \int_0^t \boldsymbol{\theta}(\tau)^\top d\mathbf{W}(\tau), \qquad t \in [0,T], \tag{3.22}$$

and $\mathcal{E}(\cdot)(t)$ is the usual Ito exponential function (see Remark C.10.3).

Remark 3.3.2. A simple exercise using Ito's product rule allows us to write $H_0(t)$ as

$$H_0(t) = 1 - \int_0^t H_0(s)\boldsymbol{\theta}(s)^\top d\mathbf{W}(s) - \int_0^t H_0(s)r(s)ds, \quad t \in [0,T]. \tag{3.23}$$

Since $\mathbf{W} \in \mathcal{M}^c_{loc}(\{\mathcal{F}_t\}, P)$, we have that

$$H_0 \in \mathcal{SM}^c(\{\mathcal{F}_t\}, P), \tag{3.24}$$

that is H_0 is a continuous $\{\mathcal{F}_t\}$-semimartingale (see Definition C.8.1 and Notation C.8.2). Furthermore, observe from (3.21) that

$$\inf_{t\in[0,T]} H_0(t) > 0 \text{ a.s.} \tag{3.25}$$

The significance of the state price density process is in the following result:

Proposition 3.3.3. *For each $x \in [0,\infty)$ and $\boldsymbol{\pi} \in \mathcal{A}(x)$, the process $\{J_0^{(x,\boldsymbol{\pi})}(t), t \in [0,T]\}$ defined as*

$$J_0^{(x,\boldsymbol{\pi})}(t) \triangleq H_0(t)X^{(x,\boldsymbol{\pi})}(t), \qquad t \in [0,T], \tag{3.26}$$

is a non-negative $\{\mathcal{F}_t\}$-supermartingale (i.e. $J_0^{(x,\boldsymbol{\pi})} \in \mathcal{SPM}(\{\mathcal{F}_t\}, P)$ and $J_0^{(x,\boldsymbol{\pi})} \geq 0$).

Proof. For ease of notation, use $X(t) = X^{(x,\boldsymbol{\pi})}(t)$ for all $t \in [0,T]$. A routine calculation using Ito's product rule establishes that

$$J_0^{(x,\boldsymbol{\pi})}(t) = x + \int_0^t H_0(s)\Big(\boldsymbol{\pi}(s)^\top\boldsymbol{\sigma}(s) - X(s)\boldsymbol{\theta}(s)^\top\Big)d\mathbf{W}(s), \qquad t \in [0,T]. \tag{3.27}$$

From $\mathbf{W} \in \mathcal{M}^c_{loc}(\{\mathcal{F}_t\}, P)$ and (3.27) we get

$$J_0^{(x,\boldsymbol{\pi})} \in \mathcal{M}^c_{loc}(\{\mathcal{F}_t\}, P), \tag{3.28}$$

that is $J_0^{(x,\boldsymbol{\pi})}$ is a continuous $\{\mathcal{F}_t\}$-local martingale. From (3.25) and $X^{(x,\boldsymbol{\pi})}(t) \geq 0$ a.s. for all $t \in [0,T]$ (from (3.17) and $\boldsymbol{\pi} \in \mathcal{A}(x)$), we have that

$$J_0^{(x,\boldsymbol{\pi})}(t) \geq 0 \text{ a.s.}, \quad t \in [0,T]. \tag{3.29}$$

As a result of $J_0^{(x,\boldsymbol{\pi})}$ being a non-negative local martingale, from Proposition C.11.3, we have that $J_0^{(x,\boldsymbol{\pi})}$ is a $\{\mathcal{F}_t\}$-supermartingale, i.e.

$$J_0^{(x,\boldsymbol{\pi})} \in \mathcal{SPM}(\{\mathcal{F}_t\}, P). \tag{3.30}$$

□

Remark 3.3.4. Note that non-negativity of the wealth process is essential for Proposition 3.3.3 to hold.

The following corollary, which is immediate from Proposition 3.3.3, is a key element to the solution of the unconstrained hedging problem.

Corollary 3.3.4.1. *For each* $x \in [0,\infty), \boldsymbol{\pi} \in \mathcal{A}(x)$,

$$E\left[H_0(T)X^{(x,\boldsymbol{\pi})}(T)\right] \le x. \tag{3.31}$$

Remark 3.3.5. Suppose that Λ is non-empty. Then for *each* $x \in \Lambda$ there exists a $\boldsymbol{\pi} \in \mathcal{A}(x)$ such that

$$X^{(x,\boldsymbol{\pi})}(T) = B \text{ a.s.} \tag{3.32}$$

Since $B > 0$ and $H_0(T) > 0$, from Corollary 3.3.4.1 and (3.32) we have

$$\Lambda \neq \emptyset \implies E\left[H_0(T)B\right] \le x, \qquad \text{for all } x \in \Lambda. \tag{3.33}$$

Remark 3.3.6. Note that the expectation $E\left[H_0(T)B\right]$ is always defined, with values in $[0,\infty]$, since the random variable $H_0(T)B$ is non-negative.

Remark 3.3.7. From (3.33) one sees that

$$\Lambda \neq \emptyset \implies E\left[H_0(T)B\right] < \infty.$$

One of the main results of this chapter (see Theorem 3.3.10) establishes the converse implication, so that we actually have the following equivalence:

$$\Lambda \neq \emptyset \iff E\left[H_0(T)B\right] < \infty.$$

Thus, if the contingent claim random variable B is stipulated so large that $E\left[H_0(T)B\right] = \infty$ then there is no possibility of hedging the claim B.

Notation 3.3.8. In the following an essential role will be played by the extended real number $\hat{u} \in [0,\infty]$ defined as follows:

$$\hat{u} \triangleq E[H_0(T)B]. \tag{3.34}$$

The following result is then immediate:

Proposition 3.3.9. *Suppose* $\hat{u} < \infty$, *and define the process process* $\{\hat{X}(t),\ t \in [0,T]\}$ *as follows:*

$$\hat{X}(t) \triangleq \frac{E[H_0(T)B|\mathcal{F}_t]}{H_0(t)}, \qquad t \in [0,T]. \tag{3.35}$$

Then the following hold:

(1) $\hat{X}(0) = \hat{u}$ *a.s.*

(2) $\hat{X}(T) = B$ *a.s.*

(3) $\hat{X}(t) \geq 0$ *a.s.,* $\quad t \in [0, T]$

(4) $H_0 \hat{X} \in \mathcal{M}^c(\{\mathcal{F}_t\}, P)$.

The main result of this chapter is as follows:

Theorem 3.3.10. *Suppose that* $\hat{u} < \infty$. *Then* $\hat{u} \in \Lambda$ *(in particular* $\Lambda \neq \emptyset$*) and* $\hat{u} = \inf \Lambda$, *that is* $\hat{u} \leq x$ *for all* $x \in \Lambda$.

Remark 3.3.11. From Theorem 3.3.10 and (3.17), one sees that $\hat{u}$ is the *least initial wealth* which guarantees that $X^{(\hat{u},\boldsymbol{\pi})}(T) = B$ a.s. for some $\boldsymbol{\pi} \in \mathcal{A}(\hat{u})$, that is $\hat{u}$ is the *price* of the contingent claim B (see Remark 3.2.4).

Proof of Theorem 3.3.10 . We are going to establish that there exists some portfolio process $\hat{\boldsymbol{\pi}} \in \mathcal{A}(\hat{u})$ such that

$$\hat{X}(t) = X^{(\hat{u},\hat{\boldsymbol{\pi}})}(t), \qquad t \in [0, T], \tag{3.36}$$

where $\hat{X}$ is defined at (3.35), and from (3.10), the process $X^{(\hat{u},\hat{\boldsymbol{\pi}})}$ satisfies the stochastic integral equation

$$\begin{aligned} X^{(\hat{u},\hat{\boldsymbol{\pi}})}(t) = \hat{u} + \int_0^t \Big\{ r(s) X^{(\hat{u},\hat{\boldsymbol{\pi}})}(s) + \hat{\boldsymbol{\pi}}(s)^\top \boldsymbol{\sigma}(s) \boldsymbol{\theta}(s) \Big\} ds \\ + \int_0^t \hat{\boldsymbol{\pi}}(s)^\top \boldsymbol{\sigma}(s) d\mathbf{W}(s), \quad t \in [0, T]. \end{aligned} \tag{3.37}$$

If the relation (3.36) indeed holds for some $\hat{\boldsymbol{\pi}} \in \mathcal{A}(\hat{u})$ then, from Proposition 3.3.9(2), we obtain

$$X^{(\hat{u},\hat{\boldsymbol{\pi}})}(T) = B \text{ a.s.} \tag{3.38}$$

Since $\hat{u}$ is the initial wealth of the process $X^{(\hat{u},\hat{\boldsymbol{\pi}})}$, this would mean, from (3.19),

$$\hat{u} \in \Lambda \text{ thus } \Lambda \neq \emptyset. \tag{3.39}$$

Furthermore, from (3.33), we have that $\hat{u} \leq \inf \Lambda$, and therefore, from (3.39),

$$\hat{u} = \inf \Lambda \in \Lambda. \tag{3.40}$$

Thus it remains to establish that (3.36) holds for some $\hat{\boldsymbol{\pi}} \in \mathcal{A}(\hat{u})$, for then Theorem 3.3.10 follows from (3.39) and (3.40).

To begin, we have from Proposition 3.3.9(4) that $H_0\hat{X} \in \mathcal{M}^c(\{\mathcal{F}_t\}, P)$. As a result, we can use the martingale representation theorem from Theorem C.12.3 to write

$$H_0(t)\hat{X}(t) = \hat{u} + \int_0^t \mathbf{\Psi}_0(s)^\top d\mathbf{W}(s), \qquad t \in [0,T], \tag{3.41}$$

for some unique $\mathbb{R}^N$-valued and $\{\mathcal{F}_t\}$-progressively measurable process $\{\mathbf{\Psi}_0(t),\ t \in [0,T]\}$ such that

$$\int_0^T ||\mathbf{\Psi}_0(t)||^2 dt < \infty \text{ a.s.} \tag{3.42}$$

Dividing both sides of (3.41) by $H_0(t)$ (recall (3.25)) yields

$$\hat{X}(t) = \frac{1}{H_0(t)}\left(\hat{u} + \int_0^t \mathbf{\Psi}_0(s)^\top d\mathbf{W}(s)\right), \quad t \in [0,T]. \tag{3.43}$$

Using Ito's product formula we can expand the right-hand side of (3.43) to find

$$\begin{aligned}\hat{X}(t) = \hat{u} &+ \int_0^t \left(\frac{\mathbf{\Psi}_0(s)^\top}{H_0(s)} + \hat{X}(s)\boldsymbol{\theta}(s)^\top\right) d\mathbf{W}(s) \\ &+ \int_0^t \left(\frac{\mathbf{\Psi}_0(s)^\top}{H_0(s)} + \hat{X}(s)\boldsymbol{\theta}(s)^\top\right)\boldsymbol{\theta}(s)ds + \int_0^t \hat{X}(s)r(s)ds\end{aligned} \tag{3.44}$$

for all $t \in [0,T]$. Now define the $\mathbb{R}^N$-valued process

$$\hat{\boldsymbol{\pi}}(t) \triangleq [\boldsymbol{\sigma}^{-1}(t)]^\top \left[\frac{\mathbf{\Psi}_0(t)}{H_0(t)} + \hat{X}(t)\boldsymbol{\theta}(t)\right], \quad t \in [0,T]. \tag{3.45}$$

Since $\mathbf{\Psi}_0 \in \mathcal{F}^*$ and $\boldsymbol{\sigma} \in \mathcal{F}^*$ (from Condition 3.1.3), one sees from (3.45) that

$$\hat{\boldsymbol{\pi}} \in \mathcal{F}^*. \tag{3.46}$$

Moreover $\boldsymbol{\theta}(\cdot)$ is uniformly bounded on $[0,T] \times \Omega$ (see Notation 3.1.7), and from (3.44) it follows that $\hat{X}(\cdot,\omega)$ is continuous on $[0,T]$ and therefore uniformly bounded on $[0,T]$ for P-almost all $\omega \in \Omega$. As a result, $\hat{X}(\cdot,\omega)\boldsymbol{\theta}(\cdot,\omega)$ is uniformly bounded on $[0,T]$ for P-almost all $\omega \in \Omega$, and so

$$\int_0^T ||\hat{X}(t)\boldsymbol{\theta}(t)||^2 dt < \infty \text{ a.s.} \tag{3.47}$$

Additionally, since $\inf_{t\in[0,T]} H_0(t) > 0$ a.s. (see (3.25)), from (3.42) we have

$$\int_0^T ||\frac{\mathbf{\Psi}_0(t)}{H_0(t)}||^2 dt < \infty \text{ a.s.} \tag{3.48}$$

Combining (3.47) and (3.48) gives

$$\int_0^T ||\frac{\mathbf{\Psi}_0(t)}{H_0(t)} + \hat{X}(t)\boldsymbol{\theta}(t)||^2 dt < \infty \text{ a.s.} \tag{3.49}$$

In view of (3.3), (3.45) and (3.49), we obtain

$$\int_0^T ||\hat{\boldsymbol{\pi}}(t)||^2 dt < \infty \text{ a.s.} \tag{3.50}$$

and from (3.50), (3.46) and (3.5) we find that

$$\hat{\boldsymbol{\pi}} \in \Pi, \tag{3.51}$$

that is $\hat{\boldsymbol{\pi}}$ is a valid portfolio process. In view of (3.44) and (3.45) we have the identity

$$\begin{aligned} \hat{X}(t) = \hat{u} &+ \int_0^t \left\{ r(s)\hat{X}(s) + \hat{\boldsymbol{\pi}}(s)^\top \boldsymbol{\sigma}(s)\boldsymbol{\theta}(s) \right\} ds \\ &+ \int_0^t \hat{\boldsymbol{\pi}}(s)^\top \boldsymbol{\sigma}(s) d\mathbf{W}(s), \quad t \in [0,T]. \end{aligned} \tag{3.52}$$

Now (3.36) follows upon comparison of (3.52) and (3.37), and it remains to verify that $\hat{\boldsymbol{\pi}} \in \mathcal{A}(\hat{u})$. From Proposition 3.3.9(3) together with (3.36) we have

$$X^{(\hat{u},\hat{\boldsymbol{\pi}})}(t) \geq 0 \text{ a.s.}, \quad t \in [0,T], \tag{3.53}$$

as required to get $\hat{\boldsymbol{\pi}} \in \mathcal{A}(\hat{u})$ (see (3.17)). We have thus established that (3.36) holds for some $\hat{\boldsymbol{\pi}} \in \mathcal{A}(\hat{u})$, and Theorem 3.3.10 follows. $\square$.

Remark 3.3.12. From Theorem 3.3.10, together with Remark 3.3.11, one sees that the price of the contingent claim B is the quantity $\hat{u}$ defined at (3.34), that is $\hat{u}$ is the *least initial wealth* from which the contingent claim can be hedged. We also note that the corresponding *hedging portfolio* is the process $\hat{\boldsymbol{\pi}} \in \Pi$ defined at (3.45). That is,

$$\hat{\boldsymbol{\pi}}(t) \triangleq [\boldsymbol{\sigma}^{-1}(t)]^\top \left[\frac{\mathbf{\Psi}_0(t)}{H_0(t)} + \hat{X}(t)\boldsymbol{\theta}(t) \right], \quad t \in [0,T], \tag{3.54}$$

in which the process $\mathbf{\Psi}_0$ is the $\mathbb{R}^N$-valued $\{\mathcal{F}_t\}$-progressively measurable integrand process obtained from application of the martingale representation theorem at (3.41). This integrand process is typically obtained from the *Clark-Ocone formula* of Malliavin calculus.

Example 3.3.13. In this example we will look at a very common hedging problem, namely the celebrated Black-Scholes problem. In this problem there is just one risky asset with price $S(t)$ (that is $N = 1$ at Conditions 3.41 and 3.1.3) and the market coefficient processes are assumed constant. That means that the price process are given by the relations

$$dS_0(t) = rS_0(t)dt, \qquad dS(t) = S(t)(bdt + \sigma dW(t)), \tag{3.55}$$

(c.f. (3.2)) in which W is a *scalar* Brownian motion and r, b and σ are real-valued constants. Furthermore, it is assumed that the contingent claim B is given by a deterministic function of the price of the stock at close of trade $t = T$, i.e. $B = \phi(S(T))$. An example of this type of contingent claim is a *European call option* for which ϕ has the form

$$\phi(S(T)) = (S(T) - q)^+, \qquad q \in [0, \infty), \tag{3.56}$$

in which, as usual,

$$(x)^+ \triangleq \max\{0, x\}, \qquad x \in \mathbb{R}, \tag{3.57}$$

and q is a real constant called the *strike price* of the option. We would like to determine the least initial wealth $\hat{u}$ and the optimal portfolio process $\hat{\pi}$ that guarantees

$$X^{(\hat{u},\hat{\pi})}(T) \geq \phi(S(T)) \quad \text{a.s.} \tag{3.58}$$

Suppose that $u(s, x)$ satisfies the partial differential equation

$$\frac{\partial u}{\partial s}(s, x) = \frac{\sigma^2 x^2}{2} \frac{\partial^2 u}{\partial x^2}(s, x), \quad 0 \leq s \leq T, \quad x \in \mathbb{R}, \tag{3.59}$$

with boundary condition

$$u(0, x) = \phi(x), \quad x \in \mathbb{R}. \tag{3.60}$$

Put

$$\hat{u} \triangleq u(T, S(0)), \tag{3.61}$$

and

$$\hat{\pi}(t) \triangleq S(t) \frac{\partial u}{\partial x}(T - t, S(t)), \quad 0 \leq t \leq T. \tag{3.62}$$

Although we shall not give the details here, one can use Theorem 3.3.10, together with the Clark-Ocone formula, to see that the price of the European call option B is given by (3.61) and the hedging portfolio is given by (3.62). Notice that these quantities are directly computable once the partial differential equation (3.59) is solved with the initial condition (3.60).

Remark 3.3.14. One can also establish the results of Example 3.3.13 quite easily by direct application of Ito's formula, without recourse to Theorem 3.3.10 and the Clark-Ocone formula, because of the simplicity of the market coefficient processes (which are constant). However, when the market coefficient processes are genuinely random then one must use the full strength of Theorem 3.3.10. This remark is important because, in the next chapter, we are going to deal with market coefficient processes which are "more random" than the market coefficients stipulated at Condition 3.2.

Chapter 4

The Unconstrained Hedging Problem in a Regime-Switching Market Model

In this chapter we address the first significant research problem of this thesis, an extension of the almost sure hedging problem to the so-called *regime-switching* financial market model. The primary problem is still the same as in Chapter 3, that is we would like to find the price of a contingent claim together with a portfolio process that promises the seller almost surely covers his/her obligation to pay off this claim at close of trade. However, in the present chapter, the market coefficient processes are determined not only by an underlying Brownian motion (as was the case for the market coefficient processes stipulated at Conditions 3.1.1 and 3.1.3 in the preceding chapter), but also by a finite-state Markov chain. As explained in more detail later in the chapter, the dependency on a finite-state Markov chain is introduced in order to model occasional but large-scale changes (or "regime-switches") in market conditions. Although the approach in this chapter follows the same basic outline summarized in Chapter 3, it will quickly become clear that the inclusion of regime-switching in the market model makes the problem significantly more challenging. The structure of the present chapter will be very similar to that of Chapter 3 so the reader can easily compare the methods of solution and see the additional aspects and challenges which arise from a regime-switching model.

4.1 Market Model

The model that will be used in this chapter is a continuous-time market model over a time horizon $[0, T]$ for a constant $T \in (0, \infty)$. The market model comprises the basic elements in Conditions 4.1.1 and 4.1.4, which should be compared with Conditions 3.1.1 and 3.1.3 respectively.

Condition 4.1.1. *A standard N-dimensional Brownian motion $\{\mathbf{W}(t); t \in [0,T]\}$ with scalar entries $W_n(t)$, where $n = 1, \ldots, N$, and a time-homogeneous continuous-time Markov chain $\{\boldsymbol{\alpha}(t); t \in [0,T]\}$ with càdlàg paths in a finite state space $S = \{1, \ldots, D\}$ and non-random initial state $\boldsymbol{\alpha}(0) \triangleq i_0 \in S$, are given on the common complete probability space $(\Omega, \mathcal{F}, P)$ and are assumed independent.*

The Markov chain $\boldsymbol{\alpha}(\cdot)$ has a $D \times D$ generator matrix $G \triangleq [g_{ij}]^D$, namely,

$$g_{ii} \triangleq -\sum_{i \neq j} g_{ij}, \quad g_{ij} \geq 0 \qquad i, j \in S, i \neq j, \tag{4.1}$$

so that $P_t \triangleq e^{tG}$ is the Markov transition matrix of $\boldsymbol{\alpha}(\cdot)$. With $\mathcal{N}(P) \triangleq \{A \in \mathcal{F} : P(A) = 0\}$, define the augmented filtrations $\{\mathcal{F}_t^{\boldsymbol{\alpha}}\}$ and $\{\mathcal{F}_t^{\boldsymbol{W}}\}$ as

$$\{\mathcal{F}_t^{\boldsymbol{\alpha}}\} \triangleq \sigma\{\boldsymbol{\alpha}(s), s \in [0,t]\} \vee \mathcal{N}(P), \qquad \{\mathcal{F}_t^{\boldsymbol{W}}\} \triangleq \sigma\{\mathbf{W}(s), s \in [0,t]\} \vee \mathcal{N}(P) \tag{4.2}$$

*for all $t \in [0,T]$. Define the **joint filtration** to be*

$$\{\mathcal{F}_t\} \triangleq \{\mathcal{F}_t^{\boldsymbol{\alpha}}\} \vee \{\mathcal{F}_t^{\boldsymbol{W}}\} \tag{4.3}$$

for all $t \in [0,T]$. In particular, $\{\mathcal{F}_t\}$ represents the information available to investors.

Remark 4.1.2. For later reference we note here that the filtrations $\{\mathcal{F}_t^{\boldsymbol{\alpha}}\}$ and $\{\mathcal{F}_t^{\mathbf{W}}\}$ defined at (4.2) are *right continuous* (see Definition C.1.4). To see this observe that the Markov chain $\boldsymbol{\alpha}(\cdot)$ is a *Feller process* (see, e.g., p. 31 of Rogers and Williams [43]) with values in S, and since the augmented filtration of any Feller process is right-continuous (see Proposition III(2.10) of Revuz and Yor [39]), one sees that $\{\mathcal{F}_t^{\boldsymbol{\alpha}}\}$ is right continuous. Similarly, since the Brownian motion $\mathbf{W}(\cdot)$ is also a Feller process, it follows that the filtration $\{\mathcal{F}_t^{\mathbf{W}}\}$ is right continuous. Furthermore, the filtration $\{\mathcal{F}_t\}$ defined at (4.3) is also right continuous. In fact, since the Markov processes $\boldsymbol{\alpha}(\cdot)$ and $\mathbf{W}(\cdot)$ are *independent* Feller processes with values in S and $\mathbb{R}^N$ respectively, it follows from Kallenberg [28] (Chapter 19, Exercise 10, p. 389) that the joint process $(\boldsymbol{\alpha}(\cdot), \mathbf{W}(\cdot))$ with values in $S \times \mathbb{R}^N$ is also a Feller process. That $\{\mathcal{F}_t\}$ is right continuous again follows from Proposition III(2.10) of Revuz and Yor [39]. The right continuity of $\{\mathcal{F}_t\}$ will play an essential technical role later in the chapter.

Notation 4.1.3. Exactly as at Notation 3.1.2, the notation $\mathcal{F}^*$ (respectively, $\mathcal{P}^*$) is used to denote the $\{\mathcal{F}_t\}$-progressively measurable (respectively, $\{\mathcal{F}_t\}$-predictable (see Definition C.1.6)) σ-algebra on $[0,T] \times \Omega$. The notation $\eta \in \mathcal{F}^*$ (respectively, $\eta \in \mathcal{P}^*$) is used to indicate that the process $\eta : [0,T]\times\Omega \to \mathbb{R}^N$ is $\{\mathcal{F}_t\}$-progressively measurable (respectively, $\{\mathcal{F}_t\}$-predictable). Of course, here the filtration $\{\mathcal{F}_t\}$ is defined by (4.3), and not by (3.1). The qualifier "a.s." always refers to the probability P on $(\Omega, \mathcal{F})$, while the qualifier "a.e." refers to the product measure $\lambda\otimes P$ on $\mathcal{B}([0,T])\times\mathcal{F}$, where λ denotes the Lebesgue measure on the Borel σ-algebra of $[0,T]$, denoted $\mathcal{B}([0,T])$.

Condition 4.1.4 which follows is formally identical to Condition 3.1.3, except that the underlying filtration is now defined at (4.3) to include the regime-switching Markov chain (rather than the filtration defined at (3.1)). For completeness we repeat the condition in full:

Condition 4.1.4. *The market comprises a single* ***risk-free asset*** *with price* $\{S_0(t); t \in [0,T]\}$ *and several* ***risky assets*** *with prices* $\{S_n(t); t \in [0,T]\}$, $n = 1,\dots,N$, *modeled by the relations*

$$dS_0(t) = r(t)S_0(t)dt, \qquad dS_n(t) = S_n(t)(b_n(t)dt + \sum_{m=1}^{N} \sigma_{nm}(t)dW_m(t)). \tag{4.4}$$

with $S_0(0) \triangleq 1$, *and* $S_n(0)$, $n = 1,\dots,N$, *being given, strictly positive constants. Additionally, assume that* $S_0(T) \geq s_0$ *a.s. for some constant* $s_0 > 0$. *The* ***risk-free interest rate*** $\{r(t)\}$, *the entries* $\{b_n(t)\}$ *of the* $\mathbb{R}^N$*-valued* ***rate of return*** $\{\boldsymbol{b}(t)\}$, *and the entries* $\{\sigma_{nm}(t)\}$ *of the* $N\times N$ *matrix-valued* ***volatility process*** $\{\boldsymbol{\sigma}(t)\}$ *are given uniformly bounded and* $\{\mathcal{F}_t\}$*-progressively measurable scalar processes on* $[0,T]\times\Omega$. *Using* $||\boldsymbol{z}||$ *for the Euclidean norm and* $\boldsymbol{z}^\top$ *for the transpose of a vector* $\boldsymbol{z} \in \mathbb{R}^N$, *we shall suppose that there exists a constant* $\kappa \in (0,\infty)$ *such that* $\boldsymbol{z}^\top\boldsymbol{\sigma}(t,\omega)\boldsymbol{\sigma}(t,\omega)^\top\boldsymbol{z} \geq \kappa||\boldsymbol{z}||^2$ *for all* $(\boldsymbol{z},t,\omega) \in \mathbb{R}^N\times[0,T]\times\Omega$.

Remark 4.1.5. From now on Conditions 4.1.1 and 4.1.4 will be assumed without explicit mention in all subsequent theorems, lemmas and propositions in the present chapter.

Remark 4.1.6. In view of the constant $\kappa \in (0,\infty)$ postulated in Condition 4.1.4 we get from Xu and Shreve (see (2.4) and (2.5) on p. 90 of Karatzas and Shreve [31]) the existence of a constant $\kappa_1 \in (1,\infty)$ such that

$$\begin{aligned}&\frac{1}{\kappa_1}\max\{||(\boldsymbol{\sigma}(t,\omega))^{-1}\boldsymbol{z}||, ||(\boldsymbol{\sigma}(t,\omega)^\top)^{-1}\boldsymbol{z}||\}\\ &\leq ||\boldsymbol{z}|| \leq \kappa_1\min\{(||(\boldsymbol{\sigma}(t,\omega))^{-1}\boldsymbol{z}||, ||(\boldsymbol{\sigma}(t,\omega)^\top)^{-1}\boldsymbol{z}||\}\end{aligned} \tag{4.5}$$

for all $(\boldsymbol{z},t,\omega) \in \mathbb{R}^N\times[0,T]\times\Omega$. This bound result will be used extensively.

Remark 4.1.7. In view of the independence of $\mathbf{W}$ and $\boldsymbol{\alpha}$ from Condition 4.1.1, the process $\{\mathbf{W}(t), t \in [0,T]\}$ is still a Brownian motion with respect to the joint filtration $\{\mathcal{F}_t\}$. As a result, the risky asset SDE in (4.4) makes sense and has the same interpretation as the SDE in (3.2).

Notation 4.1.8. Exactly as at Notation 3.1.7, define the $\mathbb{R}^N$- valued *market price of risk* $\{\boldsymbol{\theta}(t),\ t \in [0,T]\}$ by

$$\boldsymbol{\theta}(t) \triangleq \boldsymbol{\sigma}^{-1}(t)(\boldsymbol{b}(t) - r(t)\mathbf{1}) \tag{4.6}$$

for all $t \in [0,T]$ in which $\mathbf{1}$ is the N-dimensional vector of unit entries. From Condition 4.1.4 and Remark 4.1.6, the process $\boldsymbol{\theta}$ is $\{\mathcal{F}_t\}$-progressively measurable and uniformly bounded, i.e., $\kappa_{\boldsymbol{\theta}} \triangleq \sup_{(t,\omega)} ||\boldsymbol{\theta}(t,\omega)|| < \infty$.

Throughout this chapter an essential role will be played by the canonical processes associated with the Markov chain $\boldsymbol{\alpha}(\cdot)$ which are defined as follows:

Definition 4.1.9. For $i, j \in S$, $i \neq j$, put

$$R_{ij}(t) \triangleq \sum_{0<s\le t} I\{\boldsymbol{\alpha}(s-) = i\} I\{\boldsymbol{\alpha}(s) = j\}, \qquad \tilde{R}_{ij}(t) \triangleq \int_0^t g_{ij} I\{\boldsymbol{\alpha}(s) = i\} ds, \tag{4.7}$$

$$M_{ij}(t) \triangleq R_{ij}(t) - \tilde{R}_{ij}(t), \qquad t \in [0,T], \tag{4.8}$$

in which I is the indicator function given by

$$I\{\boldsymbol{\alpha}(s) = i\} \triangleq \begin{cases} 1 & \text{if } \boldsymbol{\alpha}(s) = i \\ 0 & \text{otherwise} \end{cases} \tag{4.9}$$

for each $i \in S$. For notational convenience we put $R_{ii} = \tilde{R}_{ii} = M_{ii} \triangleq 0$ for all $i \in S$. The processes $R_{ij}, \tilde{R}_{ij}, M_{ij}$, for $i, j \in S$, are called the *canonical processes* of the regime switching Markov chain $\boldsymbol{\alpha}(\cdot)$.

We next identify, for later reference, some properties of the processes formulated in Definition 4.1.9:

Remark 4.1.10. From (4.7) it is clear that $R_{ij}(t,\omega)$ counts the number of jumps from state i to state j by the Markov chain path $\boldsymbol{\alpha}(\cdot,\omega)$ over the interval $(0,t]$, from which it follows that

$$\triangle R_{ij}(t) \triangle R_{i_1 j_1}(t) = 0, \qquad t \in [0,T],$$

when $(i,j) \neq (i_1, j_1)$ (here $\triangle R_{ij}(t) \triangleq R_{ij}(t) - R_{ij}(t-)$, $t \in (0,T]$ is the discontinuous change, or jump, in the process R_{ij} at the instant t, see (C.2)). Now R_{ij} and $\tilde{R}_{ij}$ are non-decreasing $\{\mathcal{F}_t^{\boldsymbol{\alpha}}\}$-adapted processes (respectively, càdlàg and continuous), so that

$$M_{ij} \text{ is an } \{\mathcal{F}_t^{\boldsymbol{\alpha}}\}\text{-adapted càdlàg process with finite-variation paths.} \tag{4.10}$$

Furthermore, from IV(21.5) - IV(21.7) of Rogers and Williams [43] we have the following essential property of the processes M_{ij}:

Lemma 4.1.11. *M_{ij} is an $\{\mathcal{F}_t^{\boldsymbol{\alpha}}\}$-local martingale for all $i, j \in S, i \neq j$.*

From IV(21.11) of Rogers and Williams [43] we have the elementary bound

$$\gamma_{ij} \triangleq E\left[\exp(R_{ij}(T))\right] < \infty, \quad i, j \in S, i \neq j. \tag{4.11}$$

With this bound, together with the independence of $\boldsymbol{\alpha}(\cdot)$ and $\mathbf{W}(\cdot)$, we immediately strengthen Lemma 4.1.11 as follows:

Lemma 4.1.12. *M_{ij} is an $\{\mathcal{F}_t\}$-square integrable martingale for all $i, j \in S, i \neq j$.*

Furthermore, from (4.10), Lemma 4.1.12 and Lemma C.7.3 we have

Lemma 4.1.13. *M_{ij} is an $\{\mathcal{F}_t\}$-square integrable and purely discontinuous martingale for all $i, j \in S, i \neq j$.*

From Lemma 4.1.13 and the decomposition Theorem C.8.3 we then get

$$M_{ij}^c(t) = 0, \qquad M_{ij}(t) = M_{ij}^d(t), \qquad t \in [0, T], \tag{4.12}$$

for all $i, j \in S, i \neq j$. Due to the essential results of Lemmas 4.1.11-4.1.13, we call family of process M_{ij}, for $i, j \in S$, the *canonical martingales* of the Markov chain $\boldsymbol{\alpha}(\cdot)$. From (4.7) and (4.8) together with Theorem C.8.9 and the fact that $\triangle M_{ij}(t) = \triangle R_{ij}(t)$, $t \in [0, T]$ by Definition 4.1.9, we have that the *quadratic co-variation process* of the canonical martingales M_{ij} (see Definition C.8.8) are given by

$$[M_{ij}, M_{ij}](t) = R_{ij}(t) \text{ a.s.}, \quad t \in [0, T], \quad i, j \in S, i \neq j. \tag{4.13}$$

The following two Lemmas give other important results regarding the quadratic co-variation processes of the canonical martingales. These results are proved in Appendix B.1.

Lemma 4.1.14.

$$[M_{ij}, W_k](t) = 0 \ a.s., \quad t \in [0, T], \tag{4.14}$$

for $i, j \in S$ and $k = 1, \ldots, N$, where $W_k(t)$ is the k-th element of the Brownian motion $\mathbf{W}(t)$.

Lemma 4.1.15.

$$[M_{ij}, M_{i_1 j_1}](t) = 0 \ a.s., \quad t \in [0, T], \tag{4.15}$$

for $i, j, i_1, j_1 \in S$ and $(i, j) \neq (i_1, j_1)$.

Notation 4.1.16. For $i, j \in S$, $i \neq j$, define the measure $\nu_{[M_{ij}]}$ on the measurable space $([0,T] \times \Omega, \mathcal{P}^*)$ as

$$\nu_{[M_{ij}]}[A] \triangleq E\left[\int_0^T I_A(t,\omega) d[M_{ij}](t)\right] = E\left[\int_0^T I_A(t,\omega) dR_{ij}(t)\right], \quad A \in \mathcal{P}^*. \tag{4.16}$$

This is the *Doléans measure* of M_{ij}, which will be needed later in this chapter.

Before defining the problem of unconstrained hedging within a regime-switching market, we make some preliminary comments in the following remarks and text on how an agent trades in the market defined in Conditions 4.1.1 and 4.1.4, that is in a market which incorporates a regime-switching Markov chain in the information filtration:

Remark 4.1.17. To model the investment we proceed much as we did in Remark 3.1.8, that is we introduce the process $\{\pi_0(t), t \in [0,T]\}$, which denotes the *monetary (or dollar)* amount allocated by the agent to the risk-free asset. We shall always assume that this process is $\{\mathcal{F}_t\}$-progressively measurable, that is $\pi_0 \in \mathcal{F}^*$. To model investment in the N-risky assets, we define (exactly as in Remark 3.1.8) the following space Π of *portfolio processes*:

$$\Pi \triangleq \left\{\boldsymbol{\pi} : [0,T] \times \Omega \mapsto \mathbb{R}^N \middle| \boldsymbol{\pi} \in \mathcal{F}^*, \int_0^T ||\boldsymbol{\pi}(t)||^2 dt < \infty \text{ a.s.}\right\}. \tag{4.17}$$

Again, as in Remark 3.1.8, the interpretation of the elements of $\boldsymbol{\pi}$ of Π is as follows: Each scalar component $\pi_n(t)$, for $n = 1, \ldots, N$, of the $\mathbb{R}^N$-vector $\boldsymbol{\pi}(t)$, for $t \in [0,T]$, indicates the monetary (or dollar) amount allocated by the agent to the risky asset with price $S_n(t)$ at the instant $t \in [0,T]$. It then follows that the *total wealth* of the agent at any instant $t \in [0,T]$ is the sum of the amounts in the risk-free and risky assets at instant t. That is the total wealth of the agent is given by the process X defined as follows:

$$X(t) \triangleq \sum_{n=0}^{N} \pi_n(t) = \boldsymbol{\pi}(t)^\top \mathbf{1} + \pi_0(t), \qquad t \in [0,T], \tag{4.18}$$

in which $\mathbf{1}$ is the N-dimensional vector with unit entries.

To address the hedging problem with regime-switching it will prove necessary to generalize, or extend, the sense in which an agent trades in the market. In fact, it will be assumed that, at each and every instant t in the trading interval $[0,T]$, the agent (or investor) not only allocates his or her *total wealth* among the risk-free asset (with price $S_0(t)$) and the risky assets (with prices $S_n(t), n = 1, 2, \ldots, N$), but is also free to *consume* wealth through a *cumulative consumption process* which, by definition, is a member of the set of processes $\mathcal{C}$ formulated as follows:

$$\mathcal{C} \triangleq \Big\{ c : [0,T] \times \Omega \mapsto [0,\infty) \Big| c(0) = 0, c \in \mathcal{F}^*, c(T) < \infty \text{ a.s.}, \\ c(\cdot) \text{ càdlàg, non-decreasing} \Big\}. \tag{4.19}$$

A cumulative consumption process $c(t)$ indicates the *total* amount of money expended by the agent during the period $[0,t]$.

Remark 4.1.18. In this chapter, the agent does not quite follow a self funded trading strategy as in Chapter 3 (see Remark 3.1.9), for now the agent can also consume wealth through the cumulative consumption process $c \in \mathcal{C}$. However, there is still no infusion of cash to the investor from an external source, and any change in the agent's total wealth is caused by a change in the asset price less the consumption.

Remark 4.1.19. The cumulative consumption process $c \in \mathcal{C}$ is part of a broader class of processes called *cumulative income processes* (see Definition 1.3.1 in Karatzas and Shreve [31]). A cumulative income process $\Gamma(t)$ is interpreted as the cumulative wealth received by an agent on the time interval $[0,t]$. In particular, the agent is given initial wealth $\Gamma(0)$, and consumption by the agent can be captured by a decrease in $\Gamma(\cdot)$. An agent's portfolio $(\pi_0, \boldsymbol{\pi})$ is called $\Gamma(\cdot)$-*financed* if the agent's wealth is given by (4.18).

As in Chapter 3 it will be assumed that the agent begins trading with *non-negative initial wealth*, that is

$$x \in [0,\infty). \tag{4.20}$$

It is a standard result in mathematical finance (see §3.3 in Karatzas and Shreve [31]) that, with consumption included, the wealth process $\{X(t), t \in [0,T]\}$ of the agent necessarily satisfies the following stochastic integral equation:

$$X(t) = x + \int_0^t \Big\{ r(s)X(s) + \boldsymbol{\pi}(s)^\top \boldsymbol{\sigma}(s)\boldsymbol{\theta}(s) \Big\} ds \\ + \int_0^t \boldsymbol{\pi}(s)^\top \boldsymbol{\sigma}(s) d\mathbf{W}(s) - c(t), \quad t \in [0,T], \tag{4.21}$$

in which of course x is the initial wealth of the agent, $\boldsymbol{\pi} \in \Pi$ is the portfolio process according to which the wealth of the agent is allocated among the N *risky* assets, and $c \in \mathcal{C}$ is the cumulative consumption process. Similar to Remark 3.1.8, the process $\{\pi_0(t)\}$ does not appear anywhere in (4.21), and only the portfolio process $\boldsymbol{\pi} \in \Pi$ and consumption process $c \in \mathcal{C}$ appear in the equation. The reason for this, as a detailed derivation of (4.21) would make clear, is that any change in wealth is only caused by changes in asset prices and consumption (in a linear fashion). As a result, $\pi_0(t)$ is a redundant variable which can be eliminated using the relation (4.18).

Remark 4.1.20. From the stochastic integral equation (4.21), it is clear that the wealth process X is *completely determined* by the initial wealth x (which is effectively the "initial value") and both the portfolio process $\boldsymbol{\pi} \in \Pi$ and cumulative consumption process $c \in \mathcal{C}$ (which are effectively "control inputs" decided upon by the agent). From now on we shall denote the wealth process of the agent not by X, but instead by $X^{(x,c,\boldsymbol{\pi})}$, to make clear the dependence of the wealth process on the initial wealth x, the portfolio process $\boldsymbol{\pi} \in \Pi$ and cumulative consumption process $c \in \mathcal{C}$, so that (4.21) will be written as follows:

$$\begin{aligned} X^{(x,c,\boldsymbol{\pi})}(t) = x &+ \int_0^t \Big\{ r(s)X^{(x,c,\boldsymbol{\pi})}(s) + \boldsymbol{\pi}(s)^\top \boldsymbol{\sigma}(s)\boldsymbol{\theta}(s) \Big\} ds \\ &+ \int_0^t \boldsymbol{\pi}(s)^\top \boldsymbol{\sigma}(s) d\mathbf{W}(s) - c(t), \quad t \in [0,T], \end{aligned} \tag{4.22}$$

or, in the more usual notation of stochastic differential equations, as

$$dX^{(x,c,\boldsymbol{\pi})}(t) = \{r(t)X^{(x,c,\boldsymbol{\pi})}(t) + \boldsymbol{\pi}(t)^\top \boldsymbol{\sigma}(t)\boldsymbol{\theta}(t)\}dt + \boldsymbol{\pi}(t)^\top \boldsymbol{\sigma}(t) d\mathbf{W}(t) - dc(t), \tag{4.23}$$

for all $t \in [0,T]$, with the initial condition

$$X^{(x,c,\boldsymbol{\pi})}(0) = x. \tag{4.24}$$

It then follows easily from Ito's formula that the wealth process is given explicitly in terms of $x \in [0,\infty)$, $\boldsymbol{\pi} \in \Pi$ and $c \in \mathcal{C}$ by the expression

$$\begin{aligned} X^{(x,c,\boldsymbol{\pi})}(t) = xS_0(t) &+ S_0(t)\bigg\{ \int_0^t \frac{1}{S_0(\tau)} \boldsymbol{\pi}(\tau)^\top \boldsymbol{\sigma}(\tau)\boldsymbol{\theta}(\tau) d\tau \bigg\} \\ &+ S_0(t)\bigg\{ \int_0^t \frac{1}{S_0(\tau)} \boldsymbol{\pi}(\tau)^\top \boldsymbol{\sigma}(\tau) d\mathbf{W}(\tau) \bigg\} - \int_0^t \frac{1}{S_0(\tau)} dc(\tau) \end{aligned} \tag{4.25}$$

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

Remark 4.1.21. When we take consumption $c \in \mathcal{C}$ to be the *zero process* (that is $c(t) \triangleq 0$, $t \in [0,T]$) at (4.25) we get

$$\begin{aligned} X^{(x,0,\boldsymbol{\pi})}(t) = xS_0(t) &+ S_0(t)\bigg\{ \int_0^t \frac{1}{S_0(\tau)} \boldsymbol{\pi}(\tau)^\top \boldsymbol{\sigma}(\tau)\boldsymbol{\theta}(\tau) d\tau \bigg\} \\ &+ S_0(t)\bigg\{ \int_0^t \frac{1}{S_0(\tau)} \boldsymbol{\pi}(\tau)^\top \boldsymbol{\sigma}(\tau) d\mathbf{W}(\tau) \bigg\}, \end{aligned} \tag{4.26}$$

for all $t \in [0,T]$, which is identical to the right hand side of (3.14). For this reason we shall put

$$X^{(x,\boldsymbol{\pi})}(t) \triangleq X^{(x,0,\boldsymbol{\pi})}(t), \qquad t \in [0,T], \tag{4.27}$$

for all $(x,\boldsymbol{\pi}) \in [0,\infty) \times \Pi$, in order to have notational consistency with the no-consumption wealth process used throughout Chapter 3. We also see from (4.26) and (4.25) that

$$X^{(x,c,\boldsymbol{\pi})}(t) = X^{(x,0,\boldsymbol{\pi})}(t) - \int_0^t \frac{1}{S_0(\tau)} dc(\tau), \qquad t \in [0,T], \tag{4.28}$$

for all $(x,c,\boldsymbol{\pi}) \in [0,\infty) \times \mathcal{C} \times \Pi$.

4.2 Definition of the Hedging Problem

In the present section we shall define the problem of unconstrained hedging within a regime-switching market model with reference to the preliminaries outlined in the previous section. The problem that we address in this chapter is formally very similar to the problem addressed in Chapter 3. Indeed the only visible difference between the two problems is the inclusion of a regime-switching Markov chain $\boldsymbol{\alpha}(\cdot)$ in the basic filtration (compare the filtrations (3.1) in Chapter 3 and (4.3) in the present chapter). Nevertheless, we shall see that this seemingly small difference makes the hedging problem substantially more challenging than the hedging problem of Chapter 3 (without regime-switching). In particular, it will be necessary to include a *cumulative consumption process* in the formulation of the hedging problem (something which was not necessary for the problem in Chapter 3). We shall discuss the essential role of the consumption process later in this chapter.

Exactly as in Chapter 3, a *contingent claim* is a specified strictly positive $\mathcal{F}_T$-measurable random variable B, that is

$$B : (\Omega, \mathcal{F}, P) \mapsto (0,\infty), \qquad B \in \mathcal{F}_T. \tag{4.29}$$

Since B is $\mathcal{F}_T$-measurable, it incorporates all of the "randomness" in the market over the entire trading interval $t \in [0,T]$. That is, one sees a realization $B(\omega)$ of this random variable only at the close of the trade $t = T$. In the present case, the randomness of B is contributed not only by the Brownian motion process $\mathbf{W}(t)$ (as was the case in Chapter 3) but also by the finite-state Markov chain $\boldsymbol{\alpha}(t)$. It is this additional randomness from the Markov chain which adds considerably to the challenge posed by the hedging problem. The random variable B defines an obligation on the part of the agent in the following sense: starting from some initial wealth $x \in [0, \infty)$, the agent must come up with some portfolio process $\boldsymbol{\pi} \in \Pi$ and cumulative consumption process $c \in \mathcal{C}$ such that

$$X^{(x,c,\boldsymbol{\pi})}(T) = B \text{ a.s.} \tag{4.30}$$

That is, the agent starts with initial wealth $x \in [0, \infty)$ and must invest in the risky assets and consume in such a way that sufficient wealth is generated at close of trade $t = T$ for the agent to be able to pay off the contingent claim B with probability one. For both technical and financial reasons, we restrict attention to portfolio process $\boldsymbol{\pi} \in \Pi$ and cumulative consumption process $c \in \mathcal{C}$ for which the corresponding wealth process is *non-negative over the entire trading interval*. That is, for a given initial wealth $x \in [0, \infty)$, the agent is restricted to portfolio processes $\boldsymbol{\pi} \in \Pi$ and cumulative consumption processes $c \in \mathcal{C}$ in the set $\mathcal{A}(x)$ defined as follows:

$$\mathcal{A}(x) \triangleq \left\{ (c, \boldsymbol{\pi}) \in \mathcal{C} \times \Pi \middle| X^{(x,c,\boldsymbol{\pi})}(t) \geq 0 \text{ a.s. for all } t \in [0, T] \right\}. \tag{4.31}$$

The set $\mathcal{A}(x)$ is called the set of *admissible portfolio-consumption pairs*, and should be compared with the set $\mathcal{A}(x)$ at (3.17) in Chapter 3, in which there is only a portfolio process $\boldsymbol{\pi}$ without any consumption process c.

Remark 4.2.1. Observe that $\mathcal{A}(x) \neq \emptyset$ for each $x \in [0, \infty)$. To see this take $\boldsymbol{\pi} \in \Pi$ defined by $\boldsymbol{\pi} \triangleq 0$ and $c \in \mathcal{C}$ defined by $c \triangleq 0$. In accordance with Remark 4.1.17 this amounts to investing *zero wealth* in the risky assets (so the entire wealth is invested in the risk-free asset with the price S_0) with *no consumption*. With this $\boldsymbol{\pi} \in \Pi$ and $c \in \mathcal{C}$ one sees that the last two terms on the right side of (4.22) are zero so that

$$X^{(x,c,\boldsymbol{\pi})}(t) = S_0(t)x \geq 0, \qquad t \in [0, T], \tag{4.32}$$

and therefore, $(c, \boldsymbol{\pi}) \in \mathcal{A}(x)$. Thus $\mathcal{A}(x) \neq \emptyset$ for each $x \in [0, \infty)$.

Remark 4.2.2. In light of the preceding discussion, from now on we restrict ourselves to triples $(x, c, \boldsymbol{\pi})$ for which $x \in [0, \infty)$ and $(c, \boldsymbol{\pi}) \in \mathcal{A}(x)$, so that the wealth process $X^{(x,c,\boldsymbol{\pi})}$ is non-negative over the trading interval $t \in [0, T]$. Exactly as in Chapter 3 this is a reasonable requirement from the financial point of view, for it merely insists that the agent never go into debt during the trading interval (negative wealth amounts to debt). Again, from a mathematical view point, removal of the condition of non-negative wealth (essentially requiring $(c, \boldsymbol{\pi}) \in \mathcal{C} \times \Pi$ instead of $(c, \boldsymbol{\pi}) \in \mathcal{A}(x)$), which allows the agent to temporarily take on debt during the trading interval, leads to a rather intractable hedging problem, much as in the case where no regime-switching is present. As in Chapter 3 we shall take care to point out the essential role played by non-negativity of the wealth whenever this condition is used.

Remark 4.2.3. An agent who begins with initial wealth $x \in [0, \infty)$ and determines some $(c, \boldsymbol{\pi}) \in \mathcal{A}(x)$ such that (4.30) holds is said to *hedge the contingent claim* B from the initial wealth x, and the corresponding portfolio-consumption pair $(c, \boldsymbol{\pi}) \in \mathcal{A}(x)$ is called an *unconstrained hedging strategy* from the initial wealth x.

In light of the previous discussion of the hedging problem, it is clear that the initial wealth $x \in [0, \infty)$ is an essential parameter. Indeed, if the initial wealth is too small, then there may not even exist any $(c, \boldsymbol{\pi}) \in \mathcal{A}(x)$ such that (4.30) holds, that is there fails to be an unconstrained hedging strategy from the initial wealth x. One of our principal goals is to determine the *least possible* initial wealth for which there does exist a hedging strategy. To make this precise, we define the set Λ comprising of all initial wealths x for which a hedging strategy does exist, that is

$$\Lambda \triangleq \left\{ x \in [0, \infty) \middle| \exists (c, \boldsymbol{\pi}) \in \mathcal{A}(x) \quad \text{s.t. } X^{(x,c,\boldsymbol{\pi})}(T) = B \text{ a.s.} \right\}. \tag{4.33}$$

In view of (4.33) and (4.31) we have

$$\Lambda = \left\{ x \in [0, \infty) \middle| \; X^{(x,c,\boldsymbol{\pi})}(T) = B \text{ a.s. for some } (c, \boldsymbol{\pi}) \in \mathcal{C} \times \Pi \right. \\ \left. \text{such that } X^{(x,c,\boldsymbol{\pi})}(t) \geq 0 \text{ a.s. for all } t \in [0, T] \right\}. \tag{4.34}$$

Remark 4.2.4. There is of course no guarantee that Λ is non-empty. Indeed, as we shall see in Remark 4.3.12, if the contingent claim random variable is stipulated to be "unreasonably large" then Λ will be empty. However, assuming that Λ is indeed non-empty, and furthermore, attains its infimum at some $\hat{u} \in [0, \infty)$ (that is $\hat{u} \in \Lambda$ and $\hat{u} \leq x$ for all $x \in \Lambda$), it is evident that $x = \hat{u}$ is the least initial wealth from which one can hedge the contingent claim B. This least initial wealth is usually called the *price of the contingent claim* B. The most important goals of this chapter can now be stated as follows:

(a) Determine conditions on the contingent claim random variable B which ensures that Λ is non-empty

(b) With the conditions in (a) in force, establish that there exists some $\hat{u} \in \Lambda$ such that $\hat{u} \leq x$ for all $x \in \Lambda$. It then follows from (4.33) that $x = \hat{u}$ is the *least initial wealth* from which one can hedge the contingent claim B.

(c) With $x = \hat{u}$ characterize some $(c, \boldsymbol{\pi}) \in \mathcal{A}(x)$ such that (4.30) holds. This portfolio-consumption pair hedges the contingent claim B from the least initial wealth $\hat{u}$.

Remark 4.2.5. For the hedging problem addressed in Chapter 3 (i.e. defined by Conditions 3.1.1 and 3.1.3) the contingent claim B was hedged entirely with a portfolio process $\boldsymbol{\pi}$, without the use of any consumption process. With this in mind an obvious question is the following: why do we not likewise try to hedge the contingent claim in the present chapter (i.e. with Conditions 4.1.1 and 4.1.4 in force) entirely with a portfolio process, and why do we need to introduce consumption? We cannot answer this question completely until later in the chapter (see Remark 4.3.23), but at this point we can make the following observations: removal of consumption means taking $c \triangleq 0$ in (4.34), so that Λ is replaced with

$$\Lambda_1 = \left\{ x \in [0, \infty) \,\middle|\, X^{(x,0,\boldsymbol{\pi})}(T) = B \text{ a.s. for some } \boldsymbol{\pi} \in \Pi \right. \\ \left. \text{such that } X^{(x,0,\boldsymbol{\pi})}(t) \geq 0, \text{ a.s. for all } t \in [0, T] \right\}. \tag{4.35}$$

It turns out that the introduction of regime-switching in the form of Conditions 4.1.1 and 4.1.4 "complicates" the hedging problem in the sense that there is no guarantee that Λ_1 defined at (4.35) actually attains its infimum - there need not exist any $\hat{u} \in \Lambda_1$ such that $\hat{u} \leq x$ for all $x \in \Lambda_1$. On the other hand, as we shall see in the remainder of this chapter, the set Λ at (4.34), which allows for consumption, does attain its infimum. An alternative way in which to see the need for a consumption process when hedging the claim B subject to Conditions 4.1.1 and 4.1.4 is to define a variant of Λ_1 at (4.35) as follows:

$$\Lambda_2 = \left\{ x \in [0, \infty) \,\middle|\, X^{(x,0,\boldsymbol{\pi})}(T) \geq B \text{ a.s. for some } \boldsymbol{\pi} \in \Pi \right. \\ \left. \text{such that } X^{(x,0,\boldsymbol{\pi})}(t) \geq 0, \text{ a.s. for all } t \in [0, T] \right\}. \tag{4.36}$$

Notice the difference between Λ_1 and Λ_2: in the latter case we require only that the final wealth $X^{(x,0,\boldsymbol{\pi})}(T)$ almost surely *majorize* the contingent claim B (rather than be *equal* to B, as in the case of Λ_1), thus Λ_2 comprises all initial wealth x for which there exists some *admissible* $\boldsymbol{\pi} \in \Pi$ (meaning that $X^{(x,0,\boldsymbol{\pi})}(t) \geq 0$ a.s. for all $t \in [0, T]$) such that the final wealth $X^{(x,0,\boldsymbol{\pi})}(T)$ (*without* consumption) almost surely majorizes B. If Λ_2 does attain its infimal value at some $\hat{u} \in \Lambda_2$ then $\hat{u}$ is the least initial wealth such that one *super-hedges* the contingent claim B and the associated portfolio processes $\boldsymbol{\pi} \in \Pi$ is called a *super-hedging portfolio*. Since $X^{(\hat{u},0,\boldsymbol{\pi})}(T) \geq B$ for some admissible $\boldsymbol{\pi} \in \Pi$, there is an unavoidable non-negative "discrepancy" random variable $X^{(\hat{u},0,\boldsymbol{\pi})}(T) - B$. The consumption process $c \in \mathcal{C}$ is introduced to model this discrepancy random variable in the sense that the pair

$(c, \boldsymbol{\pi}) \in \mathcal{C} \times \Pi$ satisfies the relation

$$X^{(\hat{u},0,\boldsymbol{\pi})}(T) - B = \int_0^T \frac{1}{S_0(\tau)} dc(\tau) \text{ a.s.,} \tag{4.37}$$

so that from (4.37) and (4.28) we get

$$X^{(\hat{u},c,\boldsymbol{\pi})}(T) = B \text{ a.s.} \tag{4.38}$$

With this discussion in mind one sees that the hedging problem addressed in Chapter 3 is rather special, in the sense that, even when hedging with only an admissible portfolio process $\boldsymbol{\pi}$, the discrepancy random variable $X^{(\hat{u},0,\boldsymbol{\pi})}(T) - B$ is automatically equal to zero, so there is no need to introduce a compensating consumption process to model this discrepancy.

Remark 4.2.6. The need to introduce consumption when hedging contingent claims was first recognized by Cvitanic and Karatzas [10] and El Karoui and Quenez [16][17] for a hedging problem subject to Conditions 3.1.1 and 3.1.3 of Chapter 3, but with the further stipulation that the portfolio process $\boldsymbol{\pi}$ is constrained in the sense that, for some given closed convex constraint set $K \subset \mathbb{R}^N$, one requires that $\boldsymbol{\pi}(t, \omega) \in K$ a.e. They label such a market an *incomplete market*. For the hedging problem in the present chapter, the portfolio process $\boldsymbol{\pi}$ is of course unconstrained, but the introduction of regime-switching in the market model nevertheless again requires the use of a consumption process as regime-switching renders the market incomplete. We shall see later on that the construction of an appropriate consumption process is one of the main challenges in dealing with the hedging problem.

4.3 Solution to the Hedging Problem

Having outlined the problem of unconstrained hedging within a regime-switching market model (see Remark 4.2.4), we shall now address this problem. We begin by introducing the space of *Markov chain integrand processes* $\mathcal{H}$ which allows the regime-switching Markov chain to interact with the agent's wealth.

Notation 4.3.1. Recalling Definition A.1.1 we define the space of *Markov chain dual pro-*

cesses $\mathcal{H}$ as the set

$$\begin{aligned}\mathcal{H} \triangleq \Big\{\boldsymbol{\mu} = \{\mu_{ij}\}_{i,j\in S} \Big| & \mu_{ij} : [0,T] \times \Omega \to \mathbb{R}, \mu_{ii} = 0, \\ & \mu_{ij} \in L^{1/2}_{loc}(R_{ij}), i,j \in S, i \neq j, \\ & 1 + \triangle(\boldsymbol{\mu} \bullet \mathbf{M})(t) > 0 \text{ a.s.}, t \in [0,T]\Big\},\end{aligned} \tag{4.39}$$

where $L^{1/2}_{loc}(R_{ij})$ is given in Definition A.1.1 and

$$(\boldsymbol{\mu} \bullet \mathbf{M})(t) \triangleq \sum_{\substack{i,j\in S\\ i\neq j}} \int_0^t \mu_{ij}(s) dM_{ij}(s), \quad t \in [0,T], \tag{4.40}$$

is the stochastic integral of the process $\boldsymbol{\mu}(\cdot)$ with respect to the martingale $\mathbf{M}(\cdot) \triangleq \{M_{ij}(\cdot)\}_{i,j\in S}$, and is an $\{\mathcal{F}_t\}$-local martingale.

Remark 4.3.2. From Theorem C.9.3 we have that the jumps of the stochastic integral $(\boldsymbol{\mu} \bullet \mathbf{M})$ are given by

$$\triangle(\boldsymbol{\mu} \bullet \mathbf{M})(t) = \sum_{\substack{i,j\in S\\ i\neq j}} \mu_{ij}(t) \triangle M_{ij}(t), \quad t \in [0,T], \boldsymbol{\mu} \in \mathcal{H}, \tag{4.41}$$

and by Lemma C.7.3 that

$$\boldsymbol{\mu} \bullet \mathbf{M} \text{ is purely discontinuous for each } \boldsymbol{\mu} \in \mathcal{H}. \tag{4.42}$$

We then have by Theorem C.10.1, (4.40), (4.41) and (4.42) that the Doléans Dade exponential of the stochastic integral $(\boldsymbol{\mu} \bullet \mathbf{M})$ is given by

$$\begin{aligned}\mathcal{E}(\boldsymbol{\mu} \bullet \mathbf{M})(t) &= \exp\Big\{(\boldsymbol{\mu} \bullet \mathbf{M})(t)\Big\} \prod_{0<s\leq t} (1 + \triangle(\boldsymbol{\mu} \bullet \mathbf{M})(s)) \exp\Big\{ - \triangle(\boldsymbol{\mu} \bullet \mathbf{M})(s)\Big\} \\ &= \exp\Big\{ \sum_{\substack{i,j\in S\\ i\neq j}} \int_0^t \mu_{ij}(s) dM_{ij}(s)\Big\} \prod_{0<s\leq t} (1 + \sum_{\substack{i,j\in S\\ i\neq j}} \mu_{ij}(s) \triangle M_{ij}(s)) \exp\Big\{ - \sum_{\substack{i,j\in S\\ i\neq j}} \mu_{ij}(s) \triangle M_{ij}(s)\Big\}\end{aligned} \tag{4.43}$$

for all $t \in [0,T]$ and $\boldsymbol{\mu} \in \mathcal{H}$.

In Chapter 3 an essential role was played by the state price density process $H_0(t)$ introduced at Definition 3.3.1. In particular, it was seen at Proposition 3.3.3 that the process $H_0(t)X^{(x,\boldsymbol{\pi})}(t)$ is a supermartingale for every $\boldsymbol{\pi} \in \Pi$, and this supermartingale property was key to the whole approach to the hedging problem addressed in Chapter 3. We now introduce a generalization of the state price density process which will play an analogous role for the hedging problem with regime-switching:

Definition 4.3.3. The *generalized state price density process* $\{H_{\boldsymbol{\mu}}(t), t \in [0,T]\}$ for the market model defined by Conditions 4.1.1 and 4.1.4 is the process with values in $(0,\infty)$ defined as follows:

$$H_{\boldsymbol{\mu}}(t) \triangleq \exp\left\{ -\int_0^t r(s)ds \right\} \mathcal{E}(-\boldsymbol{\theta} \bullet \mathbf{W})(t)\mathcal{E}(\boldsymbol{\mu} \bullet \mathbf{M})(t), \quad t \in [0,T], \boldsymbol{\mu} \in \mathcal{H}, \tag{4.44}$$

where $(\boldsymbol{\theta} \bullet \mathbf{W})(t)$ is the stochastic integral of $\boldsymbol{\theta}(t)$ with respect to $\mathbf{W}(t)$ (see (3.22)) and $\mathcal{E}(\cdot)(t)$ is the Doléans-Dade exponential function (see Theorem C.10.1 and (4.43)).

Remark 4.3.4. The generalized state price density processes in Definition 4.3.3 is a special case of the generalized state price density introduced in Heunis [23] in order to address problems of *utility maximization* within a regime-switching market model basically identical to that formulated at Conditions 4.1.1 and 4.1.4. This generalized state price density process will also be key to the hedging problem in this chapter.

Remark 4.3.5. A simple although lengthy exercise using Ito's product rule allows us to write $H_{\boldsymbol{\mu}}(t)$ as

$$\begin{aligned} H_{\boldsymbol{\mu}}(t) = 1 - \int_0^t H_{\boldsymbol{\mu}}(s-)\boldsymbol{\theta}(s)^\top d\mathbf{W}(s) - \int_0^t H_{\boldsymbol{\mu}}(s-)r(s)ds \\ + \sum_{\substack{i,j\in S \\ i\neq j}} \int_0^t H_{\boldsymbol{\mu}}(s-)\mu_{ij}(s)dM_{ij}(s), \quad t \in [0,T], \boldsymbol{\mu} \in \mathcal{H}. \end{aligned} \tag{4.45}$$

Since $\mathbf{W} \in \mathcal{M}(\{\mathcal{F}_t\}, P)$ and $M_{ij} \in \mathcal{M}(\{\mathcal{F}_t\}, P)$ for each $i,j \in S$, we have that

$$H_{\boldsymbol{\mu}} \in \mathcal{SM}(\{\mathcal{F}_t\}, P), \tag{4.46}$$

that is $H_{\boldsymbol{\mu}}$ is an $\{\mathcal{F}_t\}$-semimartingale (recall Definition C.8.1). Also, observe from (4.39) and (3.25) that

$$\inf_{t\in[0,T]} H_{\boldsymbol{\mu}}(t) > 0 \text{ a.s.}, \quad \boldsymbol{\mu} \in \mathcal{H}. \tag{4.47}$$

Therefore, $H_{\boldsymbol{\mu}}$ is an almost surely strictly positive $\{\mathcal{F}_t\}$-semimartingale for each $\boldsymbol{\mu} \in \mathcal{H}$.

Remark 4.3.6. Note that if $\boldsymbol{\mu}(t) \triangleq 0$ for all $t \in [0,T]$ we have

$$H_{\boldsymbol{\mu}}(t) = H_0(t), \quad t \in [0,T], \tag{4.48}$$

where the process $\{H_0(t), t \in [0,T]\}$ is defined in Definition 3.3.1.

Remark 4.3.7. It is elementary to show that the process

$$z(t) \triangleq \int_0^t H_{\boldsymbol{\mu}}(s-)dc(s) + \sum_{0<s\leq t} H_{\boldsymbol{\mu}}(s-)\triangle c(s)\triangle(\boldsymbol{\mu} \bullet \mathbf{M})(s), \quad t \in [0,T], \tag{4.49}$$

is almost surely non-negative for all $t \in [0,T]$, $c \in \mathcal{C}$, and $\boldsymbol{\mu} \in \mathcal{H}$, since $c \in \mathcal{C}$ must be non-decreasing by (4.19), $\boldsymbol{\mu} \in \mathcal{H}$ must have the property $\triangle(\boldsymbol{\mu} \bullet \mathbf{M})(t) > -1$ a.s. for all $t \in [0,T]$ by (4.39), and $H_{\boldsymbol{\mu}}$ is almost surely strictly positive for all $\boldsymbol{\mu} \in \mathcal{H}$ by (4.47). This fact will be essential for the following results.

The significance of the state price density process $H_{\boldsymbol{\mu}}$ is apparent in the following proposition and corollary.

Proposition 4.3.8. *For each $x \in [0,\infty)$, $(c,\boldsymbol{\pi}) \in \mathcal{A}(x)$, and $\boldsymbol{\mu} \in \mathcal{H}$, the process $\{J_{\boldsymbol{\mu}}^{(x,c,\boldsymbol{\pi})}(t), t \in [0,T]\}$ defined as*

$$\begin{aligned} J_{\boldsymbol{\mu}}^{(x,c,\boldsymbol{\pi})}(t) \triangleq H_{\boldsymbol{\mu}}(t)X^{(x,c,\boldsymbol{\pi})}(t) &+ \int_0^t H_{\boldsymbol{\mu}}(s-)dc(s) \\ &+ \sum_{0<s\leq t} H_{\boldsymbol{\mu}}(s-)\triangle c(s)\triangle(\boldsymbol{\mu} \bullet \mathbf{M})(s), \quad t \in [0,T], \end{aligned} \tag{4.50}$$

is a non-negative $\{\mathcal{F}_t\}$-supermartingale (i.e. $J_{\boldsymbol{\mu}}^{(x,c,\boldsymbol{\pi})} \in \mathcal{SPM}(\{\mathcal{F}_t\}, P)$ and $J_{\boldsymbol{\mu}}^{(x,c,\boldsymbol{\pi})} \geq 0$).

Remark 4.3.9. Proposition 4.3.8 should be compared with Proposition 3.3.3. In fact, if one takes $c = 0$ at (4.50), that is zero consumption, then the second and third terms on the right side vanish and (4.50) reduces to (3.26).

Proof of Proposition 4.3.8. For ease of notation, set $X(t) \triangleq X^{(x,c,\boldsymbol{\pi})}(t)$ for all $t \in [0,T]$. A lengthy but straightforward calculation using Ito's product formula with (4.22) and (4.45) establishes that

$$\begin{aligned} H_{\boldsymbol{\mu}}(t)X(t) = x &+ \sum_{\substack{i,j\in S \\ i\neq j}} \int_0^t X(s-)H_{\boldsymbol{\mu}}(s-)\mu_{ij}(s)dM_{ij}(s) - \int_0^t X(s-)H_{\boldsymbol{\mu}}(s-)\boldsymbol{\theta}(s)^\top d\mathbf{W}(s) \\ &+ \int_0^t H_{\boldsymbol{\mu}}(s-)\boldsymbol{\pi}(s)^\top \boldsymbol{\sigma}(s)d\mathbf{W}(s) - \int_0^t H_{\boldsymbol{\mu}}(s-)dc(s) - \sum_{0<s\leq t} H_{\boldsymbol{\mu}}(s-)\triangle c(s)\triangle(\boldsymbol{\mu} \bullet \mathbf{M})(s), \end{aligned} \tag{4.51}$$

for all $t \in [0,T]$. By combining (4.50) and (4.51), $J_{\boldsymbol{\mu}}^{(x,c,\boldsymbol{\pi})}$ can be written as

$$\begin{aligned} J_{\boldsymbol{\mu}}^{(x,c,\boldsymbol{\pi})}(t) &= x + \int_0^t H_{\boldsymbol{\mu}}(s-)\Big(\boldsymbol{\pi}(s)^\top \boldsymbol{\sigma}(s) - X(s-)\boldsymbol{\theta}(s)^\top\Big) d\mathbf{W}(s) \\ &\quad + \sum_{\substack{i,j\in S\\ i\neq j}} \int_0^t X(s-)H_{\boldsymbol{\mu}}(s-)\mu_{ij}(s)dM_{ij}(s), \quad t\in[0,T]. \end{aligned} \tag{4.52}$$

Since $\mathbf{W}, M_{ij} \in \mathcal{M}(\{\mathcal{F}_t\}, P)$, it is true from (4.52) that

$$J_{\boldsymbol{\mu}}^{(x,c,\boldsymbol{\pi})} \in \mathcal{M}_{loc}(\{\mathcal{F}_t\}, P). \tag{4.53}$$

Since $(c,\boldsymbol{\pi}) \in \mathcal{A}(x)$, we have from (4.47) and (4.31),

$$H_{\boldsymbol{\mu}}(t)X^{(x,c,\boldsymbol{\pi})}(t) \geq 0 \text{ a.s.}, \quad t\in[0,T]. \tag{4.54}$$

Additionally, since $c \in \mathcal{C}$ and $\boldsymbol{\mu} \in \mathcal{H}$, we have from Remark 4.3.7,

$$\int_0^t H_{\boldsymbol{\mu}}(s-)dc(s) + \sum_{0<s\leq t} H_{\boldsymbol{\mu}}(s-)\triangle c(s)\triangle(\boldsymbol{\mu}\bullet\mathbf{M})(s) \geq 0 \text{ a.s.}, \quad t\in[0,T]. \tag{4.55}$$

From (4.54), (4.55) and (4.50), we have that

$$J_{\boldsymbol{\mu}}^{(x,c,\boldsymbol{\pi})}(t) \geq 0 \text{ a.s.}, \quad t\in[0,T]. \tag{4.56}$$

As a result of $J_{\boldsymbol{\mu}}^{(x,c,\boldsymbol{\pi})}$ being a non-negative local martingale, from Proposition C.11.3, we conclude

$$J_{\boldsymbol{\mu}}^{(x,c,\boldsymbol{\pi})} \in \mathcal{SPM}(\{\mathcal{F}_t\}, P). \tag{4.57}$$

□

The following corollary is immediate from Proposition 4.3.8 and Remark 4.3.7.

Corollary 4.3.9.1. *For each $\boldsymbol{\mu} \in \mathcal{H}, x \in [0,\infty)$, and $(c,\boldsymbol{\pi}) \in \mathcal{A}(x)$,*

$$E\left[H_{\boldsymbol{\mu}}(T)X^{(x,c,\boldsymbol{\pi})}(T)\right] \leq x. \tag{4.58}$$

Remark 4.3.10. Suppose that Λ is non-empty. Then, from (4.33), for *each* $x \in \Lambda$ there exists a pair $(c,\boldsymbol{\pi}) \in \mathcal{A}(x)$ such that

$$X^{(x,c,\boldsymbol{\pi})}(T) = B \text{ a.s.} \tag{4.59}$$

Using Corollary 4.3.9.1 and (4.59), we have

$$\Lambda \neq \emptyset \implies \sup_{\boldsymbol{\mu}\in\mathcal{H}} E\left[H_{\boldsymbol{\mu}}(T)B\right] \leq x, \quad \text{for all } x \in \Lambda. \tag{4.60}$$

Remark 4.3.11. Note that the supremum, $\sup_{\boldsymbol{\mu}\in\mathcal{H}} E[H_{\boldsymbol{\mu}}(T)B]$, is defined in $(0,\infty]$, since, *for each and every* $\boldsymbol{\mu}\in\mathcal{H}$, $P[H_{\boldsymbol{\mu}}(T)B>0]=1$.

Remark 4.3.12. From (4.60) one sees that

$$\Lambda \neq \emptyset \implies \sup_{\boldsymbol{\mu}\in\mathcal{H}} E[H_{\boldsymbol{\mu}}(T)B] < \infty. \tag{4.61}$$

Note that this is very similar to Remark 3.3.7 in Chapter 3; however, in the case of regime-switching, the *supremum* of $E[H_{\boldsymbol{\mu}}(T)B]$ over all $\boldsymbol{\mu}\in\mathcal{H}$ is finite.

The main result of this chapter (see Theorem 4.3.17) establishes the converse implication of (4.61), so that we actually have the following equivalence:

$$\Lambda \neq \emptyset \iff \sup_{\boldsymbol{\mu}\in\mathcal{H}} E[H_{\boldsymbol{\mu}}(T)B] < \infty, \tag{4.62}$$

that is, the contingent claim B can be hedged if and only if $\sup_{\boldsymbol{\mu}\in\mathcal{H}} E[H_{\boldsymbol{\mu}}(T)B]<\infty$. Thus, if the contingent claim random variable B is stipulated so large that

$$\sup_{\boldsymbol{\mu}\in\mathcal{H}} E[H_{\boldsymbol{\mu}}(T)B] = \infty \tag{4.63}$$

then there is no possibility of hedging the claim B.

Notation 4.3.13. In the following, an essential role will be played by the extended real number $\hat{u}\in[0,\infty]$ defined as follows:

$$\hat{u} \triangleq \sup_{\boldsymbol{\mu}\in\mathcal{H}} E[H_{\boldsymbol{\mu}}(T)B]. \tag{4.64}$$

The choice $\hat{u}$ will be our candidate for the least initial wealth which hedges the contingent claim B.

Remark 4.3.14. Since $H_{\boldsymbol{\mu}}(T)B>0$ a.s. for all $\boldsymbol{\mu}\in\mathcal{H}$, the conditional expectation $E[H_{\boldsymbol{\mu}}(T)B|\mathcal{F}_t]$ is almost surely non-negative for all $\boldsymbol{\mu}\in\mathcal{H}$ at each time $t\in[0,T]$. Furthermore, since $\inf_{t\in[0,T]} H_{\boldsymbol{\mu}}(t)>0$ a.s. for all $\boldsymbol{\mu}\in\mathcal{H}$, we have for each fixed $t\in[0,T]$ that

$$\hat{X}_{\boldsymbol{\mu}}(t) \triangleq \frac{E[H_{\boldsymbol{\mu}}(T)B|\mathcal{F}_t]}{H_{\boldsymbol{\mu}}(t)} \geq 0 \text{ a.s.}, \quad \boldsymbol{\mu}\in\mathcal{H}. \tag{4.65}$$

By Theorem C.13.2 in the Appendix, for each fixed $t\in[0,T]$ the essential-supremum (see Definition C.13.1) of the family of random variables $\{\hat{X}_{\boldsymbol{\mu}}(t),\boldsymbol{\mu}\in\mathcal{H}\}$ exists. That is

$$\operatorname*{ess-sup}_{\boldsymbol{\mu}\in\mathcal{H}} \hat{X}_{\boldsymbol{\mu}}(t) = \operatorname*{ess-sup}_{\boldsymbol{\mu}\in\mathcal{H}} \frac{E\left[H_{\boldsymbol{\mu}}(T)B \,\middle|\, \mathcal{F}_t\right]}{H_{\boldsymbol{\mu}}(t)} \text{ exists for each } t\in[0,T]. \tag{4.66}$$

Proposition 4.3.15 which follows is analogous to Proposition 3.3.9 for the case of hedging without regime-switching. However, whereas Proposition 3.3.9 is completely self-evident from the definitions, Proposition 4.3.15 is definitely not self-evident, largely on account of the essential supremum over $\boldsymbol{\mu} \in \mathcal{H}$ at (4.67) (which should be compared with (3.35)). Proposition 4.3.15 is proved in Appendix B.2.

Proposition 4.3.15. *Suppose $\hat{u} < \infty$, and define the process $\{\hat{X}(t), t \in [0,T]\}$ as follows*

$$\hat{X}(t) \triangleq \operatorname*{ess-sup}_{\boldsymbol{\mu}\in\mathcal{H}} \frac{E\left[H_{\boldsymbol{\mu}}(T)B \,\middle|\, \mathcal{F}_t\right]}{H_{\boldsymbol{\mu}}(t)}, \quad t \in [0,T], \tag{4.67}$$

where the right hand side of (4.67) exists by Remark 4.3.14. The following statements are true:

(1) $\hat{X}(0) = \hat{u}$ *a.s.*

(2) $\hat{X}(T) = B$ *a.s.*

(3) $\hat{X}(t) \geq 0$ *a.s.,* $\quad t \in [0,T]$

(4) $H_{\boldsymbol{\mu}}\hat{X} \in \mathcal{SPM}(\{\mathcal{F}_t\}, P), \quad \boldsymbol{\mu} \in \mathcal{H}$

(5) For each $0 \leq s \leq t \leq T$

$$\hat{X}(s) = \operatorname*{ess-sup}_{\boldsymbol{\mu}\in\mathcal{H}} \frac{E[H_{\boldsymbol{\mu}}(t)\hat{X}(t)|\mathcal{F}_s]}{H_{\boldsymbol{\mu}}(s)} \text{ a.s.} \tag{4.68}$$

(6) $\{\hat{X}(t), t \in [0,T]\}$ *has a càdlàg modification.*

We now state a decomposition result involving $\hat{X}$ and the generalized price state density $H_{\boldsymbol{\mu}}$ that is absolutely essential for proving our hedging results.

Lemma 4.3.16. *For each $\boldsymbol{\mu} \in \mathcal{H}$, the process $\{H_{\boldsymbol{\mu}}(t)\hat{X}(t), t \in [0,T]\}$ has the decomposition*

$$H_{\boldsymbol{\mu}}(t)\hat{X}(t) = \hat{X}(0) + \int_0^t \boldsymbol{\Psi}_{\boldsymbol{\mu}}(s)^\top d\mathbf{W}(s) + \sum_{\substack{i,j\in S\\ i\neq j}} \int_0^t \Gamma_{ij}^{\boldsymbol{\mu}}(s) dM_{ij}(s) - A_{\boldsymbol{\mu}}(t), \tag{4.69}$$

for all $t \in [0,T]$ and $\boldsymbol{\mu} \in \mathcal{H}$, where $\boldsymbol{\Psi}_{\boldsymbol{\mu}}$ is a unique, null-at-the-origin $\{\mathcal{F}_t\}$-progressively measurable process such that

$$\int_0^T ||\boldsymbol{\Psi}_{\boldsymbol{\mu}}(s)||^2 ds < \infty \text{ a.s.}, \tag{4.70}$$

$\boldsymbol{\Gamma}^{\boldsymbol{\mu}} \triangleq \{\Gamma_{ij}^{\boldsymbol{\mu}}\}_{i,j\in S}$ *is a unique, null-at-the-origin* $\{\mathcal{F}_t\}$*-predictably measurable process such that*

$$\sum_{\substack{i,j\in S\\ i\neq j}} \left(\int_0^T |\Gamma_{ij}^{\boldsymbol{\mu}}(s)|^2 d[M_{ij}](s) \right)^{1/2} < \infty \ a.s., \tag{4.71}$$

and $A_{\boldsymbol{\mu}}$ *is a unique, null-at-the-origin and non-decreasing* $\{\mathcal{F}_t\}$*-predictably measurable process such that*

$$E\left[A_{\boldsymbol{\mu}}(T)\right] < \infty. \tag{4.72}$$

Proof. Fix $\boldsymbol{\mu} \in \mathcal{H}$. From Proposition 4.3.15(4) we know that $H_{\boldsymbol{\mu}}\hat{X}$ is an $\{\mathcal{F}_t\}$-supermartingale. Using the Doob-Meyer decomposition from Theorem C.11.4, we can write any $\{\mathcal{F}_t\}$-supermartingale as the difference between a unique $\{\mathcal{F}_t\}$-local martingale, null at the origin, and a unique, non-decreasing, $\{\mathcal{F}_t\}$-predictable càdlàg process. In the case of $H_{\boldsymbol{\mu}}\hat{X}$, we can write

$$H_{\boldsymbol{\mu}}(t)\hat{X}(t) = \hat{X}(0) + \Phi_{\boldsymbol{\mu}}(t) - A_{\boldsymbol{\mu}}(t), \quad t \in [0,T], \tag{4.73}$$

where

$$\Phi_{\boldsymbol{\mu}} \in \mathcal{M}_{loc,0}(\{\mathcal{F}_t\}, P) \tag{4.74}$$

and

$$A_{\boldsymbol{\mu}} \in \mathcal{A}_0^+(\{\mathcal{F}_t\}, P) \text{ and } A_{\boldsymbol{\mu}} \in \mathcal{P}^*, \tag{4.75}$$

where the space $\mathcal{A}_0^+$ is defined in Notation C.5.5. Furthermore, since

$$E\left[H_{\boldsymbol{\mu}}(t)\hat{X}(t)\right] > -\infty, \quad t \in [0,T], \tag{4.76}$$

it follows that

$$E\left[A_{\boldsymbol{\mu}}(T)\right] < \infty, \tag{4.77}$$

again by Theorem C.11.4. We can further expand the $\{\mathcal{F}_t\}$-local martingale $\Phi_{\boldsymbol{\mu}}$ into more insightful local martingales using the martingale representation theorem (see Appendix C.12, specifically Theorem C.12.4), which in the present case states that we can write $\Phi_{\boldsymbol{\mu}}$ uniquely in the following way:

$$\Phi_{\boldsymbol{\mu}}(t) = \int_0^t \boldsymbol{\Psi}_{\boldsymbol{\mu}}(s)^\top d\mathbf{W}(s) + \sum_{\substack{i,j\in S\\ i\neq j}} \int_0^t \Gamma_{ij}^{\boldsymbol{\mu}}(s) dM_{ij}(s), \quad t \in [0,T], \tag{4.78}$$

in which

$$\boldsymbol{\Psi}_{\boldsymbol{\mu}} \in \mathcal{F}^*, \tag{4.79}$$

and

$$\mathbf{\Gamma}^{\boldsymbol{\mu}} \triangleq \{\Gamma^{\boldsymbol{\mu}}_{ij}\}_{i,j\in S},\, \mathbf{\Gamma}^{\boldsymbol{\mu}} \in \mathcal{P}^*, \tag{4.80}$$

are unique integrand processes such that

$$\int_0^T ||\mathbf{\Psi}_{\boldsymbol{\mu}}(s)||^2 ds < \infty \text{ a.s.} \quad \text{and} \quad \sum_{\substack{i,j\in S\\ i\neq j}} \left(\int_0^T |\Gamma^{\boldsymbol{\mu}}_{ij}(s)|^2 d[M_{ij}](s)\right)^{1/2} < \infty \text{ a.s.} \tag{4.81}$$

Combining (4.78) and (4.73) we can write $H_{\boldsymbol{\mu}}\hat{X}$ as

$$H_{\boldsymbol{\mu}}(t)\hat{X}(t) = \hat{X}(0) + \int_0^t \mathbf{\Psi}_{\boldsymbol{\mu}}(s)^\top d\mathbf{W}(s) + \sum_{\substack{i,j\in S\\ i\neq j}} \int_0^t \Gamma^{\boldsymbol{\mu}}_{ij}(s) dM_{ij}(s) - A_{\boldsymbol{\mu}}(t), \tag{4.82}$$

for all $t \in [0,T]$. □

The main result of this chapter is the following theorem.

Theorem 4.3.17. *Suppose that $\hat{u} < \infty$. Then $\hat{u} \in \Lambda$ (in particular $\Lambda \neq \emptyset$) and $\hat{u} = \inf \Lambda$, that is $\hat{u} \leq x$ for all $x \in \Lambda$.*

Remark 4.3.18. From Theorem 4.3.17 and (4.31), one sees that $\hat{u}$ is the *least initial wealth* which guarantees that $X^{(\hat{u},c,\boldsymbol{\pi})}(T) = B$ a.s. for some $(c,\boldsymbol{\pi}) \in \mathcal{A}(\hat{u})$, that is $\hat{u}$ is the *price* of the contingent claim B.

Proof. We are going to establish that there exists some consumption-portfolio pair $(\hat{c},\hat{\boldsymbol{\pi}}) \in \mathcal{A}(\hat{u})$ such that

$$\hat{X}(t) = X^{(\hat{u},\hat{c},\hat{\boldsymbol{\pi}})}(t), \quad t \in [0,T], \tag{4.83}$$

where, from (4.22), the process $X^{(\hat{u},\hat{c},\hat{\boldsymbol{\pi}})}$ satisfies the stochastic integral equation

$$\begin{aligned} X^{(\hat{u},\hat{c},\hat{\boldsymbol{\pi}})}(t) = \hat{u} &+ \int_0^t \left\{ r(s)X^{(\hat{u},\hat{c},\hat{\boldsymbol{\pi}})}(s) + \hat{\boldsymbol{\pi}}(s)^\top \boldsymbol{\sigma}(s)\boldsymbol{\theta}(s)\right\} ds \\ &+ \int_0^t \hat{\boldsymbol{\pi}}(s)^\top \boldsymbol{\sigma}(s) d\mathbf{W}(s) - \hat{c}(t), \quad t \in [0,T]. \end{aligned} \tag{4.84}$$

If, in fact, one can establish (4.83), we would know from Proposition 4.3.15(2) and 4.3.15(3) that there exists $(\hat{u},\hat{c},\hat{\boldsymbol{\pi}}) \in [0,\infty) \times \mathcal{A}(\hat{u})$ such that

$$X^{(\hat{u},\hat{c},\hat{\boldsymbol{\pi}})}(T) = B \text{ a.s.} \tag{4.85}$$

Since $\hat{u}$ is the initial wealth of the process $X^{(\hat{u},\hat{c},\hat{\boldsymbol{\pi}})}$, this would mean, from (4.33),

$$\hat{u} \in \Lambda \text{ and therefore } \Lambda \neq \emptyset. \tag{4.86}$$

Furthermore, from Remark 4.3.10, we have that $\hat{u} \leq \inf \Lambda$, and therefore, from (4.86),

$$\hat{u} = \inf \Lambda \in \Lambda. \tag{4.87}$$

Thus, it remains to establish (4.83), as from (4.86) and (4.87), we will have proven Theorem 4.3.17.

To begin, fix some $\boldsymbol{\mu} \in \mathcal{H}$. Isolating $\hat{X}(t)$ in (4.69) of Lemma 4.3.16 and using Proposition 4.3.15(1) we get the expression

$$\hat{X}(t) = \frac{1}{H_{\boldsymbol{\mu}}(t)}\left\{\hat{u}+\int_0^t \boldsymbol{\Psi}_{\boldsymbol{\mu}}(s)^\top d\mathbf{W}(s)+\sum_{\substack{i,j\in S\\ i\neq j}}\int_0^t \Gamma_{ij}^{\boldsymbol{\mu}}(s)dM_{ij}(s)-A_{\boldsymbol{\mu}}(t)\right\}, \quad t \in [0,T]. \tag{4.88}$$

By an easy, although lengthy, calculation using the Ito product rule, the identity (4.45), and the fact that $[M_{ij}](t) = R_{ij}(t)$ by (4.13), we can expand the right hand side of (4.88) as follows:

$$\begin{aligned}
\hat{X}(t) = \hat{u} &+ \int_0^t \left(\frac{\boldsymbol{\Psi}_{\boldsymbol{\mu}}(s)^\top}{H_{\boldsymbol{\mu}}(s)} + \hat{X}(s)\boldsymbol{\theta}(s)^\top\right)d\mathbf{W}(s)\\
&+ \int_0^t \left(\frac{\boldsymbol{\Psi}_{\boldsymbol{\mu}}(s)^\top}{H_{\boldsymbol{\mu}}(s)}\boldsymbol{\theta}(s) + \hat{X}(s)r(s) + \hat{X}(s)\|\boldsymbol{\theta}(s)\|^2\right)ds\\
&+ \sum_{\substack{i,j\in S\\ i\neq j}}\int_0^t \left(\frac{\Gamma_{ij}^{\boldsymbol{\mu}}(s)}{H_{\boldsymbol{\mu}}(s-)} - \hat{X}(s-)\mu_{ij}(s)\right)dM_{ij}(s)\\
&+ \sum_{\substack{i,j\in S\\ i\neq j}}\int_0^t \left(\frac{\triangle A_{\boldsymbol{\mu}}(s)}{H_{\boldsymbol{\mu}}(s-)} - \frac{\Gamma_{ij}^{\boldsymbol{\mu}}(s)}{H_{\boldsymbol{\mu}}(s-)} + \hat{X}(s-)\mu_{ij}(s)\right)\tilde{\mu}_{ij}(s)dR_{ij}(s)\\
&- \int_0^t \frac{1}{H_{\boldsymbol{\mu}}(s-)}dA_{\boldsymbol{\mu}}(s), \quad t \in [0,T],
\end{aligned} \tag{4.89}$$

in which we have put

$$\tilde{\mu}_{ij}(\cdot) \triangleq \mu_{ij}(\cdot)/(1+\mu_{ij}(\cdot)). \tag{4.90}$$

Remark 4.3.19. Since the Brownian motion $\mathbf{W}(t)$ is continuous, the integrands $\hat{X}(s-)$ and $H_{\boldsymbol{\mu}}(s-)$ can be written as $\hat{X}(s)$ and $H_{\boldsymbol{\mu}}(s)$, respectively, in both the $d\mathbf{W}(s)$ and ds integrals.

Since $\boldsymbol{\mu} \in \mathcal{H}$ is arbitrary we can of course take $\boldsymbol{\mu}(t) \triangleq 0$ for all $t \in [0,T]$ at (4.89) to get

$$\begin{aligned}
\hat{X}(t) = \hat{u} &+ \int_0^t \left(\frac{\boldsymbol{\Psi}_0(s)^\top}{H_0(s)} + \hat{X}(s)\boldsymbol{\theta}(s)^\top \right) d\mathbf{W}(s) \\
&+ \int_0^t \left(\frac{\boldsymbol{\Psi}_0(s)^\top}{H_0(s)}\boldsymbol{\theta}(s) + \hat{X}(s)r(s) + \hat{X}(s)\|\boldsymbol{\theta}(s)\|^2 \right) ds \\
&+ \sum_{\substack{i,j\in S \\ i\neq j}} \int_0^t \frac{\Gamma_{ij}^0(s)}{H_0(s-)} dM_{ij}(s) - \int_0^t \frac{1}{H_0(s-)} dA_0(s), \quad t \in [0,T].
\end{aligned} \tag{4.91}$$

We now define the $\mathbb{R}^N$-valued process $\hat{\boldsymbol{\pi}}$, which we will use as our candidate optimal portfolio process, as follows.

$$\hat{\boldsymbol{\pi}}(t) \triangleq [\boldsymbol{\sigma}^{-1}(t)]^\top \left[\frac{\boldsymbol{\Psi}_0(t)}{H_0(t)} + \hat{X}(t)\boldsymbol{\theta}(t) \right], \quad t \in [0,T]. \tag{4.92}$$

For $\hat{\boldsymbol{\pi}}$ to be considered a portfolio process we must show that $\hat{\boldsymbol{\pi}} \in \Pi$, as we do in the following proposition.

Proposition 4.3.20. *The process* $\{\hat{\boldsymbol{\pi}}(t), t \in [0,T]\}$ *defined by*

$$\hat{\boldsymbol{\pi}}(t) \triangleq [\boldsymbol{\sigma}^{-1}(t)]^\top \left[\frac{\boldsymbol{\Psi}_0(t)}{H_0(t)} + \hat{X}(t)\boldsymbol{\theta}(t) \right], \quad t \in [0,T], \tag{4.93}$$

is a portfolio process. That is, $\hat{\boldsymbol{\pi}} \in \Pi$.

Proof. Since $\boldsymbol{\sigma}, \boldsymbol{\Psi}_0, H_0, \hat{X}, \boldsymbol{\theta} \in \mathcal{F}^*$, one sees from (4.93) that

$$\hat{\boldsymbol{\pi}} \in \mathcal{F}^*. \tag{4.94}$$

Now, $\boldsymbol{\theta}(\cdot)$ is uniformly bounded on $[0,T] \times \Omega$ (see Notation 4.1.8). Since $\hat{X}$ has a càdlàg modification, $\hat{X}$ is bounded $\lambda \otimes P$ a.e., thus we have that $\hat{X}\boldsymbol{\theta}$ is bounded $\lambda \otimes P$ a.e., and so

$$\int_0^T ||\hat{X}(t)\boldsymbol{\theta}(t)||^2 dt < \infty \text{ a.s.} \tag{4.95}$$

Since $\inf_{t\in[0,T]} H_0(t) > 0$ a.s. by (4.47), from (4.70) in Lemma 4.3.16

$$\int_0^T ||\frac{\mathbf{\Psi}_0(t)}{H_0(t)}||^2 dt < \infty, \text{ a.s.} \tag{4.96}$$

Combining (4.95) and (4.96) gives

$$\int_0^T ||\frac{\mathbf{\Psi}_0(t)}{H_0(t)} + \hat{X}(t)\boldsymbol{\theta}(t)||^2 dt < \infty, \qquad \text{a.s.} \tag{4.97}$$

In view of (4.93), (4.97) and Remark 4.1.6, we obtain

$$\int_0^T ||\hat{\boldsymbol{\pi}}(t)||^2 dt < \infty \text{ a.s.} \tag{4.98}$$

and from (4.98), (4.94), and the definition of Π from (4.17), we find that

$$\hat{\boldsymbol{\pi}} \in \Pi. \tag{4.99}$$

That is, $\hat{\boldsymbol{\pi}}$ is a valid portfolio process. □

To continue the proof of Theorem 4.3.17 we need the following proposition, which characterizes the optimal cumulative consumption process $\hat{c}$. As the proof of Proposition 4.3.21 is quite lengthy and technical, it is left in Appendix B.3.

Proposition 4.3.21. *Define the process $\hat{c}$ as follows,*

$$\hat{c}(t) \triangleq \int_0^t \frac{1}{H_0(s-)} dA_0(s) - \sum_{\substack{i,j\in S \\ i\neq j}} \int_0^t \frac{\Gamma^0_{ij}(s)}{H_0(s-)} dM_{ij}(s), \quad t \in [0,T]. \tag{4.100}$$

Then $\hat{c}$ is a cumulative consumption process, i.e.,

$$\hat{c} \in \mathcal{C}. \tag{4.101}$$

Completing the proof of Theorem 4.3.17, in view of (4.91), Proposition 4.3.20, and Proposition 4.3.21 we have the identity

$$\begin{aligned} \hat{X}(t) = \hat{u} &+ \int_0^t \Big\{ r(s)\hat{X}(s) + \hat{\boldsymbol{\pi}}(s)^\top \boldsymbol{\sigma}(s)\boldsymbol{\theta}(s) \Big\} ds \\ &+ \int_0^t \hat{\boldsymbol{\pi}}(s)^\top \boldsymbol{\sigma}(s) d\mathbf{W}(s) - \hat{c}(t), \quad t \in [0,T], \end{aligned} \tag{4.102}$$

where $(\hat{c}, \hat{\boldsymbol{\pi}}) \in \mathcal{C} \times \Pi$. Now (4.83) follows upon comparison (4.102) and (4.84), and it remains to verify that $(\hat{c}, \hat{\boldsymbol{\pi}}) \in \mathcal{A}(\hat{u})$. From Proposition 4.3.15(3) together with (4.83) we have

$$X^{(\hat{u},\hat{c},\hat{\boldsymbol{\pi}})}(t) \geq 0 \text{ a.s.}, \quad t \in [0, T], \tag{4.103}$$

as required to get $(\hat{c}, \hat{\boldsymbol{\pi}}) \in \mathcal{A}(\hat{u})$ (see (4.31)). We have thus established that (4.83) holds for some $(\hat{c}, \hat{\boldsymbol{\pi}}) \in \mathcal{A}(\hat{u})$, and Theorem 4.3.17 follows.
□

Remark 4.3.22. From Theorem 4.3.17, together with Remark 4.3.18, one sees that the price of the contingent claim B is the quantity $\hat{u}$ defined at (4.64), that is $\hat{u}$ is the *least initial wealth* from which the contingent claim B can be hedged. We also note the corresponding *hedging portfolio* is the process $\hat{\boldsymbol{\pi}}$ defined in Proposition 4.3.20, that is

$$\hat{\boldsymbol{\pi}}(t) \triangleq [\boldsymbol{\sigma}^{-1}(t)]^{\top} \left[\frac{\boldsymbol{\Psi}_0(t)}{H_0(t)} + \hat{X}(t)\boldsymbol{\theta}(t) \right], \quad t \in [0, T]. \tag{4.104}$$

Notice that the hedging portfolio $\hat{\boldsymbol{\pi}}$ has the exact same characterization as in Chapter 3 (see (3.45)), but is now adapted to the joint filtration defined in Condition 4.1.1. Lastly we have the *cumulative consumption process* $\hat{c}$ defined in Proposition 4.3.21, that is

$$\hat{c}(t) \triangleq \int_0^t \frac{1}{H_0(s-)} dA_0(s) - \sum_{\substack{i,j \in S \\ i \neq j}} \int_0^t \frac{\Gamma_{ij}^0(s)}{H_0(s-)} dM_{ij}(s), \quad t \in [0, T]. \tag{4.105}$$

Remark 4.3.23. Suppose that Conditions 3.1.1 and 3.1.3 of Chapter 3 are in force, in place of the more general Conditions 4.1.1 and 4.1.4 of the present chapter. Then there is no regime-switching in the model, so that we can take $M_{ij} = 0$ in the consumption equation (4.105). As for the process A_0 on the right hand side of (4.105), we know that this process is given by the Doob-Meyer decomposition in Lemma 4.3.16 when $\boldsymbol{\mu} = 0$ that is

$$H_0(t)\hat{X}(t) = \hat{X}(0) + \Phi_0(t) - A_0(t), \quad t \in [0, T], \tag{4.106}$$

with $\Phi_0 \in \mathcal{M}_{loc,0}(\{\mathcal{F}_t\}, P)$ and $A_0 \in \mathcal{FV}_0(\{\mathcal{F}_t\}, P)$. But, when Conditions 3.1.1 and 3.1.3 are in force, then we know from Proposition 3.3.9(4) that $H_0\hat{X} \in \mathcal{M}(\{\mathcal{F}_t\}, P)$, and therefore by the uniqueness part of the Doob-Meyer decomposition (see Theorem C.11.4) we necessarily have $A_0 = 0$. It then follows from (4.105) that $\hat{c} = 0$, that is the consumption is necessarily zero when Conditions 3.1.1 and 3.1.3 hold.

Chapter 5

The Constrained Hedging Problem in a Regime-Switching Market Model

In this chapter we extend the hedging problems defined in Chapters 3 and 4 to the more general context of almost sure hedging in a regime-switching market model with both *convex portfolio constraints* and *margin requirements*. Just as in Chapter 4, we would like to find the price of a contingent claim together with a consumption-portfolio pair which promises that the seller of the claim almost surely covers his/her obligation at the close of trade. However, in the present chapter, the agent now trades in a regime-switching market where there are possible restrictions on investments and fees that depend on the agent's choice of portfolio. The results of this chapter generalize both the results of Chapter 4 and the paper of Cvitanic and Karatzas [10] on almost sure hedging in a standard financial market (driven solely by Brownian motion) with convex portfolio constraints. As many of the definitions and conditions required for this chapter have already been stated in Chapter 4, the reader may be asked to refer to Chapter 4 in specific instances.

5.1 Market Model

The regime-switching market model that is used in this chapter is nearly identical to the market model defined in Chapter 4. As a result, we refer to Chapter 4.1.1-4.1.16 for the appropriate conditions, remarks and results defining the regime-switching market model. In addition to the conditions defining the regime-switching market model, we impose conditions on an agent trading in the market through portfolio constraints and margin requirements. The portfolio constraints act through a convex constraint set K given in Condition 5.1.1 and the margin requirements act through a margin function g given in Condition 5.1.3.

Condition 5.1.1. *We are given a closed convex* ***portfolio constraint set*** $K \subset \mathbb{R}^N$ *with* $0 \in K$*. The agent can only invest within the constraint set* K *during the entire trading interval.*

Notation 5.1.2. For a $\mathcal{B}([0,T]) \times \mathcal{F}$-measurable mapping $\boldsymbol{\zeta} : [0,T] \times \Omega \mapsto \mathbb{R}^N$ and a set $K \subset \mathbb{R}^N$, the statement "$\boldsymbol{\zeta}(t) \in K$ a.e." means that $\boldsymbol{\zeta}(t,\omega) \in K$ for $\lambda \otimes P$-almost all $(t,\omega) \in [0,T] \times \Omega$.

Condition 5.1.3. *We are given a* ***margin function*** *(or* ***margin term****)* $g : [0,T] \times \Omega \times \mathbb{R}^N \mapsto \mathbb{R}$ *that is* $\mathcal{F}^* \times \mathcal{B}(\mathbb{R}^N)$ *measurable such that* $g(t,\omega,0) = 0$ *and* $g(t,\omega,\cdot)$ *is concave on* $\mathbb{R}^N$ *for each* $(t,\omega) \in [0,T] \times \Omega$*. Furthermore,* g *is uniformly upper-bounded and* $g(t,\omega,\cdot)$ *is* $M(t)$*-Lipschitz continuous for each* $(t,\omega) \in [0,T] \times \Omega$*. That is, there exists an* $\mathcal{F}^*$*-measurable process* $M : [0,T] \times \Omega \mapsto [0,\infty]$ *where*

$$\int_0^T |M(t)|^2 dt < \infty \ a.s., \tag{5.1}$$

such that

$$|g(t,\omega,\boldsymbol{p}) - g(t,\omega,\boldsymbol{q})| \leq M(t,\omega)\|\boldsymbol{p} - \boldsymbol{q}\|, \quad (t,\omega) \in [0,T] \times \Omega, \ \boldsymbol{p}, \boldsymbol{q} \in \mathbb{R}^N. \tag{5.2}$$

Remark 5.1.4. Requiring the margin function $g(t,\omega,\cdot)$ to be Lipschitz continuous is a reasonable condition as any concave function defined on all of $\mathbb{R}^N$ is Lipschitz continuous over every convex compact subset of $\mathbb{R}^N$ (see Proposition D.0.19).

Remark 5.1.5. The function $g(t,\omega,\cdot)$, which is generally non-linear for each $(t,\omega) \in [0,T] \times \Omega$, models payments levied on the investor by imposed margin requirements. This could include an interest rate for borrowing which is higher than the interest rate for lending, or margin payments to a broker when shorting stocks or borrowing to go long on stocks. The modeling of margin requirements by the function g, which considerably enhances the applicability and usefulness of Brownian motion market models, was introduced by Cuoco and Liu [7].

Remark 5.1.6. In any practical market it is generally safe to assume that the margin function is non-positive, that is $g(t,\omega,\boldsymbol{p}) \leq 0$ for all $(t,\omega,\boldsymbol{p}) \in [0,T] \times \Omega \times \mathbb{R}^N$, since g models a payment extracted from wealth. Having positive g would allow, for instance, an investor to be given additional money when borrowing. In this chapter, we do, however, assume g is more generally uniformly upper-bounded as it adds no additional complexity to the problem.

Remark 5.1.7. To model investment in a constrained market, we proceed in a slightly different manner than in Chapter 4. We introduce the process $\{p_0(t), t \in [0,T]\}$, which denotes the *proportion of wealth* allocated by the agent to the risk-free asset with price S_0. This process is $\{\mathcal{F}_t\}$-progressively measurable, that is $p_0 \in \mathcal{F}^*$. We model investment in the N-risky assets through the space Π of *portfolio processes*, which, with the inclusion of margin requirements, is defined as follows:

$$\Pi \triangleq \left\{ \boldsymbol{p} : [0,T] \times \Omega \to \mathbb{R}^N \middle| \boldsymbol{p} \in \mathcal{F}^*,\ \int_0^T \Big(||\boldsymbol{p}(t)||^2 + |g(t, \boldsymbol{p}(t))| \Big) dt < \infty \text{ a.s.} \right\}. \tag{5.3}$$

The vector $\boldsymbol{p} = \{p_n\}_{n=1}^N$ has the interpretation as the *proportion of wealth* the agent invests into the N risky assets. This space Π is more restrictive than the space defined in (4.17) as it requires the function $g(\cdot, \boldsymbol{p}(\cdot))$, defined in Condition 5.1.3, to be integrable, in addition to requiring $\boldsymbol{p}(\cdot)$ to be square integrable.

Consumption by the agent is modelled exactly as in Chapter 4. That is the agent is free to "consume" through a cumulative consumption process c that exists in the space $\mathcal{C}$:

$$\begin{aligned} \mathcal{C} \triangleq \Big\{ c : [0,T] \times \Omega \to [0,\infty) \Big| & c(0) = 0, c \in \mathcal{F}^*, c(T) < \infty \text{ a.s.}, \\ & c(\cdot) \text{ càdlàg, non-decreasing} \Big\}. \end{aligned} \tag{5.4}$$

The necessity of the cumulative consumption process is explained in detail in Remark 4.2.5.

The agent begins trading from a *strictly positive initial wealth* $x \in (0, \infty)$, as opposed to a non-negative initial wealth as in Chapter 4, and has *total wealth* $\{X(t), t \in [0,T]\}$, given by

$$X(t) \triangleq \sum_{n=0}^{N} \pi_n(t) = X(t)p_0(t) + X(t)\boldsymbol{p}(t)^\top \mathbf{1}, \quad t \in [0,T], \tag{5.5}$$

where $\boldsymbol{\pi} \triangleq \{\pi_n\}$ is the monetary portfolio defined as in Remark 4.1.17 (see Remark 5.1.14). Since the total wealth of the agent is defined in a similar manner as in Chapter 4, the agent follows a $\Gamma(\cdot)$-financed trading strategy; however in this case, the process Γ includes the margin function. Due to the inclusion of margin requirements, the wealth process X no longer satisfies the SDE in (4.23), but instead satisfies the stochastic integral equation

$$\begin{aligned} X(t) = x + \int_0^t X(s) \Big\{ r(s) + g(s, \boldsymbol{p}(s)) + \boldsymbol{p}(s)^\top \boldsymbol{\sigma}(s) \boldsymbol{\theta}(s) \Big\} ds \\ + \int_0^t X(s) \boldsymbol{p}(s)^\top \boldsymbol{\sigma}(s) d\mathbf{W}(s) - c(t), \quad t \in [0,T], \end{aligned} \tag{5.6}$$

in which $x \in (0, \infty)$ is the initial wealth of the agent, $\boldsymbol{p} \in \Pi$ is the portfolio process according to which the wealth of the agent is allocated among the N *risky* assets, and $c \in \mathcal{C}$ is the agent's cumulative consumption process. The pair $(c, \boldsymbol{p})$ can be interpreted as "control inputs" that control the wealth of the agent. Using this interpretation, we use the more informative notation $X^{(x,c,\boldsymbol{p})}$ to denote the agent's wealth process. The wealth of the agent can also be expressed in the more usual stochastic differential equation form

$$\begin{aligned} dX^{(x,c,\boldsymbol{p})}(t) &= X^{(x,c,\boldsymbol{p})}(t)\{r(t) + g(t, \boldsymbol{p}(t)) + \boldsymbol{p}(t)^\top \boldsymbol{\sigma}(t)\boldsymbol{\theta}(t)\}dt \\ &\quad + X^{(x,c,\boldsymbol{p})}(t)\boldsymbol{p}(t)^\top \boldsymbol{\sigma}(t) d\mathbf{W}(t) - dc(t), \end{aligned} \tag{5.7}$$

for all $t \in [0, T]$, with the initial condition

$$X^{(x,c,\boldsymbol{p})}(0) = x. \tag{5.8}$$

Remark 5.1.8. As a parallel to Remark 3.1.8 and Remark 4.1.20, the process $\{p_0(t)\}$ does not appear anywhere in (5.6), and only the portfolio process $\boldsymbol{p} \in \Pi$ and consumption process $c \in \mathcal{C}$ appear in the equation. The reason for this is that any change in wealth is caused only by changes in asset prices and consumption. As a result, $p_0(t)$ is a redundant variable which can be eliminated.

Remark 5.1.9. There are two differences when comparing the wealth SDE (5.7) to the SDE (4.23) given in Chapter 4. The first difference is the use of the proportional portfolio process $\boldsymbol{p} \in \Pi$ in place of the monetary portfolio process $\boldsymbol{\pi} \in \Pi$. Using the proportional portfolio process instead of the monetary portfolio process allows us to deal with portfolio constraint set K more naturally. One can easily convert from $\boldsymbol{\pi}$ to $\boldsymbol{p}$ by using Remark 5.1.14. The second difference between (5.7) and (4.23) is the addition of the margin function g. Since the margin function g, which usually becomes more negative as the value of the portfolio increases, is added to the interest rate process r in (5.7), one can interpret the margin term g as a "reduction in interest rate" due to penalties such as the higher cost of borrowing than lending.

The following examples describe scenarios where a portfolio constraint set K and margin function g are used.

Example 5.1.10. Consider a pension fund that is not permitted to invest in the company which it sponsors. If that company's stock is labelled stock N, then the pension fund's portfolio constraint corresponds to the set $K = \{p \in \mathbb{R}^N | p_N = 0\}$ (here p is a vector representing the proportion of wealth invested into each of the N stocks). Furthermore, pension funds are often times subject to a *no-short-selling* constraint. A no-short-selling constraint set is given by $K = \{p \in \mathbb{R}^N | p_1 \geq 0, \ldots, p_N \geq 0\}$. Combining both constraint sets above corresponds to the set $K = \{p \in \mathbb{R}^N | p_1 \geq 0, \ldots, p_{N-1} \geq 0, p_N = 0\}$.

Remark 5.1.11. The constraint set $K = \{p \in \mathbb{R}^N | p_{M+1} = 0, \ldots, p_N = 0\}$ for some $M \in \{1, \ldots, N-1\}$ corresponds to market where certain stocks are *inaccessible*. This market was called the *incomplete market* by Karatzas et al. [29] and El Karoui and Quenez [16][17] in their seminal works.

Example 5.1.12. Let $\{R(t)\}$ be a given uniformly bounded $\{\mathcal{F}_t\}$-progressively measurable process such that $R(t) \geq r(t)$ a.s. for all $t \in [0, T]$. The process $R(t)$ represents the interest the agent pays for *borrowing*, while the process $r(t)$ from Condition 4.1.4 now represents the interest the agent earns from *lending*. We call this amended market model a *market with a higher interest rate for borrowing than for lending*. We can represent such a market by using the margin function

$$g(t, \omega, \boldsymbol{p}) \triangleq [r(t, \omega) - R(t, \omega)](1 - \boldsymbol{p}^\top \mathbf{1})^-, \quad (t, \omega, \boldsymbol{p}) \in [0, T] \times \Omega \times \mathbb{R}^N, \tag{5.9}$$

where $(1 - \boldsymbol{p}^\top \mathbf{1})^- \triangleq \max\{0, \sum_{i=1}^N p_i - 1\}$, in the SDE (5.7). Cuoco and Liu (see section 3 of [7]) extended the function g beyond this simple case to model quite general margin requirements.

Remark 5.1.13. If we are given the set $K = \mathbb{R}^N$ and margin function $g = 0$, the constrained market is the same as the unconstrained market of Chapter 4.

Remark 5.1.14. Using the notation of Chapters 3 and 4, we can define

$$p_0(t, \omega) \triangleq \pi_0(t, \omega)/X(t, \omega) \text{ and } \boldsymbol{p}(t, \omega) \triangleq \boldsymbol{\pi}(t, \omega)/X(t, \omega), \quad (t, \omega) \in [0, T] \times \Omega, \tag{5.10}$$

where π_0 and $\boldsymbol{\pi}$ are the monetary portfolio processes from (4.17), if

$$\int_0^T |g(t, \omega, \boldsymbol{\pi}(t, \omega)/X(t, \omega))| dt < \infty \tag{5.11}$$

on the set $\{X \neq 0\}$. We *always* put $\boldsymbol{p}(t, \omega) \triangleq 0$ on the set $\{X = 0\}$.

5.2 Definition of the Hedging Problem

Similar to Chapter 3 and Chapter 4 a *contingent claim* is a specified strictly positive $\mathcal{F}_T$-measurable random variable B, that is

$$B : (\Omega, \mathcal{F}_T, P) \mapsto (0, \infty). \tag{5.12}$$

The random variable B defines an obligation on the part of the agent in the following sense: starting from an initial wealth $x \in (0, \infty)$, the agent must find some portfolio process $\boldsymbol{p} \in \Pi$ and cumulative consumption process $c \in \mathcal{C}$ such that

$$X^{(x,c,\boldsymbol{p})}(T) = B \text{ a.s.} \tag{5.13}$$

Just as in Chapter 4, the randomness in B is attributed to the joint effect of the Brownian motion $\mathbf{W}(t)$ and the Markov chain $\boldsymbol{\alpha}(t)$.

For reasons described in Remark 4.2.2, the hedging problem is only tractable if we assume that the agent's wealth is non-negative over the entire trading interval. In this chapter, we strengthen this condition by enforcing the agent's wealth must be *strictly positive* over the trading interval as this will be essential for proving our major results (it eliminates the need for division by 0). To enforce this condition, as well as enforce that the agent's portfolio process $\boldsymbol{p}$ be within the constraint set K, defined in Condition 5.1.1, we restrict the agent to portfolio processes $\boldsymbol{p}$ and cumulative consumption processes c to the set $\mathcal{A}(x)$:

$$\mathcal{A}(x) \triangleq \left\{ (c, \boldsymbol{p}) \in \mathcal{C} \times \Pi \middle| \boldsymbol{p}(t) \in K \text{ a.e., } X^{(x,c,\boldsymbol{p})}(t) > 0 \text{ a.s. for all } t \in [0, T] \right\}, \tag{5.14}$$

for some initial wealth $x \in (0, \infty)$. The set $\mathcal{A}(x)$ is called the *admissible set of consumption-portfolio pairs* from initial wealth x.

Remark 5.2.1. Just as in Remark 4.2.1, the set $\mathcal{A}(x)$ is non-empty for every $x \in (0, \infty)$. To see this take $\boldsymbol{p} \in \Pi$ defined by $\boldsymbol{p} \triangleq 0$ and $c \in \mathcal{C}$ defined by $c \triangleq 0$. Since $0 \in K$ by Condition 5.1.1, this is a valid choice for $\boldsymbol{p}$. With this choice of $\boldsymbol{p} \in \Pi$ and $c \in \mathcal{C}$, we have $g(t, \boldsymbol{p}) = 0$ a.s. for all $t \in [0, T]$. One sees from (5.6) and the fact that $x \in (0, \infty)$

$$X^{(x,c,\boldsymbol{p})}(t) = S_0(t)x > 0 \text{ a.s.,} \quad t \in [0, T]. \tag{5.15}$$

Therefore, $(c, \boldsymbol{p}) \in \mathcal{A}(x)$ and $\mathcal{A}(x) \neq \emptyset$ for each $x \in (0, \infty)$.

Remark 5.2.2. Starting from some initial wealth $x \in (0, \infty)$, the agent *hedges the contingent claim* B if he or she can find some admissible consumption-portfolio pair $(c, \boldsymbol{p}) \in \mathcal{A}(x)$ such that (5.13) holds. This consumption-portfolio pair is called a *(constrained) hedging strategy* from initial wealth x.

Just as in Chapter 4, if the initial wealth $x \in (0, \infty)$ is too small, then there may not exist any $(c, \boldsymbol{p}) \in \mathcal{A}(x)$ such that (5.13) holds. That is, there may not exist a hedging strategy from such an initial wealth. The goal of an agent in the hedging problem is to determine the least initial wealth $x \in (0,\infty)$ for which there does exist a hedging strategy $(c,\boldsymbol{p}) \in \mathcal{A}(x)$. To make this precise, we define the set Λ comprising of all initial wealths x for which a hedging strategy does exist. That is

$$\Lambda \triangleq \left\{ x \in (0,\infty) \middle| \exists (c,\boldsymbol{p}) \in \mathcal{A}(x) \quad \text{s.t. } X^{(x,c,\boldsymbol{p})}(T) = B \text{ a.s.} \right\}. \tag{5.16}$$

In view of (5.16) and (5.14) we have

$$\begin{aligned} \Lambda = \Bigg\{ x \in (0,\infty) \Bigg| & X^{(x,c,\boldsymbol{p})}(T) = B \text{ a.s. for some } (c,\boldsymbol{p}) \in \mathcal{C} \times \Pi \\ & \text{such that } X^{(x,c,\boldsymbol{p})}(t) > 0 \text{ a.s. for all } t \in [0,T] \text{ and } \boldsymbol{p}(t) \in K \text{ a.e.} \Bigg\}. \end{aligned} \tag{5.17}$$

Remark 5.2.3. Exactly as in Chapter 4, there is of course no guarantee that Λ is non-empty. If the contingent claim random variable is stipulated to be "unreasonably large" then we shall see later that Λ is actually the empty set. However, assuming that Λ is indeed non-empty, and furthermore, attains its infimum at some $\hat{u} \in (0,\infty)$ (that is $\hat{u} \in \Lambda$ and $\hat{u} \leq x$ for all $x \in \Lambda$), it is evident that $x = \hat{u}$ is the least initial wealth from which one can hedge the contingent claim B. This least initial wealth is usually called the *price of the contingent claim* B and the associated hedging strategy is called the *optimal hedging strategy*. The most important goals of this chapter can now be stated as follows:

(a) Determine conditions on the contingent claim random variable B which ensures that Λ is non-empty

(b) With the conditions in (a) in force, establish that there exists some $\hat{u} \in \Lambda$ such that $\hat{u} \leq x$ for all $x \in \Lambda$. It then follows from (5.16) that $x = \hat{u}$ is the *least initial wealth* from which one can hedge the contingent claim B.

(c) With $x = \hat{u}$ characterize some $(c,\boldsymbol{p}) \in \mathcal{A}(x)$ such that (5.13) holds. This consumption-portfolio pair hedges the contingent claim B from the least initial wealth $\hat{u}$.

5.3 Solution to the Hedging Problem

Having outlined the problem of hedging within a constrained regime-switching market model (see Remark 5.2.3), we shall now address this problem. We again use the space of *Markov chain dual processes* $\mathcal{H}$, given in Notation 4.3.1, which allows the regime-switching Markov chain to interact with the agent's wealth process. We also define the space of *dual processes* $\mathcal{G}$ in Notation 5.3.1 in order to handle the margin function g and constraint $\boldsymbol{p}(t) \in K$ a.e. The space $\mathcal{G} \times \mathcal{H}$ of processes plays a role similar to the space of Lagrange multipliers in classical convex optimization.

Notation 5.3.1. The *joint space of dual processes* is the space $\mathcal{G} \times \mathcal{H}$ where $\mathcal{H}$ is given in Notation 4.3.1 and the space $\mathcal{G}$ is given by

$$\begin{aligned}\mathcal{G} \triangleq \Bigg\{\boldsymbol{\nu} : [0,T] \times \Omega \mapsto \mathbb{R}^N \Bigg| \boldsymbol{\nu} \in \mathcal{F}^*, \int_0^T ||\boldsymbol{\nu}(t)||^2 dt < \infty \ \text{ a.s.}, \\ \int_0^T \sup_{\boldsymbol{p} \in K}\{g(t,\boldsymbol{p}) - \boldsymbol{p}^\top \boldsymbol{\nu}(t)\} dt < \infty \ \text{ a.s.}\Bigg\}.\end{aligned} \tag{5.18}$$

Remark 5.3.2. Since g is uniformly upper-bounded by Condition 5.1.3, it is immediate that $0 \in \mathcal{G}$.

Remark 5.3.3. In Chapters 3 and 4, an essential role in solving the hedging problem was played by state price density processes. Similarly, in this chapter a *generalized state price density process* is defined to handle portfolio constraints and margin requirements. This new generalized state price density process given in Definition 5.3.4 differs from that given in Definition 4.3.3 as it is now a function of $(\boldsymbol{\nu}, \boldsymbol{\mu}) \in \mathcal{G} \times \mathcal{H}$.

Definition 5.3.4. The *generalized state price density process* $\{H_{\boldsymbol{\nu},\boldsymbol{\mu}}(t), t \in [0,T]\}$ for the market model defined in Section 5.1 is the process with values in $(0,\infty)$ defined as follows:

$$H_{\boldsymbol{\nu},\boldsymbol{\mu}}(t) \triangleq \exp\Bigg\{ - \int_0^t [r(s) + \tilde{g}_K(s, \boldsymbol{\nu}(s))] ds \Bigg\} \mathcal{E}(-\boldsymbol{\theta}_{\boldsymbol{\nu}} \bullet \mathbf{W})(t) \mathcal{E}(\boldsymbol{\mu} \bullet \mathbf{M})(t), \tag{5.19}$$

for all $t \in [0,T]$ and $(\boldsymbol{\nu}, \boldsymbol{\mu}) \in \mathcal{G} \times \mathcal{H}$, where

$$\begin{aligned}\boldsymbol{\theta}_{\boldsymbol{\nu}}(t) &\triangleq \boldsymbol{\theta}(t) + \boldsymbol{\sigma}^{-1}(t)\boldsymbol{\nu}(t), \quad t \in [0,T], \boldsymbol{\nu} \in \mathcal{G} \\ \tilde{g}_K(t, \boldsymbol{\nu}) &\triangleq \sup_{\boldsymbol{p} \in K}\{g(t,\boldsymbol{p}) - \boldsymbol{p}^\top \boldsymbol{\nu}\}, \quad t \in [0,T], \boldsymbol{\nu} \in \mathbb{R}^N,\end{aligned} \tag{5.20}$$

and $\mathcal{E}(\cdot)(t)$ is the Doléans-Dade exponential function (see Theorem C.10.1).

Remark 5.3.5. Similar to $H_{\boldsymbol{\mu}}(t)$ defined in Definition 4.3.3, the generalized state price density $H_{\boldsymbol{\nu},\boldsymbol{\mu}}$ can be written as an $\{\mathcal{F}_t\}$-semimartingale using Ito's product rule:

$$\begin{aligned}H_{\boldsymbol{\nu},\boldsymbol{\mu}}(t) = 1 - \int_0^t H_{\boldsymbol{\nu},\boldsymbol{\mu}}(s) \boldsymbol{\theta}_{\boldsymbol{\nu}}(s)^\top d\mathbf{W}(s) - \int_0^t H_{\boldsymbol{\nu},\boldsymbol{\mu}}(s)[r(s) + \tilde{g}_K(s, \boldsymbol{\nu}(s))] ds \\ + \sum_{\substack{i,j \in S \\ i \neq j}} \int_0^t H_{\boldsymbol{\nu},\boldsymbol{\mu}}(s-) \mu_{ij}(s) dM_{ij}(s), \quad t \in [0,T], (\boldsymbol{\nu}, \boldsymbol{\mu}) \in \mathcal{G} \times \mathcal{H}.\end{aligned} \tag{5.21}$$

Furthermore, we have the fact that $H_{\boldsymbol{\nu},\boldsymbol{\mu}}$ is almost surely strictly positive, i.e.

$$\inf_{t\in[0,T]} H_{\boldsymbol{\nu},\boldsymbol{\mu}}(t) > 0 \text{ a.s.}, \quad (\boldsymbol{\nu},\boldsymbol{\mu}) \in \mathcal{G}\times\mathcal{H}. \tag{5.22}$$

Remark 5.3.6. Note that if $K \triangleq \mathbb{R}^N$ and $g(t,\cdot) \triangleq 0$ for all $t \in [0,T]$ we can put $\boldsymbol{\nu} \triangleq 0$ to get

$$H_{0,\boldsymbol{\mu}}(t) = H_{\boldsymbol{\mu}}(t) \text{ a.s.}, \quad t\in[0,T], \boldsymbol{\mu}\in\mathcal{H}, \tag{5.23}$$

where the process $\{H_{\boldsymbol{\mu}}(t), t\in[0,T]\}$, $\boldsymbol{\mu}\in\mathcal{H}$, is defined in Definition 4.3.3.

Remark 5.3.7. The generalized state price density in Definition 5.3.4 is identical to that introduced by Heunis [23] for problems of *utility maximization* in a regime-switching market model with convex portfolio constraints and margin requirements.

Remark 5.3.8. It is elementary to show that the process

$$\begin{aligned} z(t) \triangleq \int_0^t H_{\boldsymbol{\nu},\boldsymbol{\mu}}(s) X^{(x,c,\boldsymbol{p})}(s)\Big[\tilde{g}_K(s,\boldsymbol{\nu}(s)) - \Big(g(s,\boldsymbol{p}(s)) - \boldsymbol{p}(s)^\top \boldsymbol{\nu}(s)\Big)\Big] ds \\ + \int_0^t H_{\boldsymbol{\nu},\boldsymbol{\mu}}(s-)dc(s) + \sum_{0<s\le t} H_{\boldsymbol{\nu},\boldsymbol{\mu}}(s-)\triangle c(s)\triangle(\boldsymbol{\mu}\bullet\mathbf{M})(s), \quad t\in[0,T], \end{aligned} \tag{5.24}$$

is almost surely non-negative for all $t\in[0,T]$, $(c,\boldsymbol{p})\in\mathcal{A}(x)$, and $(\boldsymbol{\nu},\boldsymbol{\mu})\in\mathcal{G}\times\mathcal{H}$. Indeed, by (5.20) and the almost sure positivity of $H_{\boldsymbol{\nu},\boldsymbol{\mu}}X^{(x,c,\boldsymbol{p})}$, the first term on the right-hand side of (5.24) is positive. Since $c\in\mathcal{C}$ must be non-decreasing by (4.19), $\boldsymbol{\mu}\in\mathcal{H}$ must have the property $\triangle(\boldsymbol{\mu}\bullet\mathbf{M})(t) > -1$ a.s. for all $t\in[0,T]$ by (4.39), and $H_{\boldsymbol{\nu},\boldsymbol{\mu}}$ is almost surely strictly positive, we have that $z(t)$ is almost surely non-negative. This fact will be essential for the following results.

We would like to show results similar to Proposition 4.3.8 and Corollary 4.3.9.1 as they are again pertinent to the solution of the constrained hedging problem. As one will see, due to the inclusion of the margin function g in the wealth equation and convex conjugate $\tilde{g}_K$ in the state price density process, the statement of an equivalent proposition is slightly more complex.

Proposition 5.3.9. *For each $x\in(0,\infty)$, $(c,\boldsymbol{p})\in\mathcal{A}(x)$, and $(\boldsymbol{\nu},\boldsymbol{\mu})\in\mathcal{G}\times\mathcal{H}$, the process $\{J^{(x,c,\boldsymbol{p})}_{\boldsymbol{\nu},\boldsymbol{\mu}}(t), t\in[0,T]\}$ defined as*

$$\begin{aligned} J^{(x,c,\boldsymbol{p})}_{\boldsymbol{\nu},\boldsymbol{\mu}}(t) \triangleq \int_0^t H_{\boldsymbol{\nu},\boldsymbol{\mu}}(s) X^{(x,c,\boldsymbol{p})}(s)\Big[\tilde{g}_K(s,\boldsymbol{\nu}(s)) - \Big(g(s,\boldsymbol{p}(s)) - \boldsymbol{p}(s)^\top \boldsymbol{\nu}(s)\Big)\Big] ds \\ + H_{\boldsymbol{\nu},\boldsymbol{\mu}}(t)X^{(x,c,\boldsymbol{p})}(t) + \int_0^t H_{\boldsymbol{\nu},\boldsymbol{\mu}}(s-)dc(s) + \sum_{0<s\le t} H_{\boldsymbol{\nu},\boldsymbol{\mu}}(s-)\triangle c(s)\triangle(\boldsymbol{\mu}\bullet\mathbf{M})(s), \end{aligned} \tag{5.25}$$

for all $t \in [0,T]$, *is a non-negative* $\{\mathcal{F}_t\}$*-supermartingale (i.e.* $J^{(x,c,\boldsymbol{p})}_{\boldsymbol{\nu},\boldsymbol{\mu}} \in \mathcal{SPM}(\{\mathcal{F}_t\}, P)$ *and* $J^{(x,c,\boldsymbol{p})}_{\boldsymbol{\nu},\boldsymbol{\mu}} \geq 0$*).*

Proof. For ease of notation let $X(t,\omega) \triangleq X^{(x,c,\boldsymbol{p})}(t,\omega)$ for all $(t,\omega) \in [0,T] \times \Omega$. By a simple, yet lengthy calculation using Ito's formula, we can write the product of $H_{\boldsymbol{\nu},\boldsymbol{\mu}}(t)X(t)$ as

$$
\begin{aligned}
&H_{\boldsymbol{\nu},\boldsymbol{\mu}}(t)X(t) = x - \int_0^t H_{\boldsymbol{\nu},\boldsymbol{\mu}}(s)X(s)\boldsymbol{\theta}_{\boldsymbol{\nu}}(s)^\top d\mathbf{W}(s) \\
&- \int_0^t H_{\boldsymbol{\nu},\boldsymbol{\mu}}(s)X(s)[r(s) + \tilde{g}_K(s,\boldsymbol{\nu}(s))]ds \\
&+ \int_0^t H_{\boldsymbol{\nu},\boldsymbol{\mu}}(s)X(s)\boldsymbol{p}(s)^\top \boldsymbol{\sigma}(s)d\mathbf{W}(s) + \int_0^t H_{\boldsymbol{\nu},\boldsymbol{\mu}}(s)X(s)[r(s) + g(s,\boldsymbol{p}(s))]ds \\
&+ \int_0^t H_{\boldsymbol{\nu},\boldsymbol{\mu}}(s)X(s)\boldsymbol{p}(s)^\top \boldsymbol{\sigma}(s)[\boldsymbol{\theta}(s) - \boldsymbol{\theta}_{\boldsymbol{\nu}}(s)]ds - \int_0^t H_{\boldsymbol{\nu},\boldsymbol{\mu}}(s-)dc(s) \\
&+ \sum_{\substack{i,j \in S \\ i \neq j}} \int_0^t H_{\boldsymbol{\nu},\boldsymbol{\mu}}(s-)X(s-)\mu_{ij}(s)dM_{ij}(s) - \sum_{0<s\leq t} H_{\boldsymbol{\nu},\boldsymbol{\mu}}(s-)\triangle c(s)\triangle(\boldsymbol{\mu} \bullet \mathbf{M})(s),
\end{aligned}
\tag{5.26}
$$

for all $t \in [0,T]$. By rearranging and simplifying (5.26) with (5.25) we have

$$
\begin{aligned}
J^{(x,c,\boldsymbol{p})}_{\boldsymbol{\nu},\boldsymbol{\mu}}(t) = x &+ \int_0^t H_{\boldsymbol{\nu},\boldsymbol{\mu}}(s)X(s)[\boldsymbol{p}(s)^\top \boldsymbol{\sigma}(s) - \boldsymbol{\theta}_{\boldsymbol{\nu}}(s)^\top]d\mathbf{W}(s) \\
&+ \sum_{\substack{i,j \in S \\ i \neq j}} \int_0^t H_{\boldsymbol{\nu},\boldsymbol{\mu}}(s-)X(s-)\mu_{ij}(s)dM_{ij}(s), \quad t \in [0,T].
\end{aligned}
\tag{5.27}
$$

It's easy to see from (5.27) that $J^{(x,c,\boldsymbol{p})}_{\boldsymbol{\nu},\boldsymbol{\mu}}$ is $\{\mathcal{F}_t\}$-local martingale, and using (5.25) and Remark 5.3.8, $J^{(x,c,\boldsymbol{p})}_{\boldsymbol{\nu},\boldsymbol{\mu}}$ is non-negative. Therefore, from the fact that $J^{(x,c,\boldsymbol{p})}_{\boldsymbol{\nu},\boldsymbol{\mu}}$ is a non-negative $\{\mathcal{F}_t\}$-local martingale, we have from Proposition C.11.3

$$
J^{(x,c,\boldsymbol{p})}_{\boldsymbol{\nu},\boldsymbol{\mu}} \in \mathcal{SPM}(\{\mathcal{F}_t\}, P), \tag{5.28}
$$

that is, $J^{(x,c,\boldsymbol{p})}_{\boldsymbol{\nu},\boldsymbol{\mu}}$ is an $\{\mathcal{F}_t\}$-supermartingale. □

The following corollary is immediate from Proposition 5.3.9 and Remark 5.3.8.

Corollary 5.3.9.1. *For each* $(\boldsymbol{\nu},\boldsymbol{\mu}) \in \mathcal{G} \times \mathcal{H}, x \in (0,\infty), (c,\boldsymbol{p}) \in \mathcal{A}(x)$

$$
E\left[H_{\boldsymbol{\nu},\boldsymbol{\mu}}(T)X^{(x,c,\boldsymbol{p})}(T)\right] \leq x. \tag{5.29}
$$

Remark 5.3.10. Suppose that Λ is non-empty. Then from (5.16), for *each* $x \in \Lambda$ there exists a pair $(c, \boldsymbol{p}) \in \mathcal{A}(x)$ such that

$$X^{(x,c,\boldsymbol{p})}(T) = B \text{ a.s.} \tag{5.30}$$

Using Corollary 5.3.9.1 and (5.30), we have

$$\Lambda \neq \emptyset \implies \sup_{(\boldsymbol{\nu},\boldsymbol{\mu}) \in \mathcal{G}\times\mathcal{H}} E\left[H_{\boldsymbol{\nu},\boldsymbol{\mu}}(T)B\right] \leq x, \quad \text{for all } x \in \Lambda. \tag{5.31}$$

Remark 5.3.11. Note that $\sup_{(\boldsymbol{\nu},\boldsymbol{\mu}) \in \mathcal{G}\times\mathcal{H}} E\left[H_{\boldsymbol{\nu},\boldsymbol{\mu}}(T)B\right]$ is defined in $(0, \infty]$, since, *for each and every* pair $(\boldsymbol{\nu}, \boldsymbol{\mu}) \in \mathcal{G} \times \mathcal{H}$, we have $P[H_{\boldsymbol{\nu},\boldsymbol{\mu}}(T)B > 0] = 1$.

Remark 5.3.12. Once again, the main result of this chapter, Theorem 5.3.17, establishes the converse implication of (5.31), so that we actually have the following equivalence:

$$\Lambda \neq \emptyset \iff \sup_{(\boldsymbol{\nu},\boldsymbol{\mu}) \in \mathcal{G}\times\mathcal{H}} E\left[H_{\boldsymbol{\nu},\boldsymbol{\mu}}(T)B\right] < \infty. \tag{5.32}$$

That is, the contingent claim B can be hedged if and only if $\sup_{(\boldsymbol{\nu},\boldsymbol{\mu}) \in \mathcal{G}\times\mathcal{H}} E\left[H_{\boldsymbol{\nu},\boldsymbol{\mu}}(T)B\right] < \infty$. Thus, if the contingent claim random variable B is stipulated so large that

$$\sup_{(\boldsymbol{\nu},\boldsymbol{\mu}) \in \mathcal{G}\times\mathcal{H}} E\left[H_{\boldsymbol{\nu},\boldsymbol{\mu}}(T)B\right] = \infty \tag{5.33}$$

then there is no possibility of hedging the claim B.

Notation 5.3.13. In the following, an essential role will be played by the *strictly positive* extended real number $\hat{u} \in (0, \infty]$ defined as follows:

$$\hat{u} \triangleq \sup_{(\boldsymbol{\nu},\boldsymbol{\mu}) \in \mathcal{G}\times\mathcal{H}} E\left[H_{\boldsymbol{\nu},\boldsymbol{\mu}}(T)B\right]. \tag{5.34}$$

The choice $\hat{u}$ will be our candidate for the least initial wealth that can hedge the contingent claim B.

Remark 5.3.14. Since $H_{\boldsymbol{\nu},\boldsymbol{\mu}}(T)B > 0$ a.s. for all $(\boldsymbol{\nu}, \boldsymbol{\mu}) \in \mathcal{G}\times\mathcal{H}$, the conditional expectation $E[H_{\boldsymbol{\nu},\boldsymbol{\mu}}(T)B|\mathcal{F}_t]$ is almost surely non-negative for all $(\boldsymbol{\nu}, \boldsymbol{\mu}) \in \mathcal{G}\times\mathcal{H}$ at each time $t \in [0, T]$. Furthermore, since $\inf_{t\in[0,T]} H_{\boldsymbol{\nu},\boldsymbol{\mu}}(t) > 0$ a.s. for all $(\boldsymbol{\nu}, \boldsymbol{\mu}) \in \mathcal{G} \times \mathcal{H}$, we have for each fixed $t \in [0, T]$ that

$$\hat{X}_{\boldsymbol{\nu},\boldsymbol{\mu}}(t) \triangleq \frac{E[H_{\boldsymbol{\nu},\boldsymbol{\mu}}(T)B|\mathcal{F}_t]}{H_{\boldsymbol{\nu},\boldsymbol{\mu}}(t)} \geq 0 \text{ a.s.}, \quad (\boldsymbol{\nu}, \boldsymbol{\mu}) \in \mathcal{G} \times \mathcal{H}. \tag{5.35}$$

By Theorem C.13.2, for each fixed $t \in [0, T]$ the essential-supremum (see Definition C.13.1) of the family of random variables $\{\hat{X}_{\boldsymbol{\nu},\boldsymbol{\mu}}(t), (\boldsymbol{\nu}, \boldsymbol{\mu}) \in \mathcal{G} \times \mathcal{H}\}$ exists. That is

$$\operatorname*{ess-sup}_{(\boldsymbol{\nu},\boldsymbol{\mu}) \in \mathcal{G}\times\mathcal{H}} \hat{X}_{\boldsymbol{\nu},\boldsymbol{\mu}}(t) = \operatorname*{ess-sup}_{(\boldsymbol{\nu},\boldsymbol{\mu}) \in \mathcal{G}\times\mathcal{H}} \frac{E\left[H_{\boldsymbol{\nu},\boldsymbol{\mu}}(T)B \,\middle|\, \mathcal{F}_t\right]}{H_{\boldsymbol{\nu},\boldsymbol{\mu}}(t)} \text{ exists for each } t \in [0, T]. \tag{5.36}$$

As a parallel of Proposition 4.3.15, we have the following statement, the proof of which is technical and therefore given in Appendix B.4.

Proposition 5.3.15. *Suppose $\hat{u} < \infty$. Define the process $\{\hat{X}(t), t \in [0,T]\}$ as follows,*

$$\hat{X}(t) \triangleq \operatorname*{ess\text{-}sup}_{(\boldsymbol{\nu},\boldsymbol{\mu})\in\mathcal{G}\times\mathcal{H}} \frac{E\left[H_{\boldsymbol{\nu},\boldsymbol{\mu}}(T)B \,\middle|\, \mathcal{F}_t\right]}{H_{\boldsymbol{\nu},\boldsymbol{\mu}}(t)}, \quad t \in [0,T], \tag{5.37}$$

where the right-hand side of (5.37) exists by Remark 5.3.14. Then the following hold:

(1) $\hat{X}(0) = \hat{u}$ a.s.

(2) $\hat{X}(T) = B$ a.s.

(3) $\hat{X}(t) > 0$ a.s., $\quad t \in [0,T]$

(4) $H_{\boldsymbol{\nu},\boldsymbol{\mu}}\hat{X} \in \mathcal{SPM}(\{\mathcal{F}_t\}, P)$, $\quad (\boldsymbol{\nu},\boldsymbol{\mu}) \in \mathcal{G}\times\mathcal{H}$

(5) For each $0 \le s \le t \le T$

$$\hat{X}(s) = \operatorname*{ess\text{-}sup}_{(\boldsymbol{\nu},\boldsymbol{\mu})\in\mathcal{G}\times\mathcal{H}} \frac{E[H_{\boldsymbol{\nu},\boldsymbol{\mu}}(t)\hat{X}(t)|\mathcal{F}_s]}{H_{\boldsymbol{\nu},\boldsymbol{\mu}}(s)} \text{ a.s.} \tag{5.38}$$

(6) $\{\hat{X}(t), t \in [0,T]\}$ has a càdlàg modification

(7) For each $\{\mathcal{F}_t\}$-stopping time τ taking values in $[0,T]$, there exists a càdlàg modification of $\hat{X}$ such that

$$\hat{X}(\tau) = \operatorname*{ess\text{-}sup}_{(\boldsymbol{\nu},\boldsymbol{\mu})\in\mathcal{G}\times\mathcal{H}} \frac{E[H_{\boldsymbol{\nu},\boldsymbol{\mu}}(T)B|\mathcal{F}_\tau]}{H_{\boldsymbol{\nu},\boldsymbol{\mu}}(\tau)} \text{ a.s.} \tag{5.39}$$

We now state a decomposition result involving $\hat{X}$ and the generalized price state density $H_{\boldsymbol{\nu},\boldsymbol{\mu}}$ that is absolutely essential for proving our hedging results.

Lemma 5.3.16. *For each $(\boldsymbol{\nu},\boldsymbol{\mu}) \in \mathcal{G}\times\mathcal{H}$, the process $\{H_{\boldsymbol{\nu},\boldsymbol{\mu}}(t)\hat{X}(t), t \in [0,T]\}$ has the decomposition*

$$H_{\boldsymbol{\nu},\boldsymbol{\mu}}(t)\hat{X}(t) = \hat{X}(0) + \int_0^t \boldsymbol{\Psi}_{\boldsymbol{\nu},\boldsymbol{\mu}}(s)^\top d\mathbf{W}(s) + \sum_{\substack{i,j\in S\\ i\neq j}} \int_0^t \Gamma_{ij}^{\boldsymbol{\nu},\boldsymbol{\mu}}(s)dM_{ij}(s) - A_{\boldsymbol{\nu},\boldsymbol{\mu}}(t), \tag{5.40}$$

for all $t \in [0,T]$ *and* $(\boldsymbol{\nu},\boldsymbol{\mu}) \in \mathcal{G} \times \mathcal{H}$*, where* $\boldsymbol{\Psi}_{\boldsymbol{\nu},\boldsymbol{\mu}}$ *is a unique, null-at-the-origin* $\{\mathcal{F}_t\}$*-progressively measurable process such that*

$$\int_0^T ||\boldsymbol{\Psi}_{\boldsymbol{\nu},\boldsymbol{\mu}}(s)||^2 ds < \infty \ a.s., \tag{5.41}$$

$\boldsymbol{\Gamma}^{\boldsymbol{\nu},\boldsymbol{\mu}} \triangleq \{\Gamma_{ij}^{\boldsymbol{\nu},\boldsymbol{\mu}}\}_{i,j\in S}$ *is a unique, null-at-the-origin* $\{\mathcal{F}_t\}$*-predictably measurable process such that*

$$\sum_{\substack{i,j\in S\\ i\neq j}} \left(\int_0^T |\Gamma_{ij}^{\boldsymbol{\nu},\boldsymbol{\mu}}(s)|^2 d[M_{ij}](s) \right)^{1/2} < \infty \ a.s. \tag{5.42}$$

and $A_{\boldsymbol{\nu},\boldsymbol{\mu}}$ *is a unique, null-at-the-origin and non-decreasing* $\{\mathcal{F}_t\}$*-predictably measurable process such that*

$$E\left[A_{\boldsymbol{\nu},\boldsymbol{\mu}}(T)\right] < \infty. \tag{5.43}$$

Proof. Fix $(\boldsymbol{\nu},\boldsymbol{\mu}) \in \mathcal{G} \times \mathcal{H}$. From Proposition 5.3.15(4) we know that $H_{\boldsymbol{\nu},\boldsymbol{\mu}}\hat{X}$ is an $\{\mathcal{F}_t\}$-supermartingale. Using the Doob-Meyer decomposition from Theorem C.11.4, we can write any $\{\mathcal{F}_t\}$-supermartingale as the difference between a unique $\{\mathcal{F}_t\}$-local martingale, null at the origin, and a unique, non-decreasing, $\{\mathcal{F}_t\}$-predictable càdlàg process. In the case of $H_{\boldsymbol{\nu},\boldsymbol{\mu}}\hat{X}$, we can write

$$H_{\boldsymbol{\nu},\boldsymbol{\mu}}(t)\hat{X}(t) = \hat{X}(0) + \Phi_{\boldsymbol{\nu},\boldsymbol{\mu}}(t) - A_{\boldsymbol{\nu},\boldsymbol{\mu}}(t), \quad t \in [0,T], \tag{5.44}$$

where

$$\Phi_{\boldsymbol{\nu},\boldsymbol{\mu}} \in \mathcal{M}_{loc,0}(\{\mathcal{F}_t\}, P) \tag{5.45}$$

and

$$A_{\boldsymbol{\nu},\boldsymbol{\mu}} \in \mathcal{A}_0^+(\{\mathcal{F}_t\}, P) \text{ and } A_{\boldsymbol{\nu},\boldsymbol{\mu}} \in \mathcal{P}^*, \tag{5.46}$$

where the space $\mathcal{A}_0^+$ is defined in Notation C.5.5. Furthermore, since

$$E\left[H_{\boldsymbol{\nu},\boldsymbol{\mu}}(t)\hat{X}(t)\right] > -\infty, \quad t \in [0,T], \tag{5.47}$$

it follows that

$$E\left[A_{\boldsymbol{\nu},\boldsymbol{\mu}}(T)\right] < \infty, \tag{5.48}$$

again by Theorem C.11.4. We can further expand the $\{\mathcal{F}_t\}$-local martingale $\Phi_{\boldsymbol{\nu},\boldsymbol{\mu}}$ into more insightful local martingales using the martingale representation theorem (see Appendix

C.12, specifically Theorem C.12.4), which in the present case states that we can write $\Phi_{\boldsymbol{\nu},\boldsymbol{\mu}}$ uniquely in the following way:

$$\Phi_{\boldsymbol{\nu},\boldsymbol{\mu}}(t) = \int_0^t \boldsymbol{\Psi}_{\boldsymbol{\nu},\boldsymbol{\mu}}(s)^\top d\mathbf{W}(s) + \sum_{\substack{i,j\in S\\ i\neq j}} \int_0^t \Gamma_{ij}^{\boldsymbol{\nu},\boldsymbol{\mu}}(s)dM_{ij}(s), \quad t\in[0,T], \tag{5.49}$$

in which

$$\boldsymbol{\Psi}_{\boldsymbol{\nu},\boldsymbol{\mu}} \in \mathcal{F}^*, \tag{5.50}$$

and

$$\boldsymbol{\Gamma}^{\boldsymbol{\nu},\boldsymbol{\mu}} \triangleq \{\Gamma_{ij}^{\boldsymbol{\nu},\boldsymbol{\mu}}\}_{i,j\in S},\ \boldsymbol{\Gamma}^{\boldsymbol{\nu},\boldsymbol{\mu}} \in \mathcal{P}^*, \tag{5.51}$$

are unique integrand processes such that

$$\int_0^T ||\boldsymbol{\Psi}_{\boldsymbol{\nu},\boldsymbol{\mu}}(s)||^2 ds < \infty \text{ a.s.} \quad \text{and} \quad \sum_{\substack{i,j\in S\\ i\neq j}} \left(\int_0^T |\Gamma_{ij}^{\boldsymbol{\nu},\boldsymbol{\mu}}(s)|^2 d[M_{ij}](s)\right)^{1/2} < \infty \text{ a.s.} \tag{5.52}$$

Combining (5.49) and (5.44) we can write $H_{\boldsymbol{\nu},\boldsymbol{\mu}}\hat{X}$ as

$$H_{\boldsymbol{\nu},\boldsymbol{\mu}}(t)\hat{X}(t) = \hat{X}(0) + \int_0^t \boldsymbol{\Psi}_{\boldsymbol{\nu},\boldsymbol{\mu}}(s)^\top d\mathbf{W}(s) + \sum_{\substack{i,j\in S\\ i\neq j}} \int_0^t \Gamma_{ij}^{\boldsymbol{\nu},\boldsymbol{\mu}}(s)dM_{ij}(s) - A_{\boldsymbol{\nu},\boldsymbol{\mu}}(t), \tag{5.53}$$

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

The main hedging result of this chapter is given by the following theorem:

Theorem 5.3.17. *Suppose that $\hat{u} < \infty$. Then $\hat{u} \in \Lambda$ (in particular $\Lambda \neq \emptyset$) and $\hat{u} = \inf \Lambda$, that is $\hat{u} \leq x$ for all $x \in \Lambda$.*

Remark 5.3.18. Theorem 5.3.17 is written word-for-word the same as Theorem 4.3.17; however, due to the inclusion of portfolio constraints and the margin function g, the proof becomes far more challenging.

Remark 5.3.19. From Theorem 5.3.17 and (5.14), one sees that $\hat{u}$ is the *least initial wealth* which guarantees that $X^{(\hat{u},c,\boldsymbol{p})}(T) = B$ a.s. for some $(c,\boldsymbol{p}) \in \mathcal{A}(\hat{u})$, that is $\hat{u}$ is the *price* of the contingent claim B.

Proof. We are going to establish that there exists some admissible consumption-portfolio pair $(\hat{c},\hat{\boldsymbol{p}}) \in \mathcal{A}(\hat{u})$ such that

$$\hat{X}(t) = X^{(\hat{u},\hat{c},\hat{\boldsymbol{p}})}(t) \text{ a.s.}, \quad t \in [0,T], \tag{5.54}$$

where, from (5.6), the process $X^{(\hat{u},\hat{c},\hat{\boldsymbol{p}})}$ satisfies the stochastic integral equation

$$\begin{aligned} X^{(\hat{u},\hat{c},\hat{\boldsymbol{p}})}(t) = \hat{u} + \int_0^t X^{(\hat{u},\hat{c},\hat{\boldsymbol{p}})}(s)\Big\{r(s) + g(s,\hat{\boldsymbol{p}}(s)) + \hat{\boldsymbol{p}}(s)^\top \boldsymbol{\sigma}(s)\boldsymbol{\theta}(s)\Big\}ds \\ + \int_0^t X^{(\hat{u},\hat{c},\hat{\boldsymbol{p}})}(s)\hat{\boldsymbol{p}}(s)^\top \boldsymbol{\sigma}(s) d\mathbf{W}(s) - \hat{c}(t), \quad t \in [0,T]. \end{aligned} \tag{5.55}$$

If, in fact, $\hat{X}(\cdot)$ can be written in the form of a wealth equation, then from Proposition 5.3.15(2) and 5.3.15(3) we would know that there exists a triple $(\hat{u},\hat{c},\hat{\boldsymbol{p}}) \in (0,\infty) \times \mathcal{A}(\hat{u})$ such that

$$X^{(\hat{u},\hat{c},\hat{\boldsymbol{p}})}(T) = B \text{ a.s.} \tag{5.56}$$

Since $\hat{u}$ is the initial wealth of the wealth process $X^{(\hat{u},\hat{c},\hat{\boldsymbol{p}})}$, this would mean from (5.16)

$$\hat{u} \in \Lambda \text{ and therefore } \Lambda \neq \emptyset. \tag{5.57}$$

Furthermore, from Remark 5.3.10, we have that $\hat{u} \leq \inf \Lambda$, and therefore, from (5.57),

$$\hat{u} = \inf \Lambda \in \Lambda. \tag{5.58}$$

Thus, it remains to establish (5.54), as from (5.57) and (5.58), we will have proven Theorem 5.3.17.

To begin, fix $(\boldsymbol{\nu},\boldsymbol{\mu}) \in \mathcal{G} \times \mathcal{H}$. Isolating $\hat{X}(t)$ in the decomposition (5.40) given in Lemma 5.3.16 and using Proposition 5.3.15(1) we get the expression

$$\hat{X}(t) = \frac{1}{H_{\boldsymbol{\nu},\boldsymbol{\mu}}(t)}\Big\{\hat{u} + \int_0^t \boldsymbol{\Psi}_{\boldsymbol{\nu},\boldsymbol{\mu}}(s)^\top d\mathbf{W}(s) + \sum_{\substack{i,j\in S \\ i\neq j}} \int_0^t \Gamma_{ij}^{\boldsymbol{\nu},\boldsymbol{\mu}}(s) dM_{ij}(s) - A_{\boldsymbol{\nu},\boldsymbol{\mu}}(t)\Big\}, \tag{5.59}$$

for all $t \in [0,T]$. By using the Ito product rule, the identity (5.21), and the fact that $[M_{ij}](t) = R_{ij}(t)$ for all $t \in [0,T]$ and $i,j \in S, i \neq j$, a simple but tedious calculation allows us to expand the right hand side of (5.59) to be the following

$$
\begin{aligned}
\hat{X}(t) &= \hat{u} + \int_0^t \left(\frac{\boldsymbol{\Psi}_{\boldsymbol{\nu},\boldsymbol{\mu}}(s)^\top}{H_{\boldsymbol{\nu},\boldsymbol{\mu}}(s)} + \hat{X}(s)\boldsymbol{\theta}_{\boldsymbol{\nu}}(s)^\top \right) d\mathbf{W}(s) \\
&\quad + \int_0^t \left(\frac{\boldsymbol{\Psi}_{\boldsymbol{\nu},\boldsymbol{\mu}}(s)^\top}{H_{\boldsymbol{\nu},\boldsymbol{\mu}}(s)} \boldsymbol{\theta}_{\boldsymbol{\nu}}(s) + \hat{X}(s)[r(s) + \tilde{g}_K(s, \boldsymbol{\nu}(s))] + \hat{X}(s)\|\boldsymbol{\theta}_{\boldsymbol{\nu}}(s)\|^2 \right) ds \\
&\quad + \sum_{\substack{i,j\in S \\ i\neq j}} \int_0^t \left(\frac{\Gamma_{ij}^{\boldsymbol{\nu},\boldsymbol{\mu}}(s)}{H_{\boldsymbol{\nu},\boldsymbol{\mu}}(s-)} - \hat{X}(s-)\mu_{ij}(s) \right) dM_{ij}(s) \\
&\quad + \sum_{\substack{i,j\in S \\ i\neq j}} \int_0^t \left(\frac{\triangle A_{\boldsymbol{\nu},\boldsymbol{\mu}}(s)}{H_{\boldsymbol{\nu},\boldsymbol{\mu}}(s-)} - \frac{\Gamma_{ij}^{\boldsymbol{\nu},\boldsymbol{\mu}}(s)}{H_{\boldsymbol{\nu},\boldsymbol{\mu}}(s-)} + \hat{X}(s-)\mu_{ij}(s) \right) \tilde{\mu}_{ij}(s) dR_{ij}(s) \\
&\quad - \int_0^t \frac{1}{H_{\boldsymbol{\nu},\boldsymbol{\mu}}(s-)} dA_{\boldsymbol{\nu},\boldsymbol{\mu}}(s), \quad t \in [0,T],
\end{aligned}
\tag{5.60}
$$

where we have put

$$
\tilde{\mu}_{ij}(\cdot) \triangleq \mu_{ij}(\cdot)/(1+\mu_{ij}(\cdot)). \tag{5.61}
$$

Since $\boldsymbol{\mu} \in \mathcal{H}$ was arbitrarily chosen, we can of course take $\boldsymbol{\mu}(t) \triangleq 0$ for all $t \in [0,T]$ at (5.60) to get

$$
\begin{aligned}
\hat{X}(t) &= \hat{u} + \int_0^t \left(\frac{\boldsymbol{\Psi}_{\boldsymbol{\nu},0}(s)^\top}{H_{\boldsymbol{\nu},0}(s)} + \hat{X}(s)\boldsymbol{\theta}_{\boldsymbol{\nu}}(s)^\top \right) d\mathbf{W}(s) \\
&\quad + \int_0^t \left(\frac{\boldsymbol{\Psi}_{\boldsymbol{\nu},0}(s)^\top}{H_{\boldsymbol{\nu},0}(s)} \boldsymbol{\theta}_{\boldsymbol{\nu}}(s) + \hat{X}(s)\Big[r(s) + \tilde{g}_K(s, \boldsymbol{\nu}(s)) + \|\boldsymbol{\theta}_{\boldsymbol{\nu}}(s)\|^2 \Big] \right) ds \\
&\quad + \sum_{\substack{i,j\in S \\ i\neq j}} \int_0^t \frac{\Gamma_{ij}^{\boldsymbol{\nu},0}(s)}{H_{\boldsymbol{\nu},0}(s-)} dM_{ij}(s) - \int_0^t \frac{1}{H_{\boldsymbol{\nu},0}(s-)} dA_{\boldsymbol{\nu},0}(s), \quad t \in [0,T].
\end{aligned}
\tag{5.62}
$$

Now set

$$
\zeta_{\boldsymbol{\nu}}(t) \triangleq \hat{X}(t)\left[\boldsymbol{\theta}_{\boldsymbol{\nu}}(t) + \frac{\boldsymbol{\Psi}_{\boldsymbol{\nu},0}(t)}{H_{\boldsymbol{\nu},0}(t)\hat{X}(t)} \right], \quad t \in [0,T], \tag{5.63}
$$

and

$$
\hat{c}_{\boldsymbol{\nu}}(t) \triangleq \int_0^t \frac{1}{H_{\boldsymbol{\nu},0}(s-)} dA_{\boldsymbol{\nu},0}(s) - \sum_{\substack{i,j\in S \\ i\neq j}} \int_0^t \frac{\Gamma_{ij}^{\boldsymbol{\nu},0}(s)}{H_{\boldsymbol{\nu},0}(s-)} dM_{ij}(s), \quad t \in [0,T]. \tag{5.64}
$$

We now state the following proposition which establishes $\hat{c}_{\boldsymbol{\nu}}(t)$ is a valid cumulative consumption process for each $\boldsymbol{\nu} \in \mathcal{G}$. The proof of Proposition 5.3.20, which can be found in Appendix B.5, is very similar to that of Proposition 4.3.21 for the case without portfolio constraints and margin requirements and is given in full to illustrate the role played by the dual process $\boldsymbol{\nu} \in \mathcal{G}$.

Proposition 5.3.20. *The process $\{\hat{c}_{\boldsymbol{\nu}}(t), t \in [0,T]\}$ defined by*

$$\hat{c}_{\boldsymbol{\nu}}(t) \triangleq \int_0^t \frac{1}{H_{\boldsymbol{\nu},0}(s-)} dA_{\boldsymbol{\nu},0}(s) - \sum_{\substack{i,j \in S \\ i \neq j}} \int_0^t \frac{\Gamma_{ij}^{\boldsymbol{\nu},0}(s)}{H_{\boldsymbol{\nu},0}(s-)} dM_{ij}(s), \quad t \in [0,T], \tag{5.65}$$

is a cumulative consumption process for each $\boldsymbol{\nu} \in \mathcal{G}$. That is $\hat{c}_{\boldsymbol{\nu}} \in \mathcal{C}$ for each $\boldsymbol{\nu} \in \mathcal{G}$.

Continuing the proof of Theorem 5.3.17, putting (5.63) and (5.64) into (5.62) yields

$$\begin{aligned} \hat{X}(t) = \hat{u} + \int_0^t \zeta_{\boldsymbol{\nu}}(s)^\top d\mathbf{W}(s) + \int_0^t \hat{X}(s) \Big[r(s) + \tilde{g}_K(s, \boldsymbol{\nu}(s)) \Big] ds \\ + \int_0^t \hat{X}(s) \frac{\zeta_{\boldsymbol{\nu}}(s)^\top \boldsymbol{\theta}_{\boldsymbol{\nu}}(s)}{\hat{X}(s)} ds - \hat{c}_{\boldsymbol{\nu}}(t), \quad t \in [0,T]. \end{aligned} \tag{5.66}$$

Fixing another $\boldsymbol{\rho} \in \mathcal{G}$, we can write the process $\hat{X}$ as

$$\begin{aligned} \hat{X}(t) = \hat{u} + \int_0^t \zeta_{\boldsymbol{\rho}}(s)^\top d\mathbf{W}(s) + \int_0^t \hat{X}(s) \Big[r(s) + \tilde{g}_K(s, \boldsymbol{\rho}(s)) \Big] ds \\ + \int_0^t \hat{X}(s) \frac{\zeta_{\boldsymbol{\rho}}(s)^\top \boldsymbol{\theta}_{\boldsymbol{\rho}}(s)}{\hat{X}(s)} ds - \hat{c}_{\boldsymbol{\rho}}(t), \quad t \in [0,T]. \end{aligned} \tag{5.67}$$

By Proposition C.8.6 we can equate the continuous local martingale parts of (5.66) and (5.67), resulting in

$$\zeta_{\boldsymbol{\nu}}(t) = \zeta_{\boldsymbol{\rho}}(t), \ \lambda \otimes P \text{ a.e. on } [0,T] \times \Omega, \text{ for all } \boldsymbol{\nu}, \boldsymbol{\rho} \in \mathcal{G}. \tag{5.68}$$

As a result of (5.68), we define our candidate optimal portfolio process $\{\hat{\boldsymbol{p}}(t), t \in [0,T]\}$ as follows:

$$\begin{aligned} \hat{\boldsymbol{p}}(t) &\triangleq \frac{[\boldsymbol{\sigma}^{-1}(t)]^\top \zeta_{\boldsymbol{\nu}}(t)}{\hat{X}(t)} \\ &= [\boldsymbol{\sigma}^{-1}(t)]^\top \left[\boldsymbol{\theta}_{\boldsymbol{\nu}}(t) + \frac{\boldsymbol{\Psi}_{\boldsymbol{\nu},0}(t)}{H_{\boldsymbol{\nu},0}(t)\hat{X}(t)} \right], \quad t \in [0,T], \boldsymbol{\nu} \in \mathcal{G}, \end{aligned} \tag{5.69}$$

where we put $\hat{\boldsymbol{p}} \triangleq 0$ on the zero-measure set $\{\hat{X} = 0\}$.

Proposition 5.3.21. *The process* $\{\hat{\boldsymbol{p}}(t), t \in [0,T]\}$ *defined as*

$$\hat{\boldsymbol{p}}(t) \triangleq [\boldsymbol{\sigma}^{-1}(t)]^\top \left[\boldsymbol{\theta}_{\boldsymbol{\nu}}(t) + \frac{\boldsymbol{\Psi}_{\boldsymbol{\nu},0}(t)}{H_{\boldsymbol{\nu},0}(t)\hat{X}(t)}\right], \quad t \in [0,T], \boldsymbol{\nu} \in \mathcal{G}, \tag{5.70}$$

is a portfolio process, i.e. $\hat{\boldsymbol{p}} \in \Pi$.

Proof. Fix $\boldsymbol{\nu} = 0$. Similar to Proposition 4.3.20, we immediately have that $\hat{\boldsymbol{p}} \in \mathcal{F}^*$. From (5.70),

$$\begin{aligned}\left\{\int_0^T \|\hat{\boldsymbol{p}}(t)\|^2 dt\right\}^{1/2} &\leq \left\{\int_0^T \|[\boldsymbol{\sigma}^{-1}(t)]^\top \boldsymbol{\theta}(t)\|^2 dt\right\}^{1/2} \\ &\quad + \left\{\int_0^T \left\|[\boldsymbol{\sigma}^{-1}(t)]^\top \frac{\boldsymbol{\Psi}_{0,0}(t)}{H_{0,0}(t)\hat{X}(t)}\right\|^2 dt\right\}^{1/2} \text{ a.s.}\end{aligned} \tag{5.71}$$

From the Lemma 5.3.16, Conditions 4.1.1 and 4.1.4, (5.22), Proposition 5.3.15(6), and (5.20) we have that

$$\left\{\int_0^T \|\hat{\boldsymbol{p}}(t)\|^2 dt\right\}^{1/2} \leq A\left\{\int_0^T \|\boldsymbol{\theta}(t)\|^2 dt\right\}^{1/2} + B, \tag{5.72}$$

where A and B are non-negative real constants. From the fact that $\boldsymbol{\theta}$ is uniformly bounded,

$$\hat{\boldsymbol{p}} \in \mathcal{F}^*, \quad \int_0^T \|\hat{\boldsymbol{p}}(t)\|^2 dt < \infty \text{ a.s.} \tag{5.73}$$

For $\hat{\boldsymbol{p}} \in \Pi$ to be true, we require $\int_0^T |g(t, \hat{\boldsymbol{p}}(t))(t)| dt < \infty$ a.s. By Condition 5.1.3

$$|g(t,\omega,\hat{\boldsymbol{p}}(t,\omega))| \leq M(t,\omega)\|\hat{\boldsymbol{p}}(t,\omega)\|, \quad (t,\omega) \in [0,T] \times \Omega. \tag{5.74}$$

Since M and $\hat{\boldsymbol{p}}$ are almost surely square integrable by Condition 5.1.3 and (5.73), respectively, we have by (5.74),

$$\begin{aligned}\int_0^T |g(t,\hat{\boldsymbol{p}}(t))| dt &\leq \int_0^T M(t)\|\hat{\boldsymbol{p}}(t)\| dt \\ &\leq \left(\int_0^T |M(t)|^2 dt\right)^{1/2} \left(\int_0^T \|\hat{\boldsymbol{p}}(t)\|^2 dt\right)^{1/2} \\ &< \infty \ \text{ a.s.}\end{aligned} \tag{5.75}$$

Therefore, from (5.73) and (5.75),

$$\hat{\boldsymbol{p}} \in \Pi. \tag{5.76}$$

□

Continuing our proof of Theorem 5.3.17, we expand (5.66) with (5.69) to obtain

$$\begin{aligned}\hat{X}(t) &= \hat{u} + \int_0^t \hat{X}(s)\Big\{r(s) + \tilde{g}_K(s,\boldsymbol{\nu}(s)) + \hat{\boldsymbol{p}}(s)^\top \boldsymbol{\sigma}(s)\boldsymbol{\theta}(s)\Big\}ds \\ &+ \int_0^t \hat{X}(s)\hat{\boldsymbol{p}}(s)^\top \boldsymbol{\nu}(s)ds + \int_0^t \hat{X}(s)\hat{\boldsymbol{p}}(s)^\top \boldsymbol{\sigma}(s)d\mathbf{W}(s) - \hat{c}_{\boldsymbol{\nu}}(t), \quad t \in [0,T], \boldsymbol{\nu} \in \mathcal{G}.\end{aligned} \tag{5.77}$$

Since the margin term g is not included in (5.77), we add and subtract $\int_0^\cdot \hat{X}(s)g(s,\hat{\boldsymbol{p}}(s))ds$ in (5.77) to get

$$\begin{aligned}\hat{X}(t) &= \hat{u} + \int_0^t \hat{X}(s)\Big\{r(s) + g(s,\hat{\boldsymbol{p}}(s)) + \hat{\boldsymbol{p}}(s)^\top \boldsymbol{\sigma}(s)\boldsymbol{\theta}(s)\Big\}ds + \int_0^t \hat{X}(s)\hat{\boldsymbol{p}}(s)^\top \boldsymbol{\sigma}(s)d\mathbf{W}(s) \\ &+ \int_0^t \hat{X}(s)\Big\{\tilde{g}_K(s,\boldsymbol{\nu}(s)) - \Big[g(s,\hat{\boldsymbol{p}}(s)) - \hat{\boldsymbol{p}}(s)^\top \boldsymbol{\nu}(s)\Big]\Big\}ds - \hat{c}_{\boldsymbol{\nu}}(t), \quad t \in [0,T], \boldsymbol{\nu} \in \mathcal{G}.\end{aligned} \tag{5.78}$$

Define the process $\{\hat{c}(t), t \in [0,T]\}$, which will be our candidate optimal consumption process, to be

$$\hat{c}(t) \triangleq -\int_0^t \hat{X}(s)\Big\{\tilde{g}_K(s,\boldsymbol{\nu}(s)) - \Big[g(s,\hat{\boldsymbol{p}}(s)) - \hat{\boldsymbol{p}}(s)^\top \boldsymbol{\nu}(s)\Big]\Big\}ds + \hat{c}_{\boldsymbol{\nu}}(t), \tag{5.79}$$

for all $t \in [0,T]$ and $\boldsymbol{\nu} \in \mathcal{G}$.

Remark 5.3.22. Note that the process $\{\hat{c}(t)\}$ defined in (5.79) is not dependent on the parameter $\boldsymbol{\nu} \in \mathcal{G}$. This fact will be crucial when showing that $\hat{c}$ is a cumulative consumption process.

Using (5.78) and (5.79) we can write the process $\hat{X}$ in the form of a wealth equation

$$\begin{aligned}\hat{X}(t) &= \hat{u} + \int_0^t \hat{X}(s)\Big\{r(s) + g(s,\hat{\boldsymbol{p}}(s)) + \hat{\boldsymbol{p}}(s)^\top \boldsymbol{\sigma}(s)\boldsymbol{\theta}(s)\Big\}ds \\ &+ \int_0^t \hat{X}(s)\hat{\boldsymbol{p}}(s)^\top \boldsymbol{\sigma}(s)d\mathbf{W}(s) - \hat{c}(t), \quad t \in [0,T].\end{aligned} \tag{5.80}$$

However, for (5.80) to be considered a bona-fide wealth equation, we require $\hat{c} \in \mathcal{C}$. To show this, we first consider the following proposition.

Proposition 5.3.23. *For any $\bar{\boldsymbol{p}} \in \Pi$ there exists a $\bar{\boldsymbol{\nu}} \in \mathcal{G}$ such that*

$$\tilde{g}_K(t,\omega,\bar{\boldsymbol{\nu}}(t,\omega)) \leq g(t,\omega,\bar{\boldsymbol{p}}(t,\omega)) - \bar{\boldsymbol{p}}(t,\omega)^\top \bar{\boldsymbol{\nu}}(t,\omega), \quad (t,\omega) \in [0,T] \times \Omega. \tag{5.81}$$

Furthermore, if $\bar{\boldsymbol{p}}(t,\omega) \in K$ for some $(t,\omega) \in [0,T] \times \Omega$ then,

$$\tilde{g}_K(t,\omega,\bar{\boldsymbol{\nu}}(t,\omega)) = g(t,\omega,\bar{\boldsymbol{p}}(t,\omega)) - \bar{\boldsymbol{p}}(t,\omega)^\top \bar{\boldsymbol{\nu}}(t,\omega). \tag{5.82}$$

Proof. Fix some $\bar{\boldsymbol{p}} \in \Pi$. From Definition D.0.9, the superdifferential of $g(t,\omega,\bar{\boldsymbol{p}}(t,\omega))$ at each $(t,\omega) \in [0,T] \times \Omega$ is defined as

$$\partial g(t,\omega,\bar{\boldsymbol{p}}(t,\omega)) \triangleq \left\{ \boldsymbol{\nu} \in \mathbb{R}^N \middle| g(t,\omega,\boldsymbol{p}) \leq g(t,\omega,\bar{\boldsymbol{p}}(t,\omega)) + (\boldsymbol{p} - \bar{\boldsymbol{p}}(t,\omega))^\top \boldsymbol{\nu} \right.$$
$$\left. \text{for all } \boldsymbol{p} \in \mathbb{R}^N \right\}. \tag{5.83}$$

Since $g(t,\omega,\cdot)$ is concave for every $(t,\omega) \in [0,T] \times \Omega$, Proposition D.0.11 states that the superdifferential $\partial g(t,\omega,\boldsymbol{p})$ is a non-empty and compact set for every $\boldsymbol{p} \in \mathbb{R}^N$. Now fix $\bar{\boldsymbol{\nu}}(t,\omega) \in \partial g(t,\omega,\bar{\boldsymbol{p}}(t,\omega))$ for all $(t,\omega) \in [0,T] \times \Omega$. From (5.83) we have

$$g(t,\omega,\bar{\boldsymbol{p}}(t,\omega)) - \bar{\boldsymbol{p}}(t,\omega)^\top \bar{\boldsymbol{\nu}}(t,\omega) \geq g(t,\omega,\boldsymbol{p}) - \boldsymbol{p}^\top \bar{\boldsymbol{\nu}}(t,\omega) \tag{5.84}$$

for all $\boldsymbol{p} \in \mathbb{R}^N$, $(t,\omega) \in [0,T] \times \Omega$, and thus

$$g(t,\omega,\bar{\boldsymbol{p}}(t,\omega)) - \bar{\boldsymbol{p}}(t,\omega)^\top \bar{\boldsymbol{\nu}}(t,\omega) \geq \sup_{\boldsymbol{p} \in \mathbb{R}^N} \left\{ g(t,\omega,\boldsymbol{p}) - \boldsymbol{p}^\top \bar{\boldsymbol{\nu}}(t,\omega) \right\}, \tag{5.85}$$

for all $(t,\omega) \in [0,T] \times \Omega$. But by the definition of $\tilde{g}_K$

$$\sup_{\boldsymbol{p} \in \mathbb{R}^N} \left\{ g(t,\omega,\boldsymbol{p}) - \boldsymbol{p}^\top \bar{\boldsymbol{\nu}}(t,\omega) \right\} \geq \tilde{g}_K(t,\omega,\bar{\boldsymbol{\nu}}(t,\omega)) \triangleq \sup_{\boldsymbol{p} \in K} \left\{ g(t,\omega,\boldsymbol{p}) - \boldsymbol{p}^\top \bar{\boldsymbol{\nu}}(t,\omega) \right\}, \tag{5.86}$$

for all $(t,\omega) \in [0,T] \times \Omega$. We have from (5.85) and (5.86)

$$\tilde{g}_K(t,\omega,\bar{\boldsymbol{\nu}}(t,\omega)) \leq g(t,\omega,\bar{\boldsymbol{p}}(t,\omega)) - \bar{\boldsymbol{p}}(t,\omega)^\top \bar{\boldsymbol{\nu}}(t,\omega), \tag{5.87}$$

for all $(t,\omega) \in [0,T] \times \Omega$.

We must now check whether $\bar{\boldsymbol{\nu}}$ defined above is such that $\bar{\boldsymbol{\nu}} \in \mathcal{G}$. Take $\bar{\boldsymbol{\nu}}(t,\omega) \in \partial g(t,\omega,\bar{\boldsymbol{p}}(t,\omega))$ for each $(t,\omega) \in [0,T] \times \Omega$. Since $g(t,\omega,\cdot)$ is concave for each $(t,\omega) \in [0,T] \times \Omega$, $-g$ is a convex integrand by Definition D.0.22. Furthermore, since $g(\cdot,\cdot,\boldsymbol{p})$ is

$\mathcal{F}^*$-measurable for each $\boldsymbol{p} \in \mathbb{R}^N$ and g never takes the values $+\infty$ or $-\infty$, $-g$ is a normal convex integrand by Proposition D.0.24. Therefore, by the measurable selection given by Proposition D.0.27 and Proposition D.0.26, $\{\bar{\boldsymbol{\nu}}(t)\}$ is $\{\mathcal{F}_t\}$-progressively measurable.

Now from Proposition D.0.21 and Condition 5.1.3 we have that the supergradient $\bar{\boldsymbol{\nu}}(t,\omega)$ of the concave function $g(t,\omega,\cdot)$ is bounded by the Lipschitz constant $M(t,\omega)$ for each $(t,\omega) \in [0,T] \times \Omega$. By the square integrability of $M(t)$,

$$\int_0^T \|\bar{\boldsymbol{\nu}}(t)\|^2 dt \leq \int_0^T |M(t)|^2 dt < \infty \text{ a.s.} \tag{5.88}$$

Lastly, from the fact that $\bar{\boldsymbol{p}} \in \Pi$, (5.87) and (5.88),

$$\begin{aligned}
\int_0^T \tilde{g}_K(t,\bar{\boldsymbol{\nu}}(t))dt &\leq \int_0^T g(t,\bar{\boldsymbol{p}}(t))dt - \int_0^T \bar{\boldsymbol{p}}(t)^\top \bar{\boldsymbol{\nu}}(t)dt \\
&\leq \int_0^T |g(t,\bar{\boldsymbol{p}}(t))|dt + \int_0^T |\bar{\boldsymbol{p}}(t)^\top \bar{\boldsymbol{\nu}}(t)|dt \\
&\leq \int_0^T |g(t,\bar{\boldsymbol{p}}(t))|dt + \left(\int_0^T \|\bar{\boldsymbol{p}}(t)\|^2 dt\right)^{1/2} \left(\int_0^T \|\bar{\boldsymbol{\nu}}(t)\|^2 dt\right)^{1/2} \\
&< \infty \text{ a.s.}
\end{aligned} \tag{5.89}$$

From (5.88) and (5.89), we have that

$$\bar{\boldsymbol{\nu}} \in \mathcal{G}. \tag{5.90}$$

To show (5.82) holds for some $(t,\omega) \in [0,T] \times \Omega$ where $\bar{\boldsymbol{p}}(t,\omega) \in K$, fix some $\bar{\boldsymbol{p}}(t,\omega) \in K$. We have from (5.86) and (5.87),

$$\tilde{g}_K(t,\omega,\bar{\boldsymbol{\nu}}(t,\omega)) = g(t,\omega,\bar{\boldsymbol{p}}(t,\omega)) - \bar{\boldsymbol{p}}(t,\omega)^\top \bar{\boldsymbol{\nu}}(t,\omega). \tag{5.91}$$

□

We now show that $\hat{c}$ is in fact a cumulative consumption process.

Proposition 5.3.24. *The process $\{\hat{c}(t), t \in [0,T]\}$ defined by*

$$\hat{c}(t) \triangleq -\int_0^t \hat{X}(s)\left\{\tilde{g}_K(s,\boldsymbol{\nu}(s)) - \left[g(s,\hat{\boldsymbol{p}}(s)) - \hat{\boldsymbol{p}}(s)^\top \boldsymbol{\nu}(s)\right]\right\}ds + \hat{c}_{\boldsymbol{\nu}}(t),\ t \in [0,T], \boldsymbol{\nu} \in \mathcal{G}, \tag{5.92}$$

where $\hat{c}_{\boldsymbol{\nu}}$ is given by Proposition 5.3.20, is a cumulative consumption process, i.e. $\hat{c} \in \mathcal{C}$.

Proof. Fix $\bar{\boldsymbol{\nu}}(t,\omega) \in \partial g(t,\omega,\hat{\boldsymbol{p}}(t,\omega))$ for each $(t,\omega) \in [0,T] \times \Omega$. Using Propositions 5.3.15(3) and 5.3.23, this implies

$$\hat{X}(t,\omega)\left\{\tilde{g}_K(t,\omega,\bar{\boldsymbol{\nu}}(t,\omega)) - \left[g(t,\omega,\hat{\boldsymbol{p}}(t,\omega)) - \hat{\boldsymbol{p}}(t,\omega)^\top \bar{\boldsymbol{\nu}}(t,\omega)\right]\right\} \leq 0, \tag{5.93}$$

for $\lambda \otimes P$ almost all $(t,\omega) \in [0,T] \times \Omega$. As a result, the process

$$\tilde{c}(t) \triangleq -\int_0^t \hat{X}(s)\left\{\tilde{g}_K(s,\bar{\boldsymbol{\nu}}(s)) - \left[g(s,\hat{\boldsymbol{p}}(s)) - \hat{\boldsymbol{p}}(s)^\top \bar{\boldsymbol{\nu}}(s)\right]\right\}ds, \quad t \in [0,T], \tag{5.94}$$

is continuous, almost surely non-decreasing and null at the origin. Furthermore, since $\hat{X}$ is càdlàg, therefore $\hat{X}(t,\omega)$ is uniformly bounded on $[0,T]$ for P-almost all $\omega \in \Omega$, and since $\hat{\boldsymbol{p}} \in \Pi$ and $\bar{\boldsymbol{\nu}} \in \mathcal{G}$ by Propositions 5.3.21 and 5.3.23, respectively, we obtain

$$\tilde{c}(T) < \infty \text{ a.s.}, \tag{5.95}$$

and as a result

$$\tilde{c} \in \mathcal{C}. \tag{5.96}$$

We have from Proposition 5.3.20, (5.92), and (5.94)

$$\hat{c}(t) = \tilde{c}(t) + \hat{c}_{\bar{\boldsymbol{\nu}}}(t), \quad t \in [0,T]. \tag{5.97}$$

Since we have already shown in Proposition 5.3.20 that

$$\hat{c}_{\boldsymbol{\nu}} \in \mathcal{C}, \quad \boldsymbol{\nu} \in \mathcal{G}, \tag{5.98}$$

from (5.96), (5.97), and (5.98)

$$\hat{c} \in \mathcal{C}. \tag{5.99}$$

□

Remark 5.3.25. We can now say from (5.80) and Proposition 5.3.24 that $\hat{X}$ can be written as the wealth process $X^{(\hat{u},\hat{c},\hat{\boldsymbol{p}})}$. However, we require $(\hat{c},\hat{\boldsymbol{p}}) \in \mathcal{A}(\hat{u})$ to prove Theorem 5.3.17. We already know from Propositions 5.3.21 and 5.3.24 that $(\hat{c},\hat{\boldsymbol{p}}) \in \mathcal{C} \times \Pi$, and from Proposition 5.3.15(3) that $X^{(\hat{u},\hat{c},\hat{\boldsymbol{p}})}(t) > 0$ a.s. for all $t \in [0,T]$. Therefore, it remains to show $\hat{\boldsymbol{p}}(t) \in K$ a.e.

Proposition 5.3.26. *The process* $\hat{\mathbf{p}}$ *defined in Proposition 5.3.21 satisfies the portfolio constraints. That is* $\hat{\boldsymbol{p}}(t) \in K$ *a.e.*

Proof. Using the fact that $\hat{c}_{\boldsymbol{\nu}} \in \mathcal{C}$ for all $\boldsymbol{\nu} \in \mathcal{G}$ from Proposition 5.3.20, we have from (5.92) in Proposition 5.3.24 that

$$\hat{c}(T) + \int_0^T \hat{X}(t)\Big\{\tilde{g}_K(t, \boldsymbol{\nu}(t)) - \Big[g(t, \hat{\boldsymbol{p}}(t)) - \hat{\boldsymbol{p}}(t)^\top \boldsymbol{\nu}(t)\Big]\Big\}dt \geq 0 \text{ a.s.}, \quad \boldsymbol{\nu} \in \mathcal{G}. \quad (5.100)$$

Using the subadditivity property of the supremum in (5.20),

$$\sup_{\boldsymbol{p}\in K} g(t, \boldsymbol{p}) + \delta(\boldsymbol{\nu}) \geq \tilde{g}_K(t, \boldsymbol{\nu}) \text{ a.s.}, \quad t \in [0, T], \boldsymbol{\nu} \in \mathbb{R}^N, \quad (5.101)$$

where

$$\delta(\boldsymbol{\nu}) \triangleq \sup_{\boldsymbol{p}\in K}\{-\boldsymbol{p}^\top \boldsymbol{\nu}\}, \quad \boldsymbol{\nu} \in \mathbb{R}^N \quad (5.102)$$

is the support function of the convex set $-K$ (see Definition D.0.16). Putting (5.101) into (5.100) yields

$$\hat{c}(T) + \int_0^T \hat{X}(t)\Big[\sup_{\boldsymbol{p}\in K} g(t, \boldsymbol{p}) - g(t, \hat{\boldsymbol{p}}(t))\Big]dt + \int_0^T \hat{X}(t)\Big[\delta(\boldsymbol{\nu}(t)) + \hat{\boldsymbol{p}}(t)^\top \boldsymbol{\nu}(t)\Big]dt \geq 0 \text{ a.s.}, \quad (5.103)$$

for all $\boldsymbol{\nu} \in \mathcal{G}$. Put

$$\mathcal{B} \triangleq \Big\{(t, \omega) \in [0, T] \times \Omega \Big| \hat{\boldsymbol{p}}(t, \omega) \in K^c\Big\}. \quad (5.104)$$

From Lemma 5.4.2 in Karatzas and Shreve [31], there exists an $\{\mathcal{F}_t\}$-progressively measurable process $\tilde{\boldsymbol{\nu}} : [0, T] \times \Omega \mapsto \mathbb{R}^N$ such that

$$\|\tilde{\boldsymbol{\nu}}(t)\| \leq 1 \text{ a.s. and } |\delta(\tilde{\boldsymbol{\nu}}(t))| \leq 1 \text{ a.s.} \quad (5.105)$$

and

$$\tilde{\boldsymbol{\nu}}(t, \omega) = 0 \text{ if and only if } (t, \omega) \in \mathcal{B}^c. \quad (5.106)$$

From (5.106) we also have

$$\delta(\tilde{\boldsymbol{\nu}}(t, \omega)) + \hat{\boldsymbol{p}}(t, \omega)^\top \tilde{\boldsymbol{\nu}}(t, \omega) = 0 \text{ if and only if } (t, \omega) \in \mathcal{B}^c. \quad (5.107)$$

and

$$\delta(\tilde{\boldsymbol{\nu}}(t, \omega)) + \hat{\boldsymbol{p}}(t, \omega)^\top \tilde{\boldsymbol{\nu}}(t, \omega) < 0 \text{ if and only if } (t, \omega) \in \mathcal{B}. \quad (5.108)$$

It is easy to see from (5.105) and (5.101) that $\tilde{\boldsymbol{\nu}} \in \mathcal{G}$. Furthermore, from the positive homogeneity of the support function $\delta(\cdot)$ (see Remark D.0.17) we have

$$k\tilde{\boldsymbol{\nu}} \in \mathcal{G}, \quad k = 1, 2, \ldots \quad (5.109)$$

Again from the positive homogeneity of $\delta(\cdot)$ and (5.103), we have

$$\hat{c}(T) + \int_0^T \hat{X}(t)\left[\sup_{\boldsymbol{p}\in K} g(t,\boldsymbol{p}) - g(t,\hat{\boldsymbol{p}}(t))\right]dt + \int_0^T \hat{X}(t)\left[\delta(k\tilde{\boldsymbol{\nu}}(t)) + \hat{\boldsymbol{p}}(t)^\top k\tilde{\boldsymbol{\nu}}(t)\right]dt$$
$$= \hat{c}(T) + \int_0^T \hat{X}(t)\left[\sup_{\boldsymbol{p}\in K} g(t,\boldsymbol{p}) - g(t,\hat{\boldsymbol{p}}(t))\right]dt + k\int_0^T \hat{X}(t)\left[\delta(\tilde{\boldsymbol{\nu}}(t)) + \hat{\boldsymbol{p}}(t)^\top \tilde{\boldsymbol{\nu}}(t)\right]dt, \tag{5.110}$$

for $k = 1, 2, \ldots$. Combining (5.110) and (5.103)

$$\hat{c}(T) + \int_0^T \hat{X}(t)\left[\sup_{\boldsymbol{p}\in K} g(t,\boldsymbol{p}) - g(t,\hat{\boldsymbol{p}}(t))\right]dt \geq -k\int_0^T \hat{X}(t)\left[\delta(\tilde{\boldsymbol{\nu}}(t)) + \hat{\boldsymbol{p}}(t)^\top \tilde{\boldsymbol{\nu}}(t)\right]dt \text{ a.s.} \tag{5.111}$$

Define the set $\mathcal{D}$ as

$$\mathcal{D} \triangleq \left\{(t,\omega)\in\mathcal{B}\middle|\hat{X}(t,\omega)\left[\delta(\tilde{\boldsymbol{\nu}}(t,\omega)) + \hat{\boldsymbol{p}}(t,\omega)^\top\tilde{\boldsymbol{\nu}}(t,\omega)\right] < 0\right\}, \tag{5.112}$$

which by (5.108) and Proposition 5.3.15(3) has the property

$$\lambda\otimes P[\mathcal{D}] = \lambda\otimes P[\mathcal{B}]. \tag{5.113}$$

Assume

$$\lambda\otimes P[\mathcal{D}] > 0, \text{ or by Fubini's Theorem, } E[\textstyle\int_0^T I_{\mathcal{D}}(t,\omega)dt] > 0. \tag{5.114}$$

Put

$$Z(\omega) \triangleq \int_0^T I_{\mathcal{D}}(t,\omega)dt, \quad \omega\in\Omega, \tag{5.115}$$

so that from (5.114), $E[Z] > 0$. Define the subset

$$\Omega_{\mathcal{D}} \triangleq \left\{\omega\in\Omega\middle|Z(\omega) > 0\right\}. \tag{5.116}$$

Since $E[Z] > 0$ we have $P[\Omega_{\mathcal{D}}] > 0$. Also define the set

$$\mathcal{D}(\omega) \triangleq \left\{t\in[0,T]\middle|(t,\omega)\in\mathcal{D}\right\}. \tag{5.117}$$

We then have by (5.116) that $\lambda[\mathcal{D}(\omega)] > 0$ for each $\omega \in \Omega_{\mathcal{D}}$. Fixing $\bar{\omega} \in \Omega_{\mathcal{D}}$, we can write

$$\begin{aligned}\int_0^T \hat{X}(t,\bar{\omega})\Big[\delta(\tilde{\boldsymbol{\nu}}(t,\bar{\omega})) + \hat{\boldsymbol{p}}(t,\bar{\omega})^\top \tilde{\boldsymbol{\nu}}(t,\bar{\omega})\Big]dt \\ = \int_{\mathcal{D}(\bar{\omega})} \hat{X}(t,\bar{\omega})\Big[\delta(\tilde{\boldsymbol{\nu}}(t,\bar{\omega})) + \hat{\boldsymbol{p}}(t,\bar{\omega})^\top \tilde{\boldsymbol{\nu}}(t,\bar{\omega})\Big]dt \\ + \int_{\mathcal{D}(\bar{\omega})^c} \hat{X}(t,\bar{\omega})\Big[\delta(\tilde{\boldsymbol{\nu}}(t,\bar{\omega})) + \hat{\boldsymbol{p}}(t,\bar{\omega})^\top \tilde{\boldsymbol{\nu}}(t,\bar{\omega})\Big]dt.\end{aligned} \tag{5.118}$$

By the definition of $\mathcal{D}$ in (5.112), the first term on the right side of (5.118) is negative and from (5.107), the second term on the right side of (5.118) is zero. Therefore,

$$\int_0^T \hat{X}(t,\bar{\omega})\Big[\delta(\tilde{\boldsymbol{\nu}}(t,\bar{\omega})) + \hat{\boldsymbol{p}}(t,\bar{\omega})^\top \tilde{\boldsymbol{\nu}}(t,\bar{\omega})\Big]dt < 0. \tag{5.119}$$

Putting (5.119) into (5.111) and taking $k \to \infty$ yields

$$\hat{c}(T,\bar{\omega}) + \int_0^T \hat{X}(t,\bar{\omega})\Big[\sup_{\boldsymbol{p}\in K} g(t,\bar{\omega},\boldsymbol{p}) - g(t,\bar{\omega},\hat{\boldsymbol{p}}(t,\bar{\omega}))\Big]dt = \infty. \tag{5.120}$$

From the integrability and uniform upper-boundedness of $g(t,\hat{\boldsymbol{p}})$, and the almost sure boundedness of $\hat{X}$,

$$\hat{c}(T,\bar{\omega}) = \infty, \quad \bar{\omega} \in \Omega_{\mathcal{D}}. \tag{5.121}$$

Since $P[\Omega_{\mathcal{D}}] > 0$, there is a contradiction in (5.114) because $\hat{c} \in \mathcal{C}$ (which implies $\hat{c}(T) < \infty$ a.s.). Therefore, $\lambda \otimes P[\mathcal{D}] = 0$ and by (5.113), $\lambda \otimes P[\mathcal{B}] = 0$. Finally, we conclude from (5.104),

$$\hat{\boldsymbol{p}}(t) \in K \text{ a.e.} \tag{5.122}$$

□

Completing the proof of Theorem 5.3.17, we have from Remark 5.3.25 and Proposition 5.3.26 that there does exist a pair $(\hat{c},\hat{\boldsymbol{p}}) \in \mathcal{A}(\hat{u})$ such that

$$\hat{X}(t) = X^{(\hat{u},\hat{c},\hat{\boldsymbol{p}})}(t) \text{ a.s.}, \qquad t \in [0,T], \tag{5.123}$$

and Theorem 5.3.17 follows. □

Remark 5.3.27. From Theorem 5.3.17, together with Remark 5.3.19, one sees that the price of the contingent claim B is the quantity $\hat{u}$ defined at (5.34), that is

$$\hat{u} \triangleq \sup_{(\boldsymbol{\nu},\boldsymbol{\mu})\in\mathcal{G}\times\mathcal{H}} E\left[H_{\boldsymbol{\nu},\boldsymbol{\mu}}(T)B\right]. \tag{5.124}$$

$\hat{u}$ is the *least initial wealth* from which the contingent claim B can be hedged. By setting the dual process $\boldsymbol{\nu} \triangleq \bar{\boldsymbol{\nu}}$, where $\bar{\boldsymbol{\nu}}(t,\omega) \in \partial g(t,\omega,\hat{\boldsymbol{p}}(t,\omega))$ for each $(t,\omega) \in [0,T]\times\Omega$, the corresponding *hedging portfolio process* $\hat{\boldsymbol{p}} \in \Pi$, which by Proposition 5.3.26 has the property $\hat{\boldsymbol{p}}(t) \in K$ a.e., is given by

$$\hat{\boldsymbol{p}}(t) \triangleq [\boldsymbol{\sigma}^{-1}(t)]^\top \left[\boldsymbol{\theta}_{\bar{\boldsymbol{\nu}}}(t) + \frac{\boldsymbol{\Psi}_{\bar{\boldsymbol{\nu}},0}(t)}{H_{\bar{\boldsymbol{\nu}},0}(t)\hat{X}(t)}\right], \quad t \in [0,T]. \tag{5.125}$$

Lastly, by Proposition 5.3.23 and Proposition 5.3.24, we have the optimal *cumulative consumption process* $\hat{c} \in \mathcal{C}$ is given by

$$\hat{c}(t) = \int_0^t \frac{1}{H_{\bar{\boldsymbol{\nu}},0}(s-)} dA_{\bar{\boldsymbol{\nu}},0}(s) - \sum_{\substack{i,j\in S\\ i\neq j}} \int_0^t \frac{\Gamma_{ij}^{\bar{\boldsymbol{\nu}},0}(s)}{H_{\bar{\boldsymbol{\nu}},0}(s-)} dM_{ij}(s), \quad t \in [0,T]. \tag{5.126}$$

5.4 Conditions for Zero Consumption

We have from Theorem 5.3.17 that if $\hat{u}$ defined in (5.34) is finite, then the contingent claim B can be exactly hedged by a pair $(\hat{c},\hat{\boldsymbol{p}}) \in \mathcal{A}(\hat{u})$ where $\hat{c}$ and $\hat{\boldsymbol{p}}$ are given by (5.126) and (5.125), respectively. Furthermore, as will be seen in Proposition 5.4.1, the optimal portfolio process $\hat{\boldsymbol{p}} \in \Pi$ is in fact a super-hedging portfolio process. In other words, if an agent starting with initial wealth $\hat{u}$ defined in (5.34) trades with the strategy $(0,\hat{\boldsymbol{p}}) \in \mathcal{A}(\hat{u})$, where $\hat{\boldsymbol{p}}$ is defined in Proposition 5.3.21, they are guaranteed to have $X^{(\hat{u},0,\hat{\boldsymbol{p}})}(T) \geq B$ a.s. This means that an agent must use a portfolio that produces excess wealth to hedge the contingent claim. In view of this, one may ask: under what conditions, if any, can an agent exactly hedge a contingent claim in a regime-switching market without consuming wealth? In other words, are there conditions which assert the existence of a pair $(0,\boldsymbol{p}) \in \mathcal{A}(\hat{u})$ such that $X^{(\hat{u},0,\boldsymbol{p})}(T) = B$ a.s.? In this section we propose necessary and sufficient conditions that guarantee the existence of such an optimal hedging strategy.

Proposition 5.4.1. *Given a hedgeable contingent claim random variable B, i.e. there exists a least initial wealth $\hat{u}$ and hedging strategy $(\hat{c},\hat{\boldsymbol{p}}) \in \mathcal{A}(\hat{u})$ such that*

$$X^{(\hat{u},\hat{c},\hat{\boldsymbol{p}})}(T) = B \ a.s., \tag{5.127}$$

the consumption-portfolio pair $(0, \hat{\boldsymbol{p}})$ *is admissible and super-hedges the contingent claim* B*. In other words,* $(0, \hat{\boldsymbol{p}}) \in \mathcal{A}(\hat{u})$ *and*

$$X^{(\hat{u},0,\hat{\boldsymbol{p}})}(T) \geq B \ \textit{a.s.} \tag{5.128}$$

Proof. For ease of notation, let

$$\begin{aligned} a(t) &\triangleq r(t) + g(t, \hat{\boldsymbol{p}}(t)) + \hat{\boldsymbol{p}}(t)^\top \boldsymbol{\sigma}(t)\boldsymbol{\theta}(t), \\ b(t) &\triangleq \hat{\boldsymbol{p}}(t)^\top \boldsymbol{\sigma}(t), \\ z(t) &\triangleq a(t)dt + b(t)d\mathbf{W}(t), \quad t \in [0,T]. \end{aligned} \tag{5.129}$$

Then, from (5.6) and (5.129),

$$X^{(\hat{u},\hat{c},\hat{\boldsymbol{p}})}(t) = \hat{u} + \int_0^t X^{(\hat{u},\hat{c},\hat{\boldsymbol{p}})}(s)dz(s) - \hat{c}(t), \quad t \in [0,T], \tag{5.130}$$

and

$$X^{(\hat{u},0,\hat{\boldsymbol{p}})}(t) = \hat{u} + \int_0^t X^{(\hat{u},0,\hat{\boldsymbol{p}})}(s)dz(s), \quad t \in [0,T]. \tag{5.131}$$

Setting $Y(t) \triangleq X^{(\hat{u},0,\hat{\boldsymbol{p}})}(t) - X^{(\hat{u},\hat{c},\hat{\boldsymbol{p}})}(t), t \in [0,T]$, we have by (5.130) and (5.131)

$$Y(t) = \int_0^t Y(s)dz(s) + \hat{c}(t), \quad t \in [0,T]. \tag{5.132}$$

From Theorem C.10.4, the solution to the stochastic integral equation (5.132) is given by

$$Y(t) = \mathcal{E}(z)(t)\left\{\hat{c}(0) + \int_0^t \frac{1}{\mathcal{E}(z)(s)}d\hat{c}(s) - \int_0^t \frac{1}{\mathcal{E}(z)(s)}d[\hat{c}, z](s)\right\} \text{ a.s.}, \quad t \in [0,T], \tag{5.133}$$

where $\mathcal{E}(\cdot)$ is the Ito Exponential (see Remark C.10.3). Since $\hat{c} \in \mathcal{C}$ and z is continuous, $\hat{c}(0) = 0$ and $[\hat{c}, z](t) = 0$ a.s. for all $t \in [0,T]$. Therefore, from (5.133),

$$Y(t) = \mathcal{E}(z)(t)\int_0^t \frac{1}{\mathcal{E}(z)(s)}d\hat{c}(s) \text{ a.s.}, \quad t \in [0,T]. \tag{5.134}$$

We have by Remark C.10.3 that $\mathcal{E}(z)(t) > 0$ a.s. and by (5.4) that $\hat{c}(t) \geq 0$ a.s. for all $t \in [0,T]$. As a result, from (5.134)

$$X^{(\hat{u},0,\hat{\boldsymbol{p}})}(t) - X^{(\hat{u},\hat{c},\hat{\boldsymbol{p}})}(t) = Y(t) \geq 0 \text{ a.s.}, \quad t \in [0,T]. \tag{5.135}$$

Setting $t = T$ in (5.135) gives (5.128). Furthermore, since $X^{(\hat{u},\hat{c},\hat{\boldsymbol{p}})}(t) > 0$ a.s., we have from (5.135) that $(0, \hat{\boldsymbol{p}}) \in \mathcal{A}(\hat{u})$, as required. □

Remark 5.4.2. Conditions that guarantee the existence of a exact hedging strategy without consumption have been found in the context of a standard Brownian motion market model with convex portfolio constraints by Cvitanic and Karatzas [10]. By virtue of the similarities between our regime-switching market model and the constrained Brownian motion market model used in Cvitanic and Karatzas [10], the statement of our main result is nearly identical to Theorem 5.8.1 in Karatzas and Shreve [31]; however, instead of optimizing over a single dual process as they do in Karatzas and Shreve [31], we deal with the joint space of dual processes $\mathcal{G} \times \mathcal{H}$, which makes the proof far more challenging.

To state our result, we first need the following definition of a *K-attainable* contingent claim, originally given in Cvitanic and Karatzas [10].

Definition 5.4.3. We say that a contingent claim B is *K-attainable* if there exists a portfolio process $\boldsymbol{p} \in \Pi$ such that $(0, \boldsymbol{p}) \in \mathcal{A}(\hat{u})$ and $X^{(\hat{u},0,\boldsymbol{p})}(T) = B$ a.s.

Theorem 5.4.4. *Let $B : (\Omega, \mathcal{F}_T, P) \mapsto (0, \infty)$ be a contingent claim and assume that $\hat{u}$ defined in (5.34) is finite. Let $\hat{X}$ be defined as in (5.37) and let $(\hat{c}, \hat{\boldsymbol{p}}) \in \mathcal{A}(\hat{u})$ be such that $X^{(\hat{u},\hat{c},\hat{\boldsymbol{p}})}(t) = \hat{X}(t)$ a.s. for all $t \in [0, T]$. For a given pair $(\hat{\boldsymbol{\nu}}, \hat{\boldsymbol{\mu}}) \in \mathcal{G} \times \mathcal{H}$, the following conditions are equivalent:*

(1) *The supremum at (5.34) is attained by some $(\hat{\boldsymbol{\nu}}, \hat{\boldsymbol{\mu}}) \in \mathcal{G} \times \mathcal{H}$, i.e. $\hat{u} = E[H_{\hat{\boldsymbol{\nu}},\hat{\boldsymbol{\mu}}}(T)B]$,*

(2) *$H_{\hat{\boldsymbol{\nu}},\hat{\boldsymbol{\mu}}}\hat{X}$ is an $\{\mathcal{F}_t\}$-martingale,*

(3) $$\begin{cases} B \text{ is K-attainable, and for the associated wealth process } X^{\hat{u},0,\boldsymbol{p}}, \\ \text{the product } H_{\hat{\boldsymbol{\nu}},\hat{\boldsymbol{\mu}}}X^{\hat{u},0,\boldsymbol{p}} \text{ is an } \{\mathcal{F}_t\}\text{-martingale.} \end{cases}$$

Any of the above conditions imply

$$(4) \quad \begin{cases} \hat{c}(t) = 0 \ \text{a.s.}, \\ \tilde{g}_K(t, \hat{\boldsymbol{\nu}}(t)) - \Big[g(t, \hat{\boldsymbol{p}}(t)) - \hat{\boldsymbol{p}}(t)^\top \hat{\boldsymbol{\nu}}(t) \Big] = 0 \ \text{a.s.}, \end{cases}$$

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

Remark 5.4.5. To prove Theorem 5.4.4, we begin by showing that Theorem 5.4.4(1) and 5.4.4(2) are equivalent. From there we prove an intermediate result in Lemma 5.4.9 and finally show Theorem 5.4.4 holds.

Proposition 5.4.6. *For a pair* $(\hat{\boldsymbol{\nu}}, \hat{\boldsymbol{\mu}}) \in \mathcal{G} \times \mathcal{H}$*,* $H_{\hat{\boldsymbol{\nu}},\hat{\boldsymbol{\mu}}}\hat{X}$ *is an* $\{\mathcal{F}_t\}$*-martingale if and only if* $\hat{u} = E[H_{\hat{\boldsymbol{\nu}},\hat{\boldsymbol{\mu}}}(T)B]$*. Therefore, Theorem 5.4.4(1) and 5.4.4(2) are equivalent.*

Proof of Proposition 5.4.6. ($\Longrightarrow$) First assume $H_{\hat{\boldsymbol{\nu}},\hat{\boldsymbol{\mu}}}\hat{X}$ is an $\{\mathcal{F}_t\}$-martingale. We would then have

$$E[H_{\hat{\boldsymbol{\nu}},\hat{\boldsymbol{\mu}}}(T)\hat{X}(T)] = E[\hat{X}(0)]. \tag{5.136}$$

From Proposition 5.3.15(1) and 5.3.15(2),

$$E[H_{\hat{\boldsymbol{\nu}},\hat{\boldsymbol{\mu}}}(T)B] = \hat{u}. \tag{5.137}$$

($\Longleftarrow$) Now assume $\hat{u} = E[H_{\hat{\boldsymbol{\nu}},\hat{\boldsymbol{\mu}}}(T)B]$. From Proposition 5.3.15(4) $H_{\hat{\boldsymbol{\nu}},\hat{\boldsymbol{\mu}}}\hat{X}$ is an $\{\mathcal{F}_t\}$-supermartingale. From the supermartingale property of $H_{\hat{\boldsymbol{\nu}},\hat{\boldsymbol{\mu}}}\hat{X}$,

$$E[H_{\hat{\boldsymbol{\nu}},\hat{\boldsymbol{\mu}}}(t)\hat{X}(t)] \leq E[H_{\hat{\boldsymbol{\nu}},\hat{\boldsymbol{\mu}}}(s)\hat{X}(s)], \quad 0 \leq s \leq t \leq T. \tag{5.138}$$

But since $\hat{u} = E[H_{\hat{\boldsymbol{\nu}},\hat{\boldsymbol{\mu}}}(T)B]$, from Proposition 5.3.15(1) and 5.3.15(2)

$$E[H_{\hat{\boldsymbol{\nu}},\hat{\boldsymbol{\mu}}}(T)\hat{X}(T)] = E[\hat{X}(0)], \tag{5.139}$$

which in combination with (5.138) shows $H_{\hat{\boldsymbol{\nu}},\hat{\boldsymbol{\mu}}}\hat{X}$ is an $\{\mathcal{F}_t\}$-martingale. □

Lemma 5.4.7. *If the supremum at* (5.34) *is attained by some* $(\hat{\boldsymbol{\nu}}, \hat{\boldsymbol{\mu}}) \in \mathcal{G} \times \mathcal{H}$*, then*

$$A_{\hat{\boldsymbol{\nu}},\hat{\boldsymbol{\mu}}}(t) = 0 \ \ a.s., \quad t \in [0,T], \tag{5.140}$$

where $A_{\boldsymbol{\nu},\boldsymbol{\mu}}, (\boldsymbol{\nu}, \boldsymbol{\mu}) \in \mathcal{G} \times \mathcal{H}$*, is defined in Lemma 5.3.16.*

Proof. By Proposition 5.4.6, $H_{\hat{\boldsymbol{\nu}},\hat{\boldsymbol{\mu}}}\hat{X}$ is an $\{\mathcal{F}_t\}$-martingale. It is immediate that $A_{\hat{\boldsymbol{\nu}},\hat{\boldsymbol{\mu}}}(t) = 0$ a.s. for all $t \in [0,T]$ by the Doob-Meyer decomposition in Theorem C.11.4. □

Lemma 5.4.8. *If the supremum at* (5.34) *is attained by some* $(\hat{\boldsymbol{\nu}}, \hat{\boldsymbol{\mu}}) \in \mathcal{G} \times \mathcal{H}$*, then*

$$\int_0^t \frac{1}{H_{\hat{\boldsymbol{\nu}},\boldsymbol{\mu}}(s-)} dA_{\hat{\boldsymbol{\nu}},\boldsymbol{\mu}}(s) = \sum_{\substack{i,j\in S\\ i\neq j}} \int_0^t \frac{\Gamma_{ij}^{\hat{\boldsymbol{\nu}},0}(s)}{H_{\hat{\boldsymbol{\nu}},0}(s-)} \Big(\hat{\mu}_{ij}(s) - \mu_{ij}(s)\Big) I\{\boldsymbol{\alpha}(s-) = i\} g_{ij} ds, \tag{5.141}$$

for all $t \in [0,T]$ *and* $\boldsymbol{\mu} \in \mathcal{H}$*, where* $A_{\boldsymbol{\nu},\boldsymbol{\mu}}$ *and* $\Gamma_{ij}^{\boldsymbol{\nu},\boldsymbol{\mu}}$*,* $(\boldsymbol{\nu}, \boldsymbol{\mu}) \in \mathcal{G} \times \mathcal{H}$*,* $i,j \in S, i \neq j$*, are defined in Lemma 5.3.16, and* g_{ij}*,* $i,j \in S, i \neq j$*, are the generators of the Markov chain* $\boldsymbol{\alpha}$*.*

Proof. If we set $(\boldsymbol{\nu}, \boldsymbol{\mu}) = (\hat{\boldsymbol{\nu}}, \hat{\boldsymbol{\mu}})$ in (B.278) of Lemma B.5.2, we can use Lemma 5.4.7 to obtain

$$\int_0^t \frac{1}{H_{\hat{\boldsymbol{\nu}},0}(s-)} dA_{\hat{\boldsymbol{\nu}},0}(s) = \sum_{\substack{i,j \in S \\ i \neq j}} \int_0^t \frac{\Gamma_{ij}^{\hat{\boldsymbol{\nu}},0}(s)}{H_{\hat{\boldsymbol{\nu}},0}(s-)} \hat{\mu}_{ij}(s) d\tilde{R}_{ij}(s), \quad t \in [0,T]. \tag{5.142}$$

Putting (5.142) into (B.278) of Lemma B.5.2, where we set $\boldsymbol{\nu} = \hat{\boldsymbol{\nu}}$, results in

$$\sum_{\substack{i,j \in S \\ i \neq j}} \int_0^t \frac{\Gamma_{ij}^{\hat{\boldsymbol{\nu}},0}(s)}{H_{\hat{\boldsymbol{\nu}},0}(s-)} \hat{\mu}_{ij}(s) d\tilde{R}_{ij}(s) = \int_0^t \frac{1}{H_{\hat{\boldsymbol{\nu}},\boldsymbol{\mu}}(s-)} dA_{\hat{\boldsymbol{\nu}},\boldsymbol{\mu}}(s) + \sum_{\substack{i,j \in S \\ i \neq j}} \int_0^t \frac{\Gamma_{ij}^{\hat{\boldsymbol{\nu}},0}(s)}{H_{\hat{\boldsymbol{\nu}},0}(s-)} \mu_{ij}(s) d\tilde{R}_{ij}(s), \tag{5.143}$$

for all $t \in [0,T]$ and $\boldsymbol{\mu} \in \mathcal{H}$. Rearranging (5.143) and expanding the $d\tilde{R}_{ij}$ integral using (4.7) gives

$$\int_0^t \frac{1}{H_{\hat{\boldsymbol{\nu}},\boldsymbol{\mu}}(s-)} dA_{\hat{\boldsymbol{\nu}},\boldsymbol{\mu}}(s) = \sum_{\substack{i,j \in S \\ i \neq j}} \int_0^t \frac{\Gamma_{ij}^{\hat{\boldsymbol{\nu}},0}(s)}{H_{\hat{\boldsymbol{\nu}},0}(s-)} \Big(\hat{\mu}_{ij}(s) - \mu_{ij}(s) \Big) I\{\boldsymbol{\alpha}(s-) = i\} g_{ij} ds \tag{5.144}$$

for all $t \in [0,T]$ and $\boldsymbol{\mu} \in \mathcal{H}$. □

Lemma 5.4.9. *If the supremum at (5.34) is attained by some $(\hat{\boldsymbol{\nu}}, \hat{\boldsymbol{\mu}}) \in \mathcal{G} \times \mathcal{H}$, then*

$$\Gamma_{ij}^{\hat{\boldsymbol{\nu}},0}(t) I\{\boldsymbol{\alpha}(t-) = i\} g_{ij} = 0, \ \lambda \otimes P \ a.e. \ on \ [0,T] \times \Omega,$$

for all $i,j \in S, i \neq j$, where $\Gamma_{ij}^{\boldsymbol{\nu},\boldsymbol{\mu}}$, $(\boldsymbol{\nu}, \boldsymbol{\mu}) \in \mathcal{G} \times \mathcal{H}$, $i,j \in S, i \neq j$, is defined in Lemma 5.3.16, and g_{ij}, $i,j \in S, i \neq j$, are the generators of the Markov chain $\boldsymbol{\alpha}$.

Proof. Without loss of generality we assume the Markov chain generators g_{ij} are such that for some $i,j \in S$, $i \neq j$, $g_{ij} > 0$. Indeed, since $g_{ij} \geq 0$ for all $i,j \in S$ where $i \neq j$, if $g_{ij} = 0$ for some $i,j \in S$, $i \neq j$, then

$$\Gamma_{ij}^{\boldsymbol{\nu},0}(t) I\{\boldsymbol{\alpha}(t-) = i\} g_{ij} = 0 \quad \lambda \otimes P \text{ a.e. on } [0,T] \times \Omega, \boldsymbol{\nu} \in \mathcal{G} \tag{5.145}$$

for that choice of $i,j \in S$. Therefore, we assume $g_{ij} > 0$ for some $i,j \in S$, $i \neq j$.

From Lemma B.5.2,

$$\begin{aligned} \int_0^t \frac{1}{H_{\boldsymbol{\nu},0}(s-)} dA_{\boldsymbol{\nu},0}(s) &= \int_0^t \frac{1}{H_{\boldsymbol{\nu},\boldsymbol{\mu}}(s-)} dA_{\boldsymbol{\nu},\boldsymbol{\mu}}(s) \\ &\quad + \sum_{\substack{i,j \in S \\ i \neq j}} \int_0^t \frac{\Gamma_{ij}^{\boldsymbol{\nu},0}(s)}{H_{\boldsymbol{\nu},0}(s-)} \mu_{ij}(s) d\tilde{R}_{ij}(s), \quad t \in [0,T], \end{aligned} \tag{5.146}$$

for all $(\boldsymbol{\nu}, \boldsymbol{\mu}) \in \mathcal{G} \times \mathcal{H}$, where $\tilde{R}_{ij}$ is a canonical process of the Markov chain $\boldsymbol{\alpha}$ introduced in Definition 4.1.9. Since $A_{\boldsymbol{\nu},\boldsymbol{\mu}}$ is a non-decreasing process and $\inf_{t\in[0,T]} H_{\boldsymbol{\nu},\boldsymbol{\mu}}(t) > 0$ a.s. for every $(\boldsymbol{\nu}, \boldsymbol{\mu}) \in \mathcal{G} \times \mathcal{H}$, we have from (5.146) that

$$\int_0^t \frac{1}{H_{\boldsymbol{\nu},0}(s-)} dA_{\boldsymbol{\nu},0}(s) \geq \sum_{\substack{i,j\in S \\ i\neq j}} \int_0^t \frac{\Gamma_{ij}^{\boldsymbol{\nu},0}(s)}{H_{\boldsymbol{\nu},0}(s-)} \mu_{ij}(s) d\tilde{R}_{ij}(s), \quad t \in [0,T], \tag{5.147}$$

for every $(\boldsymbol{\nu}, \boldsymbol{\mu}) \in \mathcal{G} \times \mathcal{H}$. Fix some $\boldsymbol{\nu} \in \mathcal{G}$ and some $k, l \in S$, such that $k \neq l$ and $g_{kl} > 0$. For all other $i, j \in S$, $i \neq j$, set

$$\mu_{ij}(t) = 0, \quad t \in [0,T]. \tag{5.148}$$

Define the set

$$\mathcal{U} \triangleq \left\{ (t,\omega) \in [0,T] \times \Omega \middle| \Gamma_{kl}^{\boldsymbol{\nu},0}(t,\omega) I\{\boldsymbol{\alpha}(t-,\omega) = k\} g_{kl} > 0 \right\}, \tag{5.149}$$

where we immediately see that $\mathcal{U}$ is in the predictable σ-algebra since $\Gamma_{kl}^{\boldsymbol{\nu},0}(t) I\{\boldsymbol{\alpha}(t-) = k\} g_{kl}$ is $\{\mathcal{F}_t\}$-predictable. We would like to show

$$\lambda \otimes P[\mathcal{U}] = 0. \tag{5.150}$$

Hence, we assume

$$\lambda \otimes P[\mathcal{U}] > 0, \tag{5.151}$$

or, using Fubini's Theorem,

$$E[\int_0^T I_{\mathcal{U}}(s,\omega) ds] > 0. \tag{5.152}$$

Set

$$Z(\omega) \triangleq \int_0^T I_{\mathcal{U}}(s,\omega) ds, \quad \omega \in \Omega, \tag{5.153}$$

so from (5.153) we have $Z(\omega) \geq 0$ for all $\omega \in \Omega$. Define the subset

$$\Omega_2 \triangleq \left\{ \omega \in \Omega \middle| Z(\omega) > 0 \right\}. \tag{5.154}$$

Since $E[Z] > 0$ we have that $P[\Omega_2] > 0$. Define a sequence $\{\mu_{kl}^{(n)}\}$ for $n = 1, 2, \ldots$ as follows,

$$\mu_{kl}^{(n)}(t,\omega) \triangleq n H_{\boldsymbol{\nu},0}(t-,\omega) I_{\mathcal{U}}(t,\omega), \quad (t,\omega) \in [0,T] \times \Omega. \tag{5.155}$$

Using (5.155), we can define

$$\boldsymbol{\mu}^{(n)}(t,\omega) \triangleq \begin{cases} \mu_{kl}^{(n)}(t,\omega), & k,l \in S \\ 0, & \text{for all other } i,j \in S, i \neq j \end{cases}, \tag{5.156}$$

for all $(t,\omega) \in [0,T] \times \Omega$ and $n = 1,2,\ldots$. Since $\inf_{t\in[0,T]} H_{\boldsymbol{\nu},0} > 0$ a.s. and $H_{\boldsymbol{\nu},0}$ is continuous, $\mu_{kl}^{(n)}(t)\triangle M_{ij}(t) > -1$ a.s. for all $t \in [0,T]$, and $\mu_{kl}^{(n)} \in L_{loc}^{1/2}(R_{kl})$ for each $n = 1,2,\ldots$, where $L_{loc}^{1/2}(R_{ij})$ is given in Definition A.1.1. We clearly then have $\boldsymbol{\mu}^{(n)} \in \mathcal{H}$ for every $n = 1,2,\ldots$, where $\mathcal{H}$ is given in Notation 4.3.1.

Define the set

$$\Theta(\omega) \triangleq \left\{ t \in [0,T] \middle| I_{\mathcal{U}}(t,\omega) = 1 \right\}, \quad \omega \in \Omega. \tag{5.157}$$

From (5.154), we have that $\lambda[\Theta(\omega)] > 0$ for each $\omega \in \Omega_2$, where $\lambda[\cdot]$ is Lebesgue measure. Fix $\bar{\omega} \in \Omega_2$ and put (5.156) into (5.147). Using (5.155) and expanding the $d\tilde{R}_{ij}$ integral using (4.7) results in

$$\int_0^T \frac{1}{H_{\boldsymbol{\nu},0}(s-,\bar{\omega})} dA_{\boldsymbol{\nu},0}(s,\bar{\omega}) \geq n \int_{\Theta(\bar{\omega})} \Gamma_{kl}^{\boldsymbol{\nu},0}(s,\bar{\omega}) I\{\boldsymbol{\alpha}(s-,\bar{\omega}) = k\} g_{kl} ds, \tag{5.158}$$

where the right-hand side of (5.158) is strictly positive by (5.149). Taking $n \to \infty$ in (5.158) then gives

$$\int_0^T \frac{1}{H_{\boldsymbol{\nu},0}(s-,\bar{\omega})} dA_{\boldsymbol{\nu},0}(s,\bar{\omega}) = \infty, \quad \bar{\omega} \in \Omega_2. \tag{5.159}$$

Since $H_{\boldsymbol{\nu},0}$ is continuous, we have from (5.159) that $A_{\boldsymbol{\nu},0}(T,\bar{\omega}) = \infty$. But since $P[\Omega_2] > 0$ and by the Doob-Meyer decomposition, $A_{\boldsymbol{\nu},0}(T) < \infty$ a.s., we have a contradiction in (5.151). As a result, we have shown (5.150) holds, and since $k,l \in S$ and $\boldsymbol{\nu} \in \mathcal{G}$ were chosen arbitrarily, we have that

$$\Gamma_{ij}^{\boldsymbol{\nu},0}(t) I\{\boldsymbol{\alpha}(t-) = i\} g_{ij} \leq 0, \quad \lambda \otimes P \text{ a.e. on } [0,T] \times \Omega, \tag{5.160}$$

for all $i,j \in S, i \neq j$ and $\boldsymbol{\nu} \in \mathcal{G}$. To complete the proof of Lemma 5.4.9, we shall now show that equality in fact holds in (5.160) provided that $\boldsymbol{\nu} = \hat{\boldsymbol{\nu}}$.

Fix $\boldsymbol{\nu} = \hat{\boldsymbol{\nu}}$ and fix some $k,l \in S$ where $k \neq l$ and $g_{kl} > 0$. Define the set

$$\mathcal{V} \triangleq \left\{ (t,\omega) \in [0,T] \times \Omega \middle| \Gamma_{kl}^{\hat{\boldsymbol{\nu}},0}(t,\omega) I\{\boldsymbol{\alpha}(t-,\omega) = k\} g_{kl} < 0 \right\}. \tag{5.161}$$

We again clearly have that $\mathcal{V}$ is in the predictable σ-algebra. We would again like to show

$$\lambda \otimes P[\mathcal{V}] = 0. \tag{5.162}$$

Therefore, we assume

$$\lambda \otimes P[\mathcal{V}] > 0. \tag{5.163}$$

Define the ω-section of $\mathcal{V}$ as

$$\mathcal{V}(\omega) \triangleq \left\{ t \in [0,T] \middle| \Gamma_{kl}^{\hat{\boldsymbol{\nu}},0}(t,\omega) I\{\boldsymbol{\alpha}(t-,\omega) = k\} g_{kl} < 0 \right\}, \quad \omega \in \Omega. \tag{5.164}$$

Using (5.164) and Fubini's Theorem, we can write (5.163) as

$$E[\lambda[\mathcal{V}(\omega)]] > 0, \tag{5.165}$$

where $\lambda[\cdot]$ is Lebesgue measure. Define the set

$$\Omega_4 \triangleq \left\{ \omega \in \Omega \middle| \lambda[\mathcal{V}(\omega)] > 0 \right\}. \tag{5.166}$$

By (5.165) and (5.166),

$$P[\Omega_4] > 0. \tag{5.167}$$

Similar to $\mathcal{V}$, define the set $\mathcal{W}$ as follows,

$$\mathcal{W} \triangleq \left\{ (t,\omega) \in [0,T] \times \Omega \middle| \Gamma_{kl}^{\hat{\boldsymbol{\nu}},0}(t,\omega) I\{\boldsymbol{\alpha}(t-,\omega) = k\} g_{kl} \leq 0 \right\}. \tag{5.168}$$

$\mathcal{W}$ is also in the predictable σ-algebra, and we immediately see from (5.161) and (5.168) that

$$\mathcal{V} \subset \mathcal{W}. \tag{5.169}$$

Again define the ω-section of $\mathcal{W}$ to be

$$\mathcal{W}(\omega) \triangleq \left\{ t \in [0,T] \middle| \Gamma_{kl}^{\hat{\boldsymbol{\nu}},0}(t,\omega) I\{\boldsymbol{\alpha}(t-,\omega) = k\} g_{kl} \leq 0 \right\}, \quad \omega \in \Omega. \tag{5.170}$$

We have already shown by (5.160) that

$$E[\lambda[\mathcal{W}^c(\omega)]] = 0. \tag{5.171}$$

Therefore, we define the set

$$\Omega_5 \triangleq \left\{\omega \in \Omega \middle| \lambda[\mathcal{W}^c(\omega)] = 0\right\}, \tag{5.172}$$

which by (5.171) has the property

$$P[\Omega_5] = 1. \tag{5.173}$$

Since $\mathcal{W}(\omega) \cup \mathcal{W}^c(\omega) = [0,T]$ for all $\omega \in \Omega$, and $\lambda[\mathcal{W}^c(\omega)] = 0$ for all $\omega \in \Omega$ by (5.160),

$$\lambda[\mathcal{W}(\omega)] = T, \quad \omega \in \Omega. \tag{5.174}$$

As a result, from (5.172) and (5.174), we have

$$\Omega_5 = \left\{\omega \in \Omega \middle| \lambda[\mathcal{W}(\omega)] = T\right\}. \tag{5.175}$$

Now, for all $i, j \in S, i \neq j$ such that $(i,j) \neq (k,l)$, define

$$\mu_{ij}(t,\omega) \triangleq \hat{\mu}_{ij}(t,\omega), \quad (t,\omega) \in [0,T] \times \Omega. \tag{5.176}$$

Since $\hat{\boldsymbol{\mu}} \in \mathcal{H}$, we have that $\hat{\mu}_{ij} \in \mathcal{P}^*$ and $\hat{\mu}_{ij} \in L^{1/2}_{loc}(R_{ij})$ (see Definition A.1.1) for all $i, j \in S, i \neq j$. It is then immediate from (5.176) that

$$\mu_{ij} \in \mathcal{P}^* \text{ and } \mu_{ij} \in L^{1/2}_{loc}(R_{ij}) \text{ for all } i,j \in S, i \neq j,\ (i,j) \neq (k,l). \tag{5.177}$$

For the fixed $(k,l) \in S$, we define

$$\begin{aligned} \mu_{kl}(t,\omega) \triangleq \frac{1}{2}\Big(\hat{\mu}_{ij}(t,\omega) - 1\Big) I\{\hat{\mu}_{ij}(t,\omega) > -1\} \\ + \Big(\hat{\mu}_{ij}(t,\omega) - 1\Big) I\{\hat{\mu}_{ij}(t,\omega) \leq -1\}, \quad (t,\omega) \in [0,T] \times \Omega. \end{aligned} \tag{5.178}$$

Again, it is immediate that $\mu_{kl} \in \mathcal{P}^*$. Using the elementary bound

$$x > \frac{1}{2}(x-1) > -1, \quad \text{for all } x > -1, \tag{5.179}$$

we have from (5.178) that

$$\hat{\mu}_{kl}(t,\omega) > \mu_{kl}(t,\omega), \quad (t,\omega) \in [0,T] \times \Omega. \tag{5.180}$$

Therefore, from (5.180) and the fact that $\hat{\mu}_{kl} \in L^{1/2}_{loc}(R_{kl})$, we have

$$\mu_{kl} \in \mathcal{P}^* \text{ and } \mu_{kl} \in L^{1/2}_{loc}(R_{kl}). \tag{5.181}$$

By (5.176), (5.178) and (5.179),

$$\mu_{ij}(t,\omega) > -1, \quad (t,\omega) \in \{\hat{\mu}_{ij}(t,\omega) > -1\}, i,j \in S, i \neq j. \tag{5.182}$$

Furthermore, since $\hat{\boldsymbol{\mu}} \in \mathcal{H}$,

$$\hat{\mu}_{ij}(t,\omega) > -1, \quad (t,\omega) \in \{\triangle M_{ij}(t,\omega) = 1\}, i,j \in S, i \neq j. \tag{5.183}$$

We have from (5.182) and (5.183) that

$$\mu_{ij}(t,\omega) > -1, \quad (t,\omega) \in \{\triangle M_{ij}(t,\omega) = 1\}, i,j \in S, i \neq j. \tag{5.184}$$

As a result, setting

$$\boldsymbol{\mu}(t,\omega) \triangleq \{\mu_{ij}(t,\omega)\}_{i,j\in S}, \quad (t,\omega) \in [0,T] \times \Omega, \tag{5.185}$$

we have from (5.177), (5.181), and (5.183) that

$$\boldsymbol{\mu} \in \mathcal{H}. \tag{5.186}$$

Setting $t = T$ in (5.141) of Lemma 5.4.8 and using (5.185), we have

$$\int_0^T \frac{1}{H_{\hat{\boldsymbol{\nu}},\boldsymbol{\mu}}(s-)} dA_{\hat{\boldsymbol{\nu}},\boldsymbol{\mu}}(s) = \int_0^T \frac{\Gamma^{\hat{\boldsymbol{\nu}},0}_{kl}(s)}{H_{\hat{\boldsymbol{\nu}},0}(s-)}\Big(\hat{\mu}_{kl}(s) - \mu_{kl}(s)\Big) I\{\boldsymbol{\alpha}(s-) = k\} g_{kl} ds. \tag{5.187}$$

Fix $\bar{\omega} \in \Omega_4 \cap \Omega_5$. We can then write from (5.175) and (5.187)

$$\begin{aligned}\int_0^T & \frac{1}{H_{\hat{\boldsymbol{\nu}},\boldsymbol{\mu}}(s-,\bar{\omega})} dA_{\hat{\boldsymbol{\nu}},\boldsymbol{\mu}}(s,\bar{\omega}) \\ &= \int_{\mathcal{W}(\bar{\omega})} \frac{(\hat{\mu}_{kl}(s,\bar{\omega}) - \mu_{kl}(s,\bar{\omega}))}{H_{\hat{\boldsymbol{\nu}},0}(s-,\bar{\omega})} \Gamma^{\hat{\boldsymbol{\nu}},0}_{kl}(s,\bar{\omega}) I\{\boldsymbol{\alpha}(s-,\bar{\omega}) = k\} g_{kl} ds.\end{aligned} \tag{5.188}$$

We have from (5.169) that

$$\mathcal{V}(\bar{\omega}) \subset \mathcal{W}(\bar{\omega}), \tag{5.189}$$

and therefore,

$$\Gamma^{\hat{\boldsymbol{\nu}},0}_{kl}(s,\bar{\omega}) I\{\boldsymbol{\alpha}(s-,\bar{\omega}) = k\} g_{kl} = 0, \; s \in \mathcal{W}(\bar{\omega}) \setminus \mathcal{V}(\bar{\omega}). \tag{5.190}$$

From (5.188), (5.189), and (5.190),

$$\int_0^T \frac{1}{H_{\hat{\boldsymbol{\nu}},\boldsymbol{\mu}}(s-,\bar{\omega})} dA_{\hat{\boldsymbol{\nu}},\boldsymbol{\mu}}(s,\bar{\omega}) \\ = \int_{\mathcal{V}(\bar{\omega})} \frac{(\hat{\mu}_{kl}(s,\bar{\omega}) - \mu_{kl}(s,\bar{\omega}))}{H_{\hat{\boldsymbol{\nu}},0}(s-,\bar{\omega})} \Gamma_{kl}^{\hat{\boldsymbol{\nu}},0}(s,\bar{\omega}) I\{\boldsymbol{\alpha}(s-,\bar{\omega}) = k\} g_{kl} ds. \tag{5.191}$$

But from (5.180) and the fact that $\inf_{t\in[0,T]} H_{\hat{\boldsymbol{\nu}},0}(t) > 0$ a.s.,

$$\frac{(\hat{\mu}_{kl}(t,\bar{\omega}) - \mu_{kl}(t,\bar{\omega}))}{H_{\hat{\boldsymbol{\nu}},0}(t-,\bar{\omega})} > 0, \quad t \in [0,T], \tag{5.192}$$

and from (5.164),

$$\Gamma_{kl}^{\hat{\boldsymbol{\nu}},0}(t,\bar{\omega}) I\{\boldsymbol{\alpha}(t-,\bar{\omega}) = k\} g_{kl} < 0, \quad t \in \mathcal{V}(\bar{\omega}). \tag{5.193}$$

Since $\lambda[\mathcal{V}(\bar{\omega})] > 0$ by (5.166), we have from (5.191), (5.192), and (5.193),

$$\int_0^T \frac{1}{H_{\hat{\boldsymbol{\nu}},\boldsymbol{\mu}}(s-,\bar{\omega})} dA_{\hat{\boldsymbol{\nu}},\boldsymbol{\mu}}(s,\bar{\omega}) < 0, \quad \bar{\omega} \in \Omega_4 \cap \Omega_5. \tag{5.194}$$

But since $P[\Omega_4 \cap \Omega_5] > 0$ and $A_{\hat{\boldsymbol{\nu}},\boldsymbol{\mu}}(\cdot,\omega)$ must be non-decreasing for P-almost all $\omega \in \Omega$, by Lemma 5.3.16, (5.194) cannot be true. Therefore, there is a contradiction in (5.163) and (5.162) must be true.

Since $k, l \in S$ were arbitrarily chosen and (5.160) holds for each $i, j \in S$, $i \neq j$,

$$\Gamma_{ij}^{\hat{\boldsymbol{\nu}},0}(t) I\{\boldsymbol{\alpha}(t-) = i\} g_{ij} = 0, \quad \lambda \otimes P \text{ a.e. on } [0,T] \times \Omega \tag{5.195}$$

for all $i, j \in S$, $i \neq j$.□

Proof of Theorem 5.4.4. We have from Proposition 5.4.6 that Theorem 5.4.4(1) and Theorem 5.4.4(2) are equivalent conditions and either condition imply Lemma 5.4.9. Assume Theorem 5.4.4(2) holds. Then from Lemma 5.4.9 we have that

$$\Gamma_{ij}^{\hat{\boldsymbol{\nu}},0}(t) I\{\boldsymbol{\alpha}(t-) = i\} g_{ij} = 0, \ \lambda \otimes P \text{ a.e. on } [0,T] \times \Omega \text{ for all } i,j \in S, i \neq j. \tag{5.196}$$

By setting $\boldsymbol{\mu} = 0$ in (5.141) of Lemma 5.4.8 we obtain

$$\int_0^t \frac{1}{H_{\hat{\boldsymbol{\nu}},0}(s-)} dA_{\hat{\boldsymbol{\nu}},0}(s) = \sum_{\substack{i,j\in S \\ i\neq j}} \int_0^t \frac{\Gamma_{ij}^{\hat{\boldsymbol{\nu}},0}(s)}{H_{\hat{\boldsymbol{\nu}},0}(s-)} \hat{\mu}_{ij}(s) I\{\boldsymbol{\alpha}(s-) = i\} g_{ij} ds, \quad t \in [0,T]. \tag{5.197}$$

Putting (5.196) into (5.197) yields

$$\int_0^t \frac{1}{H_{\hat{\boldsymbol{\nu}},0}(s-)} dA_{\hat{\boldsymbol{\nu}},0}(s) = 0 \quad \text{a.s.,} \quad t \in [0,T]. \tag{5.198}$$

Putting (5.198) into the consumption process $\hat{c}_{\boldsymbol{\nu}}$ given in Proposition 5.3.20 with $\boldsymbol{\nu} = \hat{\boldsymbol{\nu}}$ yields

$$\hat{c}_{\hat{\boldsymbol{\nu}}}(t) = -\sum_{\substack{i,j \in S \\ i \neq j}} \int_0^t \frac{\Gamma_{ij}^{\hat{\boldsymbol{\nu}},0}(s)}{H_{\hat{\boldsymbol{\nu}},0}(s-)} dM_{ij}(s), \quad t \in [0,T], \tag{5.199}$$

which is an $\{\mathcal{F}_t\}$-local martingale since M_{ij} is an $\{\mathcal{F}_t\}$-martingale for all $i,j \in S$, $i \neq j$. Since $\hat{c}_{\hat{\boldsymbol{\nu}}}$ is in $\mathcal{C}$ by Proposition 5.3.20, which asserts that $\hat{c}_{\hat{\boldsymbol{\nu}}}$ is a non-decreasing process, we have by Proposition C.11.5 that the only null-at-the-origin local martingale which is non-decreasing is the zero process, meaning

$$\hat{c}_{\hat{\boldsymbol{\nu}}}(t) = 0 \quad \text{a.s.,} \quad t \in [0,T]. \tag{5.200}$$

Using (5.200), we can now write the optimal consumption process $\hat{c}$ from Proposition 5.3.24 as

$$\hat{c}(t) = -\int_0^t \hat{X}(s) \Big\{ \tilde{g}_K(s, \hat{\boldsymbol{\nu}}(s)) - \Big[g(s, \hat{\boldsymbol{p}}(s)) - \hat{\boldsymbol{p}}(s)^\top \hat{\boldsymbol{\nu}}(s) \Big] \Big\} ds, \quad t \in [0,T]. \tag{5.201}$$

Since $\hat{\boldsymbol{p}}(t) \in K$ a.e. by Proposition 5.3.26, we have from (5.20)

$$\tilde{g}_K(t, \hat{\boldsymbol{\nu}}(t)) \geq g(t, \hat{\boldsymbol{p}}(t)) - \hat{\boldsymbol{p}}(t)^\top \hat{\boldsymbol{\nu}}(t) \text{ a.s.,} \quad t \in [0,T]. \tag{5.202}$$

From (5.202) and the fact that $\hat{X}(t) > 0$ a.s. for all $t \in [0,T]$ by Proposition 5.3.15(3), we have from (5.201) that $\hat{c}$ is a non-increasing process. But $\hat{c} \in \mathcal{C}$ implying that $\hat{c}$ is a non-decreasing process as well. Since $\hat{c}$ is both a non-increasing and non-decreasing process, we must have

$$\hat{c}(t) = 0 \quad \text{a.s.,} \quad t \in [0,T], \tag{5.203}$$

which shows Theorem 5.4.4(4). Therefore, with $\hat{u} = E[H_{\hat{\boldsymbol{\nu}},\hat{\boldsymbol{\mu}}}(T)B]$, $\boldsymbol{p}(t) = \hat{\boldsymbol{p}}(t)$ as defined in Proposition 5.3.21, and $\hat{c}(t) = 0$ a.s. for $t \in [0,T]$, from Theorem 5.3.17, we have that

$$\hat{X}(t) = X^{\hat{u},0,\boldsymbol{p}}(t) \quad \text{a.s.,} \quad t \in [0,T], \tag{5.204}$$

which is a wealth process that hedges the contingent claim B. From Theorem 5.4.4(2) and (5.204) we have that

$$H_{\hat{\boldsymbol{\nu}},\hat{\boldsymbol{\mu}}} X^{\hat{u},0,\boldsymbol{p}} \text{ is an } \{\mathcal{F}_t\}\text{-martingale.} \tag{5.205}$$

Thus, Theorem 5.4.4(2) implies Theorem 5.4.4(3). Finally, suppose Theorem 5.4.4(3) holds. Then, using the martingale property of $H_{\hat{\boldsymbol{\nu}},\hat{\boldsymbol{\mu}}}X^{\hat{u},0,\boldsymbol{p}}$ gives

$$E[H_{\hat{\boldsymbol{\nu}},\hat{\boldsymbol{\mu}}}(T)B] = E[H_{\hat{\boldsymbol{\nu}},\hat{\boldsymbol{\mu}}}(T)X^{\hat{u},0,\boldsymbol{p}}(T)] = E[X^{\hat{u},0,\boldsymbol{p}}(0)] = \hat{u}. \tag{5.206}$$

Therefore, Theorem 5.4.4(3) implies Theorem 5.4.4(1), completing the proof. □

Remark 5.4.10. Theorem 5.4.4 most importantly tells us that if the price $\hat{u}$ of the contingent claim B, given by a supremum taken over the space of dual processes $\mathcal{G} \times \mathcal{H}$, is in fact attained by some $(\hat{\boldsymbol{\nu}}, \hat{\boldsymbol{\mu}}) \in \mathcal{G} \times \mathcal{H}$, one can only then find a hedging strategy $(0, \boldsymbol{p}) \in \mathcal{A}(\hat{u})$ which exactly hedges B.

Remark 5.4.11. Theorem 5.4.4 can be reworded to fit the context of Chapter 4 with no significant changes to the proof. In fact, proving this result in the unconstrained regime-switching market only requires optimizing over the space $\mathcal{H}$ instead of the joint space $\mathcal{G} \times \mathcal{H}$.

Remark 5.4.12. It may be of interest to have a converse of Theorem 5.4.4, since if one could guarantee that the required consumption is zero, then the supremum defining $\hat{u}$ is attained by some pair $(\hat{\boldsymbol{\nu}}, \hat{\boldsymbol{\mu}}) \in \mathcal{G} \times \mathcal{H}$. In general, this converse is not true; however under the specific condition that the contingent claim B is almost surely bounded, we do obtain a converse. In fact, under this boundedness condition on B, if the optimal consumption is zero, one can guarantee that the supremum at (5.34) is attained by the pair $(\hat{\boldsymbol{\nu}}, \hat{\boldsymbol{\mu}}) = (\bar{\boldsymbol{\nu}}, 0)$ where $\bar{\boldsymbol{\nu}} \in \mathcal{G}$ is defined by Proposition 5.3.23. The full statement of this result is given in Theorem 5.4.15.

We must first establish the following lemma to prove Theorem 5.4.15.

Lemma 5.4.13. *Let $B : (\Omega, \mathcal{F}_T, P) \mapsto (0, \infty)$ be a contingent claim such that $P[0 \leq B \leq \beta] = 1$ for some $\beta \in (0, \infty)$ and let $(\bar{\boldsymbol{\nu}}, \bar{\boldsymbol{\mu}}) \in \mathcal{G} \times \mathcal{H}$ be a pair such that $\mathcal{E}(-\boldsymbol{\theta}_{\bar{\nu}} \bullet \boldsymbol{W})(t)\mathcal{E}(\bar{\boldsymbol{\mu}} \bullet \boldsymbol{M})(t)$ is an $\{\mathcal{F}_t\}$-martingale. Then $H_{\bar{\boldsymbol{\nu}},\bar{\boldsymbol{\mu}}}\hat{X}$ is of class $\mathcal{D}[0,T]$ (see Definition C.2.3).*

Remark 5.4.14. One immediately sees that the set of pairs $(\bar{\boldsymbol{\nu}}, \bar{\boldsymbol{\mu}}) \in \mathcal{G} \times \mathcal{H}$ for which $\mathcal{E}(-\boldsymbol{\theta}_{\bar{\nu}} \bullet \mathbf{W})(t)\mathcal{E}(\bar{\boldsymbol{\mu}} \bullet \mathbf{M})(t)$ is an $\{\mathcal{F}_t\}$-martingale is non-empty by setting $\bar{\boldsymbol{\mu}} = 0$ and using the Novikov criterion (see Theorem C.10.5).

Proof. We have by Condition 4.1.4 that $S_0(T) \geq s_0$ a.s. for some constant $s_0 > 0$. It is also easy to see that $\tilde{g}_K(t, \boldsymbol{\nu}(t)) \geq 0$ a.s. for all $t \in [0,T]$ and $\boldsymbol{\nu} \in \mathcal{G}$. Therefore,

$$0 \leq H_{\boldsymbol{\nu},\boldsymbol{\mu}}(T)B \leq \frac{\beta}{s_0}\mathcal{E}(-\boldsymbol{\theta}_{\boldsymbol{\nu}} \bullet \mathbf{W})(T)\mathcal{E}(\boldsymbol{\mu} \bullet \mathbf{M})(T) \text{ a.s.,} \tag{5.207}$$

for all $t \in [0,T]$ and $(\boldsymbol{\nu},\boldsymbol{\mu}) \in \mathcal{G} \times \mathcal{H}$. Since $\mathcal{E}(-\boldsymbol{\theta}_{\boldsymbol{\nu}} \bullet \mathbf{W})(t)\mathcal{E}(\boldsymbol{\mu} \bullet \mathbf{M})(t) \geq 0$ a.s. for all $t \in [0,T]$ and it is an $\{\mathcal{F}_t\}$-local martingale, it is an $\{\mathcal{F}_t\}$-supermartingale by Proposition C.11.3, and

$$E[\mathcal{E}(-\boldsymbol{\theta}_{\boldsymbol{\nu}} \bullet \mathbf{W})(T)\mathcal{E}(\boldsymbol{\mu} \bullet \mathbf{M})(T)] \leq 1, \quad (\boldsymbol{\nu},\boldsymbol{\mu}) \in \mathcal{G} \times \mathcal{H}. \tag{5.208}$$

Denote by $\mathcal{S}$ the set of all stopping times taking values in $[0,T]$ and fix some $\tau \in \mathcal{S}$. Also denote by $\mathcal{D}_{\tau,\bar{\boldsymbol{\nu}},\bar{\boldsymbol{\mu}}}$ the set of processes $(\boldsymbol{\nu},\boldsymbol{\mu}) \in \mathcal{G} \times \mathcal{H}$ that agree with $(\bar{\boldsymbol{\nu}},\bar{\boldsymbol{\mu}})$ up until τ. Using Proposition 5.3.15(7), we may then write

$$\begin{aligned} 0 \leq H_{\bar{\boldsymbol{\nu}},\bar{\boldsymbol{\mu}}}(\tau)\hat{X}(\tau) &= H_{\bar{\boldsymbol{\nu}},\bar{\boldsymbol{\mu}}}(\tau) \operatorname*{ess-sup}_{(\boldsymbol{\nu},\boldsymbol{\mu}) \in \mathcal{G}\times\mathcal{H}} E\left[\frac{H_{\boldsymbol{\nu},\boldsymbol{\mu}}(T)B}{H_{\boldsymbol{\nu},\boldsymbol{\mu}}(\tau)}\middle|\mathcal{F}_\tau\right] \\ &= H_{\bar{\boldsymbol{\nu}},\bar{\boldsymbol{\mu}}}(\tau) \operatorname*{ess-sup}_{(\boldsymbol{\nu},\boldsymbol{\mu}) \in \mathcal{D}_{\tau,\bar{\boldsymbol{\nu}},\bar{\boldsymbol{\mu}}}} E\left[\frac{H_{\boldsymbol{\nu},\boldsymbol{\mu}}(T)B}{H_{\boldsymbol{\nu},\boldsymbol{\mu}}(\tau)}\middle|\mathcal{F}_\tau\right] \\ &= \operatorname*{ess-sup}_{(\boldsymbol{\nu},\boldsymbol{\mu}) \in \mathcal{D}_{\tau,\bar{\boldsymbol{\nu}},\bar{\boldsymbol{\mu}}}} E\left[H_{\boldsymbol{\nu},\boldsymbol{\mu}}(T)B\middle|\mathcal{F}_\tau\right] \text{ a.s.,} \end{aligned} \tag{5.209}$$

since $H_{\bar{\boldsymbol{\nu}},\bar{\boldsymbol{\mu}}}(\tau) = H_{\boldsymbol{\nu},\boldsymbol{\mu}}(\tau)$ for all $(\boldsymbol{\nu},\boldsymbol{\mu}) \in \mathcal{D}_{\tau,\bar{\boldsymbol{\nu}},\bar{\boldsymbol{\mu}}}$ and the conditional expectation on the right-hand side of the first line only depends on $t \in [\tau,T]$. From (5.207) and the supermartingale property of $\mathcal{E}(-\boldsymbol{\theta}_{\boldsymbol{\nu}} \bullet \mathbf{W})(t)\mathcal{E}(\boldsymbol{\mu} \bullet \mathbf{M})(t)$,

$$\begin{aligned} E\left[H_{\boldsymbol{\nu},\boldsymbol{\mu}}(T)B\middle|\mathcal{F}_\tau\right] &\leq \frac{\beta}{s_0}E\left[\mathcal{E}(-\boldsymbol{\theta}_{\boldsymbol{\nu}} \bullet \mathbf{W})(T)\mathcal{E}(\boldsymbol{\mu} \bullet \mathbf{M})(T)\middle|\mathcal{F}_\tau\right] \\ &\leq \frac{\beta}{s_0}\mathcal{E}(-\boldsymbol{\theta}_{\boldsymbol{\nu}} \bullet \mathbf{W})(\tau)\mathcal{E}(\boldsymbol{\mu} \bullet \mathbf{M})(\tau) \\ &= \frac{\beta}{s_0}\mathcal{E}(-\boldsymbol{\theta}_{\bar{\boldsymbol{\nu}}} \bullet \mathbf{W})(\tau)\mathcal{E}(\bar{\boldsymbol{\mu}} \bullet \mathbf{M})(\tau) \text{ a.s.,} \end{aligned} \tag{5.210}$$

for all $(\boldsymbol{\nu},\boldsymbol{\mu}) \in \mathcal{D}_{\tau,\bar{\boldsymbol{\nu}},\bar{\boldsymbol{\mu}}}$. Since $\mathcal{E}(-\boldsymbol{\theta}_{\bar{\boldsymbol{\nu}}} \bullet \mathbf{W})(t)\mathcal{E}(\bar{\boldsymbol{\mu}} \bullet \mathbf{M})(t)$ is an $\{\mathcal{F}_t\}$-martingale, we have from (5.209) and (5.210)

$$0 \leq H_{\bar{\boldsymbol{\nu}},\bar{\boldsymbol{\mu}}}(\tau)\hat{X}(\tau) \leq \frac{\beta}{s_0}E\left[\mathcal{E}(-\boldsymbol{\theta}_{\bar{\boldsymbol{\nu}}} \bullet \mathbf{W})(T)\mathcal{E}(\bar{\boldsymbol{\mu}} \bullet \mathbf{M})(T)\middle|\mathcal{F}_\tau\right] \text{ a.s.,} \tag{5.211}$$

for all stopping times $\tau \in \mathcal{S}$. Therefore, since the right-hand side of equation (5.211) is uniformly integrable by the Proposition C.1.10, the collection of random variables $\{H_{\bar{\boldsymbol{\nu}},\bar{\boldsymbol{\mu}}}(\tau)\hat{X}(\tau)\}_{\tau\in\mathcal{S}}$ is uniformly integrable and by Definition C.2.3, $H_{\bar{\boldsymbol{\nu}},\bar{\boldsymbol{\mu}}}\hat{X}$ is of class $\mathcal{D}[0,T]$. ☐

Theorem 5.4.15. *Let $B : (\Omega,\mathcal{F}_T,P) \mapsto (0,\infty)$ be a contingent claim such that $P[0 \leq B \leq \beta] = 1$ for some $\beta \in (0,\infty)$. Then $\hat{u} < \infty$ and Theorem 5.4.4(1)-(4) are equivalent for the pair $(\bar{\boldsymbol{\nu}},0)$, where $\bar{\boldsymbol{\nu}} \in \mathcal{G}$ is given in Proposition 5.3.23.*

Proof. We have by Condition 4.1.4 that $S_0(T) \geq s_0$ a.s. for some constant $s_0 > 0$. It is also easy to see that $\tilde{g}_K(t, \boldsymbol{\nu}(t)) \geq 0$ a.s. for all $t \in [0,T]$ and $\boldsymbol{\nu} \in \mathcal{G}$. Therefore,

$$0 \leq H_{\boldsymbol{\nu},\boldsymbol{\mu}}(T)B \leq \frac{\beta}{s_0}\mathcal{E}(-\boldsymbol{\theta}_{\boldsymbol{\nu}} \bullet \mathbf{W})(T)\mathcal{E}(\boldsymbol{\mu} \bullet \mathbf{M})(T) \text{ a.s.,} \tag{5.212}$$

for all $t \in [0,T]$ and $(\boldsymbol{\nu}, \boldsymbol{\mu}) \in \mathcal{G} \times \mathcal{H}$. Since $\mathcal{E}(-\boldsymbol{\theta}_{\boldsymbol{\nu}} \bullet \mathbf{W})(t)\mathcal{E}(\boldsymbol{\mu} \bullet \mathbf{M})(t) \geq 0$ a.s. for all $t \in [0,T]$ and it is an $\{\mathcal{F}_t\}$-local martingale,, it is an $\{\mathcal{F}_t\}$-supermartingale by Proposition C.11.3, and

$$E[\mathcal{E}(-\boldsymbol{\theta}_{\boldsymbol{\nu}} \bullet \mathbf{W})(T)\mathcal{E}(\boldsymbol{\mu} \bullet \mathbf{M})(T)] \leq 1, \quad (\boldsymbol{\nu}, \boldsymbol{\mu}) \in \mathcal{G} \times \mathcal{H}. \tag{5.213}$$

We have by (5.212) and (5.213),

$$u_{\boldsymbol{\nu},\boldsymbol{\mu}} \triangleq E[H_{\boldsymbol{\nu},\boldsymbol{\mu}}(T)B] \leq \frac{\beta}{s_0}, \quad (\boldsymbol{\nu}, \boldsymbol{\mu}) \in \mathcal{G} \times \mathcal{H}. \tag{5.214}$$

It is then easy to see that

$$\hat{u} \triangleq \sup_{(\boldsymbol{\nu},\boldsymbol{\mu}) \in \mathcal{G} \times \mathcal{H}} u_{\boldsymbol{\nu},\boldsymbol{\mu}} < \infty. \tag{5.215}$$

Now, assume that $\hat{c}(t) = 0$ a.s. for all $t \in [0,T]$. For the choice $(\bar{\boldsymbol{\nu}}, 0) \in \mathcal{G} \times \mathcal{H}$ we then have that Theorem 5.4.4(4) holds. From Theorem 5.4.4(4) and the definition of $\hat{c}$ from Proposition 5.3.24, we obtain

$$A_{\bar{\boldsymbol{\nu}},0}(t) = 0 \text{ a.s.}, \quad t \in [0,T], \tag{5.216}$$

and from the Doob-Meyer decomposition (see Theorem C.11.4) we get $H_{\bar{\boldsymbol{\nu}},0}\hat{X}$ is an $\{\mathcal{F}_t\}$-local martingale. Furthermore, we have from the Novikov criterion (see Theorem C.10.5) that $\mathcal{E}(-\boldsymbol{\theta}_{\bar{\boldsymbol{\nu}}} \bullet \mathbf{W})(t)$ is an $\{\mathcal{F}_t\}$-martingale, and therefore from Lemma 5.4.13, $H_{\bar{\boldsymbol{\nu}},0}\hat{X}$ is a class $\mathcal{D}[0,T]$ local martingale. From Proposition C.11.2 any class $\mathcal{D}[0,T]$ local martingale is in fact an $\{\mathcal{F}_t\}$-martingale. As a result we have that $H_{\bar{\boldsymbol{\nu}},0}\hat{X}$ is an $\{\mathcal{F}_t\}$-martingale and,

$$\hat{u} = E[H_{\bar{\boldsymbol{\nu}},0}(0)\hat{X}(0)] = E[H_{\bar{\boldsymbol{\nu}},0}(T)\hat{X}(T)] = E[H_{\bar{\boldsymbol{\nu}},0}(T)B], \tag{5.217}$$

which is Theorem 5.4.4(1). Since Theorem 5.4.4(1)-(3) are equivalent and each imply Theorem 5.4.4(4), we have shown the equivalence of Theorem 5.4.4(1)-(4). □

Chapter 6

Approximate Hedging in a Regime-Switching Market Model

Often times the least initial wealth $\hat{u}$ required to hedge a contingent claim B, defined in (5.34), is so high that it is unrealistic for an investor to pay such a price. This is due to the fact that the value of $\hat{u}$ is determined by taking a supremum over the very large space of dual processes $\mathcal{G} \times \mathcal{H}$. Instead of starting with initial wealth $\hat{u}$, an investor may only be able start with a much lower initial wealth $x < \hat{u}$ that they can "afford", and therefore, would like to invest in the "best" way as to minimize his/her exposure to risk when hedging the contingent claim B. We call such an approach to hedging *approximate hedging*. In this chapter we look at the approximate hedging of a given contingent claim B in a regime-switching market model with both convex portfolio constraints and margin requirements. Much of this chapter follows the work of Cvitanic [8] who solved a problem of approximate hedging in a standard Brownian motion market model with convex portfolio constraints. Cvitanic addresses the problem of minimizing an investor's exposure to risk by minimizing the convex utility function

$$V(x, c, \boldsymbol{p}) = E[B - X^{(x,c,\boldsymbol{p})}(T)]^{+} \triangleq E[\max\{B - X^{(x,c,\boldsymbol{p})}(T), 0\}] \tag{6.1}$$

over all investment strategies $(c, \boldsymbol{p}) \in \mathcal{A}(x)$ for a given initial wealth $x < \hat{u}$. Minimizing (6.1) amounts to the agent finding an "optimal" investment strategy $(\tilde{c}, \tilde{\boldsymbol{p}}) \in \mathcal{A}(x)$ which results in the least expected loss between the value of the contingent claim B and the agent's own terminal wealth. Cvitanic's method of optimizing (6.1) proves to be quite general as many of the results simply depend on the final time σ-algebra $\mathcal{F}_T$. Therefore, most of the results from Cvitanic [8] carry over to approximate hedging in a regime-switching market model. Even though this is the case, we give full treatment to the problem of approximate hedging in this chapter, without requiring the reader to refer to Cvitanic [8].

6.1 The Approximate Hedging Problem

We suppose that an agent trades in the same regime-switching market model as in Chapter 5. That is a market defined by conditions, remarks and results of Section 5.1. We again suppose that at time $t = T$, the agent is contracted to pay off a strictly positive contingent claim $B : (\Omega, \mathcal{F}_T, P) \mapsto (0, \infty)$, which in the present chapter is assumed to be a random variable in $L^1(\Omega, \mathcal{F}_T, P)$. We know from Theorem 5.3.17 that if the extended real number $\hat{u}$, defined as

$$\hat{u} \triangleq \sup_{(\boldsymbol{\nu},\boldsymbol{\mu}) \in \mathcal{G} \times \mathcal{H}} E\left[H_{\boldsymbol{\nu},\boldsymbol{\mu}}(T) B\right], \tag{6.2}$$

is finite, then there exists a hedging strategy $(\hat{c}, \hat{\boldsymbol{p}}) \in \mathcal{A}(\hat{u})$ where $\mathcal{A}(\cdot)$ is the admissible set defined in (5.14). In other words, if an investor begins trading with initial wealth $\hat{u} < \infty$, then there exists a pair $(\hat{c}, \hat{\boldsymbol{p}}) \in \mathcal{A}(\hat{u})$ such that $X^{(\hat{u},\hat{c},\hat{\boldsymbol{p}})}(T) = B$ a.s. Furthermore, $\hat{u}$ is the least initial wealth in which a hedging strategy does exist. In light of this, we assume for the rest of this chapter that B is given such that $\hat{u}$ is indeed finite.

We begin this section by extending Theorem 5.3.17 to show that the wealth process $X^{(\hat{u},\hat{c},\hat{\boldsymbol{p}})}$ starting from initial wealth $\hat{u}$ with hedging strategy $(\hat{c}, \hat{\boldsymbol{p}}) \in \mathcal{A}(\hat{u})$ defined in Chapter 5 is the "cheapest" wealth process which hedges the claim B. We can then show that an agent who begins trading with an initial wealth x that is *greater than* $\hat{u}$ can always hedge the contingent claim B "without risk", i.e. starting from initial wealth $x \geq \hat{u}$, there exists a pair $(c, \boldsymbol{p}) \in \mathcal{A}(x)$ such that $P[X^{(x,c,\boldsymbol{p})}(T) \geq B] = 1$.

Theorem 6.1.1. *Given a contingent claim random variable $B : (\Omega, \mathcal{F}_T, P) \mapsto (0, \infty)$, with $\hat{u} < \infty$, the process $\hat{X}$ defined in Proposition 5.3.15 is finite and is equal to the minimal admissible wealth process hedging the contingent claim B. More precisely, there exists a pair $(\hat{c}, \hat{\boldsymbol{p}}) \in \mathcal{A}(\hat{u})$ such that*

$$\hat{X}(t) = X^{(\hat{u},\hat{c},\hat{\boldsymbol{p}})}(t) \ \text{a.s.}, \quad t \in [0, T], \tag{6.3}$$

and, if for some $x > 0$ and some pair $(c, \boldsymbol{p}) \in \mathcal{A}(x)$ we have

$$X^{(x,c,\boldsymbol{p})}(T) \geq B \ \text{a.s.}, \tag{6.4}$$

then

$$X^{(x,c,\boldsymbol{p})}(t) \geq \hat{X}(t) \ \text{a.s.}, \quad t \in [0, T]. \tag{6.5}$$

Proof. Since $\hat{u} < \infty$ and $\hat{X}(\cdot)$ is a càdlàg process over $[0,T]$, it is almost surely bounded. That is

$$\hat{X}(t) < \infty \text{ a.s.}, \quad t \in [0,T]. \tag{6.6}$$

From Theorem 5.3.17, we know there exists a pair $(\hat{c}, \hat{\boldsymbol{p}}) \in \mathcal{A}(\hat{u})$ which hedges the contingent claim B from initial wealth $\hat{u}$ such that

$$\hat{X}(t) = X^{(\hat{u},\hat{c},\hat{\boldsymbol{p}})}(t) \text{ a.s.}, \quad t \in [0,T]. \tag{6.7}$$

Now take some admissible wealth process $X^{(x,c,\boldsymbol{p})}$ where $x \in (0,\infty)$ and $(c,\boldsymbol{p}) \in \mathcal{A}(x)$ such that

$$X^{(x,c,\boldsymbol{p})}(T) \geq B \text{ a.s.} \tag{6.8}$$

From Proposition 5.3.9 we have that $J^{(x,c,\boldsymbol{p})}_{\boldsymbol{\nu},\boldsymbol{\mu}}(\cdot)$, defined in (5.25), is a non-negative $\{\mathcal{F}_t\}$-supermartingale for each $(\boldsymbol{\nu},\boldsymbol{\mu}) \in \mathcal{G} \times \mathcal{H}$. Define the process $\{z_{\boldsymbol{\nu},\boldsymbol{\mu}}(t), t \in [0,T]\}$ by

$$\begin{aligned} z_{\boldsymbol{\nu},\boldsymbol{\mu}}(t) \triangleq \int_0^t X^{(x,c,\boldsymbol{p})}(s) H_{\boldsymbol{\nu},\boldsymbol{\mu}}(s) \Big[\tilde{g}_K(s,\boldsymbol{\nu}(s)) - \Big(g(s,\boldsymbol{p}(s)) - \boldsymbol{p}(s)^\top \boldsymbol{\nu}(s) \Big) \Big] ds \\ + \int_0^t H_{\boldsymbol{\nu},\boldsymbol{\mu}}(s-) dc(s) + \sum_{0<s\leq t} H_{\boldsymbol{\nu},\boldsymbol{\mu}}(s-) \triangle c(s) \triangle (\boldsymbol{\mu} \bullet \mathbf{M})(s), \end{aligned} \tag{6.9}$$

for all $t \in [0,T]$ and $(\boldsymbol{\nu},\boldsymbol{\mu}) \in \mathcal{G} \times \mathcal{H}$. It is easy to check that $z_{\boldsymbol{\nu},\boldsymbol{\mu}}(\cdot)$ is a non-decreasing process for each $(\boldsymbol{\nu},\boldsymbol{\mu}) \in \mathcal{G} \times \mathcal{H}$. From (6.9) and (5.25), we have

$$H_{\boldsymbol{\nu},\boldsymbol{\mu}}(t) X^{(x,c,\boldsymbol{p})}(t) = J^{(x,c,\boldsymbol{p})}_{\boldsymbol{\nu},\boldsymbol{\mu}}(t) - z_{\boldsymbol{\nu},\boldsymbol{\mu}}(t), \quad t \in [0,T], (\boldsymbol{\nu},\boldsymbol{\mu}) \in \mathcal{G} \times \mathcal{H}. \tag{6.10}$$

Using the supermartingale property of $J^{(x,c,\boldsymbol{p})}_{\boldsymbol{\nu},\boldsymbol{\mu}}$ and the non-decreasing property of $z_{\boldsymbol{\nu},\boldsymbol{\mu}}(\cdot)$, we have from (6.10) that

$$H_{\boldsymbol{\nu},\boldsymbol{\mu}} X^{(x,c,\boldsymbol{p})} \in \mathcal{SPM}(\{\mathcal{F}_t\}, P), \quad (\boldsymbol{\nu},\boldsymbol{\mu}) \in \mathcal{G} \times \mathcal{H}. \tag{6.11}$$

Using (6.8) and supermartingale property of $H_{\boldsymbol{\nu},\boldsymbol{\mu}} X^{(x,c,\boldsymbol{p})}$,

$$H_{\boldsymbol{\nu},\boldsymbol{\mu}}(t) X^{(x,c,\boldsymbol{p})}(t) \geq E[H_{\boldsymbol{\nu},\boldsymbol{\mu}}(T) B | \mathcal{F}_t] \text{ a.s.}, \quad t \in [0,T], (\boldsymbol{\nu},\boldsymbol{\mu}) \in \mathcal{G} \times \mathcal{H}, \tag{6.12}$$

and as a result,

$$\begin{aligned} X^{(x,c,\boldsymbol{p})}(t) &\geq \operatorname*{ess\text{-}sup}_{(\boldsymbol{\nu},\boldsymbol{\mu}) \in \mathcal{G} \times \mathcal{H}} \frac{E\left[H_{\boldsymbol{\nu},\boldsymbol{\mu}}(T) B \,\middle|\, \mathcal{F}_t \right]}{H_{\boldsymbol{\nu},\boldsymbol{\mu}}(t)} \\ &= \hat{X}(t) \text{ a.s.}, \quad t \in [0,T]. \end{aligned} \tag{6.13}$$

□

Remark 6.1.2. Starting with initial wealth $x \geq \hat{u}$, an agent can always find a pair $(c, \boldsymbol{p}) \in \mathcal{A}(x)$ such that $X^{(x,c,\boldsymbol{p})}(T) \geq B$ a.s. Indeed, from Proposition 5.4.1 we have

$$X^{(\hat{u},0,\hat{\boldsymbol{p}})}(t) \geq X^{(\hat{u},\hat{c},\hat{\boldsymbol{p}})}(t) \text{ a.s.}, \quad t \in [0,T], \tag{6.14}$$

and similarly, for any $x \geq \hat{u}$,

$$X^{(x,0,\hat{\boldsymbol{p}})}(t) \geq X^{(\hat{u},0,\hat{\boldsymbol{p}})}(t) \text{ a.s.}, \quad t \in [0,T]. \tag{6.15}$$

As a result, from (6.14) and (6.15), $(0, \hat{\boldsymbol{p}}) \in \mathcal{A}(x)$ and

$$X^{(x,0,\hat{\boldsymbol{p}})}(T) \geq B \text{ a.s.} \tag{6.16}$$

Remark 6.1.3. By Theorem 6.1.1 and Remark 6.1.2 one can always "hedge without risk" starting from some initial wealth $x \geq \hat{u}$. However, this is not possible for any $x < \hat{u}$ by Remark 5.2.3. That is, if $x < \hat{u}$ then $P[X^{(x,c,\boldsymbol{p})}(T) < B] > 0$ for any pair $(c, \boldsymbol{p}) \in \mathcal{A}(x)$.

By Remark 6.1.3, an agent cannot promise a trading strategy which super-hedges a contingent claim B with probability one if they begin trading from some initial wealth x *less than* $\hat{u}$. Because of this, the agent may instead be interested in finding a trading strategy that minimizes his/her losses at time $t = T$ when starting from initial wealth $x < \hat{u}$. This motivates us to define the *cost function (risk criterion) of approximate hedging*

$$V(x) \triangleq \inf_{(c,\boldsymbol{p}) \in \mathcal{A}(x)} E[\bar{B} - \bar{X}^{(x,c,\boldsymbol{p})}(T)]^+, \quad x \in (0,\infty), \tag{6.17}$$

where $E[\bar{B} - \bar{X}^{(x,c,\boldsymbol{p})}(T)]^+ \triangleq E[\max\{\bar{B} - \bar{X}^{(x,c,\boldsymbol{p})}(T), 0\}]$ and

$$\bar{B} \triangleq \frac{B}{S_0(T)}, \quad \bar{X}^{(x,c,\boldsymbol{p})}(t) \triangleq \frac{X^{(x,c,\boldsymbol{p})}(t)}{S_0(t)}, \quad t \in [0,T], \tag{6.18}$$

which at each $x \in (0,\infty)$ returns the minimum expected discounted loss over all trading strategies $(c, \boldsymbol{p}) \in \mathcal{A}(x)$. The agent would be best served to follow such a risk criterion if they cannot come up with enough capital to exactly hedge the claim B, as following such a strategy minimizes monetary loss in a tangible way (as opposed to a mean-squared cost function, for example). Therefore, the agent's goal in solving this approximate hedging problem is to find some pair $(\tilde{c}, \tilde{\boldsymbol{p}}) \in \mathcal{A}(x)$ for some initial wealth $x > 0$ they can afford so that their loss is given by

$$V(x) = \inf_{(c,\boldsymbol{p}) \in \mathcal{A}(x)} E[\bar{B} - \bar{X}^{(x,c,\boldsymbol{p})}(T)]^+ = E[\bar{B} - \bar{X}^{(x,\tilde{c},\tilde{\boldsymbol{p}})}(T)]^+. \tag{6.19}$$

It is not readily apparent that the infimum in (6.19) is attainable by some $(\tilde{c}, \tilde{\boldsymbol{p}}) \in \mathcal{A}(x)$. As a result, our goal in this chapter is to show that the infimum in (6.19) is indeed attainable and one can construct a solution $(\tilde{c}, \tilde{\boldsymbol{p}}) \in \mathcal{A}(x)$ which solves (6.19) for each $x \in (0,\infty)$.

Remark 6.1.4. If $x \geq \hat{u}$, we have $V(x) = 0$ from (6.17) and Theorem 6.1.1. We can then assume that

$$0 < x < \hat{u}. \tag{6.20}$$

We also assume that

$$X^{(y,c,\boldsymbol{p})}(T) \leq B \text{ a.s.}, \quad y \in (0, \infty), \tag{6.21}$$

since the agent can always "consume" down to the value B through the consumption process c at time $t = T$ if $y \geq \hat{u}$.

6.2 Solution to the Approximate Hedging Problem

To help us solve the optimization problem in (6.17), we define the equivalent probability measure $P^B \equiv P$

$$P^B[A] \triangleq \frac{1}{E[\bar{B}]} E[\bar{B} I_A], \quad A \in \mathcal{F}_T. \tag{6.22}$$

Using the probability measure P^B in the cost function (6.17), we get the equivalent formulation

$$V(x) = E[\bar{B}] \inf_{(c,\boldsymbol{p}) \in \mathcal{A}(x)} E^B \left[1 - \frac{X^{(x,c,\boldsymbol{p})}(T)}{B} \right]^+. \tag{6.23}$$

We approach finding the solution to (6.23) by looking at a similar, but deterministic, convex loss function

$$R(y) \triangleq (1 - y)^+ = \max\{1 - y, 0\}, \tag{6.24}$$

and consider its Legendre-Fenchel transform

$$\tilde{R}(w) \triangleq \min_{0 \leq y \leq 1} [R(y) + yw] = w \wedge 1, \quad w \geq 0, \tag{6.25}$$

where $w \wedge 1 \triangleq \min\{w, 1\}$. The minimum in (6.25) is attained by any number $I(w; b)$ of the form

$$I(w; b) \triangleq \begin{cases} 0, & w > 1 \\ 1, & 0 \leq w < 1 \\ b, & w = 1, \end{cases} \tag{6.26}$$

where $0 \leq b \leq 1$. Following this approach, we would like to put (6.23) into the form of (6.24) by denoting

$$Y^{(x,c,\boldsymbol{p})} \triangleq \frac{X^{(x,c,\boldsymbol{p})}(T)}{B} \leq 1, \quad P^B - \text{ a.s.} \tag{6.27}$$

and

$$\bar{H}_{\boldsymbol{\nu},\boldsymbol{\mu}}(t) \triangleq H_{\boldsymbol{\nu},\boldsymbol{\mu}}(t)S_0(t), \quad t \in [0,T], (\boldsymbol{\nu},\boldsymbol{\mu}) \in \mathcal{G} \times \mathcal{H}. \tag{6.28}$$

From (6.27) and (6.25), with $w \triangleq z\bar{H}_{\boldsymbol{\nu},\boldsymbol{\mu}}(T)$ for $z \geq 0$,

$$(1 - Y^{(x,c,\boldsymbol{p})})^+ \geq \tilde{R}(z\bar{H}_{\boldsymbol{\nu},\boldsymbol{\mu}}(T)) - z\bar{H}_{\boldsymbol{\nu},\boldsymbol{\mu}}(T)Y^{(x,c,\boldsymbol{p})} \quad P^B - \text{ a.s.}, \tag{6.29}$$

for some $x \in (0, \hat{u})$, $(c, \boldsymbol{p}) \in \mathcal{A}(x)$, and $z \geq 0$. Multiplying both sides of (6.29) by $E[\bar{B}]$ and taking P^B expectations gives

$$E[\bar{B}]E^B[1 - Y^{(x,c,\boldsymbol{p})}]^+ \geq E[\bar{B}]E^B[\tilde{R}(z\bar{H}_{\boldsymbol{\nu},\boldsymbol{\mu}}(T))] - zE[\bar{B}]E^B[\bar{H}_{\boldsymbol{\nu},\boldsymbol{\mu}}(T)Y^{(x,c,\boldsymbol{p})}]. \tag{6.30}$$

Since

$$E[H_{\boldsymbol{\nu},\boldsymbol{\mu}}(T)X^{(x,c,\boldsymbol{p})}(T)] \leq x, \quad (\boldsymbol{\nu},\boldsymbol{\mu}) \in \mathcal{G} \times \mathcal{H} \tag{6.31}$$

for all $x > 0$ and $(c, \boldsymbol{p}) \in \mathcal{A}(x)$ from the supermartingale property of $H_{\boldsymbol{\nu},\boldsymbol{\mu}}X^{(x,c,\boldsymbol{p})}$, we also have that

$$E[\bar{B}]E^B[\bar{H}_{\boldsymbol{\nu},\boldsymbol{\mu}}(T)Y^{(x,c,\boldsymbol{p})}] \leq x, \quad (\boldsymbol{\nu},\boldsymbol{\mu}) \in \mathcal{G} \times \mathcal{H} \tag{6.32}$$

for all $x > 0$ and $(c, \boldsymbol{p}) \in \mathcal{A}(x)$. Putting (6.32) into (6.30) gives

$$E[\bar{B}]E^B[1 - Y^{(x,c,\boldsymbol{p})}]^+ \geq E[\bar{B}]E^B[\tilde{R}(z\bar{H}_{\boldsymbol{\nu},\boldsymbol{\mu}}(T))] - xz, \tag{6.33}$$

for all $x \in (0, \hat{u})$, $(c, \boldsymbol{p}) \in \mathcal{A}(x)$, and $z \geq 0$.

Remark 6.2.1. The relation (6.33) is a type of duality relationship that has proved to be very useful in the constrained portfolio optimization studied by Cvitanic and Karatzas [9].

We can consider the maximization of the right-hand side of (6.33) to be the dual problem of our primal optimization problem (6.23). It would be ideal if we can show equality in (6.33) by minimizing left-hand side over all $(c, \boldsymbol{p}) \in \mathcal{A}(x)$ and maximizing the right-hand side over $z \in [0, \infty)$ and $\{\bar{H}_{\boldsymbol{\nu},\boldsymbol{\mu}}(T) | (\boldsymbol{\nu},\boldsymbol{\mu}) \in \mathcal{G} \times \mathcal{H}\}$, for then we can solve the dual problem and work backwards to solve for a minimizing $(\tilde{c}, \tilde{\boldsymbol{p}}) \in \mathcal{A}(x)$. However, as will be shown in Proposition 6.2.6, we cannot promise equality if we maximize the right-hand side of (6.33) over the set $\{\bar{H}_{\boldsymbol{\nu},\boldsymbol{\mu}}(T) | (\boldsymbol{\nu},\boldsymbol{\mu}) \in \mathcal{G} \times \mathcal{H}\}$. Instead, we can look at the very closely related duality relationship which is derived in the same manner as (6.33):

$$E[\bar{B}]E^B[1 - Y^{(x,c,\boldsymbol{p})}]^+ \geq E[\bar{B}]E^B[\tilde{R}(zH)] - xz, \tag{6.34}$$

where $x \in (0, \hat{u})$, $(c, \boldsymbol{p}) \in \mathcal{A}(x)$, $z \geq 0$, and $H \in \mathcal{Z}$ is defined as

$$\mathcal{Z} \triangleq \left\{ H \in L^1(\Omega, \mathcal{F}_T, P^B) \middle| H \geq 0 \ P^B \text{ a.s., } E[\bar{B}]E^B[HY^{(x,c,\boldsymbol{p})}] \leq x \text{ for all } (c, \boldsymbol{p}) \in \mathcal{A}(x) \right\}. \tag{6.35}$$

The following proposition is an essential fact about the set $\mathcal{Z}$.

Proposition 6.2.2. *The set $\mathcal{Z}$ is convex and closed in $L^1(\Omega, \mathcal{F}_T, P^B)$.*

Proof. Fix $x > 0$ and $(c, \boldsymbol{p}) \in \mathcal{A}(x)$. Take sequences $\{H_i\} \in \mathcal{Z}$ and $\{\lambda_i\} \in [0, 1]$, $i = 1, 2, \ldots, n$, such that $\sum_{i=1}^n \lambda_i = 1$. Set

$$H \triangleq \sum_{i=1}^n \lambda_i H_i \geq 0 \quad P^B - \text{ a.s.} \tag{6.36}$$

To see that $H \in \mathcal{Z}$ take

$$E[\bar{B}]E^B[HY^{(x,c,\boldsymbol{p})}] \leq \sum_{i=1}^n \lambda_i x = x. \tag{6.37}$$

Therefore, $\mathcal{Z}$ is convex.
Now take another sequence $\{H_i\} \in \mathcal{Z}$, $i = 1, 2, \ldots$, such that

$$H_i \to H \text{ in } L^1(\Omega, \mathcal{F}_T, P^B). \tag{6.38}$$

From L^1-convergence, there exists a subsequence $\{H_{n(i)}\} \subset \{H_i\}$ such that

$$H_{n(i)} \to H \;\; P^B - \text{ a.s.}, \tag{6.39}$$

and therefore,

$$H \geq 0 \quad P^B - \text{ a.s.} \tag{6.40}$$

From Fatou's Lemma

$$\begin{aligned} E[\bar{B}]E^B[HY^{(x,c,\boldsymbol{p})}] &= E[\bar{B}]E^B[\liminf_{n(i)\to\infty} H_{n(i)} Y^{(x,c,\boldsymbol{p})}] \\ &\leq \liminf_{n(i)\to\infty} E[\bar{B}]E^B[H_{n(i)} Y^{(x,c,\boldsymbol{p})}] \leq x. \end{aligned} \tag{6.41}$$

Therefore, from (6.40) and (6.41), $\mathcal{Z}$ is closed in $L^1(\Omega, \mathcal{F}_T, P^B)$. □

Remark 6.2.3. Here we will show some properties of the set $\mathcal{Z}$. By Theorem 5.3.17, there exists a triple $(\hat{u}, \hat{c}, \hat{\boldsymbol{p}}) \in (0, \infty) \times \mathcal{A}(\hat{u})$ such that

$$X^{(\hat{u},\hat{c},\hat{\boldsymbol{p}})}(T) = B \quad P - \text{ a.s.} \tag{6.42}$$

From (6.27) and (6.42) we have

$$Y^{(\hat{u},\hat{c},\hat{\boldsymbol{p}})} = 1 \quad P^B - \text{ a.s.} \tag{6.43}$$

Therefore, since $P[B > 0] = 1$, we have $E[\bar{B}] > 0$ and from (6.35)

$$E[\bar{B}]E^B[H] = E[\bar{B}H] \le \hat{u}, \quad H \in \mathcal{Z}. \tag{6.44}$$

By using the fact that $X^{(x,0,0)}(T) = S_0(T)x$, which is valid since 0 is in the convex constraint set K and $(0, 0) \in \mathcal{A}(x)$, we see that

$$E[\bar{B}]E^B[H/B] = E[H] \le 1, \quad H \in \mathcal{Z}. \tag{6.45}$$

Moreover, since

$$\begin{aligned} \infty > \hat{u} &\ge E[H_{\boldsymbol{\nu},\boldsymbol{\mu}}(T)X^{(\hat{u},\hat{c},\hat{\boldsymbol{p}})}(T)] \\ &= E[\bar{B}]E^B[\bar{H}_{\boldsymbol{\nu},\boldsymbol{\mu}}(T)Y^{(\hat{u},\hat{c},\hat{\boldsymbol{p}})}], \quad (\boldsymbol{\nu}, \boldsymbol{\mu}) \in \mathcal{G} \times \mathcal{H}, \end{aligned} \tag{6.46}$$

we have that

$$\mathcal{Z}_{\mathcal{G},\mathcal{H}} \triangleq \left\{ \bar{H}_{\boldsymbol{\nu},\boldsymbol{\mu}}(T) \middle| (\boldsymbol{\nu}, \boldsymbol{\mu}) \in \mathcal{G} \times \mathcal{H} \right\} \subset \mathcal{Z}. \tag{6.47}$$

Remark 6.2.4. As will be seen in Proposition 6.2.6, when optimizing over the larger set $\mathcal{Z}$, equality can in fact be shown in (6.34). Therefore, our agenda will be solve the dual problem (the right-hand side of (6.34)) and work backwards to find some $(\tilde{c}, \tilde{\boldsymbol{p}}) \in \mathcal{A}(x)$ that minimizes the left-hand side of (6.34), solving our optimization problem.

To simplify notation, we define

$$\tilde{J}(H; z) \triangleq E[\bar{B}]E^B[(zH) \wedge 1], \quad H \in L^1(\Omega, \mathcal{F}_T, P^B), z \ge 0. \tag{6.48}$$

Putting (6.48) into (6.34) we rewrite the duality relation as

$$E[\bar{B}]E^B[1 - Y^{(x,c,\boldsymbol{p})}]^+ \ge \tilde{J}(H; z) - xz, \quad H \in \mathcal{Z}, z \ge 0, \tag{6.49}$$

for all $x \in (0, \hat{u})$ and $(c, \boldsymbol{p}) \in \mathcal{A}(x)$.

Proposition 6.2.5. *$-\tilde{J}(\cdot, z) : L^1(\Omega, \mathcal{F}_T, P^B) \mapsto \mathbb{R}$ is a lower semi-continuous and proper convex functional for each $z \ge 0$.*

Proof. To see convexity, take $H_1, H_2 \in L^1(\Omega, \mathcal{F}_T, P^B)$, $\lambda \in [0, 1]$, and fix $z \ge 0$. Since the minimum of affine functions is concave, we have

$$\min\{z\lambda H_1 + z(1 - \lambda)H_2, 1\} \ge \lambda \min\{zH_1, 1\} + (1 - \lambda) \min\{zH_2, 1\}. \tag{6.50}$$

Taking P^B expectations on both sides of (6.50) and multiplying by $E[\bar{B}]$ gives

$$E[\bar{B}]E^B[(z\lambda H_1 + z(1-\lambda)H_2)\wedge 1] \geq \lambda E[\bar{B}]E^B[zH_1\wedge 1] + (1-\lambda)E[\bar{B}]E^B[zH_2\wedge 1], \quad (6.51)$$

which shows that $\tilde{J}(\cdot, z)$ is concave and thus $-\tilde{J}(\cdot, z)$ is convex. It is immediate that $-\tilde{J}(\cdot, z)$ is a lower semi-continuous and proper convex functional. □

The following proposition gives necessary and sufficient conditions for which equality holds in (6.49) and therefore are conditions that promise the minimization of the cost function (6.17). As a result, Proposition 6.2.6 provides a template which we will use to show the existence of an optimal trading strategy $(\tilde{c}, \tilde{\boldsymbol{p}}) \in \mathcal{A}(x)$ from a given $x < \hat{u}$.

Proposition 6.2.6. *For some $(\tilde{c}, \tilde{\boldsymbol{p}}) \in \mathcal{A}(x)$, $\tilde{z} > 0$, and $\tilde{H} \in \mathcal{Z}$,*

$$E[\bar{B}]E^B[1 - Y^{(x,\tilde{c},\tilde{\boldsymbol{p}})}]^+ = \tilde{J}(\tilde{H}; \tilde{z}) - x\tilde{z}, \quad (6.52)$$

that is the infimum in the primal problem (6.17) is attained, if and only if

$$E[\bar{B}]E^B[\tilde{H}Y^{(x,\tilde{c},\tilde{\boldsymbol{p}})}] = x \quad (6.53)$$

and

$$Y^{(x,\tilde{c},\tilde{\boldsymbol{p}})} = I(\tilde{z}\tilde{H}; \tilde{D}) = I_{\{\tilde{z}\tilde{H}<1\}} + \tilde{D}I_{\{\tilde{z}\tilde{H}=1\}} \quad P^B - \text{ a.s.}, \quad (6.54)$$

where $I(\cdot;\cdot)$ is defined in (6.26) and $\tilde{D}$ is an $\mathcal{F}_T$-measurable random variable such that

$$0 \leq \tilde{D} \leq 1 \quad P^B - \text{ a.s.} \quad (6.55)$$

Proof.($\Longrightarrow$) Assume (6.52) holds. Comparing (6.52) to (6.49), we have equality at some $(\tilde{c}, \tilde{\boldsymbol{p}}) \in \mathcal{A}(x)$, $\tilde{z} > 0$, and $\tilde{H} \in \mathcal{Z}$. We also see from (6.49) and (6.35) that if right-hand side of (6.52) is maximized then we must have

$$x = E[\bar{B}]E^B[\tilde{H}Y^{(x,\tilde{c},\tilde{\boldsymbol{p}})}]. \quad (6.56)$$

We also have from (6.25) and (6.26) that the left-hand side of (6.52), $E[\bar{B}]E^B[1-Y^{(x,\tilde{c},\tilde{\boldsymbol{p}})}]^+$, is minimized by

$$\begin{aligned} Y^{(x,\tilde{c},\tilde{\boldsymbol{p}})} = I(\tilde{z}\tilde{H}; \tilde{D}) &= \begin{cases} 0, & \tilde{z}\tilde{H} > 1 \\ 1, & 0 \leq \tilde{z}\tilde{H} < 1 \\ \tilde{D}, & \tilde{z}\tilde{H} = 1 \end{cases} \\ &= I_{\{\tilde{z}\tilde{H}<1\}} + \tilde{D}I_{\{\tilde{z}\tilde{H}=1\}}, \end{aligned} \quad (6.57)$$

for some $\mathcal{F}_T$ random variable $\tilde{D}$, such that

$$0 \le \tilde{D} \le 1 \quad P^B - \text{ a.s.} \tag{6.58}$$

($\Longleftarrow$) Assume (6.53)-(6.55) hold. From (6.25) and (6.54)

$$[1 - Y^{(x,\tilde{c},\tilde{\boldsymbol{p}})}]^+ + Y^{(x,\tilde{c},\tilde{\boldsymbol{p}})}\tilde{z}\tilde{H} = \tilde{z}\tilde{H} \wedge 1 \quad P^B - \text{ a.s.}, \tag{6.59}$$

for some $(\tilde{c}, \tilde{\boldsymbol{p}}) \in \mathcal{A}(x)$, $\tilde{z} > 0$, and $\tilde{H} \in \mathcal{Z}$. Taking P^B expectations on both sides and multiplying by $E[\tilde{B}]$ gives

$$E[\bar{B}]E^B[1 - Y^{(x,\tilde{c},\tilde{\boldsymbol{p}})}]^+ = \tilde{J}(\tilde{H}; \tilde{z}) - x\tilde{z}. \tag{6.60}$$

□

Remark 6.2.7. We assert that $\tilde{z} > 0$ because if $\tilde{z} = 0$, then from (6.27) and (6.52), we would have

$$X^{(x,\tilde{c},\tilde{\boldsymbol{p}})}(T) = B \quad P^B - \text{ a.s.} \tag{6.61}$$

But since we have assumed $x < \hat{u}$, from Remark 6.1.3, (6.61) cannot occur. Thus, we must have $\tilde{z} > 0$.

Remark 6.2.8. By Proposition 6.2.6 we see that the problem

$$\sup_{\substack{H \in \mathcal{Z} \\ z \ge 0}} \left\{ \tilde{J}(H; z) - xz \right\} \tag{6.62}$$

is in fact the dual problem to the primal problem

$$\inf_{(c,\boldsymbol{p}) \in \mathcal{A}(x)} E[\bar{B}]E^B[1 - Y^{(x,c,\boldsymbol{p})}]^+. \tag{6.63}$$

Furthermore, if (6.53)-(6.55) are satisfied, then $(\tilde{z}, \tilde{H}) \in (0, \infty) \times \mathcal{Z}$ is the solution to the dual problem and $(\tilde{c}, \tilde{\boldsymbol{p}}) \in \mathcal{A}(x)$ is the solution to the primal problem. Moreover, we can simplify the dual problem by realizing that $\tilde{H} \in \mathcal{Z}$ is optimal for the auxiliary dual problem

$$\tilde{V}(z) \triangleq \sup_{H \in \mathcal{Z}} \tilde{J}(H; z) \tag{6.64}$$

when $z = \tilde{z}$. If we let

$$X^{(x,\tilde{c},\tilde{\boldsymbol{p}})}(T) = BY^{(x,\tilde{c},\tilde{\boldsymbol{p}})} \text{ a.s.}, \tag{6.65}$$

then we can re-write (6.53) and (6.54) as

$$x = E[\tilde{H}\bar{X}^{(x,\tilde{c},\tilde{\boldsymbol{p}})}(T)] \tag{6.66}$$

and

$$X^{(x,\tilde{c},\tilde{\boldsymbol{p}})}(T) = B(I_{\{\tilde{z}\tilde{H}<1\}} + \tilde{D}I_{\{\tilde{z}\tilde{H}=1\}}) \text{ a.s.} \tag{6.67}$$

Remark 6.2.9. We will approach the optimization problem as follows: we will first find a solution to the dual problem (6.62) which amounts to finding a number $\tilde{z} > 0$ and a solution $\tilde{H} \in \mathcal{Z}$ to (6.64) with $z = \tilde{z}$. We will then find an $\mathcal{F}_T$-measurable random variable $\tilde{D}$ such that $0 \leq \tilde{D} \leq 1$ a.s., and a pair $(\tilde{c}, \tilde{\boldsymbol{p}}) \in \mathcal{A}(x)$ such that (6.53) and (6.54) hold (or equivalently (6.66) and (6.67) hold) so that (6.52) holds and the primal problem is solved.

Proposition 6.2.10. *For any given $z > 0$, there exists an optimal solution $\tilde{H} \in \mathcal{Z}$ for the auxiliary problem* (6.64)*.*

Proof. Fix $z > 0$. Let $\{H_i\} \in \mathcal{Z}$, $i = 1, 2, \ldots$, be a sequence that attains the supremum in (6.64), i.e.

$$\tilde{V}(z) = \lim_{i \to \infty} \tilde{J}(H_i; z). \tag{6.68}$$

We know by (6.44) that,

$$E^B[H] \leq \frac{\hat{u}}{E[\bar{B}]} < \infty, \quad H \in \mathcal{Z}, \tag{6.69}$$

which means that $\mathcal{Z}$ is a bounded set in $L^1(\Omega, \mathcal{F}_T, P^B)$. Therefore, we can use the Komlós theorem (see Theorem C.14.2 in the Appendix), which states that there exists a random variable $\tilde{H} \in L^1(\Omega, \mathcal{F}_T, P^B)$ and a subsequence $\{H_{n(i)}\} \subset \{H_i\}$ such that

$$G_m \triangleq \frac{1}{m} \sum_{i=1}^{m} H_{n(i)}, \quad m = 1, 2, \ldots, \tag{6.70}$$

has the property

$$\lim_{m \to \infty} G_m = \tilde{H} \quad P^B - a.s. \tag{6.71}$$

Since we have $H_{n(i)} \geq 0$ P^B a.s. for each $n(i) \in 1, 2, \ldots$, $G_m \geq 0$ P^B a.s. for each $m = 1, 2, \ldots$. Thus,

$$\lim_{m \to \infty} G_m = \tilde{H} \geq 0 \quad P^B - \text{ a.s.} \tag{6.72}$$

Now fix some $(c, \boldsymbol{p}) \in \mathcal{A}(x)$. From the definition of $\mathcal{Z}$ in (6.35),

$$\begin{aligned} x &\geq E[\bar{B}]E^B[H_{n(i)}Y^{(x,c,\boldsymbol{p})}], \quad n(i) \in 1, 2, \ldots, \\ &\geq \liminf_{m \to \infty} \frac{1}{m} \sum_{i=1}^{m} E[\bar{B}]E^B[H_{n(i)}Y^{(x,c,\boldsymbol{p})}]. \end{aligned} \tag{6.73}$$

By Fatou's lemma,

$$E[\bar{B}]E^B[\tilde{H}Y^{(x,c,\boldsymbol{p})}] \leq \liminf_{m \to \infty} \frac{1}{m} \sum_{i=1}^{m} E[\bar{B}]E^B[H_{n(i)}Y^{(x,c,\boldsymbol{p})}] \leq x. \tag{6.74}$$

Thus, by (6.72) and (6.74),

$$\tilde{H} \in \mathcal{Z}. \tag{6.75}$$

We can now write

$$\begin{aligned}\tilde{J}(\tilde{H}; z) &= E[\bar{B}]E^B[z\tilde{H} \wedge 1] \\ &= E[\bar{B}]E^B[(z \lim_{m\to\infty} \frac{1}{m}\sum_{i=1}^{m} H_{n(i)}) \wedge 1],\end{aligned} \tag{6.76}$$

and by the dominated convergence theorem

$$\tilde{J}(\tilde{H}; z) = \lim_{m\to\infty} \tilde{J}\left(\frac{1}{m}\sum_{i=1}^{m} H_{n(i)}; z\right). \tag{6.77}$$

Since $\tilde{J}(\cdot; z)$ is concave,

$$\tilde{J}(\tilde{H}; z) \geq \lim_{m\to\infty} \frac{1}{m}\sum_{i=1}^{m} \tilde{J}(H_{n(i)}; z). \tag{6.78}$$

Using Cesàro summability (see Proposition C.14.1 in the Appendix),

$$\begin{aligned}\tilde{J}(\tilde{H}; z) &\geq \lim_{m\to\infty} \frac{1}{m}\sum_{i=1}^{m} \tilde{J}(H_{n(i)}; z) \\ &= \lim_{m\to\infty} \tilde{J}(H_m; z) \\ &= \tilde{V}(z),\end{aligned} \tag{6.79}$$

but from (6.64), we have

$$\tilde{V}(z) \geq \tilde{J}(H; z), \quad H \in \mathcal{Z}. \tag{6.80}$$

Thus,

$$\tilde{V}(z) = \tilde{J}(\tilde{H}; z), \tag{6.81}$$

meaning $\tilde{H} \in \mathcal{Z}$ is optimal. □

Proposition 6.2.11. *The function $\tilde{V}(z)$ is continuous on $[0, \infty)$.*

Proof. Let $H \in \mathcal{Z}$ and assume $z_1, z_2 > 0$. We have

$$\begin{aligned}\tilde{J}(H; z_1) &= E[\bar{B}]E^B[z_1 H \wedge 1] \\ &= E[\bar{B}]E^B[z_1 H \wedge 1 + z_2 H \wedge 1 - z_2 H \wedge 1] \\ &= \tilde{J}(H; z_2) + E[\bar{B}]E^B[H(z_1 - z_2)I_{\{z_1 H<1, z_2 H<1\}} \\ &+ (Hz_1 - 1)I_{\{z_1 H<1, z_2 H\geq 1\}} + (1 - Hz_2)I_{\{z_1 H\geq 1, z_2 H<1\}}] \\ &\leq \tilde{V}(z_2) + 2E[\bar{B}](1 - \frac{z_2}{z_1})^+.\end{aligned} \tag{6.82}$$

Rearranging (6.82) with the fact that $\tilde{V}(z_1) = \tilde{J}(\tilde{H}; z_1)$ for some $\tilde{H} \in \mathcal{Z}$ (see Proposition 6.2.10), we have

$$\tilde{V}(z_1) - \tilde{V}(z_2) \leq 2E[\bar{B}](1 - \frac{z_2}{z_1})^+. \tag{6.83}$$

Taking $z_2 \to z_1$ in (6.83) shows continuity of $\tilde{V}$ when $z_1, z_2 > 0$. To show continuity at $z_2 = 0$, use the duality relation in (6.25) with (6.44) to obtain

$$\begin{aligned}\tilde{J}(\tilde{H}; z_1) = E[\bar{B}]E^B[z_1\tilde{H} \wedge 1] &\leq E[\bar{B}(1-y)^+] + yz_1E[\bar{B}]E^B[\tilde{H}] \\ &\leq (1-y)^+E[\bar{B}] + yz_1\hat{u},\end{aligned} \tag{6.84}$$

for all $0 \leq y \leq 1$ and $z_1 > 0$. Taking $y = 1$ in (6.84) gives

$$\tilde{V}(z_1) \leq z_1\hat{u}, \tag{6.85}$$

and taking $z_1 \to 0$ shows continuity of $\tilde{V}$. □

Proposition 6.2.12. *For every $0 < x < \hat{u}$, there exists $\tilde{z} > 0$ that attains the supremum*

$$\tilde{V}(\tilde{z}) - x\tilde{z} = \sup_{z \geq 0}[\tilde{V}(z) - xz] \geq \tilde{V}(z) - xz, \quad z \geq 0. \tag{6.86}$$

Proof. Denote

$$\alpha(z) \triangleq \tilde{V}(z) - xz, \quad z \geq 0 \tag{6.87}$$

We have

$$\alpha(0) = \tilde{V}(0) = 0, \tag{6.88}$$

and

$$\limsup_{z\to\infty} \alpha(z) = E[\bar{B}] + \limsup_{z\to\infty}\{-xz\} < 0. \tag{6.89}$$

From (6.88) and (6.89), the supremum of $\alpha(z)$ cannot be attained at $z = \infty$. Now, since $\alpha(\cdot)$ is continuous from Proposition 6.2.11, we either have that $\alpha(z)$ attains its supremum at some $\tilde{z} > 0$, or $\alpha(z)$ attains its supremum at $\tilde{z} = 0$ and $\alpha(z) \leq 0$ for all $z > 0$. Suppose that $\alpha(z)$ attains its supremum at $\tilde{z} = 0$. Then,

$$\tilde{V}(z) - xz \leq 0, \quad z > 0, \tag{6.90}$$

and

$$x \geq \frac{\tilde{V}(z)}{z} \geq \frac{E[\bar{B}]E^B[zH \wedge 1]}{z}, \quad z > 0, H \in \mathcal{Z}. \tag{6.91}$$

Since $(zH \wedge 1)/z = H \wedge (1/z)$, from (6.91),

$$x \geq E[\bar{B}]E^B[H \wedge \frac{1}{z}], \quad z > 0, H \in \mathcal{Z}. \tag{6.92}$$

Thus,

$$x \geq \lim_{z \to 0} E[\bar{B}]E^B[H \wedge \frac{1}{z}], \quad H \in \mathcal{Z}. \tag{6.93}$$

But from (6.69),

$$E^B[H \wedge \frac{1}{z}] \leq E^B[H] \leq \frac{\hat{u}}{E[\bar{B}]} < \infty, \quad z > 0. \tag{6.94}$$

By (6.93), (6.94), and the dominated convergence theorem,

$$x \geq E[H\bar{B}], \quad H \in \mathcal{Z}. \tag{6.95}$$

From (6.47), $\mathcal{Z}_{\mathcal{G},\mathcal{H}} \subset \mathcal{Z}$, thus, from (6.95),

$$x \geq \sup_{(\boldsymbol{\nu},\boldsymbol{\mu}) \in \mathcal{G} \times \mathcal{H}} E\left[H_{\boldsymbol{\nu},\boldsymbol{\mu}}(T)B\right] = \hat{u}, \tag{6.96}$$

which is a contradiction since $\hat{u} > x$. Therefore, $\alpha(z)$ attains its supremum at some $\tilde{z} > 0$.
□

Remark 6.2.13. From Propositions 6.2.10 and 6.2.12 we have shown the following:
1) For any given $z > 0$, there exists an $\tilde{H} \in \mathcal{Z}$ such that

$$\tilde{V}(z) = \tilde{J}(\tilde{H}; z) \geq \tilde{J}(H; z), \quad H \in \mathcal{Z}. \tag{6.97}$$

2) For every x such that $0 < x < \hat{u}$, there exists a $\tilde{z} > 0$ such that

$$\tilde{V}(\tilde{z}) - x\tilde{z} \geq \tilde{V}(z) - xz, \quad z \geq 0. \tag{6.98}$$

Combining (6.97) and (6.98) we have, for a given $0 < x < \hat{u}$,

$$\tilde{V}(\tilde{z}) - x\tilde{z} = \tilde{J}(\tilde{H}; \tilde{z}) - x\tilde{z} \geq \tilde{J}(H; z) - xz, \quad \text{for all } z \geq 0, H \in \mathcal{Z}. \tag{6.99}$$

From (6.99) and (6.49) we have

$$\begin{aligned} \inf_{(c,\boldsymbol{p}) \in \mathcal{A}(x)} E[\bar{B}]E^B[1 - Y^{(x,c,\boldsymbol{p})}]^+ &\geq \sup_{\substack{H \in \mathcal{Z} \\ z \geq 0}} \left\{ \tilde{J}(H; z) - xz \right\} \\ &= \tilde{J}(\tilde{H}; \tilde{z}) - x\tilde{z} \\ &= \tilde{V}(\tilde{z}) - x\tilde{z}. \end{aligned} \tag{6.100}$$

In summary, we know there exists some $\tilde{z} > 0$ and $\tilde{H} \in \mathcal{Z}$ which attains the maximum value of the dual problem $\tilde{J}(H; z) - xz$.

Now that we have shown the existence of the optimal variables $\tilde{z} > 0$ and $\tilde{H} \in \mathcal{Z}$ which solve the dual problem (6.62), we will show the existence of some strategy $(\tilde{c}, \tilde{\boldsymbol{p}}) \in \mathcal{A}(x)$ such that (6.53)-(6.55) hold. Introduce the space

$$\mathcal{L} \triangleq L^1(\Omega, \mathcal{F}_T, P^B) \times \mathbb{R} \tag{6.101}$$

with norm

$$\|(Z, z)\| \triangleq E[\bar{B}]E^B[|Z|] + |z|, \tag{6.102}$$

for all $Z \in L^1(\Omega, \mathcal{F}_T, P^B)$ and $z \in \mathbb{R}$. Define the subset

$$\mathcal{Q} \triangleq \left\{(zH, z) \in \mathcal{L} \middle| H \in \mathcal{Z}, z \geq 0\right\}. \tag{6.103}$$

Proposition 6.2.14. *The set $\mathcal{Q}$ is convex and closed in $\mathcal{L}$.*

Proof. Take sequences $\{q_i\} \in \mathcal{Q}$ and $\{\lambda_i\} \geq 0$ for $i = 1, 2, \ldots, n$, such that $\sum_{i=1}^n \lambda_i = 1$. Then

$$\sum_{i=1}^n \lambda_i q_i = \sum_{i=1}^n (\lambda_i z_i H_i, \lambda_i z_i), \tag{6.104}$$

for sequences $\{H_i\} \in \mathcal{Z}$ and $\{z_i\} \geq 0$, $i = 1, 2, \ldots, n$. Since the set $[0, \infty)$ is convex and the set $\mathcal{Z}$ is convex from Proposition 6.2.2, we have

$$\sum_{i=1}^n \lambda_i q_i \in \mathcal{Q}. \tag{6.105}$$

Therefore, $\mathcal{Q}$ is convex. To see that $\mathcal{Q}$ is closed in $\mathcal{L}$, take sequences $\{z_i\} \geq 0$ and $\{H_i\} \in \mathcal{Z}$ for $i = 1, 2, \ldots$, such that

$$\lim_{n\to\infty} (z_i H_i, z_i) = (Z, z), \tag{6.106}$$

for some $(Z, z) \in \mathcal{L}$. From (6.106), we obviously have

$$\lim_{i\to\infty} z_i = z \geq 0. \tag{6.107}$$

Looking at

$$\begin{aligned} E^B[|z_i H_i - z H_i|] &= E^B[|(z_i - z)H_i|] \\ &\leq |z_i - z| E^B[H_i], \quad i = 1, 2, \ldots, \end{aligned} \tag{6.108}$$

and (6.44), which states

$$E^B[H_i] \leq \frac{\hat{u}}{E[\bar{B}]}, \quad i = 1, 2, \dots, \tag{6.109}$$

we have that

$$E^B[|z_i H_i - z H_i|] \leq |z_i - z| \frac{\hat{u}}{E[\bar{B}]} \quad i = 1, 2, \dots. \tag{6.110}$$

Taking $i \to \infty$ in (6.110) with (6.107) gives

$$z H_i \to Z \text{ in } L^1(\Omega, \mathcal{F}_T, P^B). \tag{6.111}$$

If $z = 0$ in (6.111), we can choose any limit $H_i \to H \in \mathcal{Z}$ since $Z = 0$. If $z > 0$, we get from (6.111)

$$H_i \to Z/z \text{ in } L^1(\Omega, \mathcal{F}_T, P^B). \tag{6.112}$$

But since $\mathcal{Z}$ is closed in $L^1(\Omega, \mathcal{F}_T, P^B)$ by Proposition 6.2.2, $Z/z \in \mathcal{Z}$ and $\mathcal{Q}$ is shown to be closed in $\mathcal{L}$. □

To proceed, we define the functional $\tilde{U} : \mathcal{L} \to \mathbb{R}$

$$\begin{aligned} \tilde{U}(Z, z) &\triangleq -E[\bar{B}]E^B[Z \wedge 1] + xz \\ &= -\tilde{J}(Z; 1) + xz, \quad (Z, z) \in \mathcal{L}. \end{aligned} \tag{6.113}$$

Proposition 6.2.15. *$\tilde{U}(\cdot, \cdot)$ is a lower semi-continuous and proper convex function.*

Proof. Take pairs $(Z_1, z_1), (Z_2, z_2) \in \mathcal{L}$ and $\lambda \in [0, 1]$. Now,

$$\begin{aligned} \tilde{U}(\lambda(Z_1, z_1) + (1-\lambda)(Z_2, z_2)) &= \tilde{U}(\lambda Z_1 + (1-\lambda)Z_2, \lambda z_1 + (1-\lambda)z_2) \\ &= -E[\bar{B}]E^B[(\lambda Z_1 + (1-\lambda)Z_2) \wedge 1] + \lambda x z_1 + (1-\lambda)x z_2, \end{aligned} \tag{6.114}$$

but by the concavity of $\tilde{J}(\cdot; z)$ from Proposition 6.2.5,

$$\begin{aligned} \tilde{U}(\lambda(Z_1, z_1) + (1-\lambda)(Z_2, z_2)) &\leq -\lambda(E[\bar{B}]E^B[Z_1 \wedge 1] + xz_1) - (1-\lambda)(E[\bar{B}]E^B[Z_2 \wedge 1] + xz_2) \\ &= \lambda \tilde{U}(Z_1, z_1) + (1-\lambda)\tilde{U}(Z_2, z_2). \end{aligned} \tag{6.115}$$

Therefore, $\tilde{U}(\cdot, \cdot)$ is a convex function. It is easy to verify that $\tilde{U}$ is lower semi-continuous and proper. □

Since

$$\begin{aligned}\inf_{(Z,z)\in\mathcal{Q}} \tilde{U}(Z,z) &= \inf_{(Z,z)\in\mathcal{Q}} -\tilde{J}(Z;1) + xz \\ &= \tilde{J}(\tilde{H},\tilde{z}) - x\tilde{z} \\ &= \sup_{\substack{H\in\mathcal{Z}\\ z\geq 0}} \left\{ \tilde{J}(H;z) - xz \right\},\end{aligned} \tag{6.116}$$

we have that $\inf_{(Z,z)\in\mathcal{Q}} \tilde{U}(Z,z)$ has the same solution as the dual problem in (6.100). Therefore,

$$(\tilde{z}\tilde{H},\tilde{z}) \in \mathcal{Q} \text{ is optimal for the problem } \inf_{(Z,z)\in\mathcal{Q}} \tilde{U}(Z,z). \tag{6.117}$$

To characterize the pair $(\tilde{z}\tilde{H},\tilde{z})$ in (6.117) we first rewrite the dual problem in terms of a function $f : \mathcal{L} \mapsto \mathbb{R} \cup \{+\infty\}$ defined as

$$f(Z,z) \triangleq \begin{cases} \tilde{U}(Z,z), & (Z,z) \in \mathcal{Q} \\ +\infty, & \text{otherwise,} \end{cases} \tag{6.118}$$

or, more compactly as

$$f(Z,z) = \tilde{U}(Z,z) + I_{\mathcal{Q}}(Z,z), \tag{6.119}$$

where the set indicator function $I_{\mathcal{Q}}(Z,z) = 0$ when $(Z,z) \in \mathcal{Q}$ and $I_{\mathcal{Q}}(Z,z) = +\infty$, otherwise. The dual problem can be written as

$$\begin{aligned}\inf_{(Z,z)\in\mathcal{L}} f(Z,z) &= \inf_{(Z,z)\in\mathcal{Q}} \tilde{U}(Z,z) + I_{\mathcal{Q}}(Z,z) \\ &= \tilde{U}(\tilde{z}\tilde{H},\tilde{z}).\end{aligned} \tag{6.120}$$

Let

$$\mathcal{L}^* \triangleq L^\infty(\Omega, \mathcal{F}_T, P^B) \times \mathbb{R} \tag{6.121}$$

be the dual space to $\mathcal{L}$. The subdifferential of the function f at a point $(\bar{Z},\bar{z}) \in \mathcal{L}$ is defined as

$$\partial f(\bar{Z},\bar{z}) \triangleq \left\{ (Y,y) \in \mathcal{L}^* \middle| f(\bar{Z},\bar{z}) - f(Z,z) \leq \langle (Y,y), (\bar{Z}-Z, \bar{z}-z)\rangle, \text{ for all } (Z,z) \in \mathcal{L} \right\}, \tag{6.122}$$

where $\langle\cdot,\cdot\rangle$ is a bilinear operator on the duality space $(\mathcal{L}, \mathcal{L}^*, \langle\cdot,\cdot\rangle)$ (see Definition D.0.4), and using (6.119), can be written as

$$\begin{aligned}\partial f(\bar{Z},\bar{z}) &= \partial\tilde{U}(\bar{Z},\bar{z}) + \partial I_{\mathcal{Q}}(\bar{Z},\bar{z}) \\ &= \partial\tilde{U}(\bar{Z},\bar{z}) + N_c(\bar{Z},\bar{z}),\end{aligned} \tag{6.123}$$

where $N_c(\bar{Z}, \bar{z})$ is the normal cone of $\mathcal{Q}$ at the point $(\bar{Z}, \bar{z})$ which by Proposition 4.1.4 in Aubin and Ekeland [1] is given by

$$\begin{aligned} N_c(\bar{Z}, \bar{z}) &\triangleq \left\{ (Y, y) \in \mathcal{L}^* \middle| \langle (Y, y), (\bar{Z}, \bar{z}) \rangle = \max_{(Z,z)\in\mathcal{Q}} \langle (Y, y), (Z, z) \rangle \right\}, \\ &= \left\{ (Y, y) \in \mathcal{L}^* \middle| E[\bar{B}]E^B[\bar{Z}Y] + \bar{z}y = \max_{(Z,z)\in\mathcal{Q}} (E[\bar{B}]E^B[ZY] + zy) \right\}, \end{aligned} \tag{6.124}$$

and $\partial\tilde{U}(\bar{Z}, \bar{z})$ is the subdifferential of $\tilde{U}$ at the point $(\bar{Z}, \bar{z})$, which by Proposition 4.3.3 in Aubin and Ekeland [1] is given by

$$\begin{aligned} \partial\tilde{U}(\bar{Z}, \bar{z}) \triangleq \Big\{ (Y, y) \in \mathcal{L}^* \Big| \tilde{U}(\bar{Z}, \bar{z}) - \tilde{U}(Z, z) &\leq E[\bar{B}]E^B[Y(\bar{Z} - Z)] \\ &+ y(\bar{z} - z), \text{ for all } (Z, z) \in \mathcal{L} \Big\}. \end{aligned} \tag{6.125}$$

Remark 6.2.16. The bilinear form $\langle \cdot, \cdot \rangle : \mathcal{L}^* \times \mathcal{L} \mapsto \mathbb{R}$ used in (6.124) and (6.125) is given by

$$\langle (U, u), (V, v) \rangle \triangleq E[\bar{B}]E^B[VU] + vu, \quad (U, u) \in \mathcal{L}^*, (V, v) \in \mathcal{L}. \tag{6.126}$$

Proposition 6.2.17. *The pair* $(\tilde{z}\tilde{H}, \tilde{z}) \in \mathcal{Q}$ *is a solution to*

$$0 \in \partial\tilde{U}(Z, z) + N_c(Z, z). \tag{6.127}$$

In other words, there must exist a pair $(\tilde{Y}, \tilde{y}) \in \mathcal{L}^*$ *that belongs to the normal cone* $N_c(\tilde{z}\tilde{H}, \tilde{z})$ *such that* $(-\tilde{Y}, -\tilde{y})$ *belongs to the subdifferential* $\partial\tilde{U}(\tilde{z}\tilde{H}, \tilde{z})$ *at the point* $(\tilde{z}\tilde{H}, \tilde{z}) \in \mathcal{Q}$*.*

Proof. From (6.117), $(\tilde{z}\tilde{H}, \tilde{z}) \in \mathcal{Q}$ is optimal for the dual problem $\inf_{(Z,z)\in\mathcal{Q}} \tilde{U}(Z, z)$, and by (6.120), is optimal for the problem $\inf_{(Z,z)\in\mathcal{L}} f(Z, z)$. We have from Proposition D.0.13 that a point $(Z, z) \in \mathcal{L}$ is optimal if and only if

$$\begin{aligned} &0 \in \partial f(Z, z) \\ &0 \in \partial\tilde{U}(Z, z) + N_c(Z, z). \end{aligned} \tag{6.128}$$

Thus,

$$0 \in \partial\tilde{U}(\tilde{z}\tilde{H}, \tilde{z}) + N_c(\tilde{z}\tilde{H}, \tilde{z}). \tag{6.129}$$

□

Proposition 6.2.18. *Fix some* $(\tilde{Y}, \tilde{y}) \in \mathcal{L}^*$ *such that* $(\tilde{Y}, \tilde{y}) \in N_c(\tilde{z}\tilde{H}, \tilde{z})$ *and* $(-\tilde{Y}, -\tilde{y}) \in \partial\tilde{U}(\tilde{z}\tilde{H}, \tilde{z})$ *where* $(\tilde{z}\tilde{H}, \tilde{z}) \in \mathcal{Q}$ *is optimal for the dual problem in* (6.116). *Then*

$$E[\bar{B}]E^B[\tilde{H}\tilde{Y}] = x. \tag{6.130}$$

and

$$x = -\tilde{y}. \tag{6.131}$$

Proof. From (6.125) and (6.113),

$$\begin{aligned} E[\bar{B}]E^B[(\tilde{z}\tilde{H}) \wedge 1] - E[\bar{B}]E^B[Z \wedge 1] + x(z - \tilde{z}) \\ \geq E[\bar{B}]E^B[\tilde{Y}(\tilde{z}\tilde{H} - Z)] + \tilde{y}(\tilde{z} - z), \quad (Z, z) \in \mathcal{L}. \end{aligned} \tag{6.132}$$

Fixing $Z \in L^1(\Omega, \mathcal{F}_T, P^B)$ and taking $z \to +\infty$ in (6.132) yields

$$x \geq -\tilde{y}, \tag{6.133}$$

taking $z \to -\infty$ yields

$$x \leq -\tilde{y}, \tag{6.134}$$

and therefore,

$$x = -\tilde{y}. \tag{6.135}$$

From (6.124) and (6.103),

$$E[\bar{B}]E^B[\tilde{z}\tilde{H}\tilde{Y}] + \tilde{z}\tilde{y} \geq E[\bar{B}]E^B[zH\tilde{Y}] + z\tilde{y}, \quad z \geq 0, H \in \mathcal{Z}. \tag{6.136}$$

Now setting $z = \tilde{z} + \varepsilon$ for some $\varepsilon > 0$ and $H = \tilde{H}$ in (6.136) and using (6.135) gives

$$x \geq E[\bar{B}]E^B[\tilde{H}\tilde{Y}]. \tag{6.137}$$

Similarly, letting $z = \tilde{z} - \varepsilon$ for some $\varepsilon > 0$ and $H = \tilde{H}$ in (6.136) and using (6.135) gives

$$x \leq E[\bar{B}]E^B[\tilde{H}\tilde{Y}]. \tag{6.138}$$

Therefore, combining (6.137) and (6.138) yields

$$E[\bar{B}]E^B[\tilde{H}\tilde{Y}] = x. \tag{6.139}$$

□

The following proposition shows the existence of a pair $(\tilde{c}, \tilde{\boldsymbol{p}}) \in \mathcal{A}(x)$ that hedges the $\mathcal{F}_T$ random variable $B\tilde{Y}$ from initial wealth x.

Proposition 6.2.19. *There exists a pair* $(\tilde{c}, \tilde{\boldsymbol{p}}) \in \mathcal{A}(x)$ *such that*

$$X^{(x,\tilde{c},\tilde{\boldsymbol{p}})}(T) = B\tilde{Y} \text{ a.s.} \tag{6.140}$$

and

$$x = E[\tilde{H}\bar{X}^{(x,\tilde{c},\tilde{\boldsymbol{p}})}(T)]. \tag{6.141}$$

Proof. From Proposition 6.2.18 and (6.124) , we have

$$x = E[\bar{B}]E^B[\tilde{H}\tilde{Y}] \geq E[\bar{B}]E^B[H\tilde{Y}], \quad H \in \mathcal{Z}, \tag{6.142}$$

and therefore,

$$x = \sup_{H \in \mathcal{Z}} E[\bar{B}]E^B[H\tilde{Y}] = \sup_{H \in \mathcal{Z}} E[\bar{B}H\tilde{Y}]. \tag{6.143}$$

Since $E[\bar{B}H\tilde{Y}] \leq x < \hat{u} < \infty$ for all $H \in \mathcal{Z}$, we know the random variable $B\tilde{Y}$ can be hedged by Theorem 5.3.17. Denote the minimum hedging price of the contingent claim $B\tilde{Y}$ by

$$\hat{u}_{\tilde{Y}} \triangleq \sup_{(\boldsymbol{\nu},\boldsymbol{\mu}) \in \mathcal{G}\times\mathcal{H}} E[H_{\boldsymbol{\nu},\boldsymbol{\mu}}(T)B\tilde{Y}], \tag{6.144}$$

which is guaranteed to be finite by Theorem 5.3.17. Since $H_{\boldsymbol{\nu},\boldsymbol{\mu}}(T) \in \mathcal{Z}$ for all $(\boldsymbol{\nu}, \boldsymbol{\mu}) \in \mathcal{G} \times \mathcal{H}$, we have from (6.143)

$$x = \sup_{H \in \mathcal{Z}} E[\bar{B}H\tilde{Y}] \geq \sup_{(\boldsymbol{\nu},\boldsymbol{\mu}) \in \mathcal{G}\times\mathcal{H}} E[H_{\boldsymbol{\nu},\boldsymbol{\mu}}(T)B\tilde{Y}] = \hat{u}_{\tilde{Y}}. \tag{6.145}$$

From Proposition 6.1.1 and Remark 6.1.2 there exists some $(\tilde{c}, \tilde{\boldsymbol{p}}) \in \mathcal{A}(x)$ such that

$$X^{(x,\tilde{c},\tilde{\boldsymbol{p}})}(T) \geq B\tilde{Y} \text{ a.s.} \tag{6.146}$$

But since one can always consume down to the value $B\tilde{Y}$ at the close of trade, we have

$$X^{(x,\tilde{c},\tilde{\boldsymbol{p}})}(T) = B\tilde{Y} \text{ a.s.}, \tag{6.147}$$

and therefore, from (6.142) and (6.147)

$$x = E[\tilde{H}\bar{X}^{(x,\tilde{c},\tilde{\boldsymbol{p}})}(T)], \tag{6.148}$$

for some $(\tilde{c}, \tilde{\boldsymbol{p}}) \in \mathcal{A}(x)$. □

The following proposition provides a characterization for the pair $(-Y, -y) \in \partial\tilde{U}(\tilde{z}\tilde{H}, \tilde{z})$.

Proposition 6.2.20. *Let $(-Y, -y) \in \partial \tilde{U}(\tilde{z}\tilde{H}, \tilde{z})$. Then $y = -x$ and Y is of the form*

$$Y = I_{\{\tilde{z}\tilde{H}<1\}} + DI_{\{\tilde{z}\tilde{H}=1\}} \quad P^B - \ a.s. \tag{6.149}$$

for some $\mathcal{F}_T$ random variable D that satisfies

$$0 \leq D \leq 1 \quad P^B - \ a.s. \tag{6.150}$$

Proof. Fix some $(Y, y) \in \mathcal{L}^*$ such that $(-Y, -y) \in \partial \tilde{U}(\tilde{z}\tilde{H}, \tilde{z})$. Define a random variable $A \in L^\infty(\Omega, \mathcal{F}_T, P^B)$ as follows

$$A \triangleq Y - I_{\{\tilde{z}\tilde{H}<1\}}. \tag{6.151}$$

From (6.131) in Proposition 6.2.18 we already know that $y = -x$. Using (6.125) and A defined in (6.151) gives

$$E^B[I_{\{\tilde{z}\tilde{H}\geq 1\}}] - E^B[Z \wedge 1] \geq E^B[A(\tilde{z}\tilde{H} - Z)] - E^B[I_{\{\tilde{z}\tilde{H}<1\}}Z], \quad Z \in L^1(\Omega, \mathcal{F}_T, P^B). \tag{6.152}$$

Now fix $Z \in L^1(\Omega, \mathcal{F}_T, P^B)$ where

$$\{Z < 1\} = \{\tilde{z}\tilde{H} < 1\}. \tag{6.153}$$

We can write

$$E^B[A(\tilde{z}\tilde{H} - Z)] = E^B[A(\tilde{z}\tilde{H} - Z)I_{\{\tilde{z}\tilde{H}<1\}}] + E^B[A(\tilde{z}\tilde{H} - Z)I_{\{\tilde{z}\tilde{H}\geq 1\}}]. \tag{6.154}$$

From (6.152) and (6.154)

$$\begin{aligned} E^B[A(\tilde{z}\tilde{H} - Z)I_{\{\tilde{z}\tilde{H}<1\}}] + E^B[A(\tilde{z}\tilde{H} - Z)I_{\{\tilde{z}\tilde{H}\geq 1\}}] \\ \leq E^B[I_{\{\tilde{z}\tilde{H}\geq 1\}}] - E^B[Z \wedge 1] + E^B[ZI_{\{\tilde{z}\tilde{H}<1\}}]. \end{aligned} \tag{6.155}$$

Using (6.153) in (6.155)

$$E^B[A(\tilde{z}\tilde{H} - Z)I_{\{\tilde{z}\tilde{H}<1\}}] + E^B[A(\tilde{z}\tilde{H} - Z)I_{\{\tilde{z}\tilde{H}\geq 1\}}] \leq 0. \tag{6.156}$$

Now, in (6.156), set

$$Z(\omega) = \tilde{z}\tilde{H}(\omega) \quad \text{on } \omega \in \{\tilde{z}\tilde{H} \geq 1\}. \tag{6.157}$$

We would then have from (6.156)

$$E^B[A(\tilde{z}\tilde{H} - Z)I_{\{\tilde{z}\tilde{H}<1\}}] \leq 0, \tag{6.158}$$

and when expanded,

$$E^B[A(\tilde{z}\tilde{H} - Z)I_{\{\tilde{z}\tilde{H}<1\}\cap\{A\leq 0\}}] + E^B[A(\tilde{z}\tilde{H} - Z)I_{\{\tilde{z}\tilde{H}<1\}\cap\{A>0\}}] \leq 0. \tag{6.159}$$

Remark 6.2.21. We can assume without loss of generality that $\tilde{z}\tilde{H} > 0$ P^B a.s. since from Proposition 6.2.12, $\tilde{z} > 0$, and since $\{\tilde{H} = 0\}$ only on $\{B = 0\}$ and we know $P[B = 0] = 0$. If $\{\tilde{H} = 0\}$ outside of $\{B = 0\}$, then $x = 0$ by Proposition 6.2.18, which by Remark 6.1.4 is not possible.

Now take

$$Z(\omega) = \begin{cases} \tilde{z}\tilde{H}(\omega), & \omega \in \{\tilde{z}\tilde{H} < 1\} \cap \{A \leq 0\} \\ 0, & \omega \in \{\tilde{z}\tilde{H} < 1\} \cap \{A > 0\}. \end{cases} \tag{6.160}$$

From (6.159) and (6.160), we have that

$$E^B[A\tilde{z}\tilde{H}I_{\{\tilde{z}\tilde{H}<1\}\cap\{A>0\}}] \leq 0, \tag{6.161}$$

but since $A > 0$ and $\tilde{z}\tilde{H} > 0$ P^B a.s. (see Remark 6.2.21), we have

$$P^B[\{\tilde{z}\tilde{H} < 1\} \cap \{A > 0\}] = 0. \tag{6.162}$$

In (6.156), instead set

$$Z(\omega) = \tilde{z}\tilde{H}(\omega) \quad \text{on } \omega \in \{\tilde{z}\tilde{H} < 1\}. \tag{6.163}$$

This would result in

$$E^B[A(\tilde{z}\tilde{H} - Z)I_{\{\tilde{z}\tilde{H}\geq1\}\cap\{A<0\}}] + E^B[A(\tilde{z}\tilde{H} - Z)I_{\{\tilde{z}\tilde{H}\geq1\}\cap\{A\geq0\}}] \leq 0. \tag{6.164}$$

Set

$$Z(\omega) = \begin{cases} \tilde{z}\tilde{H}(\omega), & \omega \in \{\tilde{z}\tilde{H} \geq 1\} \cap \{A \geq 0\} \\ \tilde{z}\tilde{H}(\omega) + \varepsilon, & \omega \in \{\tilde{z}\tilde{H} \geq 1\} \cap \{A < 0\}, \end{cases} \tag{6.165}$$

for some $\varepsilon > 0$. Putting (6.165) into (6.164) yields,

$$\varepsilon E^B[AI_{\{\tilde{z}\tilde{H}\geq1\}\cap\{A<0\}}] \geq 0, \tag{6.166}$$

and therefore,

$$P^B[\{\tilde{z}\tilde{H} \geq 1\} \cap \{A < 0\}] = 0. \tag{6.167}$$

From (6.162) and (6.167)

$$\begin{aligned} A(\omega) &\leq 0, \quad P^B \text{ almost all } \omega \in \{\tilde{z}\tilde{H} < 1\}, \\ A(\omega) &\geq 0, \quad P^B \text{ almost all } \omega \in \{\tilde{z}\tilde{H} \geq 1\}. \end{aligned} \tag{6.168}$$

Now suppose,

$$P^B[\{\tilde{z}\tilde{H} < 1\} \cap \{A < 0\}] > 0. \tag{6.169}$$

Then by (6.168) there exists $\delta > 0$ such that

$$E^B[A(\tilde{z}\tilde{H} - 1)I_{\{\tilde{z}\tilde{H}<1\}}] > \delta. \tag{6.170}$$

For some $\varepsilon > 0$, set

$$Z(\omega) = \begin{cases} 1-\varepsilon, & \omega \in \{\tilde{z}\tilde{H} < 1\} \\ 1, & \omega \in \{\tilde{z}\tilde{H} \geq 1\}. \end{cases} \tag{6.171}$$

Putting (6.171) into (6.156) gives

$$\begin{aligned} &0 \geq E^B[A(\tilde{z}\tilde{H} - 1 + \varepsilon)I_{\{\tilde{z}\tilde{H}<1\}}] + E^B[A(\tilde{z}\tilde{H} - 1)I_{\{\tilde{z}\tilde{H}\geq 1\}}] \\ &0 \geq E^B[A(\tilde{z}\tilde{H} - 1)I_{\{\tilde{z}\tilde{H}<1\}}] + \varepsilon E^B[AI_{\{\tilde{z}\tilde{H}<1\}}] + E^B[A(\tilde{z}\tilde{H} - 1)I_{\{\tilde{z}\tilde{H}\geq 1\}}] \\ &0 \geq \delta + \varepsilon E^B[AI_{\{\tilde{z}\tilde{H}<1\}}]. \end{aligned} \tag{6.172}$$

Taking $\varepsilon \to 0$ in (6.172) gives $0 \geq \delta$ which is a contradiction. Thus,

$$P^B[\{\tilde{z}\tilde{H} < 1\} \cap \{A = 0\}] = 1. \tag{6.173}$$

From (6.156) and (6.173), we have that

$$E^B[A(\tilde{z}\tilde{H} - Z)I_{\{\tilde{z}\tilde{H}\geq 1\}}] \leq 0, \quad \{Z < 1\} = \{\tilde{z}\tilde{H} < 1\}. \tag{6.174}$$

Taking $Z(\omega) = 1$ on $\omega \in \{\tilde{z}\tilde{H} \geq 1\}$ in (6.174) gives

$$E^B[A(\tilde{z}\tilde{H} - 1)I_{\{\tilde{z}\tilde{H}>1\}}] \leq 0. \tag{6.175}$$

Since $A(\omega) \geq 0$, P^B almost all $\omega \in \{\tilde{z}\tilde{H} \geq 1\}$ and $\tilde{z}\tilde{H}(\omega) - 1 > 0$, P^B almost all $\omega \in \{\tilde{z}\tilde{H} > 1\}$, we have from (6.175),

$$P^B[\{\tilde{z}\tilde{H} > 1\} \cap \{A = 0\}] = 1. \tag{6.176}$$

From (6.152), (6.173), and (6.176),

$$E^B[I_{\{\tilde{z}\tilde{H}\geq 1\}} - I_{\{Z\geq 1\}}] \geq E^B[A(1 - Z)I_{\{\tilde{z}\tilde{H}=1\}}] - E^B[Z(I_{\{\tilde{z}\tilde{H}<1\}} - I_{\{Z<1\}})] \tag{6.177}$$

for all $Z \in L^1(\Omega, \mathcal{F}_T, P^B)$. Suppose that

$$P^B[\{\tilde{z}\tilde{H} = 1\} \cap \{A > 1\}] > 0. \tag{6.178}$$

Then there exists $\delta > 0$ such that

$$E^B[AI_{\{\tilde{z}\tilde{H}=1\}\cap\{A>1\}}] > \delta + P^B[\{\tilde{z}\tilde{H} = 1\} \cap \{A > 1\}]. \tag{6.179}$$

For $\varepsilon > 0$, setting

$$Z(\omega) = \begin{cases} 0, & \omega \in \{\tilde{z}\tilde{H} = 1\} \cap \{A > 1\} \\ 1, & \omega \in \{\tilde{z}\tilde{H} = 1\} \cap \{A \leq 1\} \\ 1 - \varepsilon, & \text{otherwise} \end{cases} \tag{6.180}$$

in (6.177) implies

$$\begin{aligned} P^B[\{\tilde{z}\tilde{H} \geq 1\}] + P^B[\{\tilde{z}\tilde{H} = 1\} \cap \{A \leq 1\}] &\geq E^B[AI_{\{\tilde{z}\tilde{H}=1\} \cap \{A>1\}}] \\ &- (1-\varepsilon)\Big(P^B[\{\tilde{z}\tilde{H} < 1\}] + P^B[\{\tilde{z}\tilde{H} = 1\} \cap \{A \neq 1\}]\Big). \end{aligned} \tag{6.181}$$

From (6.179) and (6.181), and some simplification,

$$0 > \delta - \varepsilon P^B[\{\tilde{z}\tilde{H} > 1\}] \geq \delta - \varepsilon. \tag{6.182}$$

But since $\delta > 0$, ε can be chosen so that $\varepsilon \geq \delta$ giving a contradiction, and thus

$$P^B[\{\tilde{z}\tilde{H} = 1\} \cap \{A > 1\}] = 0. \tag{6.183}$$

From (6.168), (6.173), (6.176) and (6.183),

$$\begin{cases} 0 \leq A(\omega) \leq 1, & P^B \text{ almost all } \omega \in \{\tilde{z}\tilde{H} = 1\} \\ A(\omega) = 0, & P^B \text{ almost all } \omega \in \{\tilde{z}\tilde{H} \neq 1\} \end{cases}. \tag{6.184}$$

Putting (6.184) into (6.151) gives

$$Y = I_{\{\tilde{z}\tilde{H}<1\}} + AI_{\{\tilde{z}\tilde{H}=1\}}. \tag{6.185}$$

Putting an $\mathcal{F}_T$ random variable $D \triangleq A$ in place of A in (6.185) completes the proof. $\square$

Now for the main result of this chapter.

Theorem 6.2.22. *For any initial wealth $0 < x < \hat{u} < \infty$, there exists a pair $(\tilde{c}, \tilde{\boldsymbol{p}}) \in \mathcal{A}(x)$ that solves the minimization problem*

$$\inf_{(c,\boldsymbol{p}) \in \mathcal{A}(x)} E[\bar{B} - \bar{X}^{(x,c,\boldsymbol{p})}(T)]^+. \tag{6.186}$$

The pair $(\tilde{c}, \tilde{\boldsymbol{p}}) \in \mathcal{A}(x)$ is the hedging strategy for which the terminal wealth $X^{(x,\tilde{c},\tilde{\boldsymbol{p}})}(T)$ is given by

$$X^{(x,\tilde{c},\tilde{\boldsymbol{p}})}(T) = B\Big(I_{\{\tilde{z}\tilde{H}<1\}} + \tilde{D}I_{\{\tilde{z}\tilde{H}=1\}}\Big) \ a.s. \tag{6.187}$$

Here $(\tilde{z}\tilde{H}, \tilde{z})$ is an optimal solution for the dual problem (6.116), and $\tilde{D}$ can be taken as the random variable D in Proposition 6.2.20, with (Y, y) replaced by some $(\tilde{Y}, \tilde{y}) \in \{-\partial\tilde{U}(\tilde{z}\tilde{H}, \tilde{z}) \cap N_c(\tilde{z}\tilde{H}, \tilde{z})\}$, which exists by Proposition 6.2.17.

Proof. From Proposition 6.2.19, we know there exists a pair $(\tilde{c}, \tilde{\boldsymbol{p}}) \in \mathcal{A}(x)$ that hedges the contingent claim $B\tilde{Y}$ from initial investment x where $(\tilde{Y}, \tilde{y}) \in \{-\partial\tilde{U}(\tilde{z}\tilde{H}, \tilde{z}) \cap N_c(\tilde{z}\tilde{H}, \tilde{z})\}$, $\tilde{y} = -x$, and by Proposition 6.2.20

$$\tilde{Y} = I_{\{\tilde{z}\tilde{H}<1\}} + \tilde{D}I_{\{\tilde{z}\tilde{H}=1\}} \quad P^B - \text{ a.s.} \tag{6.188}$$

for some $\mathcal{F}_T$ random variable $\tilde{D}$ that satisfies

$$0 \le \tilde{D} \le 1 \quad P^B - \text{ a.s.,} \tag{6.189}$$

where $\tilde{z} > 0$ and $\tilde{H} \in \mathcal{Z}$ exist, are characterized by Proposition 6.2.17, and are optimal for the dual problem (6.62). Therefore, (6.187) follows from the hedging of $B\tilde{Y}$. By Proposition 6.2.19

$$E[\bar{B}]E^B[\frac{1}{B}\tilde{H}X^{(x,\tilde{c},\tilde{\boldsymbol{p}})}(T)] = x. \tag{6.190}$$

If we take

$$Y^{(x,\tilde{c},\tilde{\boldsymbol{p}})} = \frac{X^{(x,\tilde{c},\tilde{\boldsymbol{p}})}(T)}{B} \text{ a.s.,} \tag{6.191}$$

then, from (6.187),

$$Y^{(x,\tilde{c},\tilde{\boldsymbol{p}})} = I_{\{\tilde{z}\tilde{H}<1\}} + \tilde{D}I_{\{\tilde{z}\tilde{H}=1\}} \text{ a.s.} \tag{6.192}$$

and from (6.190)

$$x = E[\bar{B}]E^B[\tilde{H}Y^{(x,\tilde{c},\tilde{\boldsymbol{p}})}]. \tag{6.193}$$

By Proposition 6.2.6, (6.193), and (6.192) we have solved the minimization problem (6.186) and the proof is complete. □

Remark 6.2.23. In summary of Theorem 6.2.22, we have shown that if an investor begins trading with an initial wealth $x < \hat{u}$, where $\hat{u}$ is the least initial wealth in which one can hedge the contingent claim B, then there exists a hedging strategy $(\tilde{c}, \tilde{\boldsymbol{p}}) \in \mathcal{A}(x)$ that minimizes the investor's expected net loss (6.186) at the close of trade. Furthermore, the strategy $(\tilde{c}, \tilde{\boldsymbol{p}}) \in \mathcal{A}(x)$ produces terminal wealth

$$X^{(x,\tilde{c},\tilde{\boldsymbol{p}})}(T) = B\left(I_{\{\tilde{z}\tilde{H}<1\}} + \tilde{D}I_{\{\tilde{z}\tilde{H}=1\}}\right) \text{ a.s.,} \tag{6.194}$$

where $0 \le \tilde{D} \le 1$ a.s. is some $\mathcal{F}_T$ measurable random variable and the pair $(\tilde{z}\tilde{H}, \tilde{z})$ is optimal for the dual problem (6.62) and the solution to

$$0 \in \partial\tilde{U}(Z, z) + N_c(Z, z) \tag{6.195}$$

for $\tilde{U}(\cdot,\cdot)$ defined in (6.113) and $N_c(\cdot,\cdot)$ defined in (6.124).

Chapter 7

Conclusions and Further Developments

In this thesis we have addressed the problems of almost sure hedging and approximate hedging in a market comprised of two independent sources of randomness: Brownian motion and a finite state Markov chain. These sources of randomness allow for both small scale persistent movement in asset prices (caused by supply and demand, and modeled by Brownian motion) and large scale occasional movement in the asset prices (caused, for example, by global events, and modeled by the finite-state Markov chain). Unlike the most common regime-switching market models, where the market parameters are Markov-modulated, we allow the market coefficients to be truly random processes. Using this kind of regime-switching market model is not only a significant generalization over the standard Brownian motion market model but also other regime-switching market models, since it allows for stochasticity within the market regimes.

In the almost sure hedging problem, an investor must find the *least* initial wealth so they can trade in such a way to guarantee the full payment of a given *contingent claim*. This least initial wealth and method of trading, called the *price* of the contingent claim and the *optimal hedging strategy*, respectively, are given in Chapter 3 for the standard Brownian motion market model. Although these results are well known and not new, they provide the basic martingale methods required for more general hedging problems. One of such problems is the almost sure hedging problem where the Brownian motion market model includes a regime-switching Markov chain as an additional source of randomness. The inclusion of a finite state Markov chain makes the almost sure hedging problem substantially more challenging due to the intricacies of having the market coefficients dependant on both a finite-state Markov chain and Brownian motion. Since the market model is now *incomplete*, one requires a *cumulative consumption process* and a space of *dual processes* to make the hedging problem tractable. As a result, an investor is now required to find a least initial wealth, an optimal portfolio process, and an optimal cumulative consumption using a *state price density process* indexed over the space of dual processes to exactly hedge a given contingent claim. The *optimal hedging strategy* is now used to denote a pair of portfolio and consumption processes that hedges the given contingent from the least initial wealth, and they are explicitly characterized in Chapter 4 using the Doob-Meyer decomposition and an abstract martingale representation theorem.

We further generalize the almost sure hedging problem in a regime-switching market model by including convex portfolio constraints, first proposed by Cvitanic and Karatzas [10], and margin requirements, of the kind which were first introduced by Cuoco and Liu [7] for utility maximization, in Chapter 5. In this market, a second space of dual processes, which is similar to the dual space of Lagrange multipliers in classical convex analysis, is required to deal with the effects of portfolio constraints and margin requirements. As a result, we must use a *generalized state price density process* indexed over a *joint space* of dual processes to find the least initial wealth and optimal hedging strategy to exactly hedge a given contingent claim. What makes this problem particularly challenging is the nonlinearity of the margin function over the space of portfolio processes. Again, we are able to characterize an optimal hedging strategy by use of the Doob-Meyer decomposition and the martingale representation theorem, but due to the inclusion of portfolio constraints and margin requirements, elements of convex analysis are required to guarantee that we abide by the constraints. In addition to solving the hedging problem in a constrained regime-switching market model, we propose conditions which allow an agent to hedge a contingent claim without having to consume wealth through a cumulative consumption process. We show that if the least initial wealth, defined by taking a supremum over the joint space of dual processes, is indeed attained by an *optimal pair* of dual processes, one can exactly hedge the contingent claim without consuming wealth. This result is an extension of the conditions for zero consumption given by Cvitanic and Karatzas [10] in a standard Brownian market model with only convex portfolio constraints. Since the characterization of the optimal hedging strategy relies upon a general martingale representation theorem, we cannot give computable formulae for the optimal portfolio and consumption processes. An interesting area for future research would be in the numerical approximation of these processes, or if possible, explicit solutions to these processes through a method similar to the Clark-Ocone formula. Currently, there is a method which allows one to explicitly solve for the processes given by the martingale representation theorem when the underlying filtration is generated by a Lévy process of which Brownian motion is a special kind (see Løkka [34]). However, there are no such results when the underlying filtration is generated by a continuous time Markov chain, let alone a filtration jointly generated by a Markov chain and Brownian motion. There have been promising developments in the Malliavin calculus (which is a fundamental ingredient for a Clark-Ocone formula) for processes adapted to a filtration generated by a continuous time Markov chain (see Siu [47] and Denis and Nguyen [11]); however, as of yet, there seems to be no progress in the creation of a Clark-Ocone formula for this case. Having formulae which explicitly solve for these unknown processes would greatly advance the applicability of this work to real-world problems.

The problems discussed above all involve *super-hedging* a given claim, in which the agent begins with some initial wealth which is large enough for the agent to trade in such a way as to have sufficient wealth at close of trade to *completely pay off* the claim. In the most general case of Chapter 5, the least initial wealth required to hedge a contingent claim could be extremely large since it involves the supremum taken over the typically large joint space of dual processes, which could be unacceptably high for a normal investor. Therefore, in Chapter 6 we address the problem of *approximate hedging* where an agent begins with an initial wealth they can "afford" and trades in such a way to *minimize* a piece-wise linear one-sided risk criterion which imposes a penalty when the agent falls short of being able to fund the contingent claim. Cvitanic [8] first addressed this problem in a standard Brownian motion market model with convex portfolio constraints and showed the existence of a trading strategy which in fact minimizes the risk criterion. In Chapter 6, we address the problem of approximate hedging in a regime-switching market model with convex portfolio constraints and margin requirements and show the existence of a trading strategy which minimizes the risk criterion, using nearly the identical approach as [8]. Our solution utilizes the main hedging result of Chapter 5 along with tools of general convex analysis to deal with non-smoothness of the risk criterion. Just as in the almost sure hedging problem, there are currently no methods for finding computable solutions for the optimal approximate hedging strategy. Future work would then be to study possible numerical algorithms for approximating these processes. Convex programming techniques and numerical PDE methods have been proposed as possible ways for solving these kinds of problems.

References

[1] J.P. Aubin and I. Ekeland. *Applied Nonlinear Analysis*. John Wiley and Sons, New York, 1984.

[2] L. Bachelier. Théorie de la spéculation. *Annales scientifiques de l'École normale supérieure*, 17:21–86, 1900.

[3] F. Black. Studies of the stock price volatility changes. *Proceedings of the American Statistical Association, Business and Economic Statistics Section*, pages 177–181, 1976.

[4] F. Black and M. Scholes. The pricing of options and corporate liabilities. *Journal of Political Economy*, 81(3):637–654, 1973.

[5] J. Buffington and R.J. Elliot. American options with regime switching. *International Journal of Theoretical and Applied Finance*, 5:497–514, 2002.

[6] K.L Chung and J.B. Walsh. *Markov Processes, Brownian Motion, and Time Symmetry*, volume 249. Springer-Verlag, New York, 2005.

[7] D. Cuoco and H. Liu. A martingale characterization of consumption choices and hedging costs with margin requirements. *Mathematical Finance*, 10(3):355–385, 2000.

[8] J. Cvitanic. Minimizing expected loss of hedging in incomplete and constrained markets. *SIAM Journal on Control and Optimization*, 38(4):1050–1066, 2000.

[9] J. Cvitanić and I. Karatzas. Convex duality in constrained portfolio optimization. *The Annals of Applied Probability*, 2(4):767–818, 1992.

[10] J. Cvitanic and I. Karatzas. Hedging contingent claims with constrained portfolios. *The Annals of Applied Probability*, 3(3):652–681, 1993.

[11] L. Denis and T.M. Nguyen. Malliavin calculus for markov chains using perturbations of time. *Stochastics*, 88(6):813–840, 2016.

[12] G.B. Di Masi, Y.M. Kabanov, and W.J. Runggaldier. Mean variance hedging of options on stocks with markov volatility. *Theory of Probability and its Applications*, 39:172–181, 1994.

[13] C. Donnelly and A.J. Heunis. Quadratic risk minimization in a regime-switching model with portfolio constraints. *SIAM Journal on Control and Optimization*, 50(4):2431–2461, 2012.

[14] A. Einstein. Uber die von der molekularkinetischen theorie der wärme geforderte bewegung von in ruhenden flüssigkeiten suspendierten teilchen. *Annalen der Physik und Chemie*, 17:549–560, 1905.

[15] A. Einstein. Zur theorie der brownschen bewegung. *Annalen der Physik*, 19:371–381, 1906.

[16] N. El Karoui and M.C. Quenez. Programmation dynamique et évaluation des actifs contingents en marché incomplet. *Comptes rendus de l'Académie des sciences. Série 1, Mathématique*, 331(12):851–854, 1991.

[17] N. El Karoui and M.C. Quenez. Dynamic programming and pricing of contingent claims in an incomplete market. *SIAM Journal on Control and Optimization*, 33(1):29–66, 1995.

[18] H. Follmer and D. Kramkov. Optional decompositions under constraints. *Probability Theory and Related Fields*, 109:1–25, 1997.

[19] J.D. Hamilton. A new approach to the economic analysis of non-stationary time series. *Econometrica*, 57:357–384, 1989.

[20] J.M. Harrison and D.M. Kreps. Martingales and arbitrage in multiperiod security markets. *Journal of Economic Theory*, 20:381–408, 1979.

[21] J.M. Harrison and S.R. Pliska. A stochastic calculus model of continuous trading: complete markets. *Stochastic Processes and Their Applications*, 15(3):313–316, 1983.

[22] S.L. Heston. A closed-form solution for options with stochastic volatility with applications to bond and currency options. *The Review of Financial Studies*, 6(2):327–343, 1993.

[23] A.J. Heunis. Utility maximization in a regime switching model with convex portfolio constraints and margin requirements: optimality relations and explicit solutions. *SIAM Journal on Control and Optimization*, 53(4):2608–2656, 2015.

[24] J.B. Hiriart-Urruty and C. Lemaréchal. *Fundamentals of Convex Analysis*. Springer-Verlag, Berlin Heidelberg, 2001.

[25] J.C. Hull. *Options, Futures and other Derivative Securities*. Prentice Hall, Englewood Cliffs, N.J., 1993.

[26] J. Jacod and A. Shiryaev. *Limit Theorems for Stochastic Processes*. Springer-Verlag, Berlin Heidelberg, 1987.

[27] A. Jobert and L.C.G. Rogers. Option pricing with markov-modulated dynamics. *SIAM Journal on Control and Optimization*, 44(6):2063–2078, 2006.

[28] O. Kallenberg. *Foundations of Modern Probability, 2nd Ed.* Springer-Verlag, New York, 2002.

[29] I. Karatzas, J.P. Lehoczky, S.E. Shreve, and G.L. Xu. Martingale and duality methods for utility maximization in an incomplete market. *SIAM Journal on Control and Optimization*, 29(3):702–730, 1991.

[30] I. Karatzas and S.E. Shreve. *Brownian Motion and Stochastic Calculus*. Springer-Verlag, New York, 1988.

[31] I. Karatzas and S.E. Shreve. *Methods of Mathematical Finance*. Springer-Verlag, New York, 1998.

[32] A.N. Kolmogorov and S.V. Fomin. *Introductory Real Analysis*. Dover, New York, 1975.

[33] R.S. Lipster and A.N. Shiryayev. *Theory of Martingales*. Kluwer, Dordrecht, 1989.

[34] A. Løkka. Martingale representation of functionals of Lévy processes. *Stochastic Analysis and Applications*, 22(4):867–892, 2004.

[35] H.M. Markowitz. Portfolio selection. *The Journal of Finance*, 7(1):77–91, 1952.

[36] R.C. Merton. Theory of rational option pricing. *Theory of Valuation*, pages 229–288, 1973.

[37] R.C. Merton. Option pricing when underlying stock returns are discontinuous. *Journal of Financial Economics*, 3:125–144, 1976.

[38] P.E. Protter. *Stochastic Integration and Differential Equations.* Springer-Verlag, Berlin Heidelberg, 2004.

[39] D. Revuz and M. Yor. *Continuous Martingales and Brownian Motion.* Springer-Verlag, Berlin Heidelberg, 1994.

[40] R.T. Rockafellar. Integrals which are convex functionals. *Pacific Journal of Mathematics*, 24(3):525–539, 1968.

[41] R.T. Rockafellar. Measurable dependence of convex sets and functions on parameters. *Journal of Mathematical Analysis and Applications*, 28(1):4–25, 1969.

[42] L.C.G Rogers and D. Williams. *Diffusions, Markov Processes, and Martingales: Volume 1, Foundations.* Cambridge University Press, Cambridge (UK), 2000.

[43] L.C.G Rogers and D. Williams. *Diffusions, Markov Processes and Martingales: Volume 2, Itô calculus.* Cambridge University Press, Cambridge (UK), 2000.

[44] J.A. Samuelson. Rational theory of warrant pricing. *Industrial Management Review*, 6(2):13, 1965.

[45] P. Samuelson and R.C. Merton. A complete model of warrant pricing that maximizes utility. *Industrial Management Review*, 10(2):17, 1969.

[46] M. Schwartz. New proofs of a theorem of Komlós. *Acta Mathematica Hungarica*, 47(1-2):181–185, 1986.

[47] T.K. Siu. Integration by parts and martingale representation for a markov chain. In *Abstract and Applied Analysis.* Hindawi, 2014.

[48] L.R. Sotomayor and A. Cadenillas. Explicit solutions of consumption-investment problems in financial markets with regime switching. *Mathematical Finance*, 19(2):251–279, 2009.

[49] N. Wiener. Differential space. *Journal of Mathematical Physics*, 2:131–174, 1923.

[50] G.L. Xu and S.E. Shreve. A duality method for optimal consumption and investment under short-selling prohibition. i. general market coefficients. *The Annals of Applied Probability*, pages 87–112, 1992.

[51] D.D. Yao, Zhang Q., and X.Y. Zhou. *Stochastic Processes, Optimization, and Control Theory: Applications in Financial Engineering, Queueing Networks, and Manufacturing Systems (volume in honor of Suresh Sethi)*. Springer-Verlag, New York, 2006.

[52] G. Yin, H. Liu, and Q. Zhang. Recursive algorithms for stock liquidation: A stochastic optimization approach. *SIAM Journal of Optimization*, 13:240–263, 2002.

[53] Q. Zhang. Stock trading: An optimal selling rule. *SIAM Journal on Control and Optimization*, 40:64–87, 2001.

[54] X.Y. Zhou and G. Yin. Markowitz's mean-variance portfolio selection with regime switching: A continuous-time model. *SIAM Journal on Control and Optimization*, 42(4):1466–1482, 2003.

APPENDICES

Appendix A

Supplementary Results

A.1 Spaces of Integrand Processes

Definition A.1.1. For each $i, j \in S, i \neq j$, denote $L^{1/2}_{loc}(R_{ij})$ to be the vector space of all $\mathbb{R}$-valued, $\{\mathcal{F}_t\}$-predictable processes Γ such that the process defined by the pathwise Lebesgue-Stieltjes integral $t \mapsto \{\int_0^t |\Gamma(s)|^2 dR_{ij}(s)\}^{1/2}$ is a member of $\mathcal{A}^+_{loc}(\{\mathcal{F}_t\}, P)$ (see Notation C.5.7).

Remark A.1.2. $L^{1/2}_{loc}(R_{ij})$ is the largest space of integrand processes Γ for which the stochastic integral $\Gamma \bullet M_{ij}$ is defined and is a *local martingale*.

Similarly we can define the following spaces of integrands:

Definition A.1.3. For each $i, j \in S, i \neq j$, denote $L^1_{loc}(R_{ij})$ to be the vector space of all $\mathbb{R}$-valued, $\{\mathcal{F}_t\}$-predictable processes Γ such that the process defined by the pathwise Lebesgue-Stieltjes integral $t \mapsto \int_0^t |\Gamma(s)| dR_{ij}(s)$ is a member of $\mathcal{A}^+_{loc}(\{\mathcal{F}_t\}, P)$ (see Notation C.5.7).

Definition A.1.4. For each $i, j \in S, i \neq j$, denote $L^1_{loc}(\tilde{R}_{ij})$ to be the vector space of all $\mathbb{R}$-valued, $\{\mathcal{F}_t\}$-predictable processes Γ such that the process defined by the pathwise Lebesgue-Stieltjes integral $t \mapsto \int_0^t |\Gamma(s)| d\tilde{R}_{ij}(s)$ is a member of $\mathcal{A}^+_{loc}(\{\mathcal{F}_t\}, P)$ (see Notation C.5.7).

Definition A.1.5. For each $i, j \in S, i \neq j$, denote $L^1_{loc}(\text{var}(M_{ij}))$ to be the vector space of all $\mathbb{R}$-valued, $\{\mathcal{F}_t\}$-predictable processes Γ such that the process defined by the pathwise Lebesgue-Stieltjes integral $t \mapsto \int_0^t |\Gamma(s)| \text{var}(M_{ij})(ds)$, where $\text{var}(M_{ij})$ is the variation of the process M_{ij}, is a member of $\mathcal{A}^+_{loc}(\{\mathcal{F}_t\}, P)$ (see Notation C.5.7).

The purpose of this appendix is to show that if $\{\Gamma(t), t \in [0,T]\}$ is some $\{\mathcal{F}_t\}$-predictable processes that lives in the space $L^{1/2}_{loc}(R_{ij})$, which is the usual space of integrands that are dealt with when integrating with respect to M_{ij}, we can unabashedly take the integral of Γ with respect to the processes R_{ij} and $\tilde{R}_{ij}$ and know they "make sense". This fact is used numerous times in Chapter 4 and Chapter 5.

Remark A.1.6. We recall that R_{ij} is the quadratic variation process of M_{ij}, $\tilde{R}_{ij}$ is the predictable compensator R_{ij} which we can denote as $(R_{ij})^p$, and $\text{var}(M_{ij})$ is the total variation of M_{ij} for each $i, j \in S, i \neq j$.

Theorem A.1.7. *The spaces $L^{1/2}_{loc}(R_{ij})$, $L^1_{loc}(R_{ij})$, $L^1_{loc}(\tilde{R}_{ij})$ and $L^1_{loc}(var(M_{ij}))$ are equal for $i, j \in S, i \neq j$.*

Remark A.1.8. To prove Theorem A.1.7, we will first prove a series of Propositions from which we can conclude our desired result.

Lemma A.1.9. *For a sequence $\{a_n\} \in \mathbb{R}$, $n = 1, 2, \ldots, N < \infty$ we have*

$$\left(\sum_{n=1}^{N} |a_n|^2\right)^{1/2} \leq \sum_{n=1}^{N} |a_n| \leq \sqrt{N}\left(\sum_{n=1}^{N} |a_n|^2\right)^{1/2}. \tag{A.1}$$

Proof. Using the Cauchy-Schwarz inequality, we have

$$\begin{gathered}\left(\sum_{n=1}^{N} |a_n|\right)^2 \leq \left(\sum_{n=1}^{N} 1^2\right)\left(\sum_{n=1}^{N} |a_n|^2\right) \\ \left(\sum_{n=1}^{N} |a_n|\right) \leq \sqrt{N}\left(\sum_{n=1}^{N} |a_n|^2\right)^{1/2}\end{gathered} \tag{A.2}$$

which shows the right-hand inequality of (A.1). To show the left-hand side of (A.1), suppose

$$\left(\sum_{n=1}^{k} |a_n|^2\right)^{1/2} \leq \sum_{n=1}^{k} |a_n|, \tag{A.3}$$

for some $k < N$. Then

$$\left(\sum_{n=1}^{k+1}|a_n|^2\right)^{1/2} \le \left(\sum_{n=1}^{k}|a_n|^2\right)^{1/2} + |a_{k+1}|$$
$$\le \sum_{n=1}^{k}|a_n| + |a_{k+1}| \tag{A.4}$$
$$\le \sum_{n=1}^{k+1}|a_n|$$

Thus, by induction,

$$\left(\sum_{n=1}^{N}|a_n|^2\right)^{1/2} \le \sum_{n=1}^{N}|a_n|. \tag{A.5}$$

□

Lemma A.1.10. *Suppose that a process $\{A(t), t \in [0,T]\}$ is of finite variation (i.e. $A \in \mathcal{FV}(\{\mathcal{F}_t\}, P)$). Then*

$$\mathrm{var}(A)(t) \ge \sum_{0<s\le t}|\triangle A(s)|, \qquad t \in [0,T], \tag{A.6}$$

where $\mathrm{var}(A)(t)$ *is the variation of $A(t)$.*

Proof. Put $\nu(t) \triangleq \mathrm{var}(A)(t)$, for all $t \in [0,T]$. It is easy to see that $\nu \in \mathcal{FV}_0(\{\mathcal{F}_t\}, P)$ and $\nu(\cdot)$ is non-decreasing and càdlàg. From pg. 29 in Lipster and Shiryayev [33], we have

$$\nu(t) = \nu(t-) + |\triangle A(t)|, \qquad t \in [0,T], \tag{A.7}$$

thus,

$$\triangle\nu(t) = |\triangle A(t)|, \qquad t \in [0,T]. \tag{A.8}$$

Now put

$$\nu^d(t) \triangleq \sum_{0\le s\le t}\triangle\nu(s), \tag{A.9}$$

for all $t \in [0,T]$. Since $\nu(0) = 0$ and $\nu(\cdot)$ is right continuous, we have $\triangle\nu(0) = 0$. Thus, from (A.9)

$$\nu^d(t) = \sum_{0<s\le t}\triangle\nu(s), \qquad t \in [0,T]. \tag{A.10}$$

Put

$$\nu^c(t) \triangleq \nu(t) - \nu^d(t) \tag{A.11}$$

for $t \in [0, T]$, then $\nu^c(0) = 0$. From Theorem 5 in Kolmogorov and Fomin [32], $\nu^c(\cdot)$ is continuous and non-decreasing. Then from (A.11), $\nu^c(t) \geq 0$ for all $t \in [0, T]$, and

$$\nu(t) = \nu^c(t) + \nu^d(t) \geq \nu^d(t), \qquad t \in [0, T]. \tag{A.12}$$

Now from (A.8) and (A.10)

$$\nu^d(t) = \sum_{0<s\leq t} |\triangle A(s)|, \qquad t \in [0, T]. \tag{A.13}$$

Thus, from (A.12)

$$\nu(t) \geq \sum_{0<s\leq t} |\triangle A(s)|, \qquad t \in [0, T]. \tag{A.14}$$

□

Proposition A.1.11.

$$L^{1/2}_{loc}(R_{ij}) \subseteq L^1_{loc}(R_{ij}), \tag{A.15}$$

for all $i, j \in S, i \neq j$

Proof. Fix $(i, j) \in S$ such that $i \neq j$ and $\{\Gamma(t), t \in [0, T]\}$ such that

$$\Gamma \in L^{1/2}_{loc}(R_{ij}). \tag{A.16}$$

From Definition A.1.1 we know

$$\Gamma \in \mathcal{P}^*, \qquad (|\Gamma|^2 \bullet R_{ij})^{1/2} \in \mathcal{A}^+_{loc}. \tag{A.17}$$

From the definition of $\mathcal{A}^+_{loc}(\{\mathcal{F}_t\}, P)$, there exists a sequence of $\{\mathcal{F}_t\}$-stopping times $T^{(1)}_n$ such that $T^{(1)}_n \uparrow \infty$ a.s. as $n \to \infty$ and

$$E[(|\Gamma|^2 \bullet R_{ij})^{1/2}(T^{(1)}_n)] < \infty, \quad n = 1, 2, \ldots \tag{A.18}$$

Since the jumps of the process R_{ij} are bounded (they can only take the value 0 or 1), one has that R_{ij} is locally bounded. That is, there exists a sequence of $\{\mathcal{F}_t\}$-stopping times $T^{(2)}_n$ and constants $B_n \in [0, \infty)$ such that $T^{(2)}_n \uparrow \infty$ a.s. as $n \to \infty$ and

$$R_{ij}(t \wedge T^{(2)}_n(\omega), \omega) \leq B_n \quad (t, \omega) \in [0, \infty) \times \Omega. \tag{A.19}$$

Put

$$T_n \triangleq T_n^{(1)} \wedge T_n^{(2)} \wedge n \tag{A.20}$$

then from (A.18) and (A.19), we have that T_n is an $\{\mathcal{F}_t\}$-stopping time where $T_n \uparrow \infty$ a.s. as $n \to \infty$ and

$$E[(|\Gamma|^2 \bullet R_{ij})^{1/2}(T_n)] < \infty, \quad n = 1, 2, \dots. \tag{A.21}$$

$$R_{ij}(T_n(\omega), \omega) \leq B_n, \quad \omega \in \Omega \quad n = 1, 2 \dots, \tag{A.22}$$

Now define the set

$$D_{ij}(T_n(\omega), \omega) \triangleq \big\{0 < s \leq T_n(\omega) \big| \alpha(s-, \omega) = i, \alpha(s, \omega) = j\big\}, \tag{A.23}$$

for all $i, j \in S$ where $i \neq j$ and $\omega \in \Omega$. Then

$$R_{ij}(T_n(\omega), \omega) = |D_{ij}(T_n(\omega), \omega)|, \quad \omega \in \Omega, \tag{A.24}$$

where $|\cdot|$ counts the number elements in a given set, since $R_{ij}(t, \omega)$ counts the number of jumps the Markov chain makes until some time t. Therefore, we have

$$\begin{aligned}(|\Gamma| \bullet R_{ij})(T_n(\omega), \omega) &\triangleq \int_0^{T_n(\omega)} |\Gamma(s, \omega)| dR_{ij}(s, \omega) \\ &= \sum_{s \in D_{ij}(T_n(\omega), \omega)} |\Gamma(s, \omega)| \triangle R_{ij}(s, \omega), \quad \omega \in \Omega.\end{aligned} \tag{A.25}$$

Now, $\triangle R_{ij}(s, \omega) = 1$ for all $s \in D_{ij}(T_n(\omega), \omega)$ and $\omega \in \Omega$. Using this fact along with Lemma A.1.9,

$$\begin{aligned}(|\Gamma| \bullet R_{ij})(T_n(\omega), \omega) &= \sum_{s \in D_{ij}(T_n(\omega), \omega)} |\Gamma(s, \omega)| \\ &\leq \sqrt{|D_{ij}(T_n(\omega), \omega)|} \bigg(\sum_{s \in D_{ij}(T_n(\omega), \omega)} |\Gamma(s, \omega)|^2 \bigg)^{1/2} \\ &\leq \sqrt{R_{ij}(T_n(\omega), \omega)} \bigg(\sum_{s \in D_{ij}(T_n(\omega), \omega)} |\Gamma(s, \omega)|^2 \bigg)^{1/2} \\ &\leq \sqrt{B_n} \bigg(\sum_{s \in D_{ij}(T_n(\omega), \omega)} |\Gamma(s, \omega)|^2 \bigg)^{1/2}, \quad \omega \in \Omega.\end{aligned} \tag{A.26}$$

We also have,

$$\begin{aligned}\sum_{s\in D_{ij}(T_n(\omega),\omega)}|\Gamma(s,\omega)|^2 &= \sum_{s\in D_{ij}(T_n(\omega),\omega)}|\Gamma(s,\omega)|^2\triangle R_{ij}(\omega)\\ &= \int_0^{T_n(\omega)}|\Gamma(s,\omega)|^2 dR_{ij}(s).\end{aligned} \tag{A.27}$$

From (A.26), but using more condensed notation, and (A.27)

$$\int_0^{T_n(\omega)}|\Gamma(s,\omega)|R_{ij}(ds,\omega) \leq \sqrt{B_n}\Big((|\Gamma|^2\bullet R_{ij})(T_n(\omega),\omega)\Big)^{1/2}, \quad \omega\in\Omega, \tag{A.28}$$

for $n = 1, 2, \ldots$. Taking expectations of both sides of (A.28)

$$\begin{aligned}E\Big[\int_0^{T_n(\omega)}|\Gamma(s,\omega)|R_{ij}(ds,\omega)\Big] &\leq \sqrt{B_n}E\Big[\Big(\int_0^{T_n(\omega)}|\Gamma(s,\omega)|^2R_{ij}(ds,\omega)\Big)^{1/2}\Big]\\ &< \infty\end{aligned} \tag{A.29}$$

for $n = 1, 2, \ldots$. Therefore,

$$|\Gamma|\bullet R_{ij}\in\mathcal{A}^+_{loc}, \tag{A.30}$$

and thus from (A.30) and (A.16), we conclude that

$$\Gamma\in L^1_{loc}(R_{ij}),$$

therefore,

$$L^{1/2}_{loc}(R_{ij})\subseteq L^1_{loc}(R_{ij}), \tag{A.31}$$

for $i, j \in S, i \neq j$ □.

Proposition A.1.12.

$$L^1_{loc}(\text{var}(M_{ij}))\subseteq L^{1/2}_{loc}(R_{ij}), \tag{A.32}$$

for all $i, j \in S, i \neq j$.

Proof. Fix $t \in [0, \infty)$. We have from pg. 93 in Lipster and Shiryayev [33],

$$R_{ij}(t) = [M_{ij}](t) = \sum_{s\leq t}|\triangle M_{ij}(s)|^2. \tag{A.33}$$

For some $\Gamma \in \mathcal{P}^*$, from (A.33) write

$$\begin{aligned}\left(\int_0^t |\Gamma(s)|^2 dR_{ij}(s)\right)^{1/2} &= \left(\sum_{s\le t} |\Gamma(s)|^2 \triangle R_{ij}(s)\right)^{1/2} \\ &= \left(\sum_{s\le t} |\Gamma(s)|^2 |\triangle M_{ij}(s)|^2\right)^{1/2}.\end{aligned} \tag{A.34}$$

Using Lemma A.1.9 and (A.34),

$$\left(\int_0^t |\Gamma(s)|^2 dR_{ij}(s)\right)^{1/2} \le \sum_{s\le t} |\Gamma(s)||\triangle M_{ij}(s)|. \tag{A.35}$$

Put

$$B_1(t) \triangleq \sum_{s\le t} |\triangle M_{ij}(s)|. \tag{A.36}$$

Since $M_{ij} \in \mathcal{FV}_0(\{\mathcal{F}_t\}, P)$, from Lemma A.1.10,

$$B_1(t) \le \mathrm{var}(M_{ij})(t). \tag{A.37}$$

Then

$$\begin{aligned}\sum_{s\le t} |\Gamma(s)||\triangle M_{ij}(s)| &= \int_0^t |\Gamma(s)| dB_1(s) \\ &\le \int_0^t |\Gamma(s)| \mathrm{var}(M_{ij})(ds).\end{aligned} \tag{A.38}$$

Thus, from (A.35) and (A.38) we have

$$\left(\int_0^t |\Gamma(s)|^2 R_{ij}\right)^{1/2} \le \int_0^t |\Gamma(s)| \mathrm{var}(M_{ij})(ds), \tag{A.39}$$

for all $(t, \omega) \in [0, \infty) \times \Omega$ and $\Gamma \in \mathcal{P}^*$. Now take some $\Gamma \in L^1_{loc}(\mathrm{var}(M_{ij}))$. From the definition of $L^1_{loc}(\mathrm{var}(M_{ij}))$, Definition A.1.5, we have

$$|\Gamma| \bullet \mathrm{var}(M_{ij}) \in \mathcal{A}^+_{loc}, \tag{A.40}$$

and from (A.39)

$$\left(|\Gamma|^2 \bullet R_{ij}\right)^{1/2} \in \mathcal{A}^+_{loc}, \tag{A.41}$$

thus, we can conclude from (A.40) and (A.41) that

$$\Gamma \in L^{1/2}_{loc}(R_{ij}),$$

therefore

$$L^1_{loc}(\text{var}(M_{ij})) \subseteq L^{1/2}_{loc}(R_{ij}), \tag{A.42}$$

for $i, j \in S, i \neq j$. □

Proposition A.1.13.

$$L^1_{loc}(R_{ij}) = L^1_{loc}(\tilde{R}_{ij}), \tag{A.43}$$

for $i, j \in S, i \neq j$.

To prove this statement, we need the following result:

Lemma A.1.14.

$$E\bigg[\int_0^\tau |\Gamma(s)| dR_{ij}(s)\bigg] = E\bigg[\int_0^\tau |\Gamma(s)| d\tilde{R}_{ij}(s)\bigg] \tag{A.44}$$

for all $\Gamma \in \mathcal{P}^*$ *and all* $\{\mathcal{F}_t\}$*-stopping times* τ.

Proof. Fix some $\Gamma \in \mathcal{P}^*$ and some $\{\mathcal{F}_t\}$-stopping time τ. Then

$$\int_0^\tau |\Gamma(s)| dR_{ij}(s) = \int_0^\infty |\Gamma(s)| I_{(0,\tau]}(s) dR_{ij}(s), \tag{A.45}$$

where $I_{(a,b]}(s)$ is the indicator function that returns the value 1 if $s \in (a, b] \subset \mathbb{R}$ and returns the value 0 otherwise. We obviously have that $I_{(0,\tau]}$ is $\{\mathcal{F}_t\}$-adapted and is left continuous. Thus, $I_{(0,\tau]} \in \mathcal{P}^*$. Since Γ is a predictable process, we have

$$|\Gamma| I_{(0,\tau]} \in \mathcal{P}^*. \tag{A.46}$$

From (A.46), the fact that $\tilde{R}_{ij} = (R_{ij})^p$, that is $\tilde{R}_{ij}$ is the predictable compensator of R_{ij}, and I.3.17(iii) of Jacod and Shiryayev [26]:

$$E\bigg[\int_0^\infty |\Gamma(s)| I_{(0,\tau]}(s) d\tilde{R}_{ij}(s)\bigg] = E\bigg[\int_0^\infty |\Gamma(s)| I_{(0,\tau]}(s) dR_{ij}(s)\bigg]. \tag{A.47}$$

Using the same method from (A.45),

$$\int_0^\tau |\Gamma(s)| d\tilde{R}_{ij}(s) = \int_0^\infty |\Gamma(s)| I_{(0,\tau]}(s) d\tilde{R}_{ij}(s). \tag{A.48}$$

Finally, from (A.45), (A.47), and (A.48),

$$E\left[\int_0^\tau |\Gamma(s)|dR_{ij}(s)\right] = E\left[\int_0^\tau |\Gamma(s)|d\tilde{R}_{ij}(s)\right] \tag{A.49}$$

for $i, j \in S, i \neq j$.
□

Proof of Proposition A.1.13. Take $\Gamma \in L^1_{loc}(R_{ij})$ which implies

$$\Gamma \in \mathcal{P}^* \text{ and } |\Gamma| \bullet R_{ij} \in \mathcal{A}^+_{loc}. \tag{A.50}$$

This means that there exists a sequence $\{T_n\}$, $n = 1, 2, \ldots$, of $\{\mathcal{F}_t\}$-stopping times such that $T_n \uparrow \infty$ a.s. and

$$(|\Gamma| \bullet R_{ij})(T_n) \in \mathcal{A}^+, \quad n = 1, 2, \ldots \tag{A.51}$$

which means we can write

$$E\left[\int_0^{T_n} |\Gamma(s)|dR_{ij}(s)\right] < \infty, \quad n = 1, 2, \ldots \tag{A.52}$$

and from Lemma A.1.14

$$E\left[\int_0^{T_n} |\Gamma(s)|d\tilde{R}_{ij}(s)\right] < \infty, \quad n = 1, 2, \ldots, \tag{A.53}$$

which means

$$|\Gamma| \bullet \tilde{R}_{ij} \in \mathcal{A}^+_{loc}. \tag{A.54}$$

Since Γ was chosen to be predictable, by Definition A.1.4,

$$\Gamma \in L^1_{loc}(\tilde{R}_{ij}), \tag{A.55}$$

and we can conclude

$$L^1_{loc}(R_{ij}) \subseteq L^1_{loc}(\tilde{R}_{ij}). \tag{A.56}$$

We also obtain

$$L^1_{loc}(\tilde{R}_{ij}) \subseteq L^1_{loc}(R_{ij}) \tag{A.57}$$

using the exact same reasoning and Lemma A.1.14. Thus, we have shown

$$L^1_{loc}(R_{ij}) = L^1_{loc}(\tilde{R}_{ij}). \tag{A.58}$$

□

Proposition A.1.15.

$$L^1_{loc}(\text{var}(M_{ij})) = L^1_{loc}(R_{ij}) \tag{A.59}$$

for $i, j \in S, i \neq j$.

Proof. Fix $t \in [0, \infty)$. We have from pg. 93 in Lipster and Shiryayev [33],

$$R_{ij}(t) = [M_{ij}](t) = \sum_{s \le t} |\triangle M_{ij}(s)|^2, \tag{A.60}$$

and with Lemma (A.1.10),

$$R_{ij}(t) = \sum_{s \le t} |\triangle M_{ij}(s)|^2 = \sum_{s \le t} |\triangle M_{ij}(s)| \le \text{var}(M_{ij})(t), \tag{A.61}$$

thus,

$$R_{ij}(t) \le \text{var}(M_{ij})(t), \quad t \in [0, \infty). \tag{A.62}$$

Integrating $|\Gamma|$ on both sides of (A.62) gives

$$|\Gamma| \bullet R_{ij} \le |\Gamma| \bullet \text{var}(M_{ij}). \tag{A.63}$$

Now take some $\Gamma \in L^1_{loc}(\text{var}(M_{ij}))$. We know from the definition of $L^1_{loc}(\text{var}(M_{ij}))$ that $|\Gamma| \bullet \text{var}(M_{ij}) \in \mathcal{A}^+_{loc}$. Thus, from (A.63)

$$|\Gamma| \bullet R_{ij} \in \mathcal{A}^+_{loc}. \tag{A.64}$$

From (A.64) and (A.63),

$$\Gamma \in L^1_{loc}(R_{ij}) \implies L^1_{loc}(\text{var}(M_{ij})) \subseteq L^1_{loc}(R_{ij}). \tag{A.65}$$

To the establish the converse, namely,

$$L^1_{loc}(R_{ij}) \subseteq L^1_{loc}(\text{var}(M_{ij})), \tag{A.66}$$

take $\Gamma \in L^1_{loc}(R_{ij})$. From Proposition A.1.13, this also means that $\Gamma \in L^1_{loc}(\tilde{R}_{ij})$. Thus, we have that Γ is $\{\mathcal{F}_t\}$-predictably measurable and both

$$|\Gamma| \bullet R_{ij} \in \mathcal{A}^+_{loc} \text{ and } |\Gamma| \bullet \tilde{R}_{ij} \in \mathcal{A}^+_{loc} \tag{A.67}$$

are true. As a result, we have that

$$\begin{aligned} |\Gamma| \bullet R_{ij} + |\Gamma| \bullet \tilde{R}_{ij} &\in \mathcal{A}^+_{loc} \\ |\Gamma| \bullet (R_{ij} + \tilde{R}_{ij}) &\in \mathcal{A}^+_{loc}. \end{aligned} \tag{A.68}$$

But since we already know that the martingale M_{ij} is defined as

$$M_{ij} \triangleq R_{ij} - \tilde{R}_{ij}, \tag{A.69}$$

we have

$$\begin{aligned} \mathrm{var}(M_{ij}) &= \mathrm{var}(R_{ij} - \tilde{R}_{ij}) \\ &\leq \mathrm{var}(R_{ij}) + \mathrm{var}(\tilde{R}_{ij}). \end{aligned} \tag{A.70}$$

Since R_{ij} and $\tilde{R}_{ij}$ are non-decreasing,

$$\mathrm{var}(R_{ij}) = R_{ij} \quad \text{and} \quad \mathrm{var}(\tilde{R}_{ij}) = \tilde{R}_{ij}. \tag{A.71}$$

From (A.70) and (A.71),

$$\mathrm{var}(M_{ij}) \leq R_{ij} + \tilde{R}_{ij}, \tag{A.72}$$

thus,

$$|\Gamma| \bullet \mathrm{var}(M_{ij}) \leq |\Gamma| \bullet (R_{ij} + \tilde{R}_{ij}). \tag{A.73}$$

As a result, from (A.68) we have that

$$|\Gamma| \bullet \mathrm{var}(M_{ij}) \in \mathcal{A}^{+}_{loc} \tag{A.74}$$

and since $\Gamma \in \mathcal{P}^*$, we conclude that

$$\Gamma \in L^1_{loc}(\mathrm{var}(M_{ij})),$$

therefore,

$$L^1_{loc}(R_{ij}) \subseteq L^1_{loc}(\mathrm{var}(M_{ij})). \tag{A.75}$$

As a result, from (A.75) and (A.65)

$$L^1_{loc}(\mathrm{var}(M_{ij})) = L^1_{loc}(R_{ij}) \tag{A.76}$$

for $i, j \in S, i \neq j$. $\square$.

Proof of Theorem A.1.7. From Propositions A.1.11, A.1.12, A.1.13, and A.1.15,

$$L^{1/2}_{loc}(R_{ij}) = L^1_{loc}(R_{ij}) = L^1_{loc}(\tilde{R}_{ij}) = L^1_{loc}(\mathrm{var}(M_{ij})) \tag{A.77}$$

for $i, j \in S, i \neq j$. $\square$

Appendix B

Technical Proofs

In this appendix we give proofs of several results which have been stated in the course of this thesis. We locate these often rather technical proofs in an appendix because their inclusion in the main body of the thesis obscuring the main lines of development.

B.1 Canonical Martingales of the Regime-Switching Markov Chain Results

In this section we prove results regarding the canonical martingales of the regime-switching Markov chain; specifically, Lemma 4.1.14 and Lemma 4.1.15.

Proof of Lemma 4.1.14. For $i = j$ we have $M_{ii} = 0$ and thus

$$[M_{ii}, W_k](t) = 0, \quad t \in [0, T], \tag{B.1}$$

for all $i \in S$ and $k = 1, \dots, N$. For $i, j \in S, i \neq j$, we use the quadratic co-variation formula given in Theorem C.8.9:

$$[M_{ij}, W_k](t) = \langle M_{ij}^c, W_k^c \rangle(t) + \sum_{0 \le s \le t} \triangle M_{ij}(s) \triangle W_k(s), \quad t \in [0, T]. \tag{B.2}$$

Since $M_{ij}^c = 0$ from (4.12) and W_k is continuous for each $k = 1, \dots, N$, meaning $\triangle W_k = 0$, we have from (B.2),

$$[M_{ij}, W_k](t) = 0, \quad t \in [0, T], \tag{B.3}$$

for all $i, j \in S$ and $k = 1, \dots, N$. $\square$

Proof of Lemma 4.1.15. For $i = j$ or $i_1 = j_1$ we trivially have

$$[M_{ii}, M_{i_1 i_1}](t) = 0, \quad t \in [0, T], \tag{B.4}$$

for all $i, i_1 \in S$. For $i, j \in S, i \neq j$ and $i_1, j_1 \in S, i_1 \neq j_1$ we use the quadratic co-variation formula given in Theorem C.8.9:

$$[M_{ij}, M_{i_1 j_1}](t) = \sum_{0 \le s \le t} \triangle M_{ij}(s) \triangle M_{i_1 j_1}(s), \quad t \in [0, T]. \tag{B.5}$$

However, from (4.8), we have

$$\triangle M_{ij}(t) = \triangle R_{ij}(t), \quad t \in [0, T]. \tag{B.6}$$

Thus, putting (B.6) into (B.5) gives

$$\begin{aligned} [M_{ij}, M_{i_1 j_1}](t) &= \sum_{0 \le s \le t} I[\boldsymbol{\alpha}(s-) = i] I[\boldsymbol{\alpha}(s) = j] I[\boldsymbol{\alpha}(s-) = i_1] I[\boldsymbol{\alpha}(s) = j_1] \\ &= 0 \text{ a.s.}, \quad t \in [0, T], \end{aligned} \tag{B.7}$$

for all $(i, j) \neq (i_1, j_1)$. $\square$

B.2 Optimal Wealth Process Results in the Unconstrained Regime-Switching Market Model

In this section we prove Proposition 4.3.15:

Proof of Proposition 4.3.15(1). Set $t = 0$. From (4.67)

$$\hat{X}(0) = \operatorname*{ess\text{-}sup}_{\boldsymbol{\mu} \in \mathcal{H}} \frac{E\left[H_{\boldsymbol{\mu}}(T) B \,\middle|\, \mathcal{F}_0\right]}{H_{\boldsymbol{\mu}}(0)}, \tag{B.8}$$

where the σ-algebra $\mathcal{F}_0$ is given by

$$\mathcal{F}_0 \triangleq \{\emptyset, \Omega\} \vee \mathcal{N}(P). \tag{B.9}$$

Since $H_{\boldsymbol{\mu}}(0) = 1$ a.s. for each $\boldsymbol{\mu} \in \mathcal{H}$, we have

$$\hat{X}(0) = \operatorname*{ess-sup}_{\boldsymbol{\mu} \in \mathcal{H}} E\left[H_{\boldsymbol{\mu}}(T)B\right] = \hat{u} \text{ a.s.} \tag{B.10}$$

□

Proof of Proposition 4.3.15(2). Set $t = T$. From (4.67)

$$\hat{X}(T) = \operatorname*{ess-sup}_{\boldsymbol{\mu} \in \mathcal{H}} \frac{E\left[H_{\boldsymbol{\mu}}(T)B \,\middle|\, \mathcal{F}_T\right]}{H_{\boldsymbol{\mu}}(T)}. \tag{B.11}$$

Since B is $\mathcal{F}_T$-measurable and $H_{\boldsymbol{\mu}}(T) > 0$ a.s. and $\mathcal{F}_T$-measurable for all $\boldsymbol{\mu} \in \mathcal{H}$, we have

$$\hat{X}(T) = \operatorname*{ess-sup}_{\boldsymbol{\mu} \in \mathcal{H}} \frac{H_{\boldsymbol{\mu}}(T)B}{H_{\boldsymbol{\mu}}(T)} = B \text{ a.s.} \tag{B.12}$$

□

Proof of Proposition 4.3.15(3). Since $H_{\boldsymbol{\mu}}(t) > 0$ a.s. for all $t \in [0, T]$ and $\boldsymbol{\mu} \in \mathcal{H}$, and since $B > 0$ a.s., we have that $\hat{X}(t) \geq 0$ a.s. for all $t \in [0, T]$.
□

Proof of Proposition 4.3.15(4). For ease of notation, define

$$H_{\boldsymbol{\mu}}(s, t) \triangleq \frac{H_{\boldsymbol{\mu}}(t)}{H_{\boldsymbol{\mu}}(s)}, \quad \boldsymbol{\mu} \in \mathcal{H}, \tag{B.13}$$

for $0 \leq s \leq t \leq T$, and

$$J_{\boldsymbol{\mu}}(t) \triangleq E\left[H_{\boldsymbol{\mu}}(t, T)B \,\middle|\, \mathcal{F}_t\right], \quad \boldsymbol{\mu} \in \mathcal{H}, \tag{B.14}$$

for all $t \in [0, T]$. Thus, from (4.67) and (B.14)

$$\hat{X}(t) = \operatorname*{ess-sup}_{\boldsymbol{\mu} \in \mathcal{H}} J_{\boldsymbol{\mu}}(t), \quad t \in [0, T]. \tag{B.15}$$

Fix some $\bar{t} \in [0, T]$ and $\bar{\boldsymbol{\mu}} \in \mathcal{H}$. Put

$$\mathcal{D}_{\bar{t}, \bar{\boldsymbol{\mu}}} \triangleq \Big\{ \boldsymbol{\mu} \in \mathcal{H} \,\Big|\, \mu_{ij}(s, \omega) = \bar{\mu}_{ij}(s, \omega), 0 \leq s \leq \bar{t}, \omega \in \Omega, \; i, j \in S, i \neq j \Big\}. \tag{B.16}$$

$\mathcal{D}_{\bar{t},\bar{\boldsymbol{\mu}}}$ is the set of all processes $\boldsymbol{\mu} \in \mathcal{H}$ that match the fixed $\bar{\boldsymbol{\mu}} \in \mathcal{H}$ for all times between 0 and $\bar{t}$ and all $\omega \in \Omega$. Now, if we expand $H_{\boldsymbol{\mu}}(t_1, t_2)$ for any $0 \leq t_1 \leq t_2 \leq T$, we have by (4.44) and (4.43)

$$
\begin{aligned}
H_{\boldsymbol{\mu}}(t_1, t_2) &= \exp\left\{ - \int_{t_1}^{t_2} r(s)ds \right\} \frac{\mathcal{E}(-\boldsymbol{\theta} \bullet \mathbf{W})(t_2)}{\mathcal{E}(-\boldsymbol{\theta} \bullet \mathbf{W})(t_1)} \frac{\mathcal{E}(\boldsymbol{\mu} \bullet \mathbf{M})(t_2)}{\mathcal{E}(\boldsymbol{\mu} \bullet \mathbf{M})(t_1)} \\
&= \exp\left\{ - \int_{t_1}^{t_2} r(s)ds \right\} \exp\left\{ - \int_{t_1}^{t_2} \boldsymbol{\theta}(s)^\top d\mathbf{W}(s) - \frac{1}{2} \int_{t_1}^{t_2} ||\boldsymbol{\theta}(s)||^2 ds \right\} \\
&\cdot \exp\left\{ \sum_{\substack{i,j \in S \\ i \neq j}} \int_{t_1}^{t_2} \mu_{ij}(s) dM_{ij}(s) \right\} \prod_{t_1 < s \leq t_2} (1 + \triangle(\boldsymbol{\mu} \bullet \mathbf{M})(s)) \exp\left\{ - \triangle(\boldsymbol{\mu} \bullet \mathbf{M})(s) \right\},
\end{aligned}
\tag{B.17}
$$

for all $\boldsymbol{\mu} \in \mathcal{H}$. Viewing all terms of (B.17) we see that $H_{\boldsymbol{\mu}}(t_1, t_2)$ is determined by values of $\boldsymbol{\mu}(s)$ for only $s \in [t_1, t_2]$. As a result, by setting $t_1 = \bar{t}$ and $t_2 = T$, we have from (B.17) that $H_{\boldsymbol{\mu}}(\bar{t}, T)$ is determined by $\boldsymbol{\mu}(s)$ for only $s \in [\bar{t}, T]$. Thus, by (B.14), (B.15) and (B.16),

$$
\hat{X}(\bar{t}) = \operatorname*{ess\text{-}sup}_{\boldsymbol{\mu} \in \mathcal{H}} J_{\boldsymbol{\mu}}(\bar{t}) = \operatorname*{ess\text{-}sup}_{\boldsymbol{\mu} \in \mathcal{D}_{\bar{t},\bar{\boldsymbol{\mu}}}} E\left[H_{\boldsymbol{\mu}}(\bar{t}, T) B \,\middle|\, \mathcal{F}_{\bar{t}} \right]. \tag{B.18}
$$

For a fixed $\bar{\boldsymbol{\mu}} \in \mathcal{H}$ and $\bar{t} \in [0, T]$, define the set

$$
\mathcal{J}(\bar{t}, \bar{\boldsymbol{\mu}}) \triangleq \left\{ J_{\boldsymbol{\mu}}(\bar{t}) \middle| \boldsymbol{\mu} \in \mathcal{D}_{\bar{t},\bar{\boldsymbol{\mu}}} \right\}. \tag{B.19}
$$

The following lemma is essential as it allows us to use Theorem C.13.2 on the set $\mathcal{J}(\bar{t}, \bar{\boldsymbol{\mu}})$.

Lemma B.2.1. *The set $\mathcal{J}(\bar{t}, \bar{\boldsymbol{\mu}})$ is closed under pairwise maximization in the following sense: for arbitrary elements $J_1, J_2 \in \mathcal{J}(\bar{t}, \bar{\boldsymbol{\mu}})$, there is some $J \in \mathcal{J}(\bar{t}, \bar{\boldsymbol{\mu}})$ such that*

$$
J \triangleq J_1 \vee J_2 \ a.s. \tag{B.20}
$$

Proof. Fix $\boldsymbol{\mu}_1, \boldsymbol{\mu}_2 \in \mathcal{D}_{\bar{t},\bar{\boldsymbol{\mu}}}$. To remove possible confusion with notation, we write

$$
J_{\boldsymbol{\mu}_i,\bar{t}}(\omega) \triangleq J_{\boldsymbol{\mu}_i}(\bar{t}, \omega), \quad i = 1, 2, \tag{B.21}
$$

for all $\omega \in \Omega$. Define the set

$$
A \triangleq \left\{ \omega \in \Omega \middle| J_{\boldsymbol{\mu}_1,\bar{t}}(\omega) \geq J_{\boldsymbol{\mu}_2,\bar{t}}(\omega) \right\}. \tag{B.22}
$$

Since $J_{\boldsymbol{\mu}_1}(t)$ and $J_{\boldsymbol{\mu}_2}(t)$ are $\{\mathcal{F}_t\}$-adapted, $A \in \mathcal{F}_{\bar{t}}$. Define

$$\hat{\boldsymbol{\mu}}(s,\omega) \triangleq \boldsymbol{\mu}_1(s,\omega) I_A(\omega) + \boldsymbol{\mu}_2(s,\omega) I_{A^c}(\omega), \quad (s,\omega) \in [0,T] \times \Omega. \tag{B.23}$$

Since $\boldsymbol{\mu}_1, \boldsymbol{\mu}_2 \in \mathcal{D}_{\bar{t},\bar{\boldsymbol{\mu}}}$, we have

$$\hat{\boldsymbol{\mu}}(s,\omega) = \bar{\boldsymbol{\mu}}(s,\omega), \quad (s,\omega) \in [0,\bar{t}] \times \Omega. \tag{B.24}$$

We then have from (B.23) and (B.24),

$$\hat{\boldsymbol{\mu}}(s,\omega) = \bar{\boldsymbol{\mu}}(s,\omega) I_{[0,\bar{t}]}(s) + \boldsymbol{\mu}_1(s,\omega) I_{(\bar{t},T]\times A}(s,\omega) + \boldsymbol{\mu}_2(s,\omega) I_{(\bar{t},T]\times A^c}(s,\omega). \tag{B.25}$$

for all $(s,\omega) \in [0,T]\times\Omega$. From (B.25) we see that each term is $\{\mathcal{F}_t\}$-predictably measurable, and thus, $\hat{\boldsymbol{\mu}}$ is $\{\mathcal{F}_t\}$-predictably measurable. Moreover, $\hat{\mu}_{ij}(t)\triangle M_{ij}(t) > -1$ a.s. for all $t \in [0,T]$ and $\hat{\mu}_{ij} \in L^{1/2}_{loc}(R_{ij})$ (see Definition A.1.1) for all $i, j \in S, i \neq j$ since $\bar{\boldsymbol{\mu}}, \boldsymbol{\mu}_1, \boldsymbol{\mu}_2 \in \mathcal{H}$. Thus, $\hat{\boldsymbol{\mu}} \in \mathcal{H}$, and with (B.24) we have that $\hat{\boldsymbol{\mu}} \in \mathcal{D}_{\bar{t},\bar{\boldsymbol{\mu}}}$. Therefore,

$$J_{\hat{\boldsymbol{\mu}},\bar{t}} = J_{\hat{\boldsymbol{\mu}}}(\bar{t}) \triangleq E\left[H_{\hat{\boldsymbol{\mu}}}(\bar{t},T)B \,\middle|\, \mathcal{F}_{\bar{t}}\right], \tag{B.26}$$

is a member of $\mathcal{J}(\bar{t},\bar{\boldsymbol{\mu}})$ (i.e. $J_{\hat{\boldsymbol{\mu}},\bar{t}} \in \mathcal{J}(\bar{t},\bar{\boldsymbol{\mu}})$).

Using (B.23), and the fact that $A \in \mathcal{F}_{\bar{t}}$ (as noted above), we get

$$H_{\hat{\boldsymbol{\mu}}}(\bar{t},T,\omega) = I_A(\omega) H_{\boldsymbol{\mu}_1}(\bar{t},T,\omega) + I_{A^c}(\omega) H_{\boldsymbol{\mu}_2}(\bar{t},T,\omega), \quad P-a.a.\ \omega \in \Omega. \tag{B.27}$$

Using (B.26) and (B.27) we can now write

$$J_{\hat{\boldsymbol{\mu}},\bar{t}}(\omega) = I_A(\omega) J_{\boldsymbol{\mu}_1,\bar{t}}(\omega) + I_{A^c}(\omega) J_{\boldsymbol{\mu}_2,\bar{t}}(\omega), \quad P-a.a.\ \omega \in \Omega. \tag{B.28}$$

As a result, we have

$$J_{\boldsymbol{\mu}_1,\bar{t}}(\omega) \geq J_{\boldsymbol{\mu}_2,\bar{t}}(\omega) \text{ and } J_{\hat{\boldsymbol{\mu}},\bar{t}}(\omega) = J_{\boldsymbol{\mu}_1,\bar{t}}(\omega), \quad P-a.a.\ \omega \in A. \tag{B.29}$$

Similarly, we have

$$J_{\boldsymbol{\mu}_2,\bar{t}}(\omega) \geq J_{\boldsymbol{\mu}_1,\bar{t}}(\omega) \text{ and } J_{\hat{\boldsymbol{\mu}},\bar{t}}(\omega) = J_{\boldsymbol{\mu}_2,\bar{t}}(\omega), \quad P-a.a.\ \omega \in A^c. \tag{B.30}$$

Therefore,

$$J_{\hat{\boldsymbol{\mu}},\bar{t}}(\omega) = J_{\boldsymbol{\mu}_1,\bar{t}}(\omega) \vee J_{\boldsymbol{\mu}_2,\bar{t}}(\omega), \quad P-a.a.\ \omega \in \Omega. \tag{B.31}$$

Since $J_{\hat{\boldsymbol{\mu}},\bar{t}} \in \mathcal{J}(\bar{t},\bar{\boldsymbol{\mu}})$ from (B.26), the set $\mathcal{J}(\bar{t},\bar{\boldsymbol{\mu}})$ is closed under pairwise maximization.
□

Continuing the proof of Proposition 4.3.15(4), since

$$\hat{X}(\bar{t}) = \operatorname*{ess\text{-}sup}_{\boldsymbol{\mu}\in\mathcal{D}_{\bar{t},\bar{\boldsymbol{\mu}}}} J_{\boldsymbol{\mu}}(\bar{t}), \tag{B.32}$$

we have,

$$\hat{X}(\bar{t}) = \operatorname{ess\text{-}sup} \mathcal{J}(\bar{t}, \bar{\boldsymbol{\mu}}). \tag{B.33}$$

From Theorem C.13.2, due to the pairwise maximization property of $\mathcal{J}(\bar{t}, \bar{\boldsymbol{\mu}})$, there exists a sequence $\{\zeta_n\} \in \mathcal{J}(\bar{t}, \bar{\boldsymbol{\mu}})$ for $n = 1, 2, \ldots$ such that $\zeta_n \leq \zeta_{n+1} \leq \ldots$ a.s. and

$$\begin{aligned}\lim_{n\to\infty} \zeta_n &= \operatorname{ess\text{-}sup} \mathcal{J}(\bar{t}, \bar{\boldsymbol{\mu}}) \text{ a.s.}\\ &= \hat{X}(\bar{t}) \text{ a.s.}\end{aligned} \tag{B.34}$$

This means there exists some sequence $\{\boldsymbol{\mu}_n\} \in \mathcal{D}_{\bar{t},\bar{\boldsymbol{\mu}}}$, $n = 1, 2, \ldots$, such that $J_{\boldsymbol{\mu}_n} \leq J_{\boldsymbol{\mu}_{n+1}} \leq \ldots$ a.s. and from (B.34),

$$\lim_{n\to\infty} J_{\boldsymbol{\mu}_n}(\bar{t}) = \hat{X}(\bar{t}) \text{ a.s.} \tag{B.35}$$

Now fix some s such that $0 \leq s \leq \bar{t} \leq T$, then by (B.17), and the reasoning that follows,

$$H_{\boldsymbol{\mu}}(s, \bar{t}) = H_{\bar{\boldsymbol{\mu}}}(s, \bar{t}) \text{ a.s.}, \quad \boldsymbol{\mu} \in \mathcal{D}_{\bar{t},\bar{\boldsymbol{\mu}}}. \tag{B.36}$$

Since $\{\boldsymbol{\mu}_n\} \in \mathcal{D}_{\bar{t},\bar{\boldsymbol{\mu}}}$, from (B.36)

$$H_{\boldsymbol{\mu}_n}(s, \bar{t}) = H_{\bar{\boldsymbol{\mu}}}(s, \bar{t}) \text{ a.s.}, \quad n = 1, 2, \ldots \tag{B.37}$$

Using (B.35), we can write

$$E\left[H_{\bar{\boldsymbol{\mu}}}(s, \bar{t})\hat{X}(\bar{t}) \,\middle|\, \mathcal{F}_s\right] = E\left[H_{\bar{\boldsymbol{\mu}}}(s, t) \lim_{n\to\infty} J_{\boldsymbol{\mu}_n}(\bar{t}) \,\middle|\, \mathcal{F}_s\right] \text{ a.s.}, \tag{B.38}$$

and using the monotone convergence theorem,

$$\begin{aligned}E\left[H_{\bar{\boldsymbol{\mu}}}(s, \bar{t})\hat{X}(\bar{t}) \,\middle|\, \mathcal{F}_s\right] &= \lim_{n\to\infty} E\left[H_{\bar{\boldsymbol{\mu}}}(s, t) J_{\boldsymbol{\mu}_n}(\bar{t}) \,\middle|\, \mathcal{F}_s\right]\\ &= \lim_{n\to\infty} E\left[E\left[H_{\boldsymbol{\mu}_n}(s, \bar{t}) H_{\boldsymbol{\mu}_n}(\bar{t}, T) B \,\middle|\, \mathcal{F}_{\bar{t}}\right] \,\middle|\, \mathcal{F}_s\right]\\ &= \lim_{n\to\infty} E\left[H_{\boldsymbol{\mu}_n}(s, T) B \,\middle|\, \mathcal{F}_s\right]\\ &= \lim_{n\to\infty} J_{\boldsymbol{\mu}_n}(s) \text{ a.s.}\end{aligned} \tag{B.39}$$

Therefore, by the fact that $\boldsymbol{\mu}_n \in \mathcal{D}_{\bar{t},\bar{\boldsymbol{\mu}}} \subset \mathcal{H}$ and $s \leq \bar{t}$

$$\begin{aligned} E\left[H_{\bar{\boldsymbol{\mu}}}(s,\bar{t})\hat{X}(\bar{t}) \,\middle|\, \mathcal{F}_s\right] &= \lim_{n\to\infty} J_{\boldsymbol{\mu}_n}(s) \\ &\leq \operatorname*{ess\text{-}sup}_{\boldsymbol{\mu}\in\mathcal{H}} J_{\boldsymbol{\mu}}(s) \\ &= \hat{X}(s) \text{ a.s.} \end{aligned} \tag{B.40}$$

Taking $s = 0$ in (B.40), from Proposition 4.3.15(1),

$$E\left[H_{\bar{\boldsymbol{\mu}}}(\bar{t})\hat{X}(\bar{t})\right] \leq E[\hat{X}(0)] = \hat{u} < \infty. \tag{B.41}$$

From the strict almost sure positivity of $H_{\bar{\boldsymbol{\mu}}}(\bar{t})$ and (B.41), we have $\hat{X}(\bar{t}) < \infty$ a.s. Now,

$$\begin{aligned} \frac{1}{H_{\bar{\boldsymbol{\mu}}}(s)} E\left[H_{\bar{\boldsymbol{\mu}}}(\bar{t})\hat{X}(\bar{t}) \,\middle|\, \mathcal{F}_s\right] &= E\left[H_{\bar{\boldsymbol{\mu}}}(s,\bar{t})\hat{X}(\bar{t}) \,\middle|\, \mathcal{F}_s\right] \\ &\leq \hat{X}(s) \text{ a.s.} \end{aligned} \tag{B.42}$$

for all $0 \leq s \leq \bar{t} \leq T$. Thus,

$$E\left[H_{\bar{\boldsymbol{\mu}}}(\bar{t})\hat{X}(\bar{t}) \,\middle|\, \mathcal{F}_s\right] \leq H_{\bar{\boldsymbol{\mu}}}(s)\hat{X}(s) \text{ a.s.} \tag{B.43}$$

for all $0 \leq s \leq \bar{t} \leq T$. Since $\bar{t}$ and $\bar{\boldsymbol{\mu}}$ were chosen arbitrarily, we have from (B.43) that the supermartingale property holds for all $\boldsymbol{\mu} \in \mathcal{H}$. Thus,

$$H_{\boldsymbol{\mu}}\hat{X} \in \mathcal{SPM}(\{\mathcal{F}_t\}, P), \quad \boldsymbol{\mu} \in \mathcal{H}. \tag{B.44}$$

□

Proof of Proposition 4.3.15(5). Again define

$$J_{\boldsymbol{\mu}}(t) \triangleq E\left[H_{\boldsymbol{\mu}}(t,T)B \,\middle|\, \mathcal{F}_t\right], \quad t \in [0,T], \boldsymbol{\mu} \in \mathcal{H}, \tag{B.45}$$

where,

$$H_{\boldsymbol{\mu}}(s,t) \triangleq \frac{H_{\boldsymbol{\mu}}(t)}{H_{\boldsymbol{\mu}}(s)}, \quad 0 \leq s \leq t \leq T, \boldsymbol{\mu} \in \mathcal{H}, \tag{B.46}$$

so that

$$\hat{X}(t) = \operatorname*{ess\text{-}sup}_{\boldsymbol{\mu}\in\mathcal{H}} J_{\boldsymbol{\mu}}(t), \quad t \in [0,T]. \tag{B.47}$$

Fix some s, t such that $0 \leq s \leq t \leq T$, then

$$\begin{aligned} E\left[H_{\boldsymbol{\mu}}(s,t)J_{\boldsymbol{\mu}}(t) \,\middle|\, \mathcal{F}_s\right] &= E\left[H_{\boldsymbol{\mu}}(s,t)E\left[H_{\boldsymbol{\mu}}(t,T)B \,\middle|\, \mathcal{F}_t\right] \,\middle|\, \mathcal{F}_s\right] \\ &= E\left[H_{\boldsymbol{\mu}}(s,T)B \,\middle|\, \mathcal{F}_s\right] \\ &= J_{\boldsymbol{\mu}}(s) \text{ a.s.}, \quad \boldsymbol{\mu} \in \mathcal{H} \end{aligned} \tag{B.48}$$

Using (B.47) we have

$$J_{\boldsymbol{\mu}}(t) \leq \hat{X}(t) \text{ a.s.}, \quad \boldsymbol{\mu} \in \mathcal{H}. \tag{B.49}$$

Furthermore, from (B.48) and (B.47)

$$\hat{X}(s) = \operatorname*{ess\text{-}sup}_{\boldsymbol{\mu} \in \mathcal{H}} E\left[H_{\boldsymbol{\mu}}(s,t)J_{\boldsymbol{\mu}}(t) \,\middle|\, \mathcal{F}_s\right] \text{ a.s.} \tag{B.50}$$

From (B.49) and the non-negativity of $H_{\boldsymbol{\mu}}$,

$$H_{\boldsymbol{\mu}}(s,t)J_{\boldsymbol{\mu}}(t) \leq H_{\boldsymbol{\mu}}(s,t)\hat{X}(t) \text{ a.s.}, \quad \boldsymbol{\mu} \in \mathcal{H}. \tag{B.51}$$

Taking conditional expectations and essential suprema of both sides of (B.51) (which exists by Theorem C.13.2) and then using (B.50) yields

$$\hat{X}(s) \leq \operatorname*{ess\text{-}sup}_{\boldsymbol{\mu} \in \mathcal{H}} E\left[H_{\boldsymbol{\mu}}(s,t)\hat{X}(t) \,\middle|\, \mathcal{F}_s\right] \text{ a.s.} \tag{B.52}$$

But using the supermartingale property of $H_{\boldsymbol{\mu}}\hat{X}$ from Proposition 4.3.15(4), we have

$$\hat{X}(s) \geq \operatorname*{ess\text{-}sup}_{\boldsymbol{\mu} \in \mathcal{H}} E\left[H_{\boldsymbol{\mu}}(s,t)\hat{X}(t) \,\middle|\, \mathcal{F}_s\right] \text{ a.s.} \tag{B.53}$$

Combining (B.53) and (B.52), we conclude

$$\begin{aligned} \hat{X}(s) &= \operatorname*{ess\text{-}sup}_{\boldsymbol{\mu} \in \mathcal{H}} E\left[H_{\boldsymbol{\mu}}(s,t)\hat{X}(t) \,\middle|\, \mathcal{F}_s\right] \\ &= \operatorname*{ess\text{-}sup}_{\boldsymbol{\mu} \in \mathcal{H}} \frac{E[H_{\boldsymbol{\mu}}(t)\hat{X}(t)|\mathcal{F}_s]}{H_{\boldsymbol{\mu}}(s)} \text{ a.s.} \end{aligned} \tag{B.54}$$

□

Proof of Proposition 4.3.15(6). Fix $\bar{\boldsymbol{\mu}} \in \mathcal{H}$. Define

$$z_{\bar{\boldsymbol{\mu}}}(t) \triangleq \begin{cases} H_{\bar{\boldsymbol{\mu}}}(t)\hat{X}(t), & t \in [0,T] \\ H_{\bar{\boldsymbol{\mu}}}(T)\hat{X}(T), & t \in (T,\infty) \end{cases}, \tag{B.55}$$

i.e.

$$z_{\bar{\boldsymbol{\mu}}}(t) \triangleq H_{\bar{\boldsymbol{\mu}}}(t \wedge T)\hat{X}(t \wedge T), \quad t \in [0, \infty). \tag{B.56}$$

Define $\mathcal{F}_t \triangleq \mathcal{F}_T$ for all $t \in [T, \infty)$. From Proposition 4.3.15(4), we have that

$$z_{\bar{\boldsymbol{\mu}}} \in \mathcal{SPM}(\{\mathcal{F}_t\}, P). \tag{B.57}$$

From §1.4 of Chung [6], there exists a set $\Omega^*(\bar{\boldsymbol{\mu}}) \in \mathcal{F}_\infty \equiv \mathcal{F}_T$ such that $P(\Omega^*(\bar{\boldsymbol{\mu}})) = 1$ and

$$z^+_{\bar{\boldsymbol{\mu}}}(t, \omega) \triangleq \lim_{\substack{s \in \mathbb{Q} \\ s \downarrow\downarrow t}} z_{\bar{\boldsymbol{\mu}}}(s, \omega) \tag{B.58}$$

exists in $\mathbb{R}$ for each $t \in [0, \infty)$ and $\omega \in \Omega^*(\bar{\boldsymbol{\mu}})$.

Remark B.2.2. Equation (B.58) says that for each $\omega \in \Omega^*(\bar{\boldsymbol{\mu}})$ and $t \in [0, \infty)$, there exists some $z^+_{\bar{\boldsymbol{\mu}}}(t, \omega) \in \mathbb{R}$ with the following property: for each $\epsilon > 0$ there is some $\delta = \delta(\epsilon, t, \omega) > 0$ such that

$$|z^+_{\bar{\boldsymbol{\mu}}}(t, \omega) - z_{\bar{\boldsymbol{\mu}}}(s, \omega)| < \epsilon \tag{B.59}$$

for all $s \in \mathbb{Q} \cap (t, t + \delta)$.

We have that $t \mapsto z^+_{\bar{\boldsymbol{\mu}}}(t, \omega) : [0, T) \mapsto \mathbb{R}$ is càdlàg for each $\omega \in \Omega^*(\bar{\boldsymbol{\mu}})$. Since $\hat{X}(T) = B$ a.s., we know

$$z^+_{\bar{\boldsymbol{\mu}}}(t, \omega) = H_{\bar{\boldsymbol{\mu}}}(T, \omega)B(\omega), \quad t \in [T, \infty) \quad \omega \in \Omega^*(\bar{\boldsymbol{\mu}}). \tag{B.60}$$

Put

$$z^+_{\bar{\boldsymbol{\mu}}}(t, \omega) \triangleq 0, \quad t \in [0, \infty) \quad \omega \notin \Omega^*(\bar{\boldsymbol{\mu}}) \tag{B.61}$$

Thus, due (B.60) and (B.61), $t \mapsto z^+_{\bar{\boldsymbol{\mu}}}(t, \omega) : [0, \infty) \mapsto \mathbb{R}$ is càdlàg for each $\omega \in \Omega$ (see II(65.1) in Rogers and Williams [42]). Put

$$\mathcal{G}_t \triangleq \sigma\{\mathcal{F}_{t+}, \mathcal{N}(P)\}. \tag{B.62}$$

Since $z_{\bar{\boldsymbol{\mu}}}$ is $\{\mathcal{F}_t\}$-adapted, from Lemma II (66.1) of Rogers and Williams [43], $z^+_{\bar{\boldsymbol{\mu}}}$ is $\{\mathcal{G}_t\}$-adapted. But since $\mathcal{N}(P) \in \mathcal{F}_t$ and $\mathcal{F}_{t+} = \mathcal{F}_t$ (from Remark 4.1.2), we therefore have $\mathcal{G}_t = \mathcal{F}_t$. Thus, $z^+_{\bar{\boldsymbol{\mu}}}$ is $\{\mathcal{F}_t\}$-adapted. From Lemma II (66.2) of Rogers and Williams [43], and (B.57)

$$z^+_{\bar{\boldsymbol{\mu}}} \in \mathcal{SPM}(\{\mathcal{G}_t\}, P), \tag{B.63}$$

and therefore, $z^+_{\bar{\boldsymbol{\mu}}} \in \mathcal{SPM}(\{\mathcal{F}_t\}, P)$ and is càdlàg.

From (B.55)

$$\hat{X}(t) = \frac{1}{H_{\bar{\boldsymbol{\mu}}}(t)} z_{\bar{\boldsymbol{\mu}}}(t) \text{ a.s.}, \quad t \in [0, T]. \tag{B.64}$$

Put

$$\hat{X}^+_{\bar{\boldsymbol{\mu}}}(t) \triangleq \frac{1}{H_{\bar{\boldsymbol{\mu}}}(t)} z^+_{\bar{\boldsymbol{\mu}}}(t), \quad t \in [0,T]. \tag{B.65}$$

Now $t \mapsto \frac{1}{H_{\bar{\boldsymbol{\mu}}}(t,\omega)} : [0,T] \mapsto (0,\infty)$ is càdlàg for each $\omega \in \Omega$. Since $z^+_{\bar{\boldsymbol{\mu}}}(\cdot,\omega)$ is also càdlàg for each $\omega \in \Omega$, we have that

$$t \mapsto \hat{X}^+_{\bar{\boldsymbol{\mu}}}(t,\omega) : [0,T] \mapsto \mathbb{R} \text{ is càdlàg for each } \omega \in \Omega. \tag{B.66}$$

Moreover, $\frac{1}{H_{\bar{\boldsymbol{\mu}}}}$ is $\{\mathcal{F}_t\}$-adapted and $z^+_{\bar{\boldsymbol{\mu}}}$ is $\{\mathcal{F}_t\}$-adapted meaning that $\hat{X}^+_{\bar{\boldsymbol{\mu}}}$ is $\{\mathcal{F}_t\}$-adapted. As a result,

$$H_{\bar{\boldsymbol{\mu}}} \hat{X}^+_{\bar{\boldsymbol{\mu}}} \in \mathcal{SPM}(\{\mathcal{F}_t\}, P) \tag{B.67}$$

and is a càdlàg process where

$$\hat{X}^+_{\bar{\boldsymbol{\mu}}}(T) = B \text{ a.s.} \tag{B.68}$$

From the supermartingale property of $H_{\bar{\boldsymbol{\mu}}} \hat{X}^+_{\bar{\boldsymbol{\mu}}}$ (B.67),

$$\begin{aligned} H_{\bar{\boldsymbol{\mu}}}(t)\hat{X}^+_{\bar{\boldsymbol{\mu}}}(t) &\geq E\left[H_{\bar{\boldsymbol{\mu}}}(T)B \,\middle|\, \mathcal{F}_t\right] \text{ a.s.} \\ \hat{X}^+_{\bar{\boldsymbol{\mu}}}(t) &\geq E\left[H_{\bar{\boldsymbol{\mu}}}(t,T)B \,\middle|\, \mathcal{F}_t\right] \text{ a.s.}, \quad t \in [0,T]. \end{aligned} \tag{B.69}$$

From (4.67)

$$\hat{X}(t) \geq \frac{E\left[H_{\boldsymbol{\mu}}(T)B \,\middle|\, \mathcal{F}_t\right]}{H_{\boldsymbol{\mu}}(t)} \text{ a.s.}, \quad t \in [0,T], \boldsymbol{\mu} \in \mathcal{H}. \tag{B.70}$$

Then

$$\begin{aligned} z_{\bar{\boldsymbol{\mu}}}(t) &\triangleq H_{\bar{\boldsymbol{\mu}}}(t)\hat{X}(t) \\ &\geq H_{\bar{\boldsymbol{\mu}}}(t)\frac{E\left[H_{\boldsymbol{\mu}}(T)B \,\middle|\, \mathcal{F}_t\right]}{H_{\boldsymbol{\mu}}(t)} \text{ a.s.}, \quad t \in [0,T], \boldsymbol{\mu} \in \mathcal{H}. \end{aligned} \tag{B.71}$$

Now fix a sequence $\{t_n\} \in \mathbb{Q}$ for $n = 1, 2, \ldots$ such that $t < \ldots < t_{n+1} < t_n < \ldots$ and $\lim_{n\to\infty} t_n = t$. Then, by corollary II(2.4) in Revuz and Yor [39]

$$\begin{aligned} \lim_{\substack{s\in\mathbb{Q} \\ s\downarrow\downarrow t}} E\left[BH_{\boldsymbol{\mu}}(T) \,\middle|\, \mathcal{F}_s\right] &= \lim_{n\to\infty} E\left[BH_{\boldsymbol{\mu}}(T) \,\middle|\, \mathcal{F}_{t_n}\right] \\ &= E\left[BH_{\boldsymbol{\mu}}(T) \,\middle|\, \mathcal{F}_{t+}\right] \\ &= E\left[BH_{\boldsymbol{\mu}}(T) \,\middle|\, \mathcal{F}_t\right] \text{ a.s.}, \quad t \in [0,T), \boldsymbol{\mu} \in \mathcal{H}. \end{aligned} \tag{B.72}$$

Moreover, since

$$t \mapsto \frac{H_{\bar{\boldsymbol{\mu}}}(t,\omega)}{H_{\boldsymbol{\mu}}(t,\omega)} : [0,T) \mapsto (0,\infty) \tag{B.73}$$

is càdlàg on $[0,T)$ for each $\omega \in \Omega$ and $\boldsymbol{\mu} \in \mathcal{H}$, then

$$\lim_{\substack{s \in \mathbb{Q} \\ s \downarrow\downarrow t}} \frac{H_{\bar{\boldsymbol{\mu}}}(s)}{H_{\boldsymbol{\mu}}(s)} = \frac{H_{\bar{\boldsymbol{\mu}}}(t)}{H_{\boldsymbol{\mu}}(t)}, \quad t \in [0,T), \boldsymbol{\mu} \in \mathcal{H}. \tag{B.74}$$

As a result, from (B.71), (B.72), and (B.74)

$$\begin{aligned} z^+_{\bar{\boldsymbol{\mu}}}(t) &= \lim_{n \to \infty} z_{\bar{\boldsymbol{\mu}}}(t_n) \\ &\geq \lim_{n \to \infty} \frac{H_{\bar{\boldsymbol{\mu}}}(t_n)}{H_{\boldsymbol{\mu}}(t_n)} \lim_{n \to \infty} E\left[B H_{\boldsymbol{\mu}}(T) \,\middle|\, \mathcal{F}_{t_n} \right] \\ &= \frac{H_{\bar{\boldsymbol{\mu}}}(t)}{H_{\boldsymbol{\mu}}(t)} E\left[B H_{\boldsymbol{\mu}}(T) \,\middle|\, \mathcal{F}_t \right] \\ &= H_{\bar{\boldsymbol{\mu}}}(t) E\left[B H_{\boldsymbol{\mu}}(t,T) \,\middle|\, \mathcal{F}_t \right] \text{ a.s.}, \quad t \in [0,T), \boldsymbol{\mu} \in \mathcal{H}. \end{aligned} \tag{B.75}$$

Thus

$$\hat{X}^+_{\bar{\boldsymbol{\mu}}}(t) \triangleq \frac{z^+_{\bar{\boldsymbol{\mu}}}(t)}{H_{\bar{\boldsymbol{\mu}}}(t)} \geq E\left[B H_{\boldsymbol{\mu}}(t,T) \,\middle|\, \mathcal{F}_t \right] \text{ a.s.}, \quad t \in [0,T), \boldsymbol{\mu} \in \mathcal{H}. \tag{B.76}$$

From the definition of $\hat{X}$ (see (4.67)) and (B.76), we conclude

$$\hat{X}^+_{\bar{\boldsymbol{\mu}}}(t) \geq \hat{X}(t) \text{ a.s.}, \quad t \in [0,T). \tag{B.77}$$

To show the opposite inequality, fix $t \in [0,T)$ and $\boldsymbol{\mu} \in \mathcal{H}$ and sequence $\{t_n\} \in \mathbb{Q}$ such that $t < \ldots t_{n+1} < t_n < \ldots \leq T$ and $\lim_{n \to \infty} t_n = t$. Since $\hat{X}^+_{\bar{\boldsymbol{\mu}}}$ is càdlàg by (B.66),

$$\hat{X}^+_{\bar{\boldsymbol{\mu}}}(t) = \lim_{n \to \infty} \hat{X}^+_{\bar{\boldsymbol{\mu}}}(t_n) \text{ a.s.} \tag{B.78}$$

Similarly, since $H_{\boldsymbol{\mu}}$ is càdlàg

$$H_{\boldsymbol{\mu}}(t) = \lim_{n \to \infty} H_{\boldsymbol{\mu}}(t_n) \text{ a.s.} \tag{B.79}$$

Thus,

$$\lim_{n \to \infty} H_{\boldsymbol{\mu}}(t, t_n) = 1 \text{ a.s.} \tag{B.80}$$

From the non-negativity of $H_{\boldsymbol{\mu}}$ and from (B.64) we can write

$$H_{\boldsymbol{\mu}}(t,t_n) \hat{X}(t_n) = H_{\boldsymbol{\mu}}(t,t_n) \frac{z_{\bar{\boldsymbol{\mu}}}(t_n)}{H_{\bar{\boldsymbol{\mu}}}(t_n)} \text{ a.s.} \tag{B.81}$$

It is easy to see from (B.58) that

$$\lim_{n\to\infty} z_{\bar{\boldsymbol{\mu}}}(t_n) = z^+_{\bar{\boldsymbol{\mu}}}(t) \text{ a.s.} \tag{B.82}$$

and from the càdlàg property of $\frac{1}{H_{\bar{\boldsymbol{\mu}}}}$

$$\lim_{n\to\infty} \frac{1}{H_{\bar{\boldsymbol{\mu}}}(t_n)} = \frac{1}{H_{\bar{\boldsymbol{\mu}}}(t)} \text{ a.s.} \tag{B.83}$$

Now, from (B.81), (B.80), (B.82) and (B.83),

$$\lim_{n\to\infty} H_{\boldsymbol{\mu}}(t,t_n)\hat{X}(t_n) = \frac{z^+_{\bar{\boldsymbol{\mu}}}(t)}{H_{\bar{\boldsymbol{\mu}}}(t)} = \hat{X}^+_{\bar{\boldsymbol{\mu}}}(t) \text{ a.s.} \tag{B.84}$$

Since $\hat{X}^+_{\bar{\boldsymbol{\mu}}}$ is $\{\mathcal{F}_t\}$-adapted, from (B.84) and Fatou's Lemma

$$\begin{aligned}\hat{X}^+_{\bar{\boldsymbol{\mu}}}(t) &= E\left[\hat{X}^+_{\bar{\boldsymbol{\mu}}}(t) \,\middle|\, \mathcal{F}_t\right] \\ &= E\left[\lim_{n\to\infty} H_{\boldsymbol{\mu}}(t,t_n)\hat{X}(t_n) \,\middle|\, \mathcal{F}_t\right] \\ &\le \lim_{n\to\infty} E\left[H_{\boldsymbol{\mu}}(t,t_n)\hat{X}(t_n) \,\middle|\, \mathcal{F}_t\right] \text{ a.s.}\end{aligned} \tag{B.85}$$

Now using the supermartingale property of $H_{\boldsymbol{\mu}}\hat{X}$ from Proposition 4.3.15(4), and the fact that $t < t_n$ for all $n = 1, 2, \ldots,$

$$E\left[H_{\boldsymbol{\mu}}(t_n)\hat{X}(t_n) \,\middle|\, \mathcal{F}_t\right] \le H_{\boldsymbol{\mu}}(t)\hat{X}(t) \text{ a.s.,} \tag{B.86}$$

$$E\left[H_{\boldsymbol{\mu}}(t,t_n)\hat{X}(t_n) \,\middle|\, \mathcal{F}_t\right] \le \hat{X}(t) \text{ a.s.} \tag{B.87}$$

Putting (B.87) into (B.85) gives

$$\hat{X}^+_{\bar{\boldsymbol{\mu}}}(t) \le \hat{X}(t) \text{ a.s.,} \quad t \in [0,T). \tag{B.88}$$

Therefore, from (B.77) and (B.88), we have

$$\hat{X}(t) = \hat{X}^+_{\bar{\boldsymbol{\mu}}}(t) \text{ a.s.,} \quad t \in [0,T). \tag{B.89}$$

Since $\hat{X}(T) = \hat{X}^+_{\bar{\boldsymbol{\mu}}}(T) = B$ a.s. we have by (B.89)

$$\hat{X}(t) = \hat{X}^+_{\bar{\boldsymbol{\mu}}}(t) \text{ a.s.,} \quad t \in [0,T], \tag{B.90}$$

and therefore, $\{\hat{X}(t), t \in [0,T]\}$ has a càdlàg modification.
□

B.3 Optimal Cumulative Consumption in the Unconstrained Regime-Switching Market Model

Proof of Proposition 4.3.21.

By Proposition C.8.6 we have that the continuous local martingale parts of (4.89) and (4.91) are indistinguishable. Equating the continuous local martingale parts of (4.89) and (4.91) yields

$$\int_0^t \left(\frac{\mathbf{\Psi}_0(s)^\top}{H_0(s)} + \hat{X}(s)\boldsymbol{\theta}(s)^\top \right) d\mathbf{W}(s) = \int_0^t \left(\frac{\mathbf{\Psi}_{\boldsymbol{\mu}}(s)^\top}{H_{\boldsymbol{\mu}}(s)} + \hat{X}(s)\boldsymbol{\theta}(s)^\top \right) d\mathbf{W}(s), \quad t \in [0,T], \tag{B.91}$$

for all $\boldsymbol{\mu} \in \mathcal{H}$, and therefore,

$$\frac{\mathbf{\Psi}_0(t)}{H_0(t)} = \frac{\mathbf{\Psi}_{\boldsymbol{\mu}}(t)}{H_{\boldsymbol{\mu}}(t)} \quad \lambda \otimes P \text{ a.e. on } [0,T] \times \Omega, \quad \boldsymbol{\mu} \in \mathcal{H}. \tag{B.92}$$

Combining (4.89) and (4.91) with (B.91) and (B.92) results in

$$\begin{aligned} 0 = &\sum_{\substack{i,j\in S\\ i\neq j}} \int_0^t \left(\frac{\Gamma_{ij}^{\boldsymbol{\mu}}(s)}{H_{\boldsymbol{\mu}}(s-)} - \frac{\Gamma_{ij}^{0}(s)}{H_0(s-)} - \hat{X}(s-)\mu_{ij}(s) \right) dM_{ij}(s) \\ &+ \sum_{\substack{i,j\in S\\ i\neq j}} \int_0^t \left(\hat{X}(s-)\mu_{ij}(s) - \frac{\Gamma_{ij}^{\boldsymbol{\mu}}(s)}{H_{\boldsymbol{\mu}}(s-)} \right) \tilde{\mu}_{ij}(s) dR_{ij}(s) \\ &+ \sum_{\substack{i,j\in S\\ i\neq j}} \int_0^t \frac{\triangle A_{\boldsymbol{\mu}}(s)}{H_{\boldsymbol{\mu}}(s-)} \tilde{\mu}_{ij}(s) dR_{ij}(s) \\ &+ \int_0^t \frac{1}{H_0(s-)} dA_0(s) - \int_0^t \frac{1}{H_{\boldsymbol{\mu}}(s-)} dA_{\boldsymbol{\mu}}(s), \quad t \in [0,T], \boldsymbol{\mu} \in \mathcal{H}. \end{aligned} \tag{B.93}$$

To move forward and help simplify (B.93), we state the following lemma:

Lemma B.3.1. *The process*

$$z_{ij}^{\boldsymbol{\mu}}(t) \triangleq \int_0^t \frac{\triangle A_{\boldsymbol{\mu}}(s)}{H_{\boldsymbol{\mu}}(s-)} \tilde{\mu}_{ij}(s) dR_{ij} = 0, \quad t \in [0,T], \tag{B.94}$$

for all $i, j \in S, i \neq j$, and $\boldsymbol{\mu} \in \mathcal{H}$, where $\tilde{\mu}_{ij} \triangleq \mu_{ij}/(1+\mu_{ij})$.

Proof. Fix $i, j \in S, i \neq j$ and $\boldsymbol{\mu} \in \mathcal{H}$. Define

$$G_{ij}(t) \triangleq \frac{\triangle A_{\boldsymbol{\mu}}(t)}{H_{\boldsymbol{\mu}}(t-)} \tilde{\mu}_{ij}(t), \quad t \in [0, T]. \tag{B.95}$$

We know from the Doob-Meyer decomposition in Lemma 4.3.16 that $A_{\boldsymbol{\mu}} \in \mathcal{P}^*$, and as a result, $\triangle A_{\boldsymbol{\mu}} \in \mathcal{P}^*$. We also have that $(H_{\boldsymbol{\mu}})_- \in \mathcal{P}^*$ and $\mu_{ij} \in \mathcal{P}^*$. Therefore,

$$G_{ij} \in \mathcal{P}^*. \tag{B.96}$$

Define the positive and negative parts of G_{ij} as follows

$$G_{ij}^+ \triangleq \max\{0, G_{ij}\} \tag{B.97}$$

and

$$G_{ij}^- \triangleq \max\{0, -G_{ij}\}. \tag{B.98}$$

We then see,

$$G_{ij} = G_{ij}^+ - G_{ij}^- \tag{B.99}$$

and from (B.96) both $G_{ij}^+, G_{ij}^- \in \mathcal{P}^*$. Now, substituting into $z_{ij}^{\boldsymbol{\mu}}$, we have

$$z_{ij}^{\boldsymbol{\mu}}(t) = \int_0^t G_{ij}^+(s) dR_{ij}(s) - \int_0^t G_{ij}^-(s) dR_{ij}(s), \quad t \in [0, T]. \tag{B.100}$$

Now, using the fact $M_{ij} \triangleq R_{ij} - \tilde{R}_{ij}$, and Theorem A.1.7, we can write

$$\int_0^t G_{ij}^+(s) dR_{ij}(s) = \int_0^t G_{ij}^+(s) dM_{ij}(s) + \int_0^t G_{ij}^+(s) d\tilde{R}_{ij}(s), \quad t \in [0, T]. \tag{B.101}$$

For any fixed $\omega \in \Omega$, from (B.95),

$$\left\{ 0 \leq s \leq T \middle| G_{ij}^+(s, \omega) \neq 0 \right\} \subseteq \left\{ 0 \leq s \leq T \middle| \triangle A_{\boldsymbol{\mu}}(s, \omega) \neq 0 \right\}. \tag{B.102}$$

Since $A_{\boldsymbol{\mu}}(\cdot, \omega)$ is càdlàg and non-decreasing for each $\omega \in \Omega$, the set

$$\left\{ 0 \leq s \leq T \middle| \triangle A_{\boldsymbol{\mu}}(s, \omega) \neq 0 \right\} \tag{B.103}$$

is countable for each $\omega \in \Omega$. From (B.102), we also have that

$$\left\{ 0 \leq s \leq T \middle| G_{ij}^+(s, \omega) \neq 0 \right\} \tag{B.104}$$

is countable for each $\omega \in \Omega$. From Definition 4.1.9, $\tilde{R}_{ij}(\cdot, \omega)$ is continuous on $[0, T]$ for each $\omega \in \Omega$, and thus

$$\int_0^{\cdot} G_{ij}^{+}(s, \omega) d\tilde{R}_{ij}(s, \omega) = 0, \quad \omega \in \Omega. \tag{B.105}$$

From (B.101) and (B.105),

$$\int_0^t G_{ij}^{+}(s) dR_{ij}(s) = \int_0^t G_{ij}^{+}(s) dM_{ij}(s), \quad t \in [0, T]. \tag{B.106}$$

Since $M_{ij} \in \mathcal{M}_0(\{\mathcal{F}_t\}, P)$ we have,

$$\int_0^{\cdot} G_{ij}^{+}(s) dR_{ij}(s) \in \mathcal{M}_{loc,0}(\{\mathcal{F}_t\}, P). \tag{B.107}$$

Moreover, since G_{ij}^{+} is non-negative,

$$\int_0^t G_{ij}^{+}(s) dR_{ij}(s) \geq 0, \quad t \in [0, T], \tag{B.108}$$

because R_{ij} is a non-decreasing process. Since $\int_0^{\cdot} G_{ij}^{+}(s) dR_{ij}(s)$ is non-negative by (B.108), a local martingale by (B.107), and initially takes the value 0, we have from basic measure theory that

$$\int_0^t G_{ij}^{+}(s) dR_{ij}(s) = 0, \quad t \in [0, T], \tag{B.109}$$

and similarly,

$$\int_0^t G_{ij}^{-}(s) dR_{ij}(s) = 0, \quad t \in [0, T]. \tag{B.110}$$

Finally, from (B.100), (B.109) and (B.110),

$$z_{ij}^{\boldsymbol{\mu}}(t) = 0, \quad t \in [0, T], \boldsymbol{\mu} \in \mathcal{H}, \quad i, j \in S, i \neq j. \tag{B.111}$$

□

Lemma B.3.2. *The following equality holds*

$$\begin{aligned} \int_0^t \frac{1}{H_0(s-)} dA_0(s) = \int_0^t \frac{1}{H_{\boldsymbol{\mu}}(s-)} dA_{\boldsymbol{\mu}}(s) \\ + \sum_{\substack{i,j \in S \\ i \neq j}} \int_0^t \frac{\Gamma_{ij}^0(s)}{H_0(s-)} \mu_{ij}(s) d\tilde{R}_{ij}(s), \quad t \in [0, T], \boldsymbol{\mu} \in \mathcal{H}, \end{aligned} \tag{B.112}$$

where $A_{\boldsymbol{\mu}}$ and Γ^0_{ij}, $\boldsymbol{\mu} \in \mathcal{H}$, $i, j \in S, i \neq j$, are processes from the Doob-Meyer decomposition (Lemma 4.3.16).

Proof. Using Lemma B.3.1 with (B.93) we obtain

$$\begin{aligned} 0 = &\sum_{\substack{i,j \in S \\ i \neq j}} \int_0^t \left(\frac{\Gamma^{\boldsymbol{\mu}}_{ij}(s)}{H_{\boldsymbol{\mu}}(s-)} - \frac{\Gamma^0_{ij}(s)}{H_0(s-)} - \hat{X}(s-)\mu_{ij}(s) \right) dM_{ij}(s) \\ &+ \sum_{\substack{i,j \in S \\ i \neq j}} \int_0^t \left(\hat{X}(s-)\mu_{ij}(s) - \frac{\Gamma^{\boldsymbol{\mu}}_{ij}(s)}{H_{\boldsymbol{\mu}}(s-)} \right) \tilde{\mu}_{ij}(s) dR_{ij}(s) \\ &+ \int_0^t \frac{1}{H_0(s-)} dA_0(s) - \int_0^t \frac{1}{H_{\boldsymbol{\mu}}(s-)} dA_{\boldsymbol{\mu}}(s), \quad t \in [0,T], \boldsymbol{\mu} \in \mathcal{H}, \end{aligned} \tag{B.113}$$

where $\tilde{\mu}_{ij} \triangleq \mu_{ij}/(1+\mu_{ij})$. Let

$$B(t) \triangleq \sum_{\substack{i,j \in S \\ i \neq j}} \int_0^t \Phi_{ij}(s) dR_{ij}(s), \quad t \in [0,T], \tag{B.114}$$

where we have put

$$\Phi_{ij}(t) \triangleq \tilde{\mu}_{ij}(t) \left(\hat{X}(t-)\mu_{ij}(t) - \frac{\Gamma^{\boldsymbol{\mu}}_{ij}(t)}{H_{\boldsymbol{\mu}}(t-)} \right), \quad t \in [0,T], \boldsymbol{\mu} \in \mathcal{H}, \quad i, j \in S, i \neq j. \tag{B.115}$$

Now, $(X)_- \in \mathcal{P}^*$ and $\Gamma^{\boldsymbol{\mu}}_{ij} \in \mathcal{P}^*$, therefore

$$\Phi_{ij} \in \mathcal{P}^*. \tag{B.116}$$

We have from (B.114) and (4.8),

$$\begin{aligned} B(t) &= \sum_{\substack{i,j \in S \\ i \neq j}} (\Phi_{ij} \bullet R_{ij})(t) \\ &= \sum_{\substack{i,j \in S \\ i \neq j}} (\Phi_{ij} \bullet (R_{ij} - \tilde{R}_{ij} + \tilde{R}_{ij}))(t) \\ &= \sum_{\substack{i,j \in S \\ i \neq j}} (\Phi_{ij} \bullet M_{ij})(t) + \sum_{\substack{i,j \in S \\ i \neq j}} (\Phi_{ij} \bullet \tilde{R}_{ij})(t), \quad t \in [0,T]. \end{aligned} \tag{B.117}$$

Putting (B.117) into (B.113) yields

$$
\begin{aligned}
0 = & \sum_{\substack{i,j \in S \\ i \neq j}} \int_0^t \left(\frac{1}{1+\mu_{ij}(s)} \right) \left(\frac{\Gamma^{\boldsymbol{\mu}}_{ij}(s)}{H_{\boldsymbol{\mu}}(s-)} - \hat{X}(s-)\mu_{ij}(s) \right) dM_{ij}(s) \\
& - \sum_{\substack{i,j \in S \\ i \neq j}} \int_0^t \frac{\Gamma^0_{ij}(s)}{H_0(s-)} dM_{ij}(s) + \sum_{\substack{i,j \in S \\ i \neq j}} \int_0^t \left(\hat{X}(s-)\mu_{ij}(s) - \frac{\Gamma^{\boldsymbol{\mu}}_{ij}(s)}{H_{\boldsymbol{\mu}}(s-)} \right) \tilde{\mu}_{ij}(s) d\tilde{R}_{ij}(s) \\
& + \int_0^t \frac{1}{H_0(s-)} dA_0(s) - \int_0^t \frac{1}{H_{\boldsymbol{\mu}}(s-)} dA_{\boldsymbol{\mu}}(s), \quad t \in [0,T], \boldsymbol{\mu} \in \mathcal{H}.
\end{aligned} \tag{B.118}
$$

Define processes $\{F(t), t \in [0,T]\}$ and $\{K(t), t \in [0,T]\}$ as follows:

$$
\begin{aligned}
F(t) \triangleq & \sum_{\substack{i,j \in S \\ i \neq j}} \int_0^t \left(\frac{1}{1+\mu_{ij}(s)} \right) \left(\frac{\Gamma^{\boldsymbol{\mu}}_{ij}(s)}{H_{\boldsymbol{\mu}}(s-)} - \hat{X}(s-)\mu_{ij}(s) \right) dM_{ij}(s) \\
& - \sum_{\substack{i,j \in S \\ i \neq j}} \int_0^t \frac{\Gamma^0_{ij}(s)}{H_0(s-)} dM_{ij}(s), \quad t \in [0,T], \boldsymbol{\mu} \in \mathcal{H},
\end{aligned} \tag{B.119}
$$

and

$$
\begin{aligned}
K(t) \triangleq & \sum_{\substack{i,j \in S \\ i \neq j}} \int_0^t \left(\hat{X}(s-)\mu_{ij}(s) - \frac{\Gamma^{\boldsymbol{\mu}}_{ij}(s)}{H_{\boldsymbol{\mu}}(s-)} \right) \tilde{\mu}_{ij}(s) d\tilde{R}_{ij}(s) \\
& + \int_0^t \frac{1}{H_0(s-)} dA_0(s) - \int_0^t \frac{1}{H_{\boldsymbol{\mu}}(s-)} dA_{\boldsymbol{\mu}}(s), \quad t \in [0,T], \boldsymbol{\mu} \in \mathcal{H}.
\end{aligned} \tag{B.120}
$$

It is quite easy to see from (B.119) that F is a local martingale. Since $A_{\boldsymbol{\mu}}$ is predictable for each $\boldsymbol{\mu} \in \mathcal{H}$ by Lemma 4.3.16, we have by Proposition C.9.4 and (B.120) that K is both predictable and has paths of finite variation. In summary,

$$
F \in \mathcal{M}_{loc,0}(\{\mathcal{F}_t\}, P) \qquad K \in \mathcal{FV}_0(\{\mathcal{F}_t\}, P) \cap \mathcal{P}^*, \tag{B.121}
$$

and from (B.118),

$$
F(t) = -K(t) \quad t \in [0,T]. \tag{B.122}
$$

From (B.121) and (B.122), we have

$$
F \in \mathcal{M}_{loc,0}(\{\mathcal{F}_t\}, P) \cap \mathcal{FV}_0(\{\mathcal{F}_t\}, P) \cap \mathcal{P}^*. \tag{B.123}
$$

From Proposition C.11.5 and (B.123), we have that

$$F(t) = 0 \text{ a.s.}, \quad t \in [0,T] \tag{B.124}$$

Now from (B.122) and (B.124)

$$F(t) = 0 \text{ a.s.}, \quad t \in [0,T] \tag{B.125}$$

In light of (B.124), setting the left hand side of (B.119) to zero and using Lemma 4.1.15 yields

$$\left(\frac{1}{1+\mu_{ij}(t)}\right)\left(\frac{\Gamma^{\boldsymbol{\mu}}_{ij}(t)}{H_{\boldsymbol{\mu}}(t-)} - \hat{X}(t-)\mu_{ij}(t)\right) = \frac{\Gamma^0_{ij}(t)}{H_0(t-)}, \quad \nu_{[M_{ij}]} - \text{ a.e.}, \tag{B.126}$$

for $i,j \in S, i \neq j$ and $\boldsymbol{\mu} \in \mathcal{H}$. Now, putting (B.126) into (B.120),

$$\begin{aligned} K(t) &= \int_0^t \frac{1}{H_0(s-)} dA_0(s) - \int_0^t \frac{1}{H_{\boldsymbol{\mu}}(s-)} dA_{\boldsymbol{\mu}}(s) \\ &\quad - \sum_{\substack{i,j\in S \\ i\neq j}} \int_0^t \frac{\Gamma^0_{ij}(s)}{H_0(s-)} \mu_{ij}(s) d\tilde{R}_{ij}(s), \quad t \in [0,T], \boldsymbol{\mu} \in \mathcal{H}. \end{aligned} \tag{B.127}$$

Since $K(\cdot) = 0$ from (B.125),

$$\begin{aligned} \int_0^t \frac{1}{H_0(s-)} dA_0(s) &= \int_0^t \frac{1}{H_{\boldsymbol{\mu}}(s-)} dA_{\boldsymbol{\mu}}(s) \\ &\quad + \sum_{\substack{i,j\in S \\ i\neq j}} \int_0^t \frac{\Gamma^0_{ij}(s)}{H_0(s-)} \mu_{ij}(s) d\tilde{R}_{ij}(s), \quad t \in [0,T], \boldsymbol{\mu} \in \mathcal{H}, \end{aligned} \tag{B.128}$$

which is the required result. □

To move forward, we state the following lemma, which provides an upper-bound for the process Γ^0_{ij}.

Lemma B.3.3. *$\Gamma^0_{ij}(t) \leq 0$, $\nu_{[M_{ij}]}$-almost everywhere for $i,j \in S, i \neq j$.*

Proof. Fix some $k,l \in S$ where $k \neq l$. Put $\mu_{ij}(t) \triangleq 0$ for all $t \in [0,T]$ where $(i,j) \neq (k,l)$. From Lemma B.3.2 and the positivity of $A_{\boldsymbol{\mu}}$,

$$\int_0^t \frac{1}{H_0(s-)} dA_0(s) \geq \int_0^t \frac{\Gamma^0_{kl}(s)}{H_0(s-)} \mu_{kl}(s) d\tilde{R}_{kl}(s), \quad t \in [0,T]. \tag{B.129}$$

Put

$$U \triangleq \left\{(t,\omega) \in [0,T] \times \Omega \middle| \Gamma^0_{kl}(t,\omega) > 0\right\}. \tag{B.130}$$

Since Γ^0_{kl} is $\{\mathcal{F}_t\}$-predictable,

$$U \in \mathcal{P}^*. \tag{B.131}$$

We would like to show that $\nu_{[M_{kl}]}[U] = 0$. Suppose

$$\nu_{[M_{kl}]}[U] > 0. \tag{B.132}$$

By the definition of Doléans measure from Notation 4.16,

$$\nu_{[M_{kl}]}[U] \triangleq E\left[\int_0^T I_U(t,\omega) R_{kl}(dt,\omega)\right], \tag{B.133}$$

and from I.3.17(iii) of Jacod and Shiryayev [26], we also have

$$\nu_{[M_{kl}]}[U] \triangleq E\left[\int_0^T I_U(t,\omega) \tilde{R}_{kl}(dt,\omega)\right]. \tag{B.134}$$

From (B.132) and (B.134),

$$\int_\Omega \left[\int_0^T I_U(t,\omega) \tilde{R}_{kl}(dt,\omega)\right] P(d\omega) > 0. \tag{B.135}$$

Put

$$z(\omega) \triangleq \int_0^T I_U(t,\omega) \tilde{R}_{kl}(dt,\omega), \quad \omega \in \Omega. \tag{B.136}$$

We then have by (B.135)

$$z(\omega) \geq 0 \text{ a.s.} \quad \text{and} \quad E[z] > 0. \tag{B.137}$$

Put

$$\Omega_1 \triangleq \left\{\omega \in \Omega \middle| z(\omega) > 0\right\}. \tag{B.138}$$

Since $E[z] > 0$, we have $P[\Omega_1] > 0$. Define a sequence $\{\mu^{(n)}_{kl}(t,\omega)\}$ for $n = 1, 2, \ldots$ as

$$\mu^{(n)}_{kl}(t,\omega) \triangleq n I_U(t,\omega) H_0(t-,\omega) \tag{B.139}$$

for all $(t,\omega) \in [0,T] \times \Omega$. Since each component of $\mu^{(n)}_{kl}(t,\omega)$ is predictable we have that $\mu^{(n)}_{kl} \in \mathcal{P}^*$ for each $n = 1, 2, \ldots$. Defining

$$\boldsymbol{\mu}^{(n)} \triangleq \begin{cases} \mu^{(n)}_{kl} & i = k, j = l \\ 0 & \text{for all other } i, j \in S, i \neq j \end{cases}, \tag{B.140}$$

we then have $\boldsymbol{\mu}^{(n)} \in \mathcal{P}^*$. Also from the elementary bound in (4.11) and since $H_0(t)$ is continuous on $t \in [0,T]$ with $\inf_{t\in[0,T]} H_0(t) > 0$ a.s., we have that $\mu_{kl}^{(n)} \in L_{loc}^{1/2}(R_{kl})$ and thus $\boldsymbol{\mu}^{(n)} \in \mathcal{H}$ for all $n = 1, 2, \ldots$. Putting (B.139) into (B.129) and taking $t = T$ yields

$$\int_0^T \frac{1}{H_0(s-,\omega)} dA_0(s,\omega) \geq n \int_0^T \Gamma_{kl}^0(s,\omega) I_U(s,\omega) \tilde{R}_{kl}(ds,\omega), \tag{B.141}$$

for $n = 1, 2, \ldots$. Define the set

$$\Theta(\omega) \triangleq \left\{ t \in [0,T] \middle| I_U(t,\omega) = 1 \right\} \tag{B.142}$$

for all $\omega \in \Omega$. Fix some $\bar{\omega} \in \Omega_1$. From (B.138) we have that $z(\bar{\omega}) > 0$ and therefore by (B.136) and (B.142)

$$\int_{\Theta(\bar{\omega})} \tilde{R}_{kl}(dt,\bar{\omega}) > 0. \tag{B.143}$$

Now from (B.130) and (B.142) we have that

$$\Gamma_{kl}^0(t,\bar{\omega}) > 0 \text{ for all } t \in \Theta(\bar{\omega}). \tag{B.144}$$

As a result of (B.143) and (B.144), we can conclude

$$\int_{\Theta(\bar{\omega})} \Gamma_{kl}^0(t,\bar{\omega}) \tilde{R}_{kl}(dt,\bar{\omega}) > 0. \tag{B.145}$$

Fixing some $\bar{\omega} \in \Omega_1$ in (B.141) yields

$$\int_0^T \frac{1}{H_0(s-,\bar{\omega})} dA_0(s,\bar{\omega}) \geq n \int_{\Theta(\bar{\omega})} \Gamma_{kl}^0(s,\bar{\omega}) \tilde{R}_{kl}(ds,\bar{\omega}) \tag{B.146}$$

for each $n = 1, 2, \ldots$. Putting (B.145) into (B.146) and taking $n \to \infty$ in (B.146) gives

$$\int_0^T \frac{1}{H_0(s-,\bar{\omega})} dA_0(s,\bar{\omega}) = \infty, \quad \bar{\omega} \in \Omega_1. \tag{B.147}$$

Since $H_0(t)$ is continuous on $t \in [0,T]$, therefore bounded for P-almost all ω, and $A_0(t)$ is a non-decreasing process on $t \in [0,T]$ with $A_0(T) < \infty$ a.s. from Lemma 4.3.16, there cannot be a set of non-zero probability in which $(\frac{1}{H_0} \bullet A_0)(T)$ takes the value ∞. However, $P[\Omega_1] > 0$. As a result, there is a contradiction in (B.132). Since $k, l \in S$ were arbitrarily chosen, we therefore have

$$\Gamma_{ij}^0 \leq 0 \quad \nu_{[M_{ij}]} - \text{ a.e.}, \quad i, j \in S, i \neq j. \tag{B.148}$$

□

Now that we have established Lemma B.3.3, we can complete the proof of Proposition 4.3.21. From Lemma B.3.2 and the fact that $M_{ij} \triangleq R_{ij} - \tilde{R}_{ij}$,

$$\begin{aligned}&\int_0^t \frac{1}{H_0(s-)} dA_0(s) + \sum_{\substack{i,j\in S\\ i\neq j}} \int_0^t \frac{\Gamma_{ij}^0(s)}{H_0(s-)} \mu_{ij}(s) dM_{ij}(s)\\ &= \int_0^t \frac{1}{H_{\boldsymbol{\mu}}(s-)} dA_{\boldsymbol{\mu}}(s) + \sum_{\substack{i,j\in S\\ i\neq j}} \int_0^t \frac{\Gamma_{ij}^0(s)}{H_0(s-)} \mu_{ij}(s) dR_{ij}(s), \quad t \in [0,T], \boldsymbol{\mu} \in \mathcal{H}.\end{aligned} \tag{B.149}$$

Define a sequence $\{\mu_{ij}^{(m)}\}$ for $m = 2, 3, \ldots$ where

$$\mu_{ij}^{(m)}(t,\omega) \triangleq -1 + \frac{1}{m}, \quad (t,\omega) \in [0,T] \times \Omega. \tag{B.150}$$

It is easy to see that $\boldsymbol{\mu}^{(m)} \triangleq \{\mu_{ij}^{(m)}\}_{i,j\in S} \in \mathcal{H}$.

Also define the sequence of processes $\{\hat{c}^{(m)}\}$, for $m = 2, 3, \ldots$,

$$\hat{c}^{(m)}(t) \triangleq \int_0^t \frac{1}{H_0(s-)} dA_0(s) + \sum_{\substack{i,j\in S\\ i\neq j}} \int_0^t \frac{\Gamma_{ij}^0(s)}{H_0(s-)} \mu_{ij}^{(m)}(s) dM_{ij}(s), \quad t \in [0,T]. \tag{B.151}$$

From (B.150) and (B.149)

$$\hat{c}^{(m)}(t) = \int_0^t \frac{1}{H_{\boldsymbol{\mu}^{(m)}}(s-)} dA_{\boldsymbol{\mu}^{(m)}} + \sum_{\substack{i,j\in S\\ i\neq j}} \int_0^t \frac{(-1+\frac{1}{m})\Gamma_{ij}^0(s)}{H_0(s-)}(s) dR_{ij}(s), \quad t \in [0,T], \tag{B.152}$$

for all $m = 2, 3, \ldots$. Since $A_{\boldsymbol{\mu}^{(m)}}$ and R_{ij} are non-decreasing, one sees from Lemma B.3.3 and (B.152) that, for all large integers m, one has

$$\hat{c}^{(m)}(t_2) \geq \hat{c}^{(m)}(t_1) \text{ a.s.} \tag{B.153}$$

for $T \geq t_2 \geq t_1 \geq 0$. Thus $\hat{c}^{(m)}(t)$ is non-decreasing for all large values of m. Now from (B.151) and (B.150)

$$\hat{c}^{(m)}(t) = \int_0^t \frac{1}{H_0(s-)} dA_0(s) - (1 - \frac{1}{m}) \sum_{\substack{i,j\in S\\ i\neq j}} \int_0^t \frac{\Gamma_{ij}^0(s)}{H_0(s-)} dM_{ij}(s), \quad t \in [0,T], \tag{B.154}$$

and taking $m \to \infty$

$$\lim_{m\to\infty} \hat{c}^{(m)}(t) = \int_0^t \frac{1}{H_0(s-)} dA_0(s) - \sum_{\substack{i,j\in S\\ i\neq j}} \int_0^t \frac{\Gamma^0_{ij}(s)}{H_0(s-)} dM_{ij}(s) \text{ a.s.,} \quad t \in [0,T], \tag{B.155}$$

and as a result, from (4.100),

$$\hat{c}^{(m)}(t) \to \hat{c}(t) \quad \text{as} \quad m \to \infty \tag{B.156}$$

for all $t \in [0,T]$. By taking $m \to \infty$ in (B.153), and observing (B.156), we have that

$$\hat{c}(\cdot) \text{ is non-decreasing.} \tag{B.157}$$

Now, it is immediate from (4.100) that

$$\hat{c}(0) = 0, \tag{B.158}$$

and from (4.71) and the fact that H_0 is almost surely strictly positive,

$$\tfrac{\Gamma^0_{ij}}{(H_0)_-} \in L^{1/2}_{loc}(R_{ij}) \text{ and } (\tfrac{1}{H_0} \bullet A_0) \in \mathcal{F}^*. \tag{B.159}$$

Therefore, we have

$$\hat{c} \in \mathcal{F}^*. \tag{B.160}$$

Since H_0 is almost surely strictly positive and continuous, and $A_0(T) < \infty$ a.s. by Lemma 4.3.16, we have that the first term on the right-hand side of (4.100) is finite. Now since the second term on the right-hand side of (4.100) is a local martingale by (B.159) and Remark A.1.2, we know that it is finite as well. Therefore,

$$\hat{c}(T) < \infty \text{ a.s.} \tag{B.161}$$

As a result, from (B.157), (B.158), (B.160), and (B.161), we can conclude

$$\hat{c}(t) \triangleq \int_0^t \frac{1}{H_0(s-)} dA_0(s) - \sum_{\substack{i,j\in S\\ i\neq j}} \int_0^t \frac{\Gamma^0_{ij}(s)}{H_0(s-)} dM_{ij}(s) \in \mathcal{C}. \tag{B.162}$$

□

B.4 Optimal Wealth Process Results in the Constrained Regime-Switching Market Model

In this section we prove Proposition 5.3.15:

Proof of Proposition 5.3.15(1). Set $t = 0$. From (5.37)

$$\hat{X}(0) = \operatorname*{ess-sup}_{(\boldsymbol{\nu},\boldsymbol{\mu}) \in \mathcal{G} \times \mathcal{H}} \frac{E\left[H_{\boldsymbol{\nu},\boldsymbol{\mu}}(T)B \,\middle|\, \mathcal{F}_0\right]}{H_{\boldsymbol{\nu},\boldsymbol{\mu}}(0)}, \tag{B.163}$$

where the σ-algebra $\mathcal{F}_0$ is given by

$$\mathcal{F}_0 \triangleq \{\emptyset, \Omega\} \vee \mathcal{N}(P). \tag{B.164}$$

Since $H_{\boldsymbol{\nu},\boldsymbol{\mu}}(0) = 1$ a.s. for each $(\boldsymbol{\nu}, \boldsymbol{\mu}) \in \mathcal{G} \times \mathcal{H}$, we have

$$\hat{X}(0) = \operatorname*{ess-sup}_{(\boldsymbol{\nu},\boldsymbol{\mu}) \in \mathcal{G} \times \mathcal{H}} E\left[H_{\boldsymbol{\nu},\boldsymbol{\mu}}(T)B\right] = \hat{u} \text{ a.s.} \tag{B.165}$$

□

Proof of Proposition 5.3.15(2). Set $t = T$. From (5.37),

$$\hat{X}(T) = \operatorname*{ess-sup}_{(\boldsymbol{\nu},\boldsymbol{\mu}) \in \mathcal{G} \times \mathcal{H}} \frac{E\left[H_{\boldsymbol{\nu},\boldsymbol{\mu}}(T)B \,\middle|\, \mathcal{F}_T\right]}{H_{\boldsymbol{\nu},\boldsymbol{\mu}}(T)}. \tag{B.166}$$

Since $H_{\boldsymbol{\nu},\boldsymbol{\mu}}(T)$ and B are $\mathcal{F}_T$-measurable, with $\inf_{t \in [0,T]} H_{\boldsymbol{\nu},\boldsymbol{\mu}}(t) > 0$ a.s., for all $(\boldsymbol{\nu}, \boldsymbol{\mu}) \in \mathcal{G} \times \mathcal{H}$, we have

$$\hat{X}(T) = \operatorname*{ess-sup}_{(\boldsymbol{\nu},\boldsymbol{\mu}) \in \mathcal{G} \times \mathcal{H}} \frac{H_{\boldsymbol{\nu},\boldsymbol{\mu}}(T)B}{H_{\boldsymbol{\nu},\boldsymbol{\mu}}(T)} = B \text{ a.s.} \tag{B.167}$$

□

Proof of Proposition 5.3.15(3). Since $H_{\boldsymbol{\nu},\boldsymbol{\mu}}(t) > 0$ a.s. for all $t \in [0, T]$ and $(\boldsymbol{\nu}, \boldsymbol{\mu}) \in \mathcal{G} \times \mathcal{H}$, and since $B > 0$ a.s., we have from (5.37) that $\hat{X}(t) > 0$ a.s. for all $t \in [0, T]$.
□

Proof of Proposition 5.3.15(4). For ease of notation, define

$$H_{\boldsymbol{\nu},\boldsymbol{\mu}}(s,t) \triangleq \frac{H_{\boldsymbol{\nu},\boldsymbol{\mu}}(t)}{H_{\boldsymbol{\nu},\boldsymbol{\mu}}(s)}, \quad (\boldsymbol{\nu}, \boldsymbol{\mu}) \in \mathcal{G} \times \mathcal{H}, \tag{B.168}$$

for $0 \leq s \leq t \leq T$ and

$$J_{\boldsymbol{\nu},\boldsymbol{\mu}}(t) \triangleq E\left[H_{\boldsymbol{\nu},\boldsymbol{\mu}}(t,T)B \,\middle|\, \mathcal{F}_t\right], \quad (\boldsymbol{\nu},\boldsymbol{\mu}) \in \mathcal{G} \times \mathcal{H}, \tag{B.169}$$

for all $t \in [0,T]$. From (5.37) and (B.169)

$$\hat{X}(t) = \underset{(\boldsymbol{\nu},\boldsymbol{\mu})\in\mathcal{G}\times\mathcal{H}}{\text{ess-sup}}\, J_{\boldsymbol{\nu},\boldsymbol{\mu}}(t), t \in [0,T]. \tag{B.170}$$

Fix some $\bar{t} \in [0,T]$ and $(\bar{\boldsymbol{\nu}},\bar{\boldsymbol{\mu}}) \in \mathcal{G} \times \mathcal{H}$. Put

$$\begin{aligned}\mathcal{D}_{\bar{t},\bar{\boldsymbol{\nu}},\bar{\boldsymbol{\mu}}} \triangleq \Big\{(\boldsymbol{\nu},\boldsymbol{\mu}) \in \mathcal{G} \times \mathcal{H} \Big| \boldsymbol{\nu}(s,\omega) = \bar{\boldsymbol{\nu}}(s,\omega), \mu_{ij}(s,\omega) = \bar{\mu}_{ij}(s,\omega),\\ 0 \leq s \leq \bar{t}, \omega \in \Omega, i,j \in S, i \neq j\Big\}.\end{aligned} \tag{B.171}$$

$\mathcal{D}_{\bar{t},\bar{\boldsymbol{\nu}},\bar{\boldsymbol{\mu}}}$ is the set of all processes $(\boldsymbol{\nu},\boldsymbol{\mu}) \in \mathcal{G} \times \mathcal{H}$ that match the fixed $(\bar{\boldsymbol{\nu}},\bar{\boldsymbol{\mu}}) \in \mathcal{G} \times \mathcal{H}$ for all times between 0 and $\bar{t}$ and all $\omega \in \Omega$. Now, if we expand $H_{\boldsymbol{\nu},\boldsymbol{\mu}}(t_1,t_2)$ for any $0 \leq t_1 \leq t_2 \leq T$, we have by (5.19) and (4.43)

$$\begin{aligned}H_{\boldsymbol{\nu},\boldsymbol{\mu}}(t_1,t_2) &= \exp\Big\{-\int_{t_1}^{t_2}[r(s)+\tilde{g}_K(s,\boldsymbol{\nu}(s))]ds\Big\}\frac{\mathcal{E}(-\boldsymbol{\theta}_{\boldsymbol{\nu}}\bullet\mathbf{W})(t_2)}{\mathcal{E}(-\boldsymbol{\theta}_{\boldsymbol{\nu}}\bullet\mathbf{W})(t_1)}\frac{\mathcal{E}(\boldsymbol{\mu}\bullet\mathbf{M})(t_2)}{\mathcal{E}(\boldsymbol{\mu}\bullet\mathbf{M})(t_1)}\\ &= \exp\Big\{-\int_{t_1}^{t_2}[r(s)+\tilde{g}_K(s,\boldsymbol{\nu}(s))]ds\Big\}\\ &\cdot \exp\Big\{-\int_{t_1}^{t_2}\boldsymbol{\theta}_{\boldsymbol{\nu}}(s)^\top d\mathbf{W}(s) - \frac{1}{2}\int_{t_1}^{t_2}||\boldsymbol{\theta}_{\boldsymbol{\nu}}(s)||^2 ds\Big\}\\ &\cdot \exp\Big\{\sum_{\substack{i,j\in S\\ i\neq j}}\int_{t_1}^{t_2}\mu_{ij}(s)dM_{ij}(s)\Big\}\prod_{t_1<s\leq t_2}(1+\triangle(\boldsymbol{\mu}\bullet\mathbf{M})(s))\exp\Big\{-\triangle(\boldsymbol{\mu}\bullet\mathbf{M})(s)\Big\},\end{aligned} \tag{B.172}$$

for $t \in [0,T]$ and $(\boldsymbol{\nu},\boldsymbol{\mu}) \in \mathcal{G} \times \mathcal{H}$. Viewing all terms of (B.172) we see that $H_{\boldsymbol{\nu},\boldsymbol{\mu}}(t_1,t_2)$ is determined only by values of $\boldsymbol{\nu}(s)$ and $\boldsymbol{\mu}(s)$ for $s \in [t_1,t_2]$. As a result, by setting $t_1 = \bar{t}$ and $t_2 = T$, we have from (B.172) that $H_{\boldsymbol{\nu},\boldsymbol{\mu}}(\bar{t},T)$ is determined only by values of $\boldsymbol{\nu}(s)$ and $\boldsymbol{\mu}(s)$ for $s \in [\bar{t},T]$. Thus, by (B.169), (B.170) and (B.171),

$$\hat{X}(\bar{t}) = \underset{(\boldsymbol{\nu},\boldsymbol{\mu})\in\mathcal{G}\times\mathcal{H}}{\text{ess-sup}}\, J_{\boldsymbol{\nu},\boldsymbol{\mu}}(\bar{t}) = \underset{(\boldsymbol{\nu},\boldsymbol{\mu})\in\mathcal{D}_{\bar{t},\bar{\boldsymbol{\nu}},\bar{\boldsymbol{\mu}}}}{\text{ess-sup}}\, E[H_{\boldsymbol{\nu},\boldsymbol{\mu}}(\bar{t},T)B|\mathcal{F}_{\bar{t}}]. \tag{B.173}$$

For fixed $(\bar{\boldsymbol{\nu}}, \bar{\boldsymbol{\mu}}) \in \mathcal{G} \times \mathcal{H}$ and $\bar{t} \in [0, T]$, define the set

$$\mathcal{J}(\bar{t}, \bar{\boldsymbol{\nu}}, \bar{\boldsymbol{\mu}}) \triangleq \left\{ J_{\boldsymbol{\nu},\boldsymbol{\mu}}(\bar{t}) \middle| (\boldsymbol{\nu}, \boldsymbol{\mu}) \in \mathcal{D}_{\bar{t},\bar{\boldsymbol{\nu}},\bar{\boldsymbol{\mu}}} \right\}. \tag{B.174}$$

The following lemma is essential as it allows us to use Theorem C.13.2 on the set $\mathcal{J}(\bar{t}, \bar{\boldsymbol{\nu}}, \bar{\boldsymbol{\mu}})$.

Lemma B.4.1. *The set $\mathcal{J}(\bar{t}, \bar{\boldsymbol{\nu}}, \bar{\boldsymbol{\mu}})$ is closed under pairwise maximization in the following sense: for arbitrary elements $J_1, J_2 \in \mathcal{J}(\bar{t}, \bar{\boldsymbol{\nu}}, \bar{\boldsymbol{\mu}})$, there is some $J \in \mathcal{J}(\bar{t}, \bar{\boldsymbol{\nu}}, \bar{\boldsymbol{\mu}})$ such that*

$$J \triangleq J_1 \vee J_2 \;\; a.s. \tag{B.175}$$

Proof. Fix $(\boldsymbol{\nu}_1, \boldsymbol{\mu}_1), (\boldsymbol{\nu}_2, \boldsymbol{\mu}_2) \in \mathcal{D}_{\bar{t},\bar{\boldsymbol{\nu}},\bar{\boldsymbol{\mu}}}$. To remove possible confusion with notation, we write

$$J_{\boldsymbol{\nu}_i,\boldsymbol{\mu}_i,\bar{t}}(\omega) \triangleq J_{\boldsymbol{\nu}_i,\boldsymbol{\mu}_i}(\bar{t}, \omega), \quad i = 1, 2. \tag{B.176}$$

for all $\omega \in \Omega$. Define the set

$$A \triangleq \left\{ \omega \in \Omega \middle| J_{\boldsymbol{\nu}_1,\boldsymbol{\mu}_1,\bar{t}}(\omega) \geq J_{\boldsymbol{\nu}_2,\boldsymbol{\mu}_2,\bar{t}}(\omega) \right\}. \tag{B.177}$$

Since $J_{\boldsymbol{\nu}_1,\boldsymbol{\mu}_1}$ and $J_{\boldsymbol{\nu}_2,\boldsymbol{\mu}_2}$ are $\{\mathcal{F}_t\}$-adapted, $A \in \mathcal{F}_{\bar{t}}$. Define

$$\hat{\boldsymbol{\mu}}(t, \omega) \triangleq \boldsymbol{\mu}_1(t, \omega) I_A(\omega) + \boldsymbol{\mu}_2(t, \omega) I_{A^c}(\omega), \quad (t, \omega) \in [0, T] \times \Omega, \tag{B.178}$$

and

$$\hat{\boldsymbol{\nu}}(t, \omega) \triangleq \boldsymbol{\nu}_1(t, \omega) I_A(\omega) + \boldsymbol{\nu}_2(t, \omega) I_{A^c}(\omega), \quad (t, \omega) \in [0, T] \times \Omega. \tag{B.179}$$

Since $(\boldsymbol{\nu}_1, \boldsymbol{\mu}_1), (\boldsymbol{\nu}_2, \boldsymbol{\mu}_2) \in \mathcal{D}_{\bar{t},\bar{\boldsymbol{\nu}},\bar{\boldsymbol{\mu}}}$, we have

$$(\hat{\boldsymbol{\nu}}(s, \omega), \hat{\boldsymbol{\mu}}(s, \omega)) = (\bar{\boldsymbol{\nu}}(s, \omega), \bar{\boldsymbol{\mu}}(s, \omega)), \quad (s, \omega) \in [0, \bar{t}] \times \Omega. \tag{B.180}$$

We then have from (B.178)-(B.180),

$$\hat{\boldsymbol{\mu}}(s, \omega) = \bar{\boldsymbol{\mu}}(s, \omega) I_{[0,\bar{t}]}(s) + \boldsymbol{\mu}_1(s, \omega) I_{(\bar{t},T] \times A}(s, \omega) + \boldsymbol{\mu}_2(s, \omega) I_{(\bar{t},T] \times A^c}(s, \omega), \tag{B.181}$$

and

$$\hat{\boldsymbol{\nu}}(s, \omega) = \bar{\boldsymbol{\nu}}(s, \omega) I_{[0,\bar{t}]}(s) + \boldsymbol{\nu}_1(s, \omega) I_{(\bar{t},T] \times A}(s, \omega) + \boldsymbol{\nu}_2(s, \omega) I_{(\bar{t},T] \times A^c}(s, \omega), \tag{B.182}$$

for all $(s, \omega) \in [0, T] \times \Omega$. From (B.181) we see that each term is $\{\mathcal{F}_t\}$-predictably measurable, and thus, $\hat{\boldsymbol{\mu}}$ is $\{\mathcal{F}_t\}$-predictably measurable. Moreover, $\hat{\mu}_{ij}(t) \triangle M_{ij}(t) > -1$ a.s.

for all $t \in [0,T]$ and $\hat{\mu}_{ij} \in L^{1/2}_{loc}(R_{ij})$ (see Definition A.1.1) for all $i, j \in S, i \neq j$, since $\bar{\boldsymbol{\mu}}, \boldsymbol{\mu}_1, \boldsymbol{\mu}_2 \in \mathcal{H}$. Thus, $\hat{\boldsymbol{\mu}} \in \mathcal{H}$. Similarly, from (B.182) we see that each term is $\{\mathcal{F}_t\}$-progressively measurable, and thus, $\hat{\boldsymbol{\nu}}$ is $\{\mathcal{F}_t\}$-progressively measurable. Furthermore, $\hat{\boldsymbol{\nu}}$ is square-integrable on $[0,T]$. Lastly, since

$$\tilde{g}_K(t,\omega,\hat{\boldsymbol{\nu}}(t,\omega)) = \tilde{g}_K(t,\omega,\boldsymbol{\nu}_1(t,\omega))I_A(\omega) + \tilde{g}_K(t,\omega,\boldsymbol{\nu}_2(t,\omega))I_{A^c}(\omega), \quad (t,\omega) \in [0,T] \times \Omega, \tag{B.183}$$

we easily see that $\hat{\boldsymbol{\nu}} \in \mathcal{G}$. Therefore, from (B.180), we have that $(\hat{\boldsymbol{\nu}}, \hat{\boldsymbol{\mu}}) \in \mathcal{D}_{\bar{t},\bar{\boldsymbol{\nu}},\bar{\boldsymbol{\mu}}}$, and

$$J_{\hat{\boldsymbol{\nu}},\hat{\boldsymbol{\mu}},\bar{t}} = J_{\hat{\boldsymbol{\nu}},\hat{\boldsymbol{\mu}}}(\bar{t}) \triangleq E\left[H_{\hat{\boldsymbol{\nu}},\hat{\boldsymbol{\mu}}}(\bar{t},T)B \,\middle|\, \mathcal{F}_{\bar{t}}\right], \tag{B.184}$$

is a member of $\mathcal{J}(\bar{t},\bar{\boldsymbol{\nu}},\bar{\boldsymbol{\mu}})$ (i.e. $J_{\hat{\boldsymbol{\nu}},\hat{\boldsymbol{\mu}},\bar{t}} \in \mathcal{J}(\bar{t},\bar{\boldsymbol{\nu}},\bar{\boldsymbol{\mu}})$). Using (B.181) and (B.182), and the fact that $A \in \mathcal{F}_{\bar{t}}$ (as noted above),

$$H_{\hat{\boldsymbol{\nu}},\hat{\boldsymbol{\mu}}}(\bar{t},T,\omega) = I_A(\omega)H_{\boldsymbol{\nu}_1,\boldsymbol{\mu}_1}(\bar{t},T,\omega) + I_{A^c}(\omega)H_{\boldsymbol{\nu}_2,\boldsymbol{\mu}_2}(\bar{t},T,\omega), \quad P-a.a.\ \omega \in \Omega. \tag{B.185}$$

Furthermore from (B.184) and (B.185) we can write

$$J_{\hat{\boldsymbol{\nu}},\hat{\boldsymbol{\mu}},\bar{t}}(\omega) = I_A(\omega)J_{\boldsymbol{\nu}_1,\boldsymbol{\mu}_1,\bar{t}}(\omega) + I_{A^c}(\omega)J_{\boldsymbol{\nu}_2,\boldsymbol{\mu}_2,\bar{t}}(\omega), \quad P-a.a.\ \omega \in \Omega. \tag{B.186}$$

As a result, we have

$$J_{\boldsymbol{\nu}_1,\boldsymbol{\mu}_1,\bar{t}}(\omega) \geq J_{\boldsymbol{\nu}_2,\boldsymbol{\mu}_2,\bar{t}}(\omega) \text{ and } J_{\hat{\boldsymbol{\nu}},\hat{\boldsymbol{\nu}},\bar{t}}(\omega) = J_{\boldsymbol{\nu}_1,\boldsymbol{\mu}_1,\bar{t}}(\omega) \quad P-a.a.\ \omega \in A. \tag{B.187}$$

Similarly, we have

$$J_{\boldsymbol{\nu}_2,\boldsymbol{\mu}_2,\bar{t}}(\omega) \geq J_{\boldsymbol{\nu}_1,\boldsymbol{\mu}_1,\bar{t}}(\omega) \text{ and } J_{\hat{\boldsymbol{\nu}},\hat{\boldsymbol{\mu}},\bar{t}}(\omega) = J_{\boldsymbol{\nu}_2,\boldsymbol{\mu}_2,\bar{t}}(\omega) \quad P-a.a.\ \omega \in A^c. \tag{B.188}$$

Therefore,

$$J_{\hat{\boldsymbol{\nu}},\hat{\boldsymbol{\mu}},\bar{t}}(\omega) = J_{\boldsymbol{\nu}_1,\boldsymbol{\mu}_1,\bar{t}}(\omega) \vee J_{\boldsymbol{\nu}_2,\boldsymbol{\mu}_2,\bar{t}}(\omega), \quad P-a.a.\ \omega \in \Omega. \tag{B.189}$$

Since $J_{\hat{\boldsymbol{\nu}},\hat{\boldsymbol{\mu}},\bar{t}} \in \mathcal{J}(\bar{t},\bar{\boldsymbol{\nu}},\bar{\boldsymbol{\mu}})$ from (B.184), the set $\mathcal{J}(\bar{t},\bar{\boldsymbol{\nu}},\bar{\boldsymbol{\mu}})$ is closed under pairwise maximization.
□

Continuing the proof of Proposition 5.3.15(4), since

$$\hat{X}(\bar{t}) = \operatorname*{ess-sup}_{(\boldsymbol{\nu},\boldsymbol{\mu}) \in \mathcal{D}_{\bar{t},\bar{\boldsymbol{\nu}},\bar{\boldsymbol{\mu}}}} J_{\boldsymbol{\nu},\boldsymbol{\mu}}(\bar{t}), \tag{B.190}$$

we have from (B.174),

$$\hat{X}(\bar{t}) = \operatorname{ess-sup} \mathcal{J}(\bar{t},\bar{\boldsymbol{\nu}},\bar{\boldsymbol{\mu}}). \tag{B.191}$$

From Theorem C.13.2, due to the pairwise maximization property of $\mathcal{J}(\bar{t}, \bar{\boldsymbol{\nu}}, \bar{\boldsymbol{\mu}})$, there exists a sequence $\{\zeta_n\} \in \mathcal{J}(\bar{t}, \bar{\boldsymbol{\nu}}, \bar{\boldsymbol{\mu}})$ for $n = 1, 2, \ldots$ such that $\zeta_n \leq \zeta_{n+1} \leq \ldots$ a.s. and

$$\begin{aligned} \lim_{n\to\infty} \zeta_n &= \text{ess-sup}\, \mathcal{J}(\bar{t}, \bar{\boldsymbol{\nu}}, \bar{\boldsymbol{\mu}}) \text{ a.s.} \\ &= \hat{X}(\bar{t}) \text{ a.s.} \end{aligned} \tag{B.192}$$

This means there exists some sequence $\{(\boldsymbol{\nu}_n, \boldsymbol{\mu}_n)\} \in \mathcal{D}_{\bar{t},\bar{\boldsymbol{\nu}},\bar{\boldsymbol{\mu}}}$ such that $J_{\boldsymbol{\nu}_n,\boldsymbol{\mu}_n} \leq J_{\boldsymbol{\nu}_{n+1},\boldsymbol{\mu}_{n+1}} \leq \ldots$ a.s. and from (B.192),

$$\lim_{n\to\infty} J_{\boldsymbol{\nu}_n,\boldsymbol{\mu}_n}(\bar{t}) = \hat{X}(\bar{t}) \text{ a.s.} \tag{B.193}$$

Now fix some s such that $0 \leq s \leq \bar{t} \leq T$, then by (B.172) and the reasoning that follows,

$$H_{\boldsymbol{\nu},\boldsymbol{\mu}}(s, \bar{t}) = H_{\bar{\boldsymbol{\nu}},\bar{\boldsymbol{\mu}}}(s, \bar{t}) \text{ a.s.}, \quad (\boldsymbol{\nu}, \boldsymbol{\mu}) \in \mathcal{D}_{\bar{t},\bar{\boldsymbol{\nu}},\bar{\boldsymbol{\mu}}}. \tag{B.194}$$

Since $\{(\boldsymbol{\nu}_n, \boldsymbol{\mu}_n)\} \in \mathcal{D}_{\bar{t},\bar{\boldsymbol{\nu}},\bar{\boldsymbol{\mu}}}$, from (B.194)

$$H_{\boldsymbol{\nu}_n,\boldsymbol{\mu}_n}(s, \bar{t}) = H_{\bar{\boldsymbol{\nu}},\bar{\boldsymbol{\mu}}}(s, \bar{t}) \text{ a.s.}, \quad n = 1, 2, \ldots \tag{B.195}$$

Using (B.193), we can write

$$E\left[H_{\bar{\boldsymbol{\nu}},\bar{\boldsymbol{\mu}}}(s, \bar{t})\hat{X}(\bar{t}) \,\middle|\, \mathcal{F}_s\right] = E\left[H_{\bar{\boldsymbol{\nu}},\bar{\boldsymbol{\mu}}}(s, \bar{t}) \lim_{n\to\infty} J_{\boldsymbol{\nu}_n,\boldsymbol{\mu}_n}(\bar{t}) \,\middle|\, \mathcal{F}_s\right] \text{ a.s.}, \tag{B.196}$$

and using the monotone convergence theorem

$$\begin{aligned} E\left[H_{\bar{\boldsymbol{\nu}},\bar{\boldsymbol{\mu}}}(s, \bar{t})\hat{X}(\bar{t}) \,\middle|\, \mathcal{F}_s\right] &= \lim_{n\to\infty} E\left[H_{\bar{\boldsymbol{\nu}},\bar{\boldsymbol{\mu}}}(s, \bar{t}) J_{\boldsymbol{\nu}_n,\boldsymbol{\mu}_n}(\bar{t}) \,\middle|\, \mathcal{F}_s\right] \\ &= \lim_{n\to\infty} E\left[E\left[H_{\boldsymbol{\nu}_n,\boldsymbol{\mu}_n}(s, \bar{t}) H_{\boldsymbol{\nu}_n,\boldsymbol{\mu}_n}(\bar{t}, T) B \,\middle|\, \mathcal{F}_{\bar{t}}\right] \,\middle|\, \mathcal{F}_s\right] \\ &= \lim_{n\to\infty} E\left[H_{\boldsymbol{\nu}_n,\boldsymbol{\mu}_n}(s, T) B \,\middle|\, \mathcal{F}_s\right] \\ &= \lim_{n\to\infty} J_{\boldsymbol{\nu}_n,\boldsymbol{\mu}_n}(s) \text{ a.s.} \end{aligned} \tag{B.197}$$

Therefore, by the fact that $(\boldsymbol{\nu}_n, \boldsymbol{\mu}_n) \in \mathcal{D}_{\bar{t},\bar{\boldsymbol{\nu}},\bar{\boldsymbol{\mu}}} \subset \mathcal{G} \times \mathcal{H}$ and $s \leq \bar{t}$,

$$\begin{aligned} E\left[H_{\bar{\boldsymbol{\nu}},\bar{\boldsymbol{\mu}}}(s, \bar{t})\hat{X}(\bar{t}) \,\middle|\, \mathcal{F}_s\right] &= \lim_{n\to\infty} J_{\boldsymbol{\nu}_n,\boldsymbol{\mu}_n}(s) \\ &\leq \underset{(\boldsymbol{\nu},\boldsymbol{\mu})\in\mathcal{G}\times\mathcal{H}}{\text{ess-sup}} J_{\boldsymbol{\nu},\boldsymbol{\mu}}(s) \\ &= \hat{X}(s) \text{ a.s.} \end{aligned} \tag{B.198}$$

Taking $s = 0$ in (B.198), from Proposition 5.3.15(1),

$$E\left[H_{\bar{\boldsymbol{\nu}},\bar{\boldsymbol{\mu}}}(\bar{t})\hat{X}(\bar{t})\right] \leq E[\hat{X}(0)] = \hat{u} < \infty. \tag{B.199}$$

From the strict almost sure positivity of $H_{\bar{\boldsymbol{\nu}},\bar{\boldsymbol{\mu}}}(\bar{t})$ and (B.199), we have $\hat{X}(\bar{t}) < \infty$ a.s. Now,

$$\begin{aligned}\frac{1}{H_{\bar{\boldsymbol{\nu}},\bar{\boldsymbol{\mu}}}(s)} E\left[H_{\bar{\boldsymbol{\nu}},\bar{\boldsymbol{\mu}}}(\bar{t})\hat{X}(\bar{t}) \,\middle|\, \mathcal{F}_s\right] &= E\left[H_{\bar{\boldsymbol{\nu}},\bar{\boldsymbol{\mu}}}(s,\bar{t})\hat{X}(\bar{t}) \,\middle|\, \mathcal{F}_s\right] \\ &\leq \hat{X}(s) \text{ a.s.}\end{aligned} \tag{B.200}$$

for all $0 \leq s \leq \bar{t} \leq T$. Thus,

$$E\left[H_{\bar{\boldsymbol{\nu}},\bar{\boldsymbol{\mu}}}(\bar{t})\hat{X}(\bar{t}) \,\middle|\, \mathcal{F}_s\right] \leq H_{\bar{\boldsymbol{\nu}},\bar{\boldsymbol{\mu}}}(s)\hat{X}(s) \text{ a.s.} \tag{B.201}$$

for all $0 \leq s \leq \bar{t} \leq T$. Since $\bar{t}$ and $(\bar{\boldsymbol{\nu}}, \bar{\boldsymbol{\mu}}) \in \mathcal{G} \times \mathcal{H}$ were chosen arbitrarily, we have from (B.201) that the supermartingale property holds for all $(\boldsymbol{\nu}, \boldsymbol{\mu}) \in \mathcal{G} \times \mathcal{H}$. Thus,

$$H_{\boldsymbol{\nu},\boldsymbol{\mu}}\hat{X} \in \mathcal{SPM}(\{\mathcal{F}_t\}, P), \quad (\boldsymbol{\nu}, \boldsymbol{\mu}) \in \mathcal{G} \times \mathcal{H}. \tag{B.202}$$

□

Proof of Proposition 5.3.15(5). Again define

$$J_{\boldsymbol{\nu},\boldsymbol{\mu}}(t) \triangleq E\left[H_{\boldsymbol{\nu},\boldsymbol{\mu}}(t,T)B \,\middle|\, \mathcal{F}_t\right], \quad t \in [0,T], (\boldsymbol{\nu}, \boldsymbol{\mu}) \in \mathcal{G} \times \mathcal{H}, \tag{B.203}$$

where,

$$H_{\boldsymbol{\nu},\boldsymbol{\mu}}(s,t) \triangleq \frac{H_{\boldsymbol{\nu},\boldsymbol{\mu}}(t)}{H_{\boldsymbol{\nu},\boldsymbol{\mu}}(s)}, \quad 0 \leq s \leq t \leq T, (\boldsymbol{\nu}, \boldsymbol{\mu}) \in \mathcal{G} \times \mathcal{H}, \tag{B.204}$$

so that

$$\hat{X}(t) = \operatorname*{ess\text{-}sup}_{(\boldsymbol{\nu},\boldsymbol{\mu}) \in \mathcal{G} \times \mathcal{H}} J_{\boldsymbol{\nu},\boldsymbol{\mu}}(t), \quad t \in [0,T]. \tag{B.205}$$

Fix some s, t such that $0 \leq s \leq t \leq T$, then

$$\begin{aligned}E\left[H_{\boldsymbol{\nu},\boldsymbol{\mu}}(s,t)J_{\boldsymbol{\nu},\boldsymbol{\mu}}(t) \,\middle|\, \mathcal{F}_s\right] &= E\left[H_{\boldsymbol{\nu},\boldsymbol{\mu}}(s,t)E\left[H_{\boldsymbol{\nu},\boldsymbol{\mu}}(t,T)B \,\middle|\, \mathcal{F}_t\right] \,\middle|\, \mathcal{F}_s\right] \\ &= E\left[H_{\boldsymbol{\nu},\boldsymbol{\mu}}(s,T)B \,\middle|\, \mathcal{F}_s\right] \\ &= J_{\boldsymbol{\nu},\boldsymbol{\mu}}(s) \text{ a.s.}, \quad (\boldsymbol{\nu}, \boldsymbol{\mu}) \in \mathcal{G} \times \mathcal{H}.\end{aligned} \tag{B.206}$$

Using (B.205) we have

$$J_{\boldsymbol{\nu},\boldsymbol{\mu}}(t) \leq \hat{X}(t) \text{ a.s.}, \quad (\boldsymbol{\nu}, \boldsymbol{\mu}) \in \mathcal{G} \times \mathcal{H}. \tag{B.207}$$

Furthermore, from (B.205) and (B.206),

$$\hat{X}(s) = \operatorname*{ess\text{-}sup}_{(\boldsymbol{\nu},\boldsymbol{\mu}) \in \mathcal{G} \times \mathcal{H}} E\left[H_{\boldsymbol{\nu},\boldsymbol{\mu}}(s,t)J_{\boldsymbol{\nu},\boldsymbol{\mu}}(t) \,\middle|\, \mathcal{F}_s\right] \text{ a.s.} \tag{B.208}$$

From (B.207) and the non-negativity of $H_{\boldsymbol{\nu},\boldsymbol{\mu}}$,

$$H_{\boldsymbol{\nu},\boldsymbol{\mu}}(s,t)J_{\boldsymbol{\nu},\boldsymbol{\mu}}(t) \leq H_{\boldsymbol{\nu},\boldsymbol{\mu}}(s,t)\hat{X}(t) \text{ a.s.}, \quad (\boldsymbol{\nu},\boldsymbol{\mu}) \in \mathcal{G} \times \mathcal{H}. \tag{B.209}$$

Taking conditional expectations and essential suprema of both sides of (B.209) (which exists by Theorem C.13.2) and then using (B.208) yields

$$\hat{X}(s) \leq \operatorname*{ess-sup}_{(\boldsymbol{\nu},\boldsymbol{\mu})\in\mathcal{G}\times\mathcal{H}} E\left[H_{\boldsymbol{\nu},\boldsymbol{\mu}}(s,t)\hat{X}(t) \,\middle|\, \mathcal{F}_s\right] \text{ a.s.} \tag{B.210}$$

But using the supermartingale property of $H_{\boldsymbol{\nu},\boldsymbol{\mu}}\hat{X}$ from Proposition 5.3.15(4), we have

$$\hat{X}(s) \geq \operatorname*{ess-sup}_{(\boldsymbol{\nu},\boldsymbol{\mu})\in\mathcal{G}\times\mathcal{H}} E\left[H_{\boldsymbol{\nu},\boldsymbol{\mu}}(s,t)\hat{X}(t) \,\middle|\, \mathcal{F}_s\right] \text{ a.s.} \tag{B.211}$$

Combining (B.211) and (B.210), we conclude

$$\begin{aligned} \hat{X}(s) &= \operatorname*{ess-sup}_{(\boldsymbol{\nu},\boldsymbol{\mu})\in\mathcal{G}\times\mathcal{H}} E\left[H_{\boldsymbol{\nu},\boldsymbol{\mu}}(s,t)\hat{X}(t) \,\middle|\, \mathcal{F}_s\right] \\ &= \operatorname*{ess-sup}_{(\boldsymbol{\nu},\boldsymbol{\mu})\in\mathcal{G}\times\mathcal{H}} \frac{E[H_{\boldsymbol{\nu},\boldsymbol{\mu}}(t)\hat{X}(t)|\mathcal{F}_s]}{H_{\boldsymbol{\nu},\boldsymbol{\mu}}(s)} \text{ a.s.} \end{aligned} \tag{B.212}$$

□

Proof of Proposition 5.3.15(6). Fix $(\bar{\boldsymbol{\nu}},\bar{\boldsymbol{\mu}}) \in \mathcal{G}\times\mathcal{H}$. Define

$$z_{\bar{\boldsymbol{\nu}},\bar{\boldsymbol{\mu}}}(t) \triangleq \begin{cases} H_{\bar{\boldsymbol{\nu}},\bar{\boldsymbol{\mu}}}(t)\hat{X}(t), & t \in [0,T] \\ H_{\bar{\boldsymbol{\nu}},\bar{\boldsymbol{\mu}}}(T)\hat{X}(T), & t \in (T,\infty) \end{cases}, \tag{B.213}$$

i.e.

$$z_{\bar{\boldsymbol{\nu}},\bar{\boldsymbol{\mu}}}(t) \triangleq H_{\bar{\boldsymbol{\nu}},\bar{\boldsymbol{\mu}}}(t\wedge T)\hat{X}(t\wedge T), \quad t \in [0,\infty). \tag{B.214}$$

Define $\mathcal{F}_t \triangleq \mathcal{F}_T$ for all $t \in [T,\infty)$. From Proposition 5.3.15(4), we have that

$$z_{\bar{\boldsymbol{\nu}},\bar{\boldsymbol{\mu}}} \in \mathcal{SPM}(\{\mathcal{F}_t\}, P). \tag{B.215}$$

From §1.4 of Chung [6], there exists a set $\Omega^*(\bar{\boldsymbol{\nu}},\bar{\boldsymbol{\mu}}) \in \mathcal{F}_\infty \equiv \mathcal{F}_T$ such that $P(\Omega^*(\bar{\boldsymbol{\nu}},\bar{\boldsymbol{\mu}})) = 1$ and

$$z^+_{\bar{\boldsymbol{\nu}},\bar{\boldsymbol{\mu}}}(t,\omega) \triangleq \lim_{\substack{s\in\mathbb{Q} \\ s\downarrow\downarrow t}} z_{\bar{\boldsymbol{\nu}},\bar{\boldsymbol{\mu}}}(s,\omega) \tag{B.216}$$

exists in $\mathbb{R}$ for each $t \in [0,\infty)$ and $\omega \in \Omega^*(\bar{\boldsymbol{\nu}},\bar{\boldsymbol{\mu}})$.

Remark B.4.2. Equation (B.216) says that for each $\omega \in \Omega^*(\bar{\boldsymbol{\nu}}, \bar{\boldsymbol{\mu}})$ and $t \in [0, \infty)$, there exists some $z^+_{\bar{\boldsymbol{\nu}},\bar{\boldsymbol{\mu}}}(t, \omega) \in \mathbb{R}$ with the following property: for each $\epsilon > 0$ there is some $\delta = \delta(\epsilon, t, \omega) > 0$ such that

$$|z^+_{\bar{\boldsymbol{\nu}},\bar{\boldsymbol{\mu}}}(t, \omega) - z_{\bar{\boldsymbol{\nu}},\bar{\boldsymbol{\mu}}}(s, \omega)| < \epsilon \tag{B.217}$$

for all $s \in \mathbb{Q} \cap (t, t + \delta)$.

We have that $t \mapsto z^+_{\bar{\boldsymbol{\nu}},\bar{\boldsymbol{\mu}}}(t, \omega) : [0, T) \mapsto \mathbb{R}$ is càdlàg for each $\omega \in \Omega^*(\bar{\boldsymbol{\nu}}, \bar{\boldsymbol{\mu}})$. Since $\hat{X}(T) = B$ a.s., we know

$$z^+_{\bar{\boldsymbol{\nu}},\bar{\boldsymbol{\mu}}}(t, \omega) = H_{\bar{\boldsymbol{\nu}},\bar{\boldsymbol{\mu}}}(T, \omega)B(\omega), \quad t \in [T, \infty) \quad \omega \in \Omega^*(\bar{\boldsymbol{\nu}}, \bar{\boldsymbol{\mu}}). \tag{B.218}$$

Put

$$z^+_{\bar{\boldsymbol{\nu}},\bar{\boldsymbol{\mu}}}(t, \omega) \triangleq 0, \quad t \in [0, \infty) \quad \omega \notin \Omega^*(\bar{\boldsymbol{\nu}}, \bar{\boldsymbol{\mu}}). \tag{B.219}$$

Thus, due to (B.218) and (B.219), $t \mapsto z^+_{\bar{\boldsymbol{\nu}},\bar{\boldsymbol{\mu}}}(t, \omega) : [0, \infty) \mapsto \mathbb{R}$ is càdlàg for each $\omega \in \Omega$ (see II(65.1) in Rogers and Williams [42]). Define

$$\mathcal{G}_t \triangleq \sigma\{\mathcal{F}_{t+}, \mathcal{N}(P)\}. \tag{B.220}$$

Since $z_{\bar{\boldsymbol{\nu}},\bar{\boldsymbol{\mu}}}$ is $\{\mathcal{F}_t\}$-adapted, from Lemma II (66.1) of Rogers and Williams [43], $z^+_{\bar{\boldsymbol{\nu}},\bar{\boldsymbol{\mu}}}$ is $\{\mathcal{G}_t\}$-adapted. But since $\mathcal{N}(P) \in \mathcal{F}_t$ and $\mathcal{F}_{t+} = \mathcal{F}_t$ (from Remark 4.1.2), we therefore have $\mathcal{G}_t = \mathcal{F}_t$. Thus, $z^+_{\bar{\boldsymbol{\nu}},\bar{\boldsymbol{\mu}}}$ is $\{\mathcal{F}_t\}$-adapted. From Lemma II (66.2) of Rogers and Williams [43], and (B.215)

$$z^+_{\bar{\boldsymbol{\nu}},\bar{\boldsymbol{\mu}}} \in \mathcal{SPM}(\{\mathcal{G}_t\}, P), \tag{B.221}$$

and therefore, $z^+_{\bar{\boldsymbol{\nu}},\bar{\boldsymbol{\mu}}} \in \mathcal{SPM}(\{\mathcal{F}_t\}, P)$ and is càdlàg.

From (B.213)

$$\hat{X}(t) = \frac{1}{H_{\bar{\boldsymbol{\nu}},\bar{\boldsymbol{\mu}}}(t)} z_{\bar{\boldsymbol{\nu}},\bar{\boldsymbol{\mu}}}(t) \text{ a.s.,} \quad t \in [0, T]. \tag{B.222}$$

Put

$$\hat{X}^+_{\bar{\boldsymbol{\nu}},\bar{\boldsymbol{\mu}}}(t) \triangleq \frac{1}{H_{\bar{\boldsymbol{\nu}},\bar{\boldsymbol{\mu}}}(t)} z^+_{\bar{\boldsymbol{\nu}},\bar{\boldsymbol{\mu}}}(t), \quad t \in [0, T]. \tag{B.223}$$

Now $t \mapsto \frac{1}{H_{\bar{\boldsymbol{\nu}},\bar{\boldsymbol{\mu}}}(t,\omega)} : [0, T] \mapsto (0, \infty)$ is càdlàg for each $\omega \in \Omega$. Since $z^+_{\bar{\boldsymbol{\nu}},\bar{\boldsymbol{\mu}}}(\cdot, \omega)$ is also càdlàg for each $\omega \in \Omega$, we have that

$$t \to \hat{X}^+_{\bar{\boldsymbol{\nu}},\bar{\boldsymbol{\mu}}}(t, \omega) : [0, T] \to \mathbb{R} \text{ is càdlàg for each } \omega \in \Omega. \tag{B.224}$$

Moreover, $\frac{1}{H_{\bar{\boldsymbol{\nu}},\bar{\boldsymbol{\mu}}}}$ is $\{\mathcal{F}_t\}$-adapted and $z^+_{\bar{\boldsymbol{\nu}},\bar{\boldsymbol{\mu}}}$ is $\{\mathcal{F}_t\}$-adapted meaning that $\hat{X}^+_{\bar{\boldsymbol{\nu}},\bar{\boldsymbol{\mu}}}$ is $\{\mathcal{F}_t\}$-adapted. As a result

$$H_{\bar{\boldsymbol{\nu}},\bar{\boldsymbol{\mu}}} \hat{X}^+_{\bar{\boldsymbol{\nu}},\bar{\boldsymbol{\mu}}} \in \mathcal{SPM}(\{\mathcal{F}_t\}, P) \tag{B.225}$$

and is a càdlàg process where

$$\hat{X}^{+}_{\bar{\boldsymbol{\nu}},\bar{\boldsymbol{\mu}}}(T) = B \text{ a.s.} \tag{B.226}$$

From the supermartingale property of $H_{\bar{\boldsymbol{\nu}},\bar{\boldsymbol{\mu}}}\hat{X}^{+}_{\bar{\boldsymbol{\nu}},\bar{\boldsymbol{\mu}}}$ in (B.225),

$$\begin{aligned} H_{\bar{\boldsymbol{\nu}},\bar{\boldsymbol{\mu}}}(t)\hat{X}^{+}_{\bar{\boldsymbol{\nu}},\bar{\boldsymbol{\mu}}}(t) &\geq E\left[H_{\bar{\boldsymbol{\nu}},\bar{\boldsymbol{\mu}}}(T)B \,\middle|\, \mathcal{F}_t\right] \text{ a.s.} \\ \hat{X}^{+}_{\bar{\boldsymbol{\nu}},\bar{\boldsymbol{\mu}}}(t) &\geq E\left[H_{\bar{\boldsymbol{\nu}},\bar{\boldsymbol{\mu}}}(t,T)B \,\middle|\, \mathcal{F}_t\right] \text{ a.s.}, \quad t \in [0,T]. \end{aligned} \tag{B.227}$$

From (5.37)

$$\hat{X}(t) \geq \frac{E\left[H_{\boldsymbol{\nu},\boldsymbol{\mu}}(T)B \,\middle|\, \mathcal{F}_t\right]}{H_{\boldsymbol{\nu},\boldsymbol{\mu}}(t)} \text{ a.s.}, \quad t \in [0,T], (\boldsymbol{\nu},\boldsymbol{\mu}) \in \mathcal{G} \times \mathcal{H}. \tag{B.228}$$

Then

$$\begin{aligned} z_{\bar{\boldsymbol{\nu}},\bar{\boldsymbol{\mu}}}(t) &\triangleq H_{\bar{\boldsymbol{\nu}},\bar{\boldsymbol{\mu}}}(t)\hat{X}(t) \\ &\geq H_{\bar{\boldsymbol{\nu}},\bar{\boldsymbol{\mu}}}(t)\frac{E\left[H_{\boldsymbol{\nu},\boldsymbol{\mu}}(T)B \,\middle|\, \mathcal{F}_t\right]}{H_{\boldsymbol{\nu},\boldsymbol{\mu}}(t)} \text{ a.s.}, \quad t \in [0,T], (\boldsymbol{\nu},\boldsymbol{\mu}) \in \mathcal{G} \times \mathcal{H}. \end{aligned} \tag{B.229}$$

Now fix a sequence $\{t_n\} \in \mathbb{Q}$ for $n = 1, 2, \ldots$ such that $t < \ldots < t_{n+1} < t_n < \ldots$ and $\lim_{n\to\infty} t_n = t$. Then, by corollary II(2.4) in Revuz and Yor [39]

$$\begin{aligned} \lim_{\substack{s\in\mathbb{Q} \\ s\downarrow\downarrow t}} E\left[BH_{\boldsymbol{\nu},\boldsymbol{\mu}}(T) \,\middle|\, \mathcal{F}_s\right] &= \lim_{n\to\infty} E\left[BH_{\boldsymbol{\nu},\boldsymbol{\mu}}(T) \,\middle|\, \mathcal{F}_{t_n}\right] \\ &= E\left[BH_{\boldsymbol{\nu},\boldsymbol{\mu}}(T) \,\middle|\, \mathcal{F}_{t+}\right] \\ &= E\left[BH_{\boldsymbol{\nu},\boldsymbol{\mu}}(T) \,\middle|\, \mathcal{F}_t\right] \text{ a.s.}, \quad t \in [0,T), (\boldsymbol{\nu},\boldsymbol{\mu}) \in \mathcal{G} \times \mathcal{H}. \end{aligned} \tag{B.230}$$

Moreover, since

$$t \mapsto \frac{H_{\bar{\boldsymbol{\nu}},\bar{\boldsymbol{\mu}}}(t,\omega)}{H_{\boldsymbol{\nu},\boldsymbol{\mu}}(t,\omega)} : [0,T] \mapsto (0,\infty) \tag{B.231}$$

is càdlàg on $[0,T)$ for each $\omega \in \Omega$ and $(\boldsymbol{\nu},\boldsymbol{\mu}) \in \mathcal{G} \times \mathcal{H}$, then

$$\lim_{\substack{s\in\mathbb{Q} \\ s\downarrow\downarrow t}} \frac{H_{\bar{\boldsymbol{\nu}},\bar{\boldsymbol{\mu}}}(s)}{H_{\boldsymbol{\nu},\boldsymbol{\mu}}(s)} = \frac{H_{\bar{\boldsymbol{\nu}},\bar{\boldsymbol{\mu}}}(t)}{H_{\boldsymbol{\nu},\boldsymbol{\mu}}(t)}, \quad t \in [0,T), (\boldsymbol{\nu},\boldsymbol{\mu}) \in \mathcal{G} \times \mathcal{H}. \tag{B.232}$$

As a result, from (B.216), (B.229), (B.230), and (B.232)

$$\begin{aligned}
z^+_{\bar{\boldsymbol{\nu}},\bar{\boldsymbol{\mu}}}(t) &= \lim_{n\to\infty} z_{\bar{\boldsymbol{\nu}},\bar{\boldsymbol{\mu}}}(t_n) \\
&\geq \lim_{n\to\infty} \frac{H_{\bar{\boldsymbol{\nu}},\bar{\boldsymbol{\mu}}}(t_n)}{H_{\boldsymbol{\nu},\boldsymbol{\mu}}(t_n)} \lim_{n\to\infty} E\left[BH_{\boldsymbol{\nu},\boldsymbol{\mu}}(T)\,\middle|\,\mathcal{F}_{t_n}\right] \\
&= \frac{H_{\bar{\boldsymbol{\nu}},\bar{\boldsymbol{\mu}}}(t)}{H_{\boldsymbol{\nu},\boldsymbol{\mu}}(t)} E\left[BH_{\boldsymbol{\nu},\boldsymbol{\mu}}(T)\,\middle|\,\mathcal{F}_t\right] \\
&= H_{\bar{\boldsymbol{\nu}},\bar{\boldsymbol{\mu}}}(t) E\left[BH_{\boldsymbol{\nu},\boldsymbol{\mu}}(t,T)\,\middle|\,\mathcal{F}_t\right] \text{ a.s.,} \quad t \in [0,T), (\boldsymbol{\nu},\boldsymbol{\mu}) \in \mathcal{G}\times\mathcal{H}.
\end{aligned} \tag{B.233}$$

Thus

$$\hat{X}^+_{\bar{\boldsymbol{\nu}},\bar{\boldsymbol{\mu}}}(t) \triangleq \frac{z^+_{\bar{\boldsymbol{\nu}},\bar{\boldsymbol{\mu}}}(t)}{H_{\bar{\boldsymbol{\nu}},\bar{\boldsymbol{\mu}}}(t)} \geq E\left[BH_{\boldsymbol{\nu},\boldsymbol{\mu}}(t,T)\,\middle|\,\mathcal{F}_t\right] \text{ a.s.,} \quad t \in [0,T), (\boldsymbol{\nu},\boldsymbol{\mu}) \in \mathcal{G}\times\mathcal{H}. \tag{B.234}$$

From the definition $\hat{X}$ (see (5.37)) and (B.234), we conclude

$$\hat{X}^+_{\bar{\boldsymbol{\nu}},\bar{\boldsymbol{\mu}}}(t) \geq \hat{X}(t) \text{ a.s.,} \quad t \in [0,T). \tag{B.235}$$

To show the opposite inequality, fix $t \in [0,T)$ and $(\boldsymbol{\nu},\boldsymbol{\mu}) \in \mathcal{G}\times\mathcal{H}$ and sequence $\{t_n\} \in \mathbb{Q}$ such that $t < \ldots t_{n+1} < t_n < \ldots \leq T$ and $\lim_{n\to\infty} t_n = t$. Since $\hat{X}^+_{\bar{\boldsymbol{\nu}},\bar{\boldsymbol{\mu}}}$ is càdlàg from (B.224),

$$\hat{X}^+_{\bar{\boldsymbol{\nu}},\bar{\boldsymbol{\mu}}}(t) = \lim_{n\to\infty} \hat{X}^+_{\bar{\boldsymbol{\nu}},\bar{\boldsymbol{\mu}}}(t_n) \text{ a.s.} \tag{B.236}$$

Similarly, since $H_{\boldsymbol{\nu},\boldsymbol{\mu}}$ is càdlàg

$$H_{\boldsymbol{\nu},\boldsymbol{\mu}}(t) = \lim_{n\to\infty} H_{\boldsymbol{\nu},\boldsymbol{\mu}}(t_n) \text{ a.s.} \tag{B.237}$$

Thus,

$$\lim_{n\to\infty} H_{\boldsymbol{\nu},\boldsymbol{\mu}}(t,t_n) = 1 \text{ a.s.} \tag{B.238}$$

From the non-negativity of $H_{\boldsymbol{\nu},\boldsymbol{\mu}}$ and from (B.222) we can write

$$H_{\boldsymbol{\nu},\boldsymbol{\mu}}(t,t_n)\hat{X}(t_n) = H_{\boldsymbol{\nu},\boldsymbol{\mu}}(t,t_n) \frac{z_{\bar{\boldsymbol{\nu}},\bar{\boldsymbol{\mu}}}(t_n)}{H_{\bar{\boldsymbol{\nu}},\bar{\boldsymbol{\mu}}}(t_n)} \text{ a.s.} \tag{B.239}$$

Again from (B.216) we have that

$$\lim_{n\to\infty} z_{\bar{\boldsymbol{\nu}},\bar{\boldsymbol{\mu}}}(t_n) = z^+_{\bar{\boldsymbol{\nu}},\bar{\boldsymbol{\mu}}}(t) \text{ a.s.} \tag{B.240}$$

and from the càdlàg property of $\frac{1}{H_{\bar{\boldsymbol{\nu}},\bar{\boldsymbol{\mu}}}}$

$$\lim_{n\to\infty} \frac{1}{H_{\bar{\boldsymbol{\nu}},\bar{\boldsymbol{\mu}}}(t_n)} = \frac{1}{H_{\bar{\boldsymbol{\nu}},\bar{\boldsymbol{\mu}}}(t)} \text{ a.s.} \tag{B.241}$$

Now, from (B.239), (B.238), (B.240) and (B.241),

$$\lim_{n\to\infty} H_{\boldsymbol{\nu},\boldsymbol{\mu}}(t,t_n)\hat{X}(t_n) = \frac{z^{+}_{\bar{\boldsymbol{\nu}},\bar{\boldsymbol{\mu}}}(t)}{H_{\bar{\boldsymbol{\nu}},\bar{\boldsymbol{\mu}}}(t)} = \hat{X}^{+}_{\bar{\boldsymbol{\nu}},\bar{\boldsymbol{\mu}}}(t) \text{ a.s.} \tag{B.242}$$

Since $\hat{X}^{+}_{\bar{\boldsymbol{\nu}},\bar{\boldsymbol{\mu}}}$ is $\{\mathcal{F}_t\}$-adapted, from (B.242) and Fatou's Lemma

$$\begin{aligned} \hat{X}^{+}_{\bar{\boldsymbol{\nu}},\bar{\boldsymbol{\mu}}}(t) &= E\left[\hat{X}^{+}_{\bar{\boldsymbol{\nu}},\bar{\boldsymbol{\mu}}}(t) \,\middle|\, \mathcal{F}_t\right] \\ &= E\left[\lim_{n\to\infty} H_{\boldsymbol{\nu},\boldsymbol{\mu}}(t,t_n)\hat{X}(t_n) \,\middle|\, \mathcal{F}_t\right] \\ &\le \lim_{n\to\infty} E\left[H_{\boldsymbol{\nu},\boldsymbol{\mu}}(t,t_n)\hat{X}(t_n) \,\middle|\, \mathcal{F}_t\right] \text{ a.s.} \end{aligned} \tag{B.243}$$

Now using the supermartingale property of $H_{\boldsymbol{\nu},\boldsymbol{\mu}}\hat{X}$, from Proposition 5.3.15(4), and the fact that $t < t_n$ for all $n = 1, 2, \ldots$,

$$E\left[H_{\boldsymbol{\nu},\boldsymbol{\mu}}(t_n)\hat{X}(t_n) \,\middle|\, \mathcal{F}_t\right] \le H_{\boldsymbol{\nu},\boldsymbol{\mu}}(t)\hat{X}(t) \text{ a.s.,} \tag{B.244}$$

$$E\left[H_{\boldsymbol{\nu},\boldsymbol{\mu}}(t,t_n)\hat{X}(t_n) \,\middle|\, \mathcal{F}_t\right] \le \hat{X}(t) \text{ a.s.} \tag{B.245}$$

Putting (B.245) into (B.243) gives

$$\hat{X}^{+}_{\bar{\boldsymbol{\nu}},\bar{\boldsymbol{\mu}}}(t) \le \hat{X}(t) \text{ a.s.,} \quad t \in [0,T). \tag{B.246}$$

Therefore, from (B.235) and (B.246), we have

$$\hat{X}(t) = \hat{X}^{+}_{\bar{\boldsymbol{\nu}},\bar{\boldsymbol{\mu}}}(t) \text{ a.s.,} \quad t \in [0,T). \tag{B.247}$$

Since $\hat{X}(T) = \hat{X}^{+}_{\bar{\boldsymbol{\nu}},\bar{\boldsymbol{\mu}}}(T) = B$ a.s. we have by (B.247),

$$\hat{X}(t) = \hat{X}^{+}_{\bar{\boldsymbol{\nu}},\bar{\boldsymbol{\mu}}}(t) \text{ a.s.,} \quad t \in [0,T], \tag{B.248}$$

and therefore, $\{\hat{X}(t), t \in [0,T]\}$ has a càdlàg modification.
□

Proof of Proposition 5.3.15(7): Fix some $(\boldsymbol{\nu}, \boldsymbol{\mu}) \in \mathcal{G} \times \mathcal{H}$. Since $H_{\boldsymbol{\nu},\boldsymbol{\mu}}\hat{X}$ is an $\{\mathcal{F}_t\}$-supermartingale, we have from the optional sampling theorem (Theorem C.3.13) that for any stopping time τ taking values in $[0, T]$

$$H_{\boldsymbol{\nu},\boldsymbol{\mu}}(\tau)\hat{X}(\tau) \geq E[H_{\boldsymbol{\nu},\boldsymbol{\mu}}(T)B|\mathcal{F}_\tau] \text{ a.s.} \tag{B.249}$$

By dividing both sides of (B.249) by $H_{\boldsymbol{\nu},\boldsymbol{\mu}}(\tau)$ and taking the essential supremum over $\mathcal{G}\times\mathcal{H}$ (which exists by Theorem C.13.2), we obtain

$$\hat{X}(\tau) \geq \operatorname*{ess-sup}_{(\boldsymbol{\nu},\boldsymbol{\mu})\in\mathcal{G}\times\mathcal{H}} \frac{E[H_{\boldsymbol{\nu},\boldsymbol{\mu}}(T)B|\mathcal{F}_\tau]}{H_{\boldsymbol{\nu},\boldsymbol{\mu}}(\tau)} \text{ a.s.} \tag{B.250}$$

To show the opposite inequality, define for any stopping time τ taking values in $[0, T]$,

$$\tilde{X}(\tau) \triangleq \operatorname*{ess-sup}_{(\boldsymbol{\nu},\boldsymbol{\mu})\in\mathcal{G}\times\mathcal{H}} \frac{E[H_{\boldsymbol{\nu},\boldsymbol{\mu}}(T)B|\mathcal{F}_\tau]}{H_{\boldsymbol{\nu},\boldsymbol{\mu}}(\tau)}. \tag{B.251}$$

By using Proposition 5.3.15(5), it is easy to see that for stopping times ρ and τ satisfying $0 \leq \rho \leq \tau \leq T$ almost surely, we have

$$\tilde{X}(\rho) = \operatorname*{ess-sup}_{(\boldsymbol{\nu},\boldsymbol{\mu})\in\mathcal{G}\times\mathcal{H}} \frac{E[H_{\boldsymbol{\nu},\boldsymbol{\mu}}(\tau)\tilde{X}(\tau)|\mathcal{F}_\rho]}{H_{\boldsymbol{\nu},\boldsymbol{\mu}}(\rho)} \text{ a.s.,} \tag{B.252}$$

and therefore,

$$H_{\boldsymbol{\nu},\boldsymbol{\mu}}(\rho)\tilde{X}(\rho) \geq E[H_{\boldsymbol{\nu},\boldsymbol{\mu}}(\tau)\tilde{X}(\tau)|\mathcal{F}_\rho] \text{ a.s.,} \quad (\boldsymbol{\nu}, \boldsymbol{\mu}) \in \mathcal{G} \times \mathcal{H}. \tag{B.253}$$

It is immediate that $\hat{X}(t) = \tilde{X}(t)$ a.s. for each deterministic $t \in [0, T]$. Furthermore, we also have that $\hat{X}(\tau) = \tilde{X}(\tau)$ a.s. for each stopping time τ taking only finitely many values. Let τ be an arbitrary stopping time taking values in $[0, T]$ and construct a sequence of stopping times $\{\tau_n\}$, for $n = 1, 2, \ldots$, each element of which takes only finitely many values and such that $\tau_n \downarrow \tau$ almost surely. The right continuity of $\hat{X}$, Fatou's lemma, and (B.253) imply,

$$\begin{aligned}
H_{\boldsymbol{\nu},\boldsymbol{\mu}}(\tau)\hat{X}(\tau) &= E[H_{\boldsymbol{\nu},\boldsymbol{\mu}}(\tau)\hat{X}(\tau)|\mathcal{F}_\tau] \\
&= E[\lim_{n\to\infty} H_{\boldsymbol{\nu},\boldsymbol{\mu}}(\tau_n)\hat{X}(\tau_n)|\mathcal{F}_\tau] \\
&\leq \lim_{n\to\infty} E[H_{\boldsymbol{\nu},\boldsymbol{\mu}}(\tau_n)\hat{X}(\tau_n)|\mathcal{F}_\tau] \\
&= \lim_{n\to\infty} E[H_{\boldsymbol{\nu},\boldsymbol{\mu}}(\tau_n)\tilde{X}(\tau_n)|\mathcal{F}_\tau] \\
&\leq H_{\boldsymbol{\nu},\boldsymbol{\mu}}(\tau)\tilde{X}(\tau) \text{ a.s.}
\end{aligned} \tag{B.254}$$

Dividing both sides of (B.254) by $H_{\boldsymbol{\nu},\boldsymbol{\mu}}(\tau)$ and using (B.251) gives

$$\hat{X}(\tau) \leq \operatorname*{ess-sup}_{(\boldsymbol{\nu},\boldsymbol{\mu})\in\mathcal{G}\times\mathcal{H}} \frac{E[H_{\boldsymbol{\nu},\boldsymbol{\mu}}(T)B|\mathcal{F}_\tau]}{H_{\boldsymbol{\nu},\boldsymbol{\mu}}(\tau)} \text{ a.s.} \tag{B.255}$$

Therefore, from (B.250) and (B.255),

$$\hat{X}(\tau) = \operatorname*{ess-sup}_{(\boldsymbol{\nu},\boldsymbol{\mu})\in\mathcal{G}\times\mathcal{H}} \frac{E[H_{\boldsymbol{\nu},\boldsymbol{\mu}}(T)B|\mathcal{F}_\tau]}{H_{\boldsymbol{\nu},\boldsymbol{\mu}}(\tau)} \text{ a.s.} \tag{B.256}$$

□

B.5 Optimal Cumulative Consumption in the Constrained Regime-Switching Market Model

Proof of Proposition 5.3.20.

Fix $\boldsymbol{\nu} \in \mathcal{G}$. By Proposition C.8.6 we have that the continuous local martingale parts of (5.60) and (5.62) are indistinguishable. Equating the continuous local martingale parts of (5.60) and (5.62) yields

$$\int_0^t \left(\frac{\boldsymbol{\Psi}_{\boldsymbol{\nu},0}(s)^\top}{H_{\boldsymbol{\nu},0}(s)} + \hat{X}(s)\boldsymbol{\theta}_{\boldsymbol{\nu}}(s)^\top\right) d\mathbf{W}(s) = \int_0^t \left(\frac{\boldsymbol{\Psi}_{\boldsymbol{\nu},\boldsymbol{\mu}}(s)^\top}{H_{\boldsymbol{\nu},\boldsymbol{\mu}}(s)} + \hat{X}(s)\boldsymbol{\theta}_{\boldsymbol{\nu}}(s)^\top\right) d\mathbf{W}(s), \tag{B.257}$$

for all $t \in [0,T]$ and $\boldsymbol{\mu} \in \mathcal{H}$, and therefore,

$$\frac{\boldsymbol{\Psi}_{\boldsymbol{\nu},0}(t)}{H_{\boldsymbol{\nu},0}(t)} = \frac{\boldsymbol{\Psi}_{\boldsymbol{\nu},\boldsymbol{\mu}}(t)}{H_{\boldsymbol{\nu},\boldsymbol{\mu}}(t)} \quad \lambda \otimes P \text{ a.e. on } [0,T]\times\Omega,\ \boldsymbol{\mu} \in \mathcal{H}. \tag{B.258}$$

Combining (5.60) and (5.62) with (B.257) and (B.258) results in

$$\begin{aligned} 0 = &\sum_{\substack{i,j\in S\\ i\neq j}} \int_0^t \left(\frac{\Gamma_{ij}^{\boldsymbol{\nu},\boldsymbol{\mu}}(s)}{H_{\boldsymbol{\nu},\boldsymbol{\mu}}(s-)} - \frac{\Gamma_{ij}^{\boldsymbol{\nu},0}(s)}{H_{\boldsymbol{\nu},0}(s-)} - \hat{X}(s-)\mu_{ij}(s)\right) dM_{ij}(s) \\ &+ \sum_{\substack{i,j\in S\\ i\neq j}} \int_0^t \left(\hat{X}(s-)\mu_{ij}(s) - \frac{\Gamma_{ij}^{\boldsymbol{\nu},\boldsymbol{\mu}}(s)}{H_{\boldsymbol{\nu},\boldsymbol{\mu}}(s-)}\right) \tilde{\mu}_{ij}(s) dR_{ij}(s) \\ &+ \sum_{\substack{i,j\in S\\ i\neq j}} \int_0^t \frac{\triangle A_{\boldsymbol{\nu},\boldsymbol{\mu}}(s)}{H_{\boldsymbol{\nu},\boldsymbol{\mu}}(s-)} \tilde{\mu}_{ij}(s) dR_{ij}(s) \\ &+ \int_0^t \frac{1}{H_{\boldsymbol{\nu},0}(s-)} dA_{\boldsymbol{\nu},0}(s) - \int_0^t \frac{1}{H_{\boldsymbol{\nu},\boldsymbol{\mu}}(s-)} dA_{\boldsymbol{\nu},\boldsymbol{\mu}}(s), \quad t \in [0,T], \boldsymbol{\mu} \in \mathcal{H} \end{aligned} \tag{B.259}$$

To move forward and help simplify (B.259), we state the following lemma:

Lemma B.5.1. *The process*

$$z_{ij}^{\boldsymbol{\nu},\boldsymbol{\mu}}(t) \triangleq \int_0^t \frac{\triangle A_{\boldsymbol{\nu},\boldsymbol{\mu}}(s)}{H_{\boldsymbol{\nu},\boldsymbol{\mu}}(s-)}\tilde{\mu}_{ij}(s)dR_{ij}(s) = 0, \quad t \in [0,T], \tag{B.260}$$

for all $i,j \in S, i \neq j$, *and* $(\boldsymbol{\nu},\boldsymbol{\mu}) \in \mathcal{G} \times \mathcal{H}$, *where* $\tilde{\mu}_{ij}$ *is given by* (5.61).

Proof. Fix $i,j \in S, i \neq j$ and $(\boldsymbol{\nu},\boldsymbol{\mu}) \in \mathcal{G} \times \mathcal{H}$. Define

$$G_{ij}(t) \triangleq \frac{\triangle A_{\boldsymbol{\nu},\boldsymbol{\mu}}(t)}{H_{\boldsymbol{\nu},\boldsymbol{\mu}}(t-)}\tilde{\mu}_{ij}(t), \quad t \in [0,T]. \tag{B.261}$$

We know from the Doob-Meyer decomposition in Lemma 5.3.16 that $A_{\boldsymbol{\nu},\boldsymbol{\mu}} \in \mathcal{P}^*$, and as a result, $\triangle A_{\boldsymbol{\nu},\boldsymbol{\mu}} \in \mathcal{P}^*$. We also have that $(H_{\boldsymbol{\nu},\boldsymbol{\mu}})_- \in \mathcal{P}^*$ and $\mu_{ij} \in \mathcal{P}^*$. Therefore,

$$G_{ij} \in \mathcal{P}^*. \tag{B.262}$$

Define the positive and negative parts of G_{ij} as follows

$$G_{ij}^+ \triangleq \max\{0, G_{ij}\} \tag{B.263}$$

and

$$G_{ij}^- \triangleq \max\{0, -G_{ij}\}. \tag{B.264}$$

We then see,

$$G_{ij} = G_{ij}^+ - G_{ij}^- \tag{B.265}$$

and from (B.262) both $G_{ij}^+, G_{ij}^- \in \mathcal{P}^*$. Now, substituting into $z_{ij}^{\boldsymbol{\nu},\boldsymbol{\mu}}$, we have

$$z_{ij}^{\boldsymbol{\nu},\boldsymbol{\mu}}(t) = \int_0^t G_{ij}^+(s)dR_{ij}(s) - \int_0^t G_{ij}^-(s)dR_{ij}(s), \quad t \in [0,T] \tag{B.266}$$

Now, using the fact $M_{ij} \triangleq R_{ij} - \tilde{R}_{ij}$, and Theorem A.1.7, we can write

$$\int_0^t G_{ij}^+(s)dR_{ij}(s) = \int_0^t G_{ij}^+(s)dM_{ij}(s) + \int_0^t G_{ij}^+(s)d\tilde{R}_{ij}(s), \quad t \in [0,T]. \tag{B.267}$$

For any fixed $\omega \in \Omega$, from (B.261),

$$\left\{0 \leq s \leq T \middle| G_{ij}^+(s,\omega) \neq 0\right\} \subseteq \left\{0 \leq s \leq T \middle| \triangle A_{\boldsymbol{\nu},\boldsymbol{\mu}}(s,\omega) \neq 0\right\}. \tag{B.268}$$

Since $A_{\boldsymbol{\nu},\boldsymbol{\mu}}(\cdot,\omega)$ is càdlàg and non-decreasing for each $\omega \in \Omega$, the set

$$\left\{0 \leq s \leq T \middle| \triangle A_{\boldsymbol{\nu},\boldsymbol{\mu}}(s,\omega) \neq 0\right\} \tag{B.269}$$

is countable for each $\omega \in \Omega$. From (B.268), we also have that

$$\left\{0 \leq s \leq T \middle| G^{+}_{ij}(s,\omega) \neq 0\right\} \tag{B.270}$$

is countable for each $\omega \in \Omega$. From Definition 4.1.9, $\tilde{R}_{ij}(\cdot,\omega)$ is continuous on $[0,T]$ for each $\omega \in \Omega$, and thus

$$\int_0^{\cdot} G^{+}_{ij}(s,\omega)d\tilde{R}_{ij}(s,\omega) = 0, \quad \omega \in \Omega. \tag{B.271}$$

From (B.267) and (B.271),

$$\int_0^t G^{+}_{ij}(s)dR_{ij}(s) = \int_0^t G^{+}_{ij}(s)dM_{ij}(s), \quad t \in [0,T]. \tag{B.272}$$

Since $M_{ij} \in \mathcal{M}_0(\{\mathcal{F}_t\},P)$ we also have,

$$\int_0^{\cdot} G^{+}_{ij}(s)dR_{ij}(s) \in \mathcal{M}_{loc,0}(\{\mathcal{F}_t\},P). \tag{B.273}$$

Moreover, since G^{+}_{ij} is non-negative,

$$\int_0^t G^{+}_{ij}(s)dR_{ij}(s) \geq 0, \quad t \in [0,T], \tag{B.274}$$

because R_{ij} is a non-decreasing process. Since $\int_0^{\cdot} G^{+}_{ij}(s)dR_{ij}(s)$ is non-negative from (B.274), a local martingale from (B.273), and initially takes the value 0, we have from basic measure theory that

$$\int_0^t G^{+}_{ij}(s)dR_{ij}(s) = 0, \quad t \in [0,T], \tag{B.275}$$

and similarly,

$$\int_0^t G^{-}_{ij}(s)dR_{ij}(s) = 0, \quad t \in [0,T]. \tag{B.276}$$

Finally, from (B.266), (B.275) and (B.276),

$$z^{\boldsymbol{\nu},\boldsymbol{\mu}}_{ij}(t) = 0, \quad t \in [0,T], \quad i,j \in S, i \neq j. \tag{B.277}$$

□

Lemma B.5.2. *The following equality holds,*

$$\int_0^t \frac{1}{H_{\boldsymbol{\nu},0}(s-)} dA_{\boldsymbol{\nu},0}(s) = \int_0^t \frac{1}{H_{\boldsymbol{\nu},\boldsymbol{\mu}}(s-)} dA_{\boldsymbol{\nu},\boldsymbol{\mu}}(s) + \sum_{\substack{i,j\in S\\ i\neq j}} \int_0^t \frac{\Gamma_{ij}^{\boldsymbol{\nu},0}(s)}{H_{\boldsymbol{\nu},0}(s-)} \mu_{ij}(s) d\tilde{R}_{ij}(s), \quad t\in[0,T], (\boldsymbol{\nu},\boldsymbol{\mu})\in\mathcal{G}\times\mathcal{H}, \tag{B.278}$$

where $A_{\boldsymbol{\nu},\boldsymbol{\mu}}$ and $\Gamma_{ij}^{\boldsymbol{\nu},0}$, $(\boldsymbol{\nu},\boldsymbol{\mu})\in\mathcal{G}\times\mathcal{H}$, $i,j\in S, i\neq j$, are processes from the Doob-Meyer decomposition (Lemma 5.3.16).

Proof. Using Lemma B.5.1 with (B.259) we obtain

$$\begin{aligned} 0 = &\sum_{\substack{i,j\in S\\ i\neq j}} \int_0^t \left(\frac{\Gamma_{ij}^{\boldsymbol{\nu},\boldsymbol{\mu}}(s)}{H_{\boldsymbol{\nu},\boldsymbol{\mu}}(s-)} - \frac{\Gamma_{ij}^{\boldsymbol{\nu},0}(s)}{H_{\boldsymbol{\nu},0}(s-)} - \hat{X}(s-)\mu_{ij}(s)\right) dM_{ij}(s) \\ &+ \sum_{\substack{i,j\in S\\ i\neq j}} \int_0^t \left(\hat{X}(s-)\mu_{ij}(s) - \frac{\Gamma_{ij}^{\boldsymbol{\nu},\boldsymbol{\mu}}(s)}{H_{\boldsymbol{\nu},\boldsymbol{\mu}}(s-)}\right) \tilde{\mu}_{ij}(s) dR_{ij}(s) \\ &+ \int_0^t \frac{1}{H_{\boldsymbol{\nu},0}(s-)} dA_{\boldsymbol{\nu},0}(s) - \int_0^t \frac{1}{H_{\boldsymbol{\nu},\boldsymbol{\mu}}(s-)} dA_{\boldsymbol{\nu},\boldsymbol{\mu}}(s), \quad t\in[0,T], (\boldsymbol{\nu},\boldsymbol{\mu})\in\mathcal{G}\times\mathcal{H}, \end{aligned} \tag{B.279}$$

where $\tilde{\mu}_{ij} \triangleq \mu_{ij}/(1+\mu_{ij})$. Let

$$B(t) \triangleq \sum_{\substack{i,j\in S\\ i\neq j}} \int_0^t \Phi_{ij}(s) dR_{ij}(s), \quad t\in[0,T], \tag{B.280}$$

where we have put

$$\Phi_{ij}(t) \triangleq \tilde{\mu}_{ij}(t)\left(\hat{X}(t-)\mu_{ij}(t) - \frac{\Gamma_{ij}^{\boldsymbol{\nu},\boldsymbol{\mu}}(t)}{H_{\boldsymbol{\nu},\boldsymbol{\mu}}(t-)}\right), \ t\in[0,T], (\boldsymbol{\nu},\boldsymbol{\mu})\in\mathcal{G}\times\mathcal{H}, i,j\in S, i\neq j. \tag{B.281}$$

Now, $(X)_- \in \mathcal{P}^*$ and $\Gamma_{ij}^{\boldsymbol{\nu},\boldsymbol{\mu}} \in \mathcal{P}^*$, therefore

$$\Phi_{ij} \in \mathcal{P}^*. \tag{B.282}$$

From (B.280) and (4.8),

$$
\begin{aligned}
B(t) &= \sum_{\substack{i,j\in S\\ i\neq j}} (\Phi_{ij} \bullet R_{ij})(t) \\
&= \sum_{\substack{i,j\in S\\ i\neq j}} (\Phi_{ij} \bullet (R_{ij} - \tilde{R}_{ij} + \tilde{R}_{ij}))(t) \\
&= \sum_{\substack{i,j\in S\\ i\neq j}} (\Phi_{ij} \bullet M_{ij})(t) + \sum_{\substack{i,j\in S\\ i\neq j}} (\Phi_{ij} \bullet \tilde{R}_{ij})(t), \quad t \in [0,T].
\end{aligned} \tag{B.283}
$$

Putting (B.283) into (B.279) yields

$$
\begin{aligned}
0 = &\sum_{\substack{i,j\in S\\ i\neq j}} \int_0^t \left(\frac{1}{1+\mu_{ij}(s)}\right)\left(\frac{\Gamma_{ij}^{\boldsymbol{\nu},\boldsymbol{\mu}}(s)}{H_{\boldsymbol{\nu},\boldsymbol{\mu}}(s-)} - \hat{X}(s-)\mu_{ij}(s)\right) dM_{ij}(s) \\
&- \sum_{\substack{i,j\in S\\ i\neq j}} \int_0^t \frac{\Gamma_{ij}^{\boldsymbol{\nu},0}(s)}{H_{\boldsymbol{\nu},0}(s-)} dM_{ij}(s) + \sum_{\substack{i,j\in S\\ i\neq j}} \int_0^t \left(\hat{X}(s-)\mu_{ij}(s) - \frac{\Gamma_{ij}^{\boldsymbol{\nu},\boldsymbol{\mu}}(s)}{H_{\boldsymbol{\nu},\boldsymbol{\mu}}(s-)}\right)\tilde{\mu}_{ij}(s) d\tilde{R}_{ij}(s) \\
&+ \int_0^t \frac{1}{H_{\boldsymbol{\nu},0}(s-)} dA_{\boldsymbol{\nu},0}(s) - \int_0^t \frac{1}{H_{\boldsymbol{\nu},\boldsymbol{\mu}}(s-)} dA_{\boldsymbol{\nu},\boldsymbol{\mu}}(s), \quad t \in [0,T], \boldsymbol{\mu} \in \mathcal{H}.
\end{aligned} \tag{B.284}
$$

Define processes $\{F(t), t \in [0,T]\}$ and $\{K(t), t \in [0,T]\}$ as follows:

$$
\begin{aligned}
F(t) \triangleq &\sum_{\substack{i,j\in S\\ i\neq j}} \int_0^t \left(\frac{1}{1+\mu_{ij}(s)}\right)\left(\frac{\Gamma_{ij}^{\boldsymbol{\nu},\boldsymbol{\mu}}(s)}{H_{\boldsymbol{\nu},\boldsymbol{\mu}}(s-)} - \hat{X}(s-)\mu_{ij}(s)\right) dM_{ij}(s) \\
&- \sum_{\substack{i,j\in S\\ i\neq j}} \int_0^t \frac{\Gamma_{ij}^{\boldsymbol{\nu},0}(s)}{H_{\boldsymbol{\nu},0}(s-)} dM_{ij}(s), \quad t \in [0,T], (\boldsymbol{\nu},\boldsymbol{\mu}) \in \mathcal{G}\times\mathcal{H},
\end{aligned} \tag{B.285}
$$

and

$$
\begin{aligned}
K(t) \triangleq &\sum_{\substack{i,j\in S\\ i\neq j}} \int_0^t \left(\hat{X}(s-)\mu_{ij}(s) - \frac{\Gamma_{ij}^{\boldsymbol{\nu},\boldsymbol{\mu}}(s)}{H_{\boldsymbol{\nu},\boldsymbol{\mu}}(s-)}\right)\tilde{\mu}_{ij}(s) d\tilde{R}_{ij}(s) \\
&+ \int_0^t \frac{1}{H_{\boldsymbol{\nu},0}(s-)} dA_{\boldsymbol{\nu},0}(s) - \int_0^t \frac{1}{H_{\boldsymbol{\nu},\boldsymbol{\mu}}(s-)} dA_{\boldsymbol{\nu},\boldsymbol{\mu}}(s), \quad t \in [0,T], (\boldsymbol{\nu},\boldsymbol{\mu}) \in \mathcal{G}\times\mathcal{H}.
\end{aligned} \tag{B.286}
$$

It is quite easy to see from (B.285) that F is a local martingale. Since $A_{\boldsymbol{\nu},\boldsymbol{\mu}}$ is predictable for each $(\boldsymbol{\nu},\boldsymbol{\mu}) \in \mathcal{G} \times \mathcal{H}$ by Lemma 5.3.16, we have by Proposition C.9.4 and (B.286) that K is both predictable and has paths of finite variation. In summary,

$$F \in \mathcal{M}_{loc,0}(\{\mathcal{F}_t\}, P) \qquad K \in \mathcal{FV}_0(\{\mathcal{F}_t\}, P) \cap \mathcal{P}^*, \tag{B.287}$$

and from (B.284),

$$F(t) = -K(t) \quad t \in [0,T]. \tag{B.288}$$

From (B.287) and (B.288), we have

$$F \in \mathcal{M}_{loc,0}(\{\mathcal{F}_t\}, P) \cap \mathcal{FV}_0(\{\mathcal{F}_t\}, P) \cap \mathcal{P}^*. \tag{B.289}$$

From Proposition C.11.5 and (B.289), we have that

$$F(t) = 0 \text{ a.s.} \quad t \in [0,T]. \tag{B.290}$$

Now from (B.288) and (B.290)

$$K(t) = 0 \text{ a.s.} \quad t \in [0,T]. \tag{B.291}$$

In light of (B.290) and (B.285), setting the left hand side of (B.285) to zero and using Lemma 4.1.15 yields

$$\left(\frac{1}{1+\mu_{ij}(t)}\right)\left(\frac{\Gamma_{ij}^{\boldsymbol{\nu},\boldsymbol{\mu}}(t)}{H_{\boldsymbol{\nu},\boldsymbol{\mu}}(t-)} - \hat{X}(t-)\mu_{ij}(t)\right) = \frac{\Gamma_{ij}^{\boldsymbol{\nu},0}(t)}{H_{\boldsymbol{\nu},0}(t-)}, \quad \nu_{[M_{ij}]} - \text{ a.e.}, \tag{B.292}$$

for $i, j \in S, i \neq j$ and $(\boldsymbol{\nu},\boldsymbol{\mu}) \in \mathcal{G} \times \mathcal{H}$. Now, putting (B.292) into (B.286),

$$\begin{aligned} K(t) = &\int_0^t \frac{1}{H_{\boldsymbol{\nu},0}(s-)} dA_{\boldsymbol{\nu},0}(s) - \int_0^t \frac{1}{H_{\boldsymbol{\nu},\boldsymbol{\mu}}(s-)} dA_{\boldsymbol{\nu},\boldsymbol{\mu}}(s) \\ &- \sum_{\substack{i,j\in S\\ i\neq j}} \int_0^t \frac{\Gamma_{ij}^{\boldsymbol{\nu},0}(s)}{H_{\boldsymbol{\nu},0}(s-)} \mu_{ij}(s) d\tilde{R}_{ij}(s), \quad t \in [0,T], (\boldsymbol{\nu},\boldsymbol{\mu}) \in \mathcal{G} \times \mathcal{H}. \end{aligned} \tag{B.293}$$

Since $K(\cdot) = 0$ from (B.291),

$$\begin{aligned} \int_0^t \frac{1}{H_{\boldsymbol{\nu},0}(s-)} dA_{\boldsymbol{\nu},0}(s) = &\int_0^t \frac{1}{H_{\boldsymbol{\nu},\boldsymbol{\mu}}(s-)} dA_{\boldsymbol{\nu},\boldsymbol{\mu}}(s) \\ &+ \sum_{\substack{i,j\in S\\ i\neq j}} \int_0^t \frac{\Gamma_{ij}^{\boldsymbol{\nu},0}(s)}{H_{\boldsymbol{\nu},0}(s-)} \mu_{ij}(s) d\tilde{R}_{ij}(s), \quad t \in [0,T], (\boldsymbol{\nu},\boldsymbol{\mu}) \in \mathcal{G} \times \mathcal{H}, \end{aligned} \tag{B.294}$$

which is the required result. □

To move forward, we state the following lemma which provides an upper-bound for the process $\Gamma^{\boldsymbol{\nu},0}_{ij}$.

Lemma B.5.3. *$\Gamma^{\boldsymbol{\nu},0}_{ij}(t) \leq 0$, $\nu_{[M_{ij}]}$-almost everywhere for $i, j \in S, i \neq j$ and $\boldsymbol{\nu} \in \mathcal{G}$.*

Proof. Fix some $k, l \in S$ where $k \neq l$. Put $\mu_{ij}(t) \triangleq 0$ for all $t \in [0, T]$ where $(i, j) \neq (k, l)$. From Lemma B.5.2 and the positivity of $A_{\boldsymbol{\nu},\boldsymbol{\mu}}$,

$$\int_0^t \frac{1}{H_{\boldsymbol{\nu},0}(s-)} dA_{\boldsymbol{\nu},0}(s) \geq \int_0^t \frac{\Gamma^{\boldsymbol{\nu},0}_{kl}(s)}{H_{\boldsymbol{\nu},0}(s-)} \mu_{kl}(s) d\tilde{R}_{kl}(s), \quad t \in [0, T]. \tag{B.295}$$

Put

$$U \triangleq \left\{ (t, \omega) \in [0, T] \times \Omega \middle| \Gamma^{\boldsymbol{\nu},0}_{kl}(t, \omega) > 0 \right\}. \tag{B.296}$$

Since $\Gamma^{\boldsymbol{\nu},0}_{kl}$ is $\{\mathcal{F}_t\}$-predictable,

$$U \in \mathcal{P}^*. \tag{B.297}$$

We would like to show that $\nu_{[M_{kl}]}[U] = 0$. Suppose

$$\nu_{[M_{kl}]}[U] > 0. \tag{B.298}$$

By the definition of Doléans measure from Notation 4.1.16,

$$\nu_{[M_{kl}]}[U] \triangleq E\left[\int_0^T I_U(t, \omega) R_{kl}(dt, \omega) \right], \tag{B.299}$$

from I.3.17(iii) of Jacod and Shiryayev [26], we also have

$$\nu_{[M_{kl}]}[U] \triangleq E\left[\int_0^T I_U(t, \omega) \tilde{R}_{kl}(dt, \omega) \right]. \tag{B.300}$$

From (B.298) and (B.300),

$$\int_\Omega \left[\int_0^T I_U(t, \omega) \tilde{R}_{kl}(dt, \omega) \right] P(d\omega) > 0. \tag{B.301}$$

Put

$$z(\omega) \triangleq \int_0^T I_U(t, \omega) \tilde{R}_{kl}(dt, \omega), \quad \omega \in \Omega \tag{B.302}$$

we then have by (B.301)

$$z(\omega) \geq 0 \text{ a.s.} \quad \text{and} \quad E[z] > 0. \tag{B.303}$$

Put

$$\Omega_1 \triangleq \left\{ \omega \in \Omega \middle| z(\omega) > 0 \right\}. \tag{B.304}$$

Since $E[z] > 0$, we have $P[\Omega_1] > 0$. Define a sequence $\{\mu_{kl}^{(n)}(t,\omega)\}$ for $n = 1, 2, \ldots$ as

$$\mu_{kl}^{(n)}(t,\omega) \triangleq n I_U(t,\omega) H_{\boldsymbol{\nu},0}(t-,\omega) \tag{B.305}$$

for all $(t,\omega) \in [0,T] \times \Omega$. Since each component of $\mu_{kl}^{(n)}(t,\omega)$ is predictable we have that $\mu_{kl}^{(n)} \in \mathcal{P}^*$ for each $n = 1, 2, \ldots$. Defining

$$\boldsymbol{\mu}^{(n)} \triangleq \begin{cases} \mu_{kl}^{(n)} & i = k, j = l \\ 0 & \text{for all other } i, j \in S, i \neq j \end{cases}, \tag{B.306}$$

we then have $\boldsymbol{\mu}^{(n)} \in \mathcal{P}^*$. Also from the elementary bound in (4.11) and since $H_{\boldsymbol{\nu},0}(t)$ is continuous on $t \in [0,T]$ with $\inf_{t\in[0,T]} H_{\boldsymbol{\nu},0}(t) > 0$ a.s., we have that $\mu_{kl}^{(n)} \in L_{loc}^{1/2}(R_{kl})$ and thus $\boldsymbol{\mu}^{(n)} \in \mathcal{H}$ for all $n = 1, 2, \ldots$. Putting (B.305) into (B.295) and taking $t = T$ yields

$$\int_0^T \frac{1}{H_{\boldsymbol{\nu},0}(s-,\omega)} dA_{\boldsymbol{\nu},0}(s,\omega) \geq n \int_0^T \Gamma_{kl}^{\boldsymbol{\nu},0}(s,\omega) I_U(s,\omega) \tilde{R}_{kl}(ds,\omega). \tag{B.307}$$

for $n = 1, 2, \ldots$ Define the set

$$\Theta(\omega) \triangleq \left\{ t \in [0,T] \middle| I_U(t,\omega) = 1 \right\} \tag{B.308}$$

for all $\omega \in \Omega$. Fix some $\bar{\omega} \in \Omega_1$. From (B.304) we have that $z(\bar{\omega}) > 0$ and therefore by (B.302) and (B.308)

$$\int_{\Theta(\bar{\omega})} \tilde{R}_{kl}(dt,\bar{\omega}) > 0. \tag{B.309}$$

Now from (B.296) and (B.308) we have that

$$\Gamma_{kl}^{\boldsymbol{\nu},0}(t,\bar{\omega}) > 0 \text{ for all } t \in \Theta(\bar{\omega}). \tag{B.310}$$

As a result of (B.309) and (B.310), we can conclude

$$\int_{\Theta(\bar{\omega})} \Gamma_{kl}^{\boldsymbol{\nu},0}(t,\bar{\omega}) \tilde{R}_{kl}(dt,\bar{\omega}) > 0. \tag{B.311}$$

Fixing some $\bar{\omega} \in \Omega_1$ in (B.307) yields

$$\int_0^T \frac{1}{H_{\boldsymbol{\nu},0}(s-,\bar{\omega})} dA_{\boldsymbol{\nu},0}(s,\bar{\omega}) \geq n \int_{\Theta(\bar{\omega})} \Gamma_{kl}^{\boldsymbol{\nu},0}(s,\bar{\omega}) \tilde{R}_{kl}(ds,\bar{\omega}) \tag{B.312}$$

for each $n = 1, 2, \ldots$. Putting (B.311) into (B.312) and taking $n \to \infty$ in (B.312) gives

$$\int_0^T \frac{1}{H_{\boldsymbol{\nu},0}(s-,\bar{\omega})} dA_{\boldsymbol{\nu},0}(s,\bar{\omega}) = \infty, \quad \bar{\omega} \in \Omega_1. \tag{B.313}$$

Since $H_{\boldsymbol{\nu},0}(t)$ is continuous on $t \in [0,T]$, therefore bounded for P-almost all ω, and $A_{\boldsymbol{\nu},0}(t)$ is a non-decreasing process on $t \in [0,T]$ with $A_{\boldsymbol{\nu},0}(T) < \infty$ a.s. from (5.44), there cannot be a set of non-zero probability in which $(\frac{1}{H_{\boldsymbol{\nu},0}} \bullet A_{\boldsymbol{\nu},0})(T)$ takes the value ∞. However, $P[\Omega_1] > 0$. As a result, there is a contradiction in (B.298). Since $k, l \in S$ were arbitrarily chosen, we therefore have

$$\Gamma_{ij}^{\boldsymbol{\nu},0} \leq 0 \quad \nu_{[M_{ij}]} - \text{ a.e.}, \quad \boldsymbol{\nu} \in \mathcal{G} \quad i,j \in S, i \neq j. \tag{B.314}$$

□

Now that we have established Lemma B.5.3, we can complete the proof of Proposition 5.3.20. From Lemma B.5.2 and the fact that $M_{ij} \triangleq R_{ij} - \tilde{R}_{ij}$,

$$\begin{aligned} &\int_0^t \frac{1}{H_{\boldsymbol{\nu},0}(s-)} dA_{\boldsymbol{\nu},0}(s) + \sum_{\substack{i,j\in S \\ i\neq j}} \int_0^t \frac{\Gamma_{ij}^{\boldsymbol{\nu},0}(s)}{H_{\boldsymbol{\nu},0}(s-)} \mu_{ij}(s) dM_{ij}(s) \\ &= \int_0^t \frac{1}{H_{\boldsymbol{\nu},\boldsymbol{\mu}}(s-)} dA_{\boldsymbol{\nu},\boldsymbol{\mu}}(s) + \sum_{\substack{i,j\in S \\ i\neq j}} \int_0^t \frac{\Gamma_{ij}^{\boldsymbol{\nu},0}(s)}{H_{\boldsymbol{\nu},0}(s-)} \mu_{ij}(s) dR_{ij}(s), \quad t \in [0,T], \boldsymbol{\mu} \in \mathcal{H}. \end{aligned} \tag{B.315}$$

Define a sequence $\{\mu_{ij}^{(m)}\}$ for $m = 2, 3, \ldots$ where

$$\mu_{ij}^{(m)}(t,\omega) \triangleq -1 + \frac{1}{m}, \quad (t,\omega) \in [0,T] \times \Omega. \tag{B.316}$$

It is easy to see that $\boldsymbol{\mu}^{(m)} \triangleq \{\mu_{ij}^{(m)}\}_{i,j\in S} \in \mathcal{H}$ for all $m = 2, 3, \ldots$

Also define the sequence of processes $\{\hat{c}_{\boldsymbol{\nu}}^{(m)}\}$, for $m = 2, 3, \ldots$,

$$\hat{c}_{\boldsymbol{\nu}}^{(m)}(t) \triangleq \int_0^t \frac{1}{H_{\boldsymbol{\nu},0}(s-)} dA_{\boldsymbol{\nu},0}(s) + \sum_{\substack{i,j\in S \\ i\neq j}} \int_0^t \frac{\Gamma_{ij}^{\boldsymbol{\nu},0}(s)}{H_{\boldsymbol{\nu},0}(s-)} \mu_{ij}^{(m)}(s) dM_{ij}(s), \quad t \in [0,T]. \tag{B.317}$$

From (B.316) and (B.315)

$$\hat{c}_{\boldsymbol{\nu}}^{(m)}(t) = \int_0^t \frac{1}{H_{\boldsymbol{\nu},\boldsymbol{\mu}^{(m)}}(s-)} dA_{\boldsymbol{\nu},\boldsymbol{\mu}^{(m)}}(s) + \sum_{\substack{i,j\in S\\ i\neq j}} \int_0^t \frac{(-1+\frac{1}{m})\Gamma_{ij}^{\boldsymbol{\nu},0}(s)}{H_{\boldsymbol{\nu},0}(s-)} dR_{ij}(s), \quad t\in[0,T], \tag{B.318}$$

for all $m = 2, 3, \ldots$. Since $A_{\boldsymbol{\nu},\boldsymbol{\mu}^{(m)}}$ and R_{ij} are non-decreasing, one sees from Lemma B.5.3 and (B.318) that, for all large integers m, one has

$$\hat{c}_{\boldsymbol{\nu}}^{(m)}(t_2) \geq \hat{c}_{\boldsymbol{\nu}}^{(m)}(t_1) \text{ a.s.} \tag{B.319}$$

for $T \geq t_2 \geq t_1 \geq 0$. Thus $\hat{c}_{\boldsymbol{\nu}}^{(m)}(t)$ is non-decreasing for all large values of m. Now from (B.317) and (B.316)

$$\hat{c}_{\boldsymbol{\nu}}^{(m)}(t) = \int_0^t \frac{1}{H_{\boldsymbol{\nu},0}(s-)} dA_{\boldsymbol{\nu},0}(s) - (1-\frac{1}{m}) \sum_{\substack{i,j\in S\\ i\neq j}} \int_0^t \frac{\Gamma_{ij}^{\boldsymbol{\nu},0}(s)}{H_{\boldsymbol{\nu},0}(s-)} dM_{ij}(s), \quad t\in[0,T], \tag{B.320}$$

and taking $m\to\infty$

$$\lim_{m\to\infty} \hat{c}_{\boldsymbol{\nu}}^{(m)}(t) = \int_0^t \frac{1}{H_{\boldsymbol{\nu},0}(s-)} dA_{\boldsymbol{\nu},0}(s) - \sum_{\substack{i,j\in S\\ i\neq j}} \int_0^t \frac{\Gamma_{ij}^{\boldsymbol{\nu},0}(s)}{H_{\boldsymbol{\nu},0}(s-)} dM_{ij}(s) \text{ a.s.}, \quad t\in[0,T]. \tag{B.321}$$

As a result, from (5.79),

$$\hat{c}_{\boldsymbol{\nu}}^{(m)}(t) \to \hat{c}_{\boldsymbol{\nu}}(t) \text{ a.s.} \quad \text{as} \quad m\to\infty \tag{B.322}$$

for all $t\in[0,T]$. By taking $m\to\infty$ in (B.319), and observing (B.322), we have that

$$\hat{c}_{\boldsymbol{\nu}}(\cdot) \text{ is non-decreasing.} \tag{B.323}$$

Now, it is immediate from (5.79) that

$$\hat{c}_{\boldsymbol{\nu}}(0) = 0, \tag{B.324}$$

and from (5.42) and the fact that $H_{\boldsymbol{\nu},0}$ is almost surely strictly positive,

$$\tfrac{\Gamma_{ij}^{\boldsymbol{\nu},0}}{(H_{\boldsymbol{\nu},0})_-} \in L_{loc}^{1/2}(R_{ij}) \text{ and } (\tfrac{1}{H_{\boldsymbol{\nu},0}} \bullet A_{\boldsymbol{\nu},0}) \in \mathcal{F}^*. \tag{B.325}$$

Therefore, we have

$$\hat{c}_{\boldsymbol{\nu}} \in \mathcal{F}^*. \tag{B.326}$$

Since $H_{\boldsymbol{\nu},0}$ is almost surely strictly positive and continuous, and $A_{\boldsymbol{\nu},0}(T) < \infty$ a.s. by Lemma 5.3.16, we have that the first term on the right-hand side of (5.79) is finite. Now since the second term on the right-hand side of (5.79) is a local martingale by (B.325) and Remark A.1.2, we know that it is finite as well. Therefore,

$$\hat{c}_{\boldsymbol{\nu}}(T) < \infty \text{ a.s.} \tag{B.327}$$

As a result, from (B.323), (B.324), (B.326), and (B.327), we can conclude

$$\hat{c}_{\boldsymbol{\nu}}(t) \triangleq \int_0^t \frac{1}{H_{\boldsymbol{\nu},0}(s-)} dA_{\boldsymbol{\nu},0}(s) - \sum_{\substack{i,j\in S \\ i\neq j}} \int_0^t \frac{\Gamma_{ij}^{\boldsymbol{\nu},0}(s)}{H_{\boldsymbol{\nu},0}(s-)} dM_{ij}(s) \in \mathcal{C}, \tag{B.328}$$

for all $\boldsymbol{\nu} \in \mathcal{G}$. □

Appendix C

Standard Definitions and Results

C.1 General Definitions and Conventions for Stochastic Processes

Definition C.1.1. A process $X = \{X(t) : t \in [0,T]\}$ is called *càdlàg* (continu à droite avec des limites à gauche) if the mappings $t \mapsto X(\omega, t)$ are right-continuous with finite left-hand limits on $[0, \infty)$ for all $\omega \in \Omega$. A process $X = \{X(t) : t \in [0,T]\}$ is called *càg* (continu à gauche) if the mappings $t \mapsto X(\omega, t)$ are left-continuous on $[0, \infty)$ for all $\omega \in \Omega$.

Remark C.1.2. If X is *càdlàg* then we define the process $(X)_- \triangleq \{X(t-) : t \in [0,T]\}$ as

$$X(0-) \triangleq X(0) \qquad \text{and} \qquad X(t-) \triangleq \lim_{\substack{s \to t \\ s < t}} X(s), \qquad \forall t \in (0,T] \tag{C.1}$$

and we also define the process $\triangle X(t) = \{\triangle X(t) : t \in [0,T]\}$ as

$$\triangle X(t) \triangleq X(t) - X(t-), \qquad \forall t \in [0,T]. \tag{C.2}$$

Definition C.1.3. The *raw filtration* $\{\mathcal{F}_t^X\}$ generated by the stochastic process $X = \{X(t) : t \in [0,T]\}$ is

$$\mathcal{F}_t^X \triangleq \sigma\{X(s) : s \in [0,t]\} \qquad \forall t \in [0,T]. \tag{C.3}$$

The following definition is from Karatzas and Shreve [30] Chapter 1, page 4:

Definition C.1.4. Let $\{\mathcal{F}_t\}_{t\in[0,T]}$ be a filtration. We define

$$\mathcal{F}_{t+} \triangleq \bigcap_{\epsilon>0} \mathcal{F}_{t+\epsilon} \tag{C.4}$$

to be the *σ-algebra of events immediately after* $t \in [0,T]$. We say that the filtration $\{\mathcal{F}_t\}$ is *right-continuous* if

$$\mathcal{F}_t = \mathcal{F}_{t+} \tag{C.5}$$

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

Definition C.1.5. A *filtered probability space* is a pair $((\Omega,\mathcal{F},P),\{\mathcal{F}_t\})$ consisting of a probability space $(\Omega,\mathcal{F},P)$ and filtration $\{\mathcal{F}_t\}_{t\in[0,T]}$ on $\mathcal{F}$.

Definition C.1.6. The *predictable σ-algebra* $\mathcal{P}$ is the σ-algebra generated by all $\{\mathcal{F}_t\}$-adapted càg processes.

Definition C.1.7. A process $X = \{X(t) : t \in [0,T]\}$ defined on a filtered probability space $((\Omega,\mathcal{F},P),\{\mathcal{F}_t\})$ is *non-decreasing* if the mappings $t \mapsto X(t,\omega)$ are non-decreasing on $[0,\infty)$ for all $\omega \in \Omega$.

Definition C.1.8. A collection of random variables $\mathcal{C}$ on a probability space $(\Omega,\mathcal{F},P)$ is *uniformly integrable* when the following holds: corresponding to each $\epsilon \in (0,\infty)$ there is some $c(\epsilon) \in [0,\infty)$ such that

$$\sup_{X\in\mathcal{C}} E[|X| I_{\{|X|\geq c\}}] < \epsilon, \tag{C.6}$$

for all $c \in [c(\epsilon),\infty)$.

Proposition C.1.9. *Suppose that $\mathcal{C}$ is a given collection of random variables defined on the probability space $(\Omega,\mathcal{F},P)$. If there is some non-negative random variable Y on Ω such that $E[Y] < \infty$ and $|X| \leq Y$ a.s. for all $X \in \mathcal{C}$, then $\mathcal{C}$ is uniformly integrable.*

Proposition C.1.10. *Supposed that X is a random variable on the probability space $(\Omega,\mathcal{F},P)$ such that $E|X| < \infty$, and $\{\mathcal{G}_\lambda, \lambda \in \Lambda\}$ is a collection of sub σ-algebras of $\mathcal{F}$. For each $\lambda \in \Lambda$ put $Y_\lambda \triangleq E[X|\mathcal{G}_\lambda]$. Then the collection of random variables $\{Y_\lambda, \lambda \in \Lambda\}$ is uniformly integrable.*

C.2 Stopping Time Results

Definition C.2.1. Suppose that $\{\mathcal{F}_t, t \in [0,T]\}$ is a given filtration on a probability space $(\Omega,\mathcal{F},P)$, then a mapping $\tau : (\Omega,\mathcal{F},P) \mapsto [0,T]$ is called a continuous-parameter *stopping time* with respect to $\{\mathcal{F}_t\}$ when $\{\tau \leq t\} \in \mathcal{F}_t$ for all $t \in [0,T]$.

Definition C.2.2. Suppose that $\tau : (\Omega, \mathcal{F}, P) \mapsto [0, T]$ is a stopping time with respect to the filtration $\{\mathcal{F}_t, t \in [0, T]\}$. We denote by $\mathcal{F}_\tau$ the collection of sets $A \subset \Omega$ having the property that $A \cap \{\tau \leq t\} \in \mathcal{F}_t$ for each $t \in [0, T]$. We call the σ-algebra $\mathcal{F}_\tau$ the *pre-σ-algebra* generated by τ.

Definition C.2.3. Let $\mathcal{S}$ be the set of all stopping times taking values in $[0, T]$. We say that an $\{\mathcal{F}_t\}$-adapted process Y is of *class* $\mathcal{D}[0, T]$ if the family of random variables $\{Y(\tau)\}_{\tau \in \mathcal{S}}$ is uniformly integrable.

C.3 Spaces of Martingales

Definition C.3.1. A real-valued, $\{\mathcal{F}_t\}$-adapted process $M = \{M(t) : t \in [0, T]\}$ on $(\Omega, \mathcal{F}, P)$ such that

- $E[|M(t)|] < \infty$ for all $t \in [0, T]$ and
- $E[M(t)|\mathcal{F}_s] = M(s)$ P-a.s., for all $0 \leq s \leq t \leq T$

is called an $\{\mathcal{F}_t\}$-*martingale*. We denote the set of all $\{\mathcal{F}_t\}$-martingales $\mathcal{M}((\Omega, \mathcal{F}, P), \{\mathcal{F}_t\})$.

If there is no ambiguity about the measurable space on which the space of martingales is defined, we use the notation $\mathcal{M}(\{\mathcal{F}_t\}, P)$ instead of $\mathcal{M}((\Omega, \mathcal{F}, P), \{\mathcal{F}_t\})$. This notation continues for the following spaces of processes.

Notation C.3.2. $\mathcal{M}_0(\{\mathcal{F}_t\}, P)$ denotes the set of $M \in \mathcal{M}(\{\mathcal{F}_t\}, P)$ which are P-a.s. null at the origin.

Notation C.3.3. $\mathcal{M}^c(\{\mathcal{F}_t\}, P)$ denotes the set of $M \in \mathcal{M}(\{\mathcal{F}_t\}, P)$ whose sample paths are continuous.

Notation C.3.4. $\mathcal{M}_0^c(\{\mathcal{F}_t\}, P)$ denotes the set of $M \in \mathcal{M}^c(\{\mathcal{F}_t\}, P)$ which are P-a.s. null at the origin.

Definition C.3.5. A martingale M is *square-integrable* if $E\left[|M(t)|^2\right] < \infty$, for all $t \in [0, T]$.

Definition C.3.6. A martingale M is L^2-*bounded* if $\sup_{t \in [0,T]} E\left[|M(t)|^2\right] < \infty$.

Remark C.3.7. As we are dealing with a finite time interval $[0, T]$, then a martingale M is square integrable if and only if it is L^2-bounded.

Notation C.3.8. $\mathcal{M}^2(\{\mathcal{F}_t\}, P)$ denotes the set of $M \in \mathcal{M}(\{\mathcal{F}_t\}, P)$ which are square integrable.

Notation C.3.9. $\mathcal{M}_0^2(\{\mathcal{F}_t\}, P)$ denotes the set of $M \in \mathcal{M}_0(\{\mathcal{F}_t\}, P)$ which are square integrable.

Definition C.3.10. A real-valued, $\{\mathcal{F}_t\}$-adapted process $M = \{M(t) : t \in [0, T]\}$ on $(\Omega, \mathcal{F}, P)$ such that

- $E[|M(t)|] < \infty$ for all $t \in [0, T]$ and
- $E[M(t)|\mathcal{F}_s] \leq M(s)$ P-a.s., for all $0 \leq s \leq t \leq T$

is called an $\{\mathcal{F}_t\}$-*supermartingale*. We denote the set of all $\{\mathcal{F}_t\}$-supermartingales $\mathcal{SPM}((\Omega, \mathcal{F}, P), \{\mathcal{F}_t\})$ or $\mathcal{SPM}(\{\mathcal{F}_t\}, P)$.

Definition C.3.11. A real-valued, $\{\mathcal{F}_t\}$-adapted process $M = \{M(t) : t \in [0, T]\}$ on $(\Omega, \mathcal{F}, P)$ such that

- $E[|M(t)|] < \infty$ for all $t \in [0, T]$ and
- $E[M(t)|\mathcal{F}_s] \geq M(s)$ P-a.s., for all $0 \leq s \leq t \leq T$

is called an $\{\mathcal{F}_t\}$-*submartingale*. We denote the set of all $\{\mathcal{F}_t\}$-submartingales $\mathcal{SBM}((\Omega, \mathcal{F}, P), \{\mathcal{F}_t\})$ or $\mathcal{SBM}(\{\mathcal{F}_t\}, P)$.

Remark C.3.12. If $M \in \mathcal{SBM}((\Omega, \mathcal{F}, P), \{\mathcal{F}_t\})$ and $M \in \mathcal{SPM}((\Omega, \mathcal{F}, P), \{\mathcal{F}_t\})$, then $M \in \mathcal{M}((\Omega, \mathcal{F}, P), \{\mathcal{F}_t\})$.

The following is the famous Optional Sampling Theorem by Joseph Doob.

Theorem C.3.13. *Suppose that $\{(X_t, \mathcal{F}_t), t \in [0, T]\}$ is a right-continuous supermartingale and τ and ρ are $\{\mathcal{F}_t\}$-stopping times on $(\Omega, \mathcal{F}, P)$ such that $\rho(\omega) \leq \tau(\omega) \leq T$ for all $\omega \in \Omega$. Then, $E[|X_\tau|] \leq \infty$ and $E[X_\tau|\mathcal{F}_\rho] \leq X_\rho$ a.s.*

C.4 Spaces of Local Martingales

Notation C.4.1. For a sequence $\{T^m\}_{m \in \mathbb{N}}$ of $\{\mathcal{F}_t\}$-stopping times, we write

$$T^m \uparrow T \quad P - a.s. \tag{C.7}$$

to mean that

- $0 \leq T^m(\omega) \leq T^{m+1}(\omega)$ for all $\omega \in \Omega$ for all $m \in \mathbb{N}$; and
- there exists $M(\omega) \in \mathbb{N}$ such that $T^m(\omega) = T$, for all $m \geq M(\omega)$ and for all $\omega \in \Omega$

If there is no ambiguity about the measure, then we will write $T^m \uparrow T$ a.s.

Definition C.4.2. A real-valued process $\{\mathcal{F}_t\}$-adapted process $M = \{M(t), t \in [0, T]\}$ on $(\Omega, \mathcal{F}, P)$ where there exists a sequence $\{T^m\}_{m \in \mathbb{N}}$ of $\{\mathcal{F}_t\}$-stopping times such that

- $T^m \uparrow T \quad P - a.s.$
- $\{M(t \wedge T^m), t \in [0, T]\} \in \mathcal{M}(\{\mathcal{F}_t\}, P)$ for each $m \in \mathbb{N}$,

is called an $\{\mathcal{F}_t\}$-*local martingale*. We denote the set of all $\{\mathcal{F}_t\}$-*local martingales* $\mathcal{M}_{loc}(\{\mathcal{F}_t\}, P)$. We say that the sequence $\{T^m\}_{m \in \mathbb{N}}$ of $\{\mathcal{F}_t\}$-stopping times is a *localizing sequence* for M.

Notation C.4.3. $\mathcal{M}_{loc,0}(\{\mathcal{F}_t\}, P)$ denotes the set of $M \in \mathcal{M}_{loc}(\{\mathcal{F}_t\}, P)$ which are P-a.s. null at the origin.

Notation C.4.4. $\mathcal{M}^2_{loc}(\{\mathcal{F}_t\}, P)$ denotes the set of $M \in \mathcal{M}_{loc}(\{\mathcal{F}_t\}, P)$ which are locally square integrable. We say that M is a *locally square-integrable local martingale.*

Notation C.4.5. $\mathcal{M}^c_{loc}(\{\mathcal{F}_t\}, P)$ denotes the set of $M \in \mathcal{M}_{loc}(\{\mathcal{F}_t\}, P)$ whose sample paths are continuous. We say that M is a *continuous local martingale.*

C.5 Spaces of Finite Variation Processes

Definition C.5.1. A process $\{A(t), t \in [0, T]\}$ is a process of *finite variation* if it is an $\{\mathcal{F}_t\}$-adapted, càdlàg process such that $t \mapsto A(t, \omega)$ is of finite variation for each $\omega \in \Omega$. In other words, for all $(t, \omega) \in [0, T] \times \Omega$ the variation $V_A(t, \omega)$ of $s \mapsto A(\omega, s)$ over $(0, t]$ is finite:

$$V_A(t, \omega) \triangleq \sup \sum_{i=1}^{n} |A(\omega, s_i) - A(\omega, s_{i-1})| < \infty. \tag{C.8}$$

The supremum is taken over all partitions $0 = s_0 < s_1 < \ldots < s_n = t$ of $[0, t]$.

Notation C.5.2. We denote by $\mathcal{FV}(\{\mathcal{F}_t\}, P)$ the set of all real-valued, $\{\mathcal{F}_t\}$-adapted, càdlàg processes A on $(\Omega, \mathcal{F}, P)$ which are of finite variation.

Notation C.5.3. We denote by $\mathcal{FV}_0(\{\mathcal{F}_t\}, P)$ the set of all $A \in \mathcal{FV}(\{\mathcal{F}_t\}, P)$ (Notation C.5.2) which are null at the origin.

Notation C.5.4. We denote by $\mathcal{A}^+(\{\mathcal{F}_t\}, P)$ the set of all real-valued, $\{\mathcal{F}_t\}$-adapted, càdlàg processes A on $(\Omega, \mathcal{F}, P)$ which are non-decreasing and integrable, i.e. $E[A(T)] < \infty$. When the filtration and probability measure are obvious, the notation $\mathcal{A}^+$ is used.

Notation C.5.5. We denote by $\mathcal{A}_0^+(\{\mathcal{F}_t\}, P)$ the set of all $A \in \mathcal{A}^+(\{\mathcal{F}_t\}, P)$ (Notation C.5.4) which are null at the origin. When the filtration and probability measure are obvious, the notation $\mathcal{A}_0^+$ is used.

Notation C.5.6. Let $\mathcal{FV}_{loc}(\{\mathcal{F}_t\}, P)$ denote the set of processes such that for each $A \in \mathcal{FV}_{loc}(\{\mathcal{F}_t\}, P)$ there exists a sequence of $\{\mathcal{F}_t\}$-stopping times $\{T^m\}_{m\in\mathbb{N}}$ (depending on A) such that $T^m \uparrow T \quad P-a.s.$ and each stopped process $A(t \wedge T^m) \in \mathcal{FV}(\{\mathcal{F}_t\}, P)$ (Notation C.5.2).

Notation C.5.7. Let $\mathcal{A}_{loc}^+(\{\mathcal{F}_t\}, P)$ denote the set of processes such that for each $A \in \mathcal{A}_{loc}^+(\{\mathcal{F}_t\}, P)$ there exists a sequence of $\{\mathcal{F}_t\}$-stopping times $\{T^m\}_{m\in\mathbb{N}}$ (depending on A) such that $T^m \uparrow T \quad P-a.s.$ and each stopped process $A(t \wedge T^m) \in \mathcal{A}^+(\{\mathcal{F}_t\}, P)$ (Notation C.5.4). When the filtration and probability measure are obvious, the notation $\mathcal{A}_{loc}^+$ is used.

C.6 Quadratic Co-variation and Variation Processes

The following theorem is from Jacod and Shiryaev [26], Theorem I.4.2

Theorem C.6.1. *For each pair $N, M \in \mathcal{M}_{loc}^2(\{\mathcal{F}_t\}, P)$, there exists a real-valued, càdlàg, $\{\mathcal{F}_t\}$-adapted, finite variation process $\langle N, M\rangle$, which is unique up to indistinguishability, such that*

1. *$\langle N, M\rangle(0) = 0$ a.s.*
2. *$\langle N, M\rangle$ is $\{\mathcal{F}_t\}$-predictable*
3. *$NM - \langle N, M\rangle \in \mathcal{M}_{loc}(\{\mathcal{F}_t\}, P)$.*

Moreover,

$$\langle N, M\rangle = \frac{1}{4}(\langle N+M, N+M\rangle - \langle N-M, N-M\rangle). \tag{C.9}$$

Remark C.6.2. We call $\langle N, M\rangle$ the *angle-bracket co-variation process* of N and M.

Remark C.6.3. For any $M \in \mathcal{M}_{loc}^2(\{\mathcal{F}_t\}, P)$, the process $\langle M, M\rangle$ is called the *angle-bracket quadratic variation process* of M. We often write $\langle M\rangle$ for $\langle M, M\rangle$.

From Jacod and Shiryaev [26], equation I.4.46 and Proposition I.4.50 and Rogers and Williams [43], Theorem VI.36.6 and Theorem VI.37.8, we have the following theorem.

Theorem C.6.4. *For each pair $N, M \in \mathcal{M}_{loc}(\{\mathcal{F}_t\}, P)$, there exists a càdlàg, $\{\mathcal{F}_t\}$-adapted process $[N, M]$ of finite variation, which is unique up to indistinguishability, such that*

1. $[N, M](0) = 0$ *a.s.*
2. $\triangle[N, M](t) = \triangle N(t)\triangle M(t)$ *for all* $t > 0$
3. $NM - [N, M] \in \mathcal{M}_{loc}(\{\mathcal{F}_t\}, P)$

Moreover,

$$[N, M] = \frac{1}{4}([N + M, N + M] - [N - M, N - M]). \tag{C.10}$$

Remark C.6.5. We call $[N, M]$ the *square-bracket quadratic co-variation process* of N and M.

Remark C.6.6. For any $M \in \mathcal{M}_{loc}(\{\mathcal{F}_t\}, P)$, the process $[M, M]$ is called the *square-bracket quadratic variation process* of M. We often write $[M]$ for $[M, M]$.

Remark C.6.7. The square-bracket quadratic co-variation process $[N, M]$ exists for all local martingales N, M. This is the main reason for preferring $[N, M]$ to $\langle N, M\rangle$; the angle-bracket co-variation process $\langle N, M\rangle$ only exists for locally square-integrable local martingales. Furthermore, the square-bracket quadratic variation process $[M]$ is invariant under absolutely continuous changes of measure (see Jacod and Shiryaev [26], Theorem III.3.13), unlike the angle-bracket quadratic variation process $\langle M\rangle$.

Remark C.6.8. If $M \in \mathcal{M}^c_{loc}(\{\mathcal{F}_t\}, P)$, then

$$[M] = \langle M\rangle. \tag{C.11}$$

C.7 Purely Discontinuous Local Martingales

From Jacod and Shiryaev [26], Definition I.4.11 we have the following definitions.

Definition C.7.1. Two local martingales N and M are called *orthogonal* if their product $L = NM$ is a local martingale.

Definition C.7.2. A local martingale M is called a *purely discontinuous local martingale* if $M(0) = 0$ and if it is orthogonal to all continuous local martingales. We denote the set of purely discontinuous local martingales by $\mathcal{M}^d_{loc}(\{\mathcal{F}_t\}, P)$.

From Jacod and Shiryaev [26], Lemma I.4.14(b) we have the following result.

Lemma C.7.3. *A local martingale that belongs to $\mathcal{FV}(\{\mathcal{F}_t\}, P)$ is a purely discontinuous local martingale.*

C.8 Decomposition of Semimartingales

Definition C.8.1. A real-valued process $\{S(t), t \in [0,T]\}$ on $(\Omega, \mathcal{F}, P)$ which can be written in the form

$$S = S(0) + M + A, \tag{C.12}$$

for some $M \in \mathcal{M}_{loc,0}(\{\mathcal{F}_t\}, P)$ and $A \in \mathcal{FV}_0(\{\mathcal{F}_t\}, P)$ is called an $\{\mathcal{F}_t\}$-*semimartingale*. The set of all $\{\mathcal{F}_t\}$-semimartingales is denoted $\mathcal{SM}(\{\mathcal{F}_t\}, P)$. We call M the *local martingale part* of the semimartingale S and call A the *finite variation part* of the semimartingale S.

Notation C.8.2. Denote by $\mathcal{SM}^c(\{\mathcal{F}_t\}, P)$ the set of $S \in \mathcal{SM}(\{\mathcal{F}_t\}, P)$ whose sample paths are continuous. We say that S is a continuous semimartingale.

From Jacod and Shiryaev [26], Theorem I.4.18 we have the following theorem.

Theorem C.8.3. *Any local martingale M admits a unique (up to indistinguishability) decomposition*

$$M = M(0) + M^c + M^d, \tag{C.13}$$

where $M^c(0) = M^d(0) = 0$, M^c is a continuous local martingale and M^d is a purely discontinuous local martingale.

Remark C.8.4. We call M^c the *continuous part* of the local martingale M and we call M^d the *purely discontinuous part* of the local martingale M.

Remark C.8.5. Given some $S \in \mathcal{SM}(\{\mathcal{F}_t\}, P)$, one may ask whether the decomposition of S given by C.12 is unique. That is, for any two arbitrary decompositions

$$S = S(0) + M + A \tag{C.14}$$

and

$$S = S(0) + \tilde{M} + \tilde{A} \tag{C.15}$$

where $M, \tilde{M} \in \mathcal{M}_{loc,0}(\{\mathcal{F}_t\}, P)$ and $A, \tilde{A} \in \mathcal{FV}_0(\{\mathcal{F}_t\}, P)$, is it true that M and $\tilde{M}$ are indistinguishable, and likewise, that A and $\tilde{A}$ are indistinguishable? In general, this is not the case. However, from Jacod and Shiryaev [26], Proposition I.4.27, if we use Theorem C.8.3 to further decompose M and $\tilde{M}$ into their continuous and purely discontinuous local martingale parts, i.e.

$$M = M(0) + M^c + M^d, \tag{C.16}$$

and

$$\tilde{M} = \tilde{M}(0) + \tilde{M}^c + \tilde{M}^d, \tag{C.17}$$

where $M^c(0) = M^d(0) = \tilde{M}^c(0) = \tilde{M}^d(0) = 0$, M^c, $\tilde{M}^c$ are continuous local martingales and M^d, $\tilde{M}^d$ are purely discontinuous local martingales, we see from the following proposition that the semimartingale S is unique in a weak sense.

Proposition C.8.6. *Let $S \in \mathcal{SM}(\{\mathcal{F}_t\}, P)$. Then for any two arbitrary decompositions*

$$S = S(0) + M + A \tag{C.18}$$

and

$$S = S(0) + \tilde{M} + \tilde{A}, \tag{C.19}$$

the continuous local martingale parts of M and $\tilde{M}$ given by M^c and $\tilde{M}^c$, respectively (see Remark C.8.5), are indistinguishable.

Remark C.8.7. Given any $S \in \mathcal{SM}(\{\mathcal{F}_t\}, P)$ denote by S^c the unique (up to indistinguishability) member of $\mathcal{M}^c_{loc,0}(\{\mathcal{F}_t\}, P)$ such that for any decomposition

$$S = S(0) + M + A \tag{C.20}$$

for $M \in \mathcal{M}_{loc,0}(\{\mathcal{F}_t\}, P)$ and $A \in \mathcal{FV}_0(\{\mathcal{F}_t\}, P)$, the continuous part of the local martingale M and S^c are indistinguishable.

From Jacod and Shiryaev [26], Definition I.4.45, we have the following definition.

Definition C.8.8. The square-bracket quadratic co-variation process of two semimartingales X and Y is

$$[X, Y] \triangleq X(t)Y(t) - X(0)Y(0) - \int_0^t X(s-)dY(s) - \int_0^t Y(s-)dX(s) \tag{C.21}$$

which is defined uniquely up to indistinguishability.

From Jacod and Shiryaev [26], Definition I.4.52, we have the following theorem.

Theorem C.8.9. *Let X, Y be semimartingales and let X^c, Y^c denote their continuous local martingale parts, respectively. Then*

$$[X,Y](t) = \langle X^c, Y^c\rangle(t) + \sum_{0\le s\le t} \triangle X(s)\triangle Y(s) \tag{C.22}$$

C.9 Stochastic Integration Results

From Rogers and Williams [43], Theorem IV.38.3, we have the following *integration-by-parts formula* for semimartingales, also called *Ito's product rule.*

Theorem C.9.1. *Let X and Y be semimartingales. Then*

$$X(t)Y(t) = X(0)Y(0) + \int_0^t X(s-)dY(s) + \int_0^t Y(s-)dX(s) + [X,Y](t) \tag{C.23}$$

From Rogers and Williams [43], Theorem VI.39.1, we have the following *Ito's Formula* for semimartingales.

Theorem C.9.2. *Let $f : \mathbb{R}^N \to \mathbb{R}$ be a function which has continuous derivatives up to order two. Suppose $\boldsymbol{X} = (X_1, \dots, X_n)$ is a semimartingale in $\mathbb{R}^N$. Then*

$$\begin{aligned} f(\boldsymbol{X}(t)) - f(\boldsymbol{X}(0)) &= \sum_{i=1}^n \int_0^t \frac{\partial f}{\partial X_i}\boldsymbol{X}(s-)d\boldsymbol{X}(s) \\ &+ \frac{1}{2}\sum_{i=1}^n\sum_{j=1}^n \int_0^t \frac{\partial^2 f}{\partial X_i X_j}\boldsymbol{X}(s-)d\langle (X_i)^c, (X_j)^c\rangle(s) \\ &+ \sum_{0\le s\le t}\left(f(\boldsymbol{X}(s)) - f(\boldsymbol{X}(s-)) - \sum_{i=1}^n \frac{\partial f}{\partial X_i}\boldsymbol{X}(s-)\triangle X_i(s)\right), \end{aligned} \tag{C.24}$$

$(X_i)^c$ denoting the continuous local martingale part of the semimartingale X_i.

From Rogers and Williams [43], Theorem IV.27.6.iv, we have the following result

Theorem C.9.3. *Let $M \in \mathcal{M}_0^2(\{\mathcal{F}_t\}, P)$. Define the space*

$$L^2(M) \triangleq \left\{ H : [0,T] \times \Omega \to \mathbb{R} \middle| H \text{ is predictable and } E\left[\int_0^T H^2(s)d[M](s)\right] < \infty \right\}. \tag{C.25}$$

Then for $H \in L^2(M)$, we have for all $t \in [0,T]$,

$$\triangle\left(\int_0^t H(s)dM(s)\right) = H(t)\triangle M(t) \text{ a.s.} \tag{C.26}$$

The next result comes from Proposition I.3.5 in Jacod and Shiryayev [26] and deals with integration with respect to a process $A \in \mathcal{A}^+(\{\mathcal{F}_t\}, P)$ (see Notation C.5.4).

Proposition C.9.4. *Let $A \in \mathcal{A}^+(\{\mathcal{F}_t\}, P)$ and let H be a non-negative $\{\mathcal{F}_t\}$-progressively measurable processes such that the process*

$$B(t) \triangleq \int_0^t H(s)dA(s) < \infty \text{ a.s.}, \quad t \in [0,T]. \tag{C.27}$$

Then $B \in \mathcal{A}^+(\{\mathcal{F}_t\}, P)$. Moreover, if A and H are $\{\mathcal{F}_t\}$-predictable, then B is $\{\mathcal{F}_t\}$-predictable.

C.10 Ito-Doléans-Dade Exponential Results

From Jacod and Shiryaev [26], Theorem I.4.61, we have the following theorem defining the Doléans-Dade exponential.

Theorem C.10.1. *Let X be an $\{\mathcal{F}_t\}$-semimartingale. Then*

$$Y(t) = 1 + \int_0^t Y(s-)dX(s), \tag{C.28}$$

has one and only one (up to indistinguishability) càdlàg $\{\mathcal{F}_t\}$-adapted solution. This solution is a semimartingale, and is given by

$$Y(t) = \exp\left\{X(t) - \frac{1}{2}\langle X^c, X^c\rangle(t)\right\} \prod_{0\le s\le t} (1 + \triangle X(s)) \exp\left\{ - \triangle X(s)\right\}, \tag{C.29}$$

where the (possibly) infinite product is absolutely convergent. Furthermore,

a) If $X \in \mathcal{FV}(\{\mathcal{F}_t\}, P)$, then $Y \in \mathcal{FV}(\{\mathcal{F}_t\}, P)$

b) If $X \in \mathcal{M}_{loc}(\{\mathcal{F}_t\}, P)$, then $Y \in \mathcal{M}_{loc}(\{\mathcal{F}_t\}, P)$.

Remark C.10.2. We will use the notation $\mathcal{E}(X)(t)$ to represent $Y(t)$ in (C.29), that is $Y(t) = \mathcal{E}(X)(t)$, and we call $\mathcal{E}(X)$ the *Doléans-Dade exponential* of the semimartingale X.

Remark C.10.3. From (C.29), $\mathcal{E}(X)$ is strictly positive if and only if $\triangle X(t) > -1$ a.s. for all $t \geq 0$. In particular, if X is continuous then,

$$\mathcal{E}(X)(t) = \exp\left\{X(t) - \frac{1}{2}[X,X](t)\right\}, \tag{C.30}$$

and therefore,

$$\mathcal{E}(X)(t) > 0 \text{ a.s.} \tag{C.31}$$

In this case, $\mathcal{E}(X)$ is called the *Ito exponential* of the continuous semimartingale X.

From Protter [38], Chapter V, Section 9, Theorem 52, we have the following theorem which deals with solving a common linear SDE.

Theorem C.10.4. *Let H be a semimartingale and let Z be a continuous semimartingale with $Z(0) = 0$. The solution to the stochastic integral equation*

$$X(t) = H(t) + \int_0^t X(s-)dZ(s), \quad t \in [0,T], \tag{C.32}$$

is given by

$$X(t) = \mathcal{E}(Z)(t)\left\{H(0) + \int_{0^+}^t \frac{1}{\mathcal{E}(Z)(s)} d(H(s) - [Z,H](s))\right\}, \quad t \in [0,T]. \tag{C.33}$$

From Protter [38], Chapter III, Section 8, Theorem 45, we have *Novikov's Criterion*, which gives conditions for the Doléans-Dade exponential of a continuous local martingale to be a martingale.

Theorem C.10.5. *Let M be a continuous local martingale and suppose that*

$$E\left[\exp\left\{\frac{1}{2}[M,M](\infty)\right\}\right] < \infty, \tag{C.34}$$

Then $\mathcal{E}(M)$ is a uniformly integrable martingale.

C.11 Martingale Results

The following propositions give conditions for a martingale to be a uniformly integrable martingale and for a local martingale to be a martingale.

Proposition C.11.1. *Let M be a càdlàg martingale. M is of class $\mathcal{D}[0,T]$ if and only if it is uniformly integrable.*

Proposition C.11.2. *Let M be a local martingale. If M is of class $\mathcal{D}[0,T]$ then M is in fact a martingale.*

From Karatzas and Shreve [30] Problem 1.5.19(ii), we have conditions which make a local martingale a supermartingale.

Proposition C.11.3. *Suppose that X is an $\{\mathcal{F}_t\}$-local martingale with $X(t) \geq 0$ a.s. for each $t \in [0,\infty)$. Then X is a $\{\mathcal{F}_t\}$-supermartingale.*

From Protter [38], Theorem 13, Chapter III, page 115 is the following result called the *Doob-Meyer decomposition.*

Theorem C.11.4. *Let Z be a càdlàg supermartingale. Then Z has the decomposition*

$$Z = Z_0 + M - A \tag{C.35}$$

where M is a local martingale and A is an increasing process which is predictable, and $M_0 = A_0 = 0$. Such a decomposition is unique. Moreover if $\lim_{t\to\infty} E\left[Z(t)\right] > -\infty$, then $E\left[A(\infty)\right] < \infty$.

The following is an essential result about predictable local martingales from I.3.16 of Jacod and Shiryaev [26].

Proposition C.11.5. *Any $\{\mathcal{F}_t\}$-predictable local martingale X which belongs to $\mathcal{FV}_0(\{\mathcal{F}_t\},P)$ is equal to 0 (up to a set of measure zero).*

C.12 Martingale Representation Theorems

Definition C.12.1. Suppose that $\{\mathbf{W}(t), t \in [0,T]\}$ is a given d-dimensional Brownian motion on a complete probability space $(\Omega, \mathcal{F}, P)$. The standard *Wiener filtration* is the filtration given by

$$\mathcal{F}_t \triangleq \sigma\{\mathbf{W}(s), 0 \leq s \leq t\} \vee \mathcal{N}(P), \quad t \in [0,T], \tag{C.36}$$

where $\mathcal{N}(P) \triangleq \{A \in \mathcal{F} : P(A) = 0\}$ are the P-measure zero sets in $\mathcal{F}$.

The following is the famous Ito representation theorem.

Theorem C.12.2. *Suppose that* $\{\mathbf{W}(t), t \in [0,T]\}$ *is a given* d*-dimensional Brownian motion on a complete probability space* $(\Omega, \mathcal{F}, P)$ *and* $\{\mathcal{F}_t\}$ *is the standard Wiener filtration in* $(\Omega, \mathcal{F}, P)$*. For each* $X \in L^2(\Omega, \mathcal{F}, P)$ *there exists some* d*-dimensional* $\{\mathcal{F}_t\}$ *progressively measurable process* $\mathbf{\Phi}$ *such that*

$$E\left[\sum_{k=1}^{d}\int_0^t |\Phi^k(s)|^2 ds\right] < \infty, \quad t \in [0,T], \tag{C.37}$$

and

$$X = E[X] + \sum_{k=1}^{d}\int_0^T \Phi^k(t) dW^k(t) \quad a.s., \tag{C.38}$$

where Φ^k *and* W^k, $k = 1, \ldots, d$, *are the* k*-th scalar components of* $\mathbf{\Phi}$ *and* $\boldsymbol{W}$ *respectively. Furthermore,* $\mathbf{\Phi}$ *is unique* $\lambda \otimes P$ *almost everywhere on* $[0,T] \times \Omega$.

Theorem C.12.2 can be extended to $X \in \mathcal{M}^2(\{\mathcal{F}_t\}, P)$ quite easily as stated by the following corollary.

Corollary C.12.2.1. *Suppose that* $\{\mathbf{W}(t), t \in [0,T]\}$ *is a given* d*-dimensional Brownian motion on a complete probability space* $(\Omega, \mathcal{F}, P)$ *and* $\{\mathcal{F}_t\}$ *is the standard Wiener filtration in* $(\Omega, \mathcal{F}, P)$*. If* $X \in \mathcal{M}^2(\{\mathcal{F}_t\}, P)$ *then there exists some* d*-dimensional* $\{\mathcal{F}_t\}$ *progressively measurable process* $\mathbf{\Phi}$ *such that*

$$E\left[\sum_{k=1}^{d}\int_0^t |\Phi^k(s)|^2 ds\right] < \infty, \quad t \in [0,T], \tag{C.39}$$

and

$$X(t) = E[X(0)] + \sum_{k=1}^{d}\int_0^t \Phi^k(s) dW^k(s) \;\; a.s., \quad t \in [0,T], \tag{C.40}$$

where Φ^k *and* W^k, $k = 1, \ldots, d$, *are the* k*-th scalar components of* $\mathbf{\Phi}$ *and* $\boldsymbol{W}$ *respectively. Furthermore,* $\mathbf{\Phi}$ *is unique* $\lambda \otimes P$ *almost everywhere on* $[0,T] \times \Omega$.

The results of Corollary C.12.2.1 are extended to the class of $\{\mathcal{F}_t\}$ local martingales by the following theorem.

Theorem C.12.3. *Suppose that* $\{\mathbf{W}(t), t \in [0,T]\}$ *is a given d-dimensional Brownian motion on a complete probability space* $(\Omega, \mathcal{F}, P)$ *and* $\{\mathcal{F}_t\}$ *is the standard Wiener filtration in* $(\Omega, \mathcal{F}, P)$*. If* $X \in \mathcal{M}_{loc,0}(\{\mathcal{F}_t\}, P)$ *then there exists some d-dimensional* $\{\mathcal{F}_t\}$ *progressively measurable process* $\mathbf{\Phi}$ *such that*

$$\int_0^t \|\mathbf{\Phi}(s)\|^2 ds < \infty \ \text{a.s.}, \quad t \in [0,T], \tag{C.41}$$

and

$$X(t) = \sum_{k=1}^{d} \int_0^t \Phi^k(s) dW^k(s) \ \text{a.s.}, \quad t \in [0,T], \tag{C.42}$$

Furthermore, $\mathbf{\Phi}$ *is unique* $\lambda \otimes P$ *almost everywhere on* $[0,T] \times \Omega$*.*

Using the abstract martingale representation theorem of Jacod and Shiryaev (see Theorem III.4.29 of Jacod and Shiryaev [26]) we can extend Theorem C.12.3 to the case where the filtration $\{\mathcal{F}_t\}$ is generated by both a Brownian motion and a Markov chain.

Theorem C.12.4. *Suppose that* $\{\mathbf{W}(t), t \in [0,T]\}$ *is a given d-dimensional Brownian motion and* $\{\boldsymbol{\alpha}(t), t \in [0,T]\}$ *is a given continuous-time Markov chain with state space* S *on a complete probability space* $(\Omega, \mathcal{F}, P)$ *and* $\{\mathcal{F}_t\}$ *is the joint filtration given in Condition 4.1.1. If* $X \in \mathcal{M}_{loc,0}(\{\mathcal{F}_t\}, P)$ *then there exists some some d-dimensional* $\{\mathcal{F}_t\}$ *progressively measurable process* $\mathbf{\Phi}$ *and some* $S \times S$*-dimensional* $\{\mathcal{F}_t\}$ *predictably measurable process* $\mathbf{\Gamma}$ *such that*

$$\int_0^t \|\mathbf{\Phi}(s)\|^2 ds < \infty \ \text{a.s. and} \ \sum_{\substack{i,j \in S \\ i \neq j}} \left(\int_0^t |\Gamma_{ij}(s)|^2 d[M_{ij}](s) \right)^{1/2} < \infty \ \text{a.s.}, \tag{C.43}$$

for all $t \in [0,T]$*, and*

$$X(t) = \sum_{k=1}^{d} \int_0^t \Phi^k(s) dW^k(s) + \sum_{\substack{i,j \in S \\ i \neq j}} \int_0^t \Gamma_{ij}(s) dM_{ij}(s) \ \text{a.s.}, \quad t \in [0,T], \tag{C.44}$$

where M_{ij} *are the canonical martingales of the Markov chain* $\boldsymbol{\alpha}$ *(see Definition 4.1.9). Furthermore,* $\mathbf{\Phi}$ *is unique* $\lambda \otimes P$ *almost everywhere on* $[0,T] \times \Omega$ *and* Γ_{ij} *is* $\nu_{[M_{ij}]}$ *unique for all* $i, j \in S, i \neq j$*.*

C.13 Essential Supremum of a Family of Random Variables

This section is taken from Appendix A of Karatzas and Shreve [31].

From Karatzas and Shreve Definition A.1 [31]:

Definition C.13.1. Let $(\Omega, \mathcal{F}, P)$ be a probability space and let $\mathcal{X}$ be a non-empty family of non-negative random variables defined on $(\Omega, \mathcal{F}, P)$. The *essential supremum* of $\mathcal{X}$, denoted by ess-sup $\mathcal{X}$, is a random variable X^* satisfying:

(i) $\forall X \in \mathcal{X}$, $X \leq X^*$ a.s.

(ii) if Y is a random variable satisfying $X \leq Y$ a.s. for all $X \in \mathcal{X}$, then $X^* \leq Y$ a.s.

Because random variables are defined only up to P-almost surely, it is in general not meaningful to speak of an "ω by ω" supremum $\sup\{X(\omega); X \in \mathcal{X}\}$. The essential supremum substitutes for this concept.

From Karatzas and Shreve Theorem A.3 [31]:

Theorem C.13.2. *Let $\mathcal{X}$ be a non-empty family of non-negative random variables. Then $X^* = \text{ess-sup}\,\mathcal{X}$ exists. Furthermore, if $\mathcal{X}$ is closed under pairwise maximization, i.e., $X, Y \in \mathcal{X}$ implies $X \vee Y \in \mathcal{X}$, then there is a non-decreasing sequence $\{Z_n\}_{n=1}^{\infty}$ of random variables in $\mathcal{X}$ satisfying $X^* = \lim_{n\to\infty} Z_n$ almost surely.*

C.14 Komlós Theorem

Proposition C.14.1. *Given a real number a and a sequence of real numbers $\{a_i\}$, $i = 1, 2, \ldots$, such that*

$$\lim_{n\to\infty} a_i = a, \tag{C.45}$$

then $\{a_i\}$ is Cesàro Summable and converges Cesàro to a, that is,

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} a_i = a. \tag{C.46}$$

From Theorem 6 of Schwartz [46]

Theorem C.14.2. *Consider a probability space* $(\Omega, \mathcal{F}, P)$. *Suppose* $\{X_n\}$ *is a sequence of random variables on* $(\Omega, \mathcal{F}, P)$ *satisfying* $\sup_n E|X_n| < \infty$ *a.s. Then there exists a subsequence* $\{X_n^0\}$ *and a random variable* $\beta \in L^1(\Omega, \mathcal{F}, P)$ *such that for each further subsequence* $\{X_n'\}$,

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} X_i' = \beta \text{ a.s.} \tag{C.47}$$

Remark C.14.3. Theorem C.14.2 asserts that for any $L^1(\Omega, \mathcal{F}, P)$ bounded sequence, a subsequence can be extracted so that every further subsequence converges Cesàro to the same limit.

Appendix D

Elementary Convex Analysis Theory

Definition D.0.1. Let $\mathbb{U}$ be an arbitrary vector space over $\mathbb{R}$. A set $A \subset \mathbb{U}$ is said to be *convex* if for every finite set of elements $\{a_i\} \subset A$ where $i = 1, \ldots, n$ and non-negative real constants $\{\lambda_i\}$, $i = 1, \ldots, n$, such that $\lambda_1 + \ldots + \lambda_n = 1$, we have

$$\sum_{i=1}^{n} \lambda_i a_i \in A. \tag{D.1}$$

Definition D.0.2. Let A be a convex subset of $\mathbb{U}$ and let f be a mapping $f : A \mapsto \bar{\mathbb{R}}$. f is said to be *convex* when we have

$$f(\lambda u + (1-\lambda)v) \leq \lambda f(u) + (1-\lambda) f(v), \tag{D.2}$$

for all $\lambda \in [0,1]$ and for all $u, v \in A$ such that the right hand side of (D.2) is defined. A mapping $g : A \mapsto \bar{\mathbb{R}}$ is said to be *concave* when $-g(\cdot)$ is convex.

Definition D.0.3. Let A be a convex subset of $\mathbb{U}$ and let f be a convex mapping $f : A \mapsto \bar{\mathbb{R}}$. The set of values $u \in A$ for which $f(u) < +\infty$ is a convex set and is called the *effective domain* of f. The effective domain of f is denoted $\operatorname{dom} f$.

Definition D.0.4. Given two vector spaces $\mathbb{U}$ and $\mathbb{Y}$, a mapping $\alpha : \mathbb{U} \times \mathbb{Y} \mapsto \mathbb{R}$ is a *bilinear form* on $\mathbb{U} \times \mathbb{Y}$ when
(a) The mapping $u \mapsto \alpha(u, y) : \mathbb{U} \mapsto \mathbb{R}$ is linear for each fixed $y \in \mathbb{Y}$,
(b) The mapping $y \mapsto \alpha(u, y) : \mathbb{Y} \mapsto \mathbb{R}$ is linear for each fixed $u \in \mathbb{U}$.
When there is just one designated bilinear form α on $\mathbb{U} \times \mathbb{Y}$, then the notation $\langle \cdot, \cdot \rangle$ is typically used to denote α, so that in particular $\alpha(u, y)$ is denoted by $\langle u, y \rangle$ for $u \in \mathbb{U}$ and $y \in \mathbb{Y}$. With a fixed bilinear form $\langle \cdot, \cdot \rangle$ the triple $(\mathbb{U}, \mathbb{Y}, \langle \cdot, \cdot \rangle)$ is called a *duality system.*

Example D.0.5. A bilinear form $\langle \cdot, \cdot \rangle$ on the space $\mathbb{R}^N \times \mathbb{R}^N$ is given by

$$\langle u, v \rangle \triangleq u^\top v, \quad u, v \in \mathbb{R}^N, \tag{D.3}$$

which is the inner-product on $\mathbb{R}^N$.

Definition D.0.6. Fix any arbitrary duality system $(\mathbb{U}, \mathbb{Y}, \langle \cdot, \cdot \rangle)$ and take a function $f : \mathbb{U} \mapsto \bar{\mathbb{R}}$ (f is not necessarily convex). Then the mapping $f^* : \mathbb{Y} \mapsto \bar{\mathbb{R}}$ defined as

$$f^*(y) \triangleq \sup_{u \in \mathbb{U}} \{\langle u, y \rangle - f(u)\}, \quad y \in \mathbb{Y}, \tag{D.4}$$

is called the *convex conjugate function* of f.

Remark D.0.7. It is immediate from Definition D.0.6 that f^* is a convex function regardless of whether or not f is a convex function.

Definition D.0.8. Fix any arbitrary duality system $(\mathbb{U}, \mathbb{Y}, \langle \cdot, \cdot \rangle)$ and a convex set $A \subset \mathbb{U}$. The direction $y \in \mathbb{Y}$ is said to be *normal* to A at $u \in A$ when

$$\langle y, x - u \rangle \leq 0 \quad \text{for all } x \in A. \tag{D.5}$$

The set of all such directions is called the *normal cone* to A at u, denoted $N_c^A(u)$,

$$N_c^A(u) \triangleq \{y \in \mathbb{Y} | \langle y, x - u \rangle \leq 0 \text{ for all } x \in A\}. \tag{D.6}$$

Definition D.0.9. Fix any arbitrary duality system $(\mathbb{U}, \mathbb{Y}, \langle \cdot, \cdot \rangle)$ and take a function $f : \mathbb{U} \mapsto \bar{\mathbb{R}}$ (f is not necessarily convex). Then the set $\partial f(u) \subset \mathbb{Y}$ given by

$$\partial f(u) \triangleq \{y \in \mathbb{Y} | f(x) \geq f(u) + \langle y, x - u \rangle, \text{ for all } x \in \mathbb{U}\}, \tag{D.7}$$

defined at each $u \in \mathbb{U}$ is called the *subdifferential* of f at u. An element $y \in \partial f(u)$ is called a *subgradient* of f at u. Similarly, if we define the set $\partial f(u) \subset \mathbb{Y}$ by

$$\partial f(u) \triangleq \{y \in \mathbb{Y} | f(x) \leq f(u) + \langle y, x - u \rangle, \text{ for all } x \in \mathbb{U}\}, \tag{D.8}$$

at each $u \in \mathbb{U}$, it is called the *superdifferential* of f at u. An element $y \in \partial f(u)$ is called a *supergradient* of f at u.

Example D.0.10. The subdifferential of a function $f : \mathbb{R}^N \mapsto \mathbb{R}$ is given by

$$\partial f(u) \triangleq \{y \in \mathbb{R}^N | f(x) \geq f(u) + y^\top (x - u), \text{ for all } x \in \mathbb{R}^N\}, \ u \in \mathbb{R}^N. \tag{D.9}$$

Proposition D.0.11. *If $f : A \mapsto \bar{\mathbb{R}}$ is a convex (concave) function over a convex subset $A \subset \mathbb{U}$, then the subdifferential (superdifferential) $\partial f(u)$ is a non-empty compact set in the weak topology for each $u \in A$.*

Example D.0.12. Fix any arbitrary duality system $(\mathbb{U}, \mathbb{Y}, \langle \cdot, \cdot \rangle)$ and define the *set indicator function* of the convex set $A \subset \mathbb{U}$, $I_A : \mathbb{U} \mapsto \{0, +\infty\}$ by

$$I_A(u) \triangleq \begin{cases} 0, & u \in A \\ +\infty, & \text{otherwise.} \end{cases} \tag{D.10}$$

It is quite easy to determine that the subdifferential $\partial I_A(u)$ is given by the normal cone of the set A at u, i.e.

$$\partial I_A(u) = N_c^A(u), \quad u \in A. \tag{D.11}$$

Proposition D.0.13. *Fix any arbitrary duality system $(\mathbb{U}, \mathbb{Y}, \langle \cdot, \cdot \rangle)$, a convex subset $A \subset \mathbb{U}$, and a convex (concave) function $f : A \mapsto \bar{\mathbb{R}}$. The point $u \in A$ is the minimum (maximum) of the convex (concave) function f if and only if $0 \in \partial f(u)$.*

We now state definitions and results which deal with convex functions defined on $\mathbb{R}^N$.

Definition D.0.14. A convex function $f : \mathbb{R}^N \mapsto (-\infty, +\infty]$ is called a *proper* convex function (f is not identically $-\infty$). A convex function is called *lower semi-continuous* at $x \in \mathbb{R}^N$ if for every sequence $\{x_i\} \subset \mathbb{R}^N$, $i = 1, 2, \ldots$ which converges in norm to x, we have

$$f(x) \leq \liminf_{i \to \infty} f(x_i). \tag{D.12}$$

A convex function is called *closed* if for each $\alpha \in \mathbb{R}$, the set

$$\{x \in \mathbb{R}^N | f(x) \leq \alpha\} \tag{D.13}$$

is closed.

Proposition D.0.15. *A proper convex function is closed if and only if it is lower semi-continuous.*

From §5.4 in Karatzas and Shreve [31]

Definition D.0.16. For a given closed, convex subset $K \neq 0$ of $\mathbb{R}^N$, the mapping $\delta : \mathbb{R}^N \mapsto \mathbb{R} \cup \{+\infty\}$ given by

$$\delta(\nu) \triangleq \sup_{p \in K} \{-p^\top \nu\}, \quad \nu \in \mathbb{R}^N, \tag{D.14}$$

is the *support function of the set* $-K$. It is a closed and proper convex function, which is finite on its *effective domain*

$$\tilde{K} \triangleq \{\nu \in \mathbb{R}^N | \delta(\nu) < +\infty\}, \tag{D.15}$$

a convex cone, called the *barrier cone* of $-K$. In particular, $0 \in \tilde{K}$ and $\delta(0) = 0$.

Remark D.0.17. The function δ is *positively homogeneous*,

$$\delta(\alpha\nu) = \alpha\delta(\nu), \quad \nu \in \mathbb{R}^N, \alpha \geq 0, \tag{D.16}$$

and *subadditive*,

$$\delta(\nu + \mu) \leq \delta(\nu) + \delta(\mu), \quad \nu, \mu \in \mathbb{R}^N. \tag{D.17}$$

Proposition D.0.18. *$p \in K$ if and only if $\delta(\nu) + p^\top \nu \geq 0$ for all $\nu \in \tilde{K}$.*

The next proposition is adapted from Theorem 3.1.2 in Hiriart-Urrut and Lemaréchal [24] and deals with the local Lipschitz continuity of a convex function.

Proposition D.0.19. *Let $f : \mathbb{R}^N \mapsto (-\infty, \infty)$ be a convex function and let S be a convex compact subset of $\mathbb{R}^N$. Then there exists $L = L(S) \geq 0$ such that*

$$|f(x) - f(x')| \leq L\|x - x'\| \text{ for all } x \text{ and } x' \text{ in } S. \tag{D.18}$$

That is, f is Lipschitz continuous on S.

The following definition and proposition are adapted from pages 164-165 of §4.1 in Hiriart-Urrut and Lemaréchal [24] .

Definition D.0.20. Let $f : \mathbb{R}^N \mapsto \mathbb{R}$ be a *finite valued* convex function, let x and d be fixed in $\mathbb{R}^N$, and consider the difference quotient of f at x in the direction d:

$$q(t) \triangleq \frac{f(x + td) - f(x)}{t}, \quad \text{for } t > 0. \tag{D.19}$$

The *directional derivative* of f at x in the direction d is

$$f'(x, d) \triangleq \lim_{t \to 0+} q(t) = \inf_{t > 0} q(t). \tag{D.20}$$

Proposition D.0.21. *For fixed $x \in \mathbb{R}^N$, the directional derivative $f'(x, \cdot)$ has the property,*

$$|f'(x, d)| \leq L\|d\|, \quad \text{for all } d \in \mathbb{R}^N, \tag{D.21}$$

where L is the Lipschitz constant given in Proposition D.0.19. Furthermore, for any subgradient $s \in \partial f(x)$,

$$\|s\| \leq L. \tag{D.22}$$

The following statements which deal with *normal convex integrands* and *measurable selectors* are from J.T. Rockafellar [40] [41].

Definition D.0.22. Let $(E, \mathcal{S}, \mu)$ be a measure space. A *convex integrand* on $E \times \mathbb{R}^N$ is a function

$$f : E \times \mathbb{R}^N \mapsto (-\infty, \infty] \tag{D.23}$$

where $f(e, x)$ is convex in $x \in \mathbb{R}^N$ for each $e \in E$.

Definition D.0.23. A convex integrand f is called a *normal convex integrand* if $f(e, x)$ is proper and lower-semicontinuous in x for each $e \in E$, and there exists a *countable* collection U of $\mathcal{S}$ measurable functions $u : E \mapsto \mathbb{R}^N$ having the following properties:

(a) for each $u \in U$, $f(e, u(e))$ is $\mathcal{S}$ measurable,

(b) for each $e \in E$, $U_e \cap \operatorname{dom} f_e$ is dense in $\operatorname{dom} f_e$ where,

$$U_e \triangleq \{u(e) | u \in U\}. \tag{D.24}$$

Here $\operatorname{dom} f_e$ denotes the effective domain of the convex function $f_e(x) \triangleq f(e, x)$.

Proposition D.0.24. *Suppose f is a convex integrand such that $f(e, x)$ is $\mathcal{S}$ measurable for each fixed $x \in \mathbb{R}^N$, and such that, for each $e \in E$, $f(e, x)$ is lower semi-continuous in x and has interior points in its effective domain* $\operatorname{dom} f_e$. *Then f is a normal convex integrand.*

Definition D.0.25. A multi-valued mapping $K : E \mapsto \mathbb{R}^N$ is called *measurable* if, for every closed subset $V \in \mathbb{R}^N$, the set

$$K^{-1}(V) \triangleq \{e \in E | K(e) \cap V \neq \emptyset\} \tag{D.25}$$

is in $\mathcal{S}$.

Proposition D.0.26. *Let $K : E \mapsto \mathbb{R}^N$ be a measurable multi-valued mapping such that $K(e)$ is a non-empty closed set for every $e \in E$. Then there exists a measurable selector for K, i.e. an $\mathcal{S}$ measurable function $u : E \mapsto \mathbb{R}^N$ such that $u(e) \in K(e)$ for every $e \in E$.*

Proposition D.0.27. *Let $f : E \times \mathbb{R}^N \mapsto (-\infty, \infty]$ be a normal convex integrand, and for each $e \in E$ and $x \in \mathbb{R}^N$, let $\partial f(e, x)$ be the subdifferential of $f(e, \cdot)$ at x, i.e.,*

$$\partial f(e, x) \triangleq \{\nu \in \mathbb{R}^N | f(e, y) \geq f(e, x) + (y - x)^\top \nu, \text{ for all } y \in \mathbb{R}^N\}. \tag{D.26}$$

Then for any $\mathcal{S}$ measurable function $u : E \mapsto \mathbb{R}^N$, the multi-valued mapping

$$K : e \mapsto \partial f(e, u(e)) \tag{D.27}$$

is measurable.

Glossary

Notation	Description	Page List
$[0, T]$	finite time horizon	19
$\mathbf{W}$	N-dimensional Brownian motion	19, 34
$(\Omega, \mathcal{F}, P)$	complete probability space	19
$\mathcal{N}(P)$	P-null sets	19, 34
$\{\mathcal{F}_t\}$	filtration generated by driving processes	20, 34
$\mathcal{F}^*$	progressively measurable σ-algebra	20, 34
$\lambda \otimes P$	product measure	20
λ	Lebesgue measure	20
$\mathcal{B}([0, T])$	Borel σ-algebra on $[0, T]$	20
S_0	price of risk-free asset	20, 35
S_n	price of the n-th risky asset	20, 35
r	risk-free interest rate process	20, 35
$\boldsymbol{b}$	rate of return process	20, 35
$\boldsymbol{\sigma}$	volatility process	20, 35
$\boldsymbol{\theta}$	market price of risk	21, 36
Π	space of portfolio processes	21, 38, 60
$X^{(x,\boldsymbol{\pi})}, X^{(x,c,\boldsymbol{\pi})}, X^{(x,c,\boldsymbol{p})}$	investor wealth process	22, 40, 61
B	contingent claim random variable	23, 41, 62, 99
$\mathcal{A}$	set of admissible trading strategies	24, 42, 63
Λ	set of initial wealths	24, 43, 64
$H_0, H_{\boldsymbol{\mu}}, H_{\boldsymbol{\nu},\boldsymbol{\mu}}$	state price density process	25, 47, 65
$\boldsymbol{\alpha}$	Markov chain	34
S	state space of Markov chain	34
G	generator of $\boldsymbol{\alpha}$	34
$\mathcal{P}^*$	predictably measurable σ-algebra	34

Notation	Description	Page List
$R, \tilde{R}$	canonical processes of $\boldsymbol{\alpha}$	36
M	canonical martingales of $\boldsymbol{\alpha}$	36
$I\{\cdot\}$	indicator function	36
$\nu_{[M_{ij}]}$	Doléans measure of M_{ij}	38
$\mathcal{C}$	space of cumulative consumption processes	39, 60
$\mathcal{H}$	space of Markov chain dual processes	45
K	portfolio constraint set	59
$g(t)$	margin function	59
$\mathcal{G} \times \mathcal{H}$	joint space of dual processes	65
$\tilde{g}_K(t, \cdot)$	convex conjugate of $g(t, \cdot)$ on K	65
$\partial g(t, \cdot)$	superdifferential of $g(t, \cdot)$	77
$V(\cdot)$	cost function of approximate hedging	101
P^B	P equivalent probability measure	102
$\mathcal{Z}$	convex subspace of $L_1(\Omega, \mathcal{F}_T, P^B)$	103
$\mathcal{L}$	the space $L_1(\Omega, \mathcal{F}_T, P^B) \times \mathbb{R}$	112
$\mathcal{Q}$	convex subspace of $\mathcal{L}$	112
$\mathcal{L}^*$	dual space of $\mathcal{L}$	114
$N_c(\cdot, \cdot)$	normal cone of $\mathcal{Q}$	115
$\partial\tilde{U}(\cdot, \cdot)$	subdifferential of convex functional $\tilde{U}$	115
$L^{1/2}_{loc}(R_{ij})$	space of Markov chain integrand processes	132
$\triangle X(t)$	jump of the process X at time t	188
$\mathcal{M}(\{\mathcal{F}_t\}, P)$	space of $\{\mathcal{F}_t\}$-martingales	190
$\mathcal{SPM}(\{\mathcal{F}_t\}, P)$	space of $\{\mathcal{F}_t\}$-supermartingales	191
$\mathcal{M}_{loc}(\{\mathcal{F}_t\}, P)$	space of $\{\mathcal{F}_t\}$-local martingales	192
$\mathcal{FV}(\{\mathcal{F}_t\}, P)$	space of $\{\mathcal{F}_t\}$-adapted finite variation processes	192
$\mathcal{A}^+(\{\mathcal{F}_t\}, P)$	space of $\{\mathcal{F}_t\}$-adapted non-decreasing and integrable processes	193
$\mathcal{FV}_{loc}(\{\mathcal{F}_t\}, P)$	processes that are locally in $\mathcal{FV}(\{\mathcal{F}_t\}, P)$	193
$\mathcal{A}^+_{loc}(\{\mathcal{F}_t\}, P)$	processes that are locally in $\mathcal{A}^+(\{\mathcal{F}_t\}, P)$	193
$\langle X, Y\rangle(t)$	angle bracket quadratic co-variation process of $X(t)$ and $Y(t)$	193
$[X, Y](t)$	square bracket quadratic co-variation process of $X(t)$ and $Y(t)$	194
$\mathcal{SM}(\{\mathcal{F}_t\}, P)$	space of $\{\mathcal{F}_t\}$-semimartingales	195

Notation	Description	Page List
$\mathcal{E}(\cdot)(t)$	Doléans Dade exponential	199
ess-sup	essential supremum of a family of random variables	203
$(\mathbb{U}, \mathbb{Y}, \langle \cdot, \cdot \rangle)$	duality system	205
$\delta(\cdot)$	support function of $-K$	207

Index