Convex Stochastic Control and Conjugate Duality in a Problem of Unconstrained Utility Maximization Under a Regime Switching Model

by

Aaron Xin Situ

A thesis
presented to the University of Waterloo
in fulfilment of the
thesis requirement for the degree of
Master of Quantitative Finance
in
Mathematics

Waterloo, Ontario, Canada, 2015

© Aaron Xin Situ 2015

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

In this thesis, we examine a problem of convex stochastic optimal control applied to mathematical finance. The goal is to maximize the *expected utility from wealth at close of trade* (or terminal wealth) under a *regime switching model.* The presence of regime switching constitutes a definite challenge, and in order to keep the analysis tractable we therefore adopt a market model which is in other respects quite simple, and in particular does not involve margin payments, inter-temporal consumption or portfolio constraints. The asset prices will be modeled by classical Itô processes, and the market parameters will be dependent on the underlying Brownian Motion as well as a finite-state Markov Chain which represents the "regime switching" aspect of the market model. We use *conjugate duality* to construct a *dual optimization problem* and establish *optimality relations* between (putative) solutions of the dual and primal problems. We then apply these optimality relations to two specific types of utility functions, namely the *power utility* and *logarithmic utility* functions, and for these utility functions we obtain the optimal portfolios in *completely explicit and implementable form.*

Acknowledgments

I would like to express my sincere gratitude to Professor Andrew J. Heunis for his guidance throughout the process of completing this thesis. His expertise and patience has been extremely helpful from the start. It has been my pleasure to have someone who is incredibly knowledgeable and enthusiastic about this topic as my supervisor.

Table of Contents

1 Introduction

We address a problem of stochastic optimal control, with the goal of trading over a finite time-horizon to maximize expected utility from *wealth at close of trade.* The market model comprises a single risk-free asset and a finite number of risky assets, the prices of which are modeled by continuous-time Itô processes driven by a standard multidimensional Brownian motion. In addition, the problem includes *regime switching* in the market model in the following sense: besides the standard Brownian motion driving the stock price models a further source of randomness in the model is a finite-state continuous-time *Markov chain* which is independent of the Brownian motion, it being stipulated that the market parameters (i.e. the risk-free interest rate, mean rate of return and volatility of stocks) are adapted to the *joint filtration* of the Brownian motion and the Markov chain. Each state of the Markov chain models a market "mode" or "regime state"; a simple but apt example is that of a two-state Markov chain, in which one state represents a "bull market", with generally rising prices, while the other state represents a "bear market", with generally falling prices. Market models in which the basic randomness derives from a Brownian motion together with a finite-state Markov chain are known as "regime switching models", and such models are *incomplete.* In these models the Brownian motion effectively drives persistent, short-duration and small-scale micro-economic changes in the market parameters, while the finite-state Markov chain drives occasional long-duration, large-scale macro-economic changes. The independence of the Brownian motion and the Markov chain amounts to the reasonable assumption that the micro- and macro-economic effects are independent.

Regime switching market models have received considerable attention for the pricing of options, but are not as widely used for portfolio optimization. Two works which nevertheless address portfolio optimization in the setting of regime switching models are those of Zhou and Yin [18], who apply stochastic LQ-control to the problem of *mean-variance portfolio selection* (which involves the minimization of a quadratic loss function), and Sotomayor and Cadenillas [16], who use dynamic programming to maximize *discounted expected utility from consumption* on an infinite horizon. In both [18] and [16] the regime switching is incorporated in a fairly simple dependency structure, namely at each time instant the market parameters are *completely determined* by the state of the regime-state Markov chain at that same instant (the market parameters are then said to be *Markov modulated*). The optimization problems are tackled *directly* as primal problems, the approach of [18] being based on stochastic LQ-control, whereas the approach of [16] is based on dynamic programming. In contrast, in this thesis we shall proceed by the method of *convex duality*, the crux of which is to construct an associated *dual optimization problem*, together with *optimality relations* between putative solutions of the given (i.e. "primal") optimization problem and the dual optimization problem. Our route to convex duality is motivated by a simple and elegant "calculus of variations approach" of Bismut [1], which was also essential for addressing quadratic minimization in the setting of regime switching (see Donnelly *et-al* [4] and [5]). The essence of this approach is to suppress the portfolio as the basic "problem variable", and write the utility maximization problem as a problem of *calculus of variations* which involves the minimization of a cost functional over a vector space of Itô processes. This latter problem is well suited to application of the duality theory of Bismut [1], which yields a dual problem along with the associated optimality relations. This approach was used by Donnelly ([4] and [5]) for a problem of mean-variance portfolio selection with portfolio constraints in a market model which includes regime switching. At a very basic level the approach of the present work is similar to that of [5]. However, we can no longer work in the setting of "square-integrable" semimartingales (which are the appropriate price processes for the mean-square problem of [4] and [5]), and must instead deal with price process which are semimartingales without any inherent integrability properties. To compensate for this we shall instead exploit the "one-sidedness" or "non-negativity" which is natural to problems of utility maximization.

The present work is divided into two main parts. The first part comprises Section 2 - Section 6 as follows: In Section 2 we define the market model and the *primal optimization problem* (see problem (2.27) which follows). In Section 3 we re-formulate the *primal problem* in calculus-of-variations form as an equivalent minimization problem over a set of Itô processes (see (3.36) which follows). In Section 4 we construct a vector space of *dual variables*, each dual variable essentially corresponding to a specific type of semimartingale (see (4.22) which follows). In Section 5, following Bismut [1], we construct a *dual optimization problem* associated with the primal optimization problem and establish an equivalence between existence of solutions of the primal and dual problems with zero duality gap and a set of *preliminary optimality relations* (see Proposition 5.4 which follows). In Section 6 we refine the preliminary optimality relations obtained in Section 5 to a much more tractable and useful set of optimality relations (see Proposition 6.13 which follows). The second part of the thesis comprises the concrete examples of Section 7. In this section we will use the refined optimality relation established in Section 6 to get *explicit* optimal portfolios for the *primal problem* defined in Section 2 in the particular cases of the *power* utility function (see Section 7.1) and the *logarithmic* utility function (see Section 7.2). In each of these cases we identify a *candidate* for the optimal portfolio and we will use the refined optimality relations obtained in Section 6 as a *verification tool* to verify that the candidate portfolio is indeed optimal. This use of optimality relations as a verification tool is motivated by the principle of *totally unhedgeable coefficients* of Karatzas and Shreve [10] discussed at Remark 7.7. The results that we obtain complement and extend the recent results of Sotomayor and Cadenillas [16] as we discuss at Remark 7.10 and Remark 7.17. In order to avoid obscuring the main lines of development, we have relegated most of the proofs to the Appendix in Section A. Finally, the Appendix in Section B contains background information and results that will be used throughout this thesis, and which are included here to enhance readability of the thesis. Section **??** is a glossary of the notations used in the thesis.

Remark 1.1. Portfolio optimization is traditionally implemented with reference to one of two possible *preference structures*, namely mean-variance minimization (also known as quadratic loss minimization or mean-variance hedging) on the one hand, and utility maximization on the other hand. These preference structures have rather distinct goals. In the case of quadratic loss minimization one wants to minimize the *mean-square discrepancy* between the actual wealth at close of trade and a specified contingent claim. In this sense quadratic loss minimization is a form of *approximate hedging* of the contingent claim, in that the goal is to approximate the value of the contingent claim at close of trade. Utility maximization on the other hand is concerned with *wealth maximization*, as is ensured by the monotonic increase of the utility function, and approximation of contingent claims is not relevant to the goal of utility maximization. Any debate about the relative advantages and disadvantages of the two preference structures is not likely to be very meaningful as the structures have rather different goals. For example, a pension fund which is responsible for paying out contingent claims upon retirement is likely to be more interested in the approximate hedging of claims provided by quadratic loss minimization, while an investor who simply wants to maximize wealth is likely to be more interested in utility maximization. The application of the basic framework adopted in this thesis to problems of mean-variance minimization has been quite thoroughly investigated by Donnelly *et-al* [4] and [5], and the present thesis represents a first attempt at applying the same framework to utility maximization.

Remark 1.2. There are several aspects of utility maximization which are not addressed in this thesis, in particular portfolio constraints and intertemporal consumption. This is simply to limit the present work to a scope which is appropriate for a Master's thesis.

Remark 1.3. The very elegant and powerful calculus of variations approach of Bismut [1] is the essential motivation for the approach of this thesis, as it is for much work involving conjugate duality in stochastic control applied to mathematical finance (see for example Karatzas, Lehoczky, Shreve and Xu [9] and Xu and Shreve [17]) and was extended to the setting of regime switching models with the goal of quadratic risk minimization by Donnelly *et-al* [4] and [5]. In the the calculus of variations approach of Bismut [1] the basic dual variable is always a semimartingale, and as such this approach may be used to address problems with portfolio constraints, as well as problems with intertemporal consumption, since in all of these problems the basic dual variable is quite naturally a semimartingale. One class of problems to which the approach of Bismut [1] does not apply involves *state constraints*, such as a specified lower bound on wealth at close of trade. This because state constraints, or constraints on the wealth process, are really *indirect* constraints on the portfolio process since the wealth process is controlled *only indirectly* by the portfolio process through the stochastic differential equation which relates the wealth to the portfolio (i.e. the wealth equation, see (2.20) which follows). As a consequence of this indirect relationship state constraints demand a dual variable which is a "compound object" that is "more complex" than just a semimartingale and includes a semimartingale as one of its components. It is not appropriate to discuss these matters in detail here, and it suffices to say that state constraints present something of a formidable challenge and require a non-trivial extension of the calculus of variations approach of Bismut [1]. The thesis of Ramchandani [14] presents an extension of the results of Donnelly *et-al* [4] and [5] to problems with a very simple state constraint (a stipulated almost-sure lower-bound on the wealth at close of trade). The problem addressed in the present thesis is one for which the dual variables are naturally semimartingales (see Section 4 in particular) and therefore the calculus of variations approach of Bismut [1], as extended to regime switching models in [4] and [5], is adopted here.

2 Formulation of market model and the primal problem

We begin this section with the specification of conditions defining a market model with regime switching. This market model is essentially identical to that postulated in the works of Donnelly [4] and Donnelly *et-al* [5], which address a problem different from that of this thesis, namely quadratic loss minimization, but in the same regime-switching market model that is used here.

Condition 2.1. We are given a complete probability space $(\Omega, \mathcal{F}, \mathbb{P})$ (meaning that all subsets of a $\mathcal{F}$-measurable set of zero probability are necessarily $\mathcal{F}$-measurable) and a time horizon $[0, T]$, for a fixed finite "close-of-trade" instant $T \in (0, \infty)$, over which the market model is defined and over which all trades take place. Defined on this probability space is an $\mathbb{R}^N$-valued Brownian Motion

$$\{\mathbf{W}(t); t \in [0, T]\} \tag{2.1}$$

together with a time-homogeneous continuous-time Markov Chain

$$\{\boldsymbol{\alpha}(t),\ t \in [0, T]\} \tag{2.2}$$

with cádlág paths in a finite state space S and given non-random initial state i_0, that is

$$\boldsymbol{\alpha}(t) \in S := \{1, \ldots, D\} \quad \& \quad \boldsymbol{\alpha}(0) := i_0 \in S. \tag{2.3}$$

Remark 2.2. The Markov chain $\boldsymbol{\alpha}$ of Condition 2.1 has a $D \times D$ *generator matrix*

$$[\boldsymbol{G}] = [g_{ij}]_{i,j=1}^{D} \tag{2.4}$$

with the property that, for all i,$j \in S$,

$$g_{ii} = -\sum_{j \neq i} g_{ij} \quad \& \quad g_{ij} \geq 0, \quad i \neq j. \tag{2.5}$$

The significance of the generator matrix $[\boldsymbol{G}]$ is that exponential-matrix function

$$P_t := e^{t[\boldsymbol{G}]}, \quad t \in [0, T], \tag{2.6}$$

defines the Markov *transition function* of $\boldsymbol{\alpha}$ (P_t is actually defined for all $t \in [0, \infty)$, but is here limited to the interval $[0, T]$ as this is the only time-period of interest).

Condition 2.3. The Brownian motion $\mathbf{W}$ and finite-state Markov chain $\boldsymbol{\alpha}$ in Condition 2.1 are independent.

◁

Define the set of all $\mathbb{P}$-null events as

$$\mathcal{N}(\mathbb{P}) := \{A \in \mathcal{F} : \mathbb{P}(A) = 0\} \tag{2.7}$$

and define the filtration $\{\mathcal{F}_t,\ t \in [0, T]\}$ by

$$\mathcal{F}_t := \sigma\{(\boldsymbol{\alpha}(s), \mathbf{W}(s)), s \in [0, t]\} \vee \mathcal{N}(\mathbb{P}), \tag{2.8}$$

The σ-algebra $\mathcal{F}_t$ will be interpreted as the information available to investors up to and including the time instant $t \in [0, T]$.

Remark 2.4. Note that the fitration at (2.8) is necessarily right continuous just by virtue of inclusion of the null events $\mathcal{N}(\mathbb{P})$. This is because the process $(\boldsymbol{\alpha}, \mathbf{W})$ is necessarily a Feller process (see the argument in Section 2 of Donnelly *et-al* [5]).

Remark 2.5. In the market models that we shall introduce shortly, the role of the Brownian motion $\mathbf{W}$ will be to "drive" persistent, short-duration and small-scale micro-economic changes in the market parameters, while the role of the finite-state Markov chain $\boldsymbol{\alpha}$ will be to drive occasional long-duration, large-scale macro-economic changes. This latter class of large-scale changes constitutes *regime switching*, and a primary goal of this thesis is to account for *regime switching* in market models as well as the effect which regime switching has on the computation of optimal portfolios. The independence of the Brownian motion and the Markov chain at Condition 2.3 amounts to the reasonable assumption that the micro- and macro-economic effects are independent. In most works on portfolio optimization the large-scale macro-economic changes are completely unaccounted for, and it is only the driving Brownian motion $\mathbf{W}$ which appears in the market models as the driver of small-scale micro-economic changes. Portfolio optimization in which one tries to also model large-scale macro-economic changes by a finite state Markov chain has received much less attention, and the main prior works seem to be Zhou and Yin [18] (for unconstrained quadratic loss minimization with regime switching), Donnelly [4] and [5] (for quadratic loss minimization with regime switching and convex portfolio constraints), and Sotomayor *et-al* [16] (for utility maximization subject to rather stringent conditions to be discussed later). The results of this thesis will in fact complement and extend the results of Sotomayor *et-al* [16], as we shall discuss later (see Remark 2.6, Remark 2.9, Remark 7.4, Remark 7.10 and Remark 7.17).

Remark 2.6. The problem addressed by Zhou and Yin [18] does not involve portfolio constraints, and consequently the authors can use a simple and elegant (but very problem-specific) completion of squares approach. Likewise, the problem of utility maximization adressed by Sotomayor *et-al* [16] again does not involve portfolio constraints. Of course in this case one cannot use completion of squares (which is specific to quadratic loss functions only) and Sotomayor *et-al* [16] adopt an approach based dynamic programming and the Bellman equation. It is important to note in each of the works [18] and [16] the primal problem is analyzed directly, and neither work exploits the power of conjugate duality.

We next specify *market parameters*, namely an interest rate process, a mean rate of return on stocks, and a volatility process:

Condition 2.7. We are given on $(\Omega, \mathcal{F}, \mathbb{P})$ a real-valued *interest rate process* $\{r(t),\ t \in [0,T]\}$, together with an $\mathbb{R}^N$-valued process $\{\mathbf{b}(t),\ t \in [0,T]\}$ (with $\mathbb{R}$-valued components $\{b_n(t),\ t \in [0,T]\}$, $n = 1,\dots,N$), called the *mean rate of returns on stocks*, and an N by N matrix-valued process $\{[\boldsymbol{\sigma}(t)],\ t \in [0,T]\}$ (with $\mathbb{R}$-valued components $\{\sigma_{nm}(t),\ t \in [0,T]\}$, $m, n = 1,\dots,N$), called the *volatility process*. The process r, the entries b_n of $\mathbf{b}$, and the entries σ_{nm} of $[\boldsymbol{\sigma}]$ (where $n, m = 1, 2, 3..., N$) , are stipulated to be *uniformly bounded and $\mathcal{F}_t$-progressively measurable* $\mathbb{R}$-valued processes on the joint set $\Omega \times [0,T]$. Moreover, there exists a constant $k \in (0,\infty)$ such that

$$\mathbf{z}^\top[\boldsymbol{\sigma}(\omega,t)][\boldsymbol{\sigma}(\omega,t)]^T\mathbf{z} \geq k\|\mathbf{z}\|^2 \quad \forall(\mathbf{z},\omega,t) \in \mathbb{R}^N \times \Omega \times [0,T]. \tag{2.9}$$

◁

Remark 2.8. The stipulation at Condition 2.7 that the market parameter processes r, $\mathbf{b}$ and $[\boldsymbol{\sigma}]$ are $\mathcal{F}_t$-progressively measurable on $\Omega \times [0,T]$ amounts to the non-anticipative condition that, at each instant $t \in [0,T]$, these processes are completely determined by the histories $\{\mathbf{W}(s),\ 0 \leq s \leq t\}$ of the Brownian motion $\mathbf{W}$ and $\{\boldsymbol{\alpha}(s),\ 0 \leq s \leq t\}$ of the Markov chain $\boldsymbol{\alpha}$ over the interval $0 \leq s \leq t$. This condition stipulates a general progressively measurable dependence of the market parameters on the

Markov chain $\boldsymbol{\alpha}$ and the Brownian motion $\mathbf{W}$. This dependency structure was introduced by Donnelly [4]. In this way regime switching is built into the market parameters. Furthermore, it is a consequence of Condition 2.7 that the matrix $[\boldsymbol{\sigma}(\omega,t)]$ is nonsingular for all $(\omega,t)\in\Omega\times[0,T]$.

Remark 2.9. In Sotomayor and Cadenillas [16] and Zhou and Yin [18] it is stipulated the market parameters must be *Markov modulated and time-invariant*, meaning that the interest rate process r must have the special form $r(t)=\tilde{r}(\boldsymbol{\alpha}(t))$, for some specified constants $\tilde{r}(i)$, $i\in S$, and similarly for each b_n and σ_{nm}. Effectively this means that, at each $t\in[0,T]$, the values of $r(t)$, $\mathbf{b}(t)$ and $[\boldsymbol{\sigma}](t)$ are *completely determined* by the value $\boldsymbol{\alpha}(t)$ of the Markov chain $\boldsymbol{\alpha}$ at the same instant t, and are not at all determined by the Browian motion $\mathbf{W}$. In particular, the values of $r(t)$, $\mathbf{b}(t)$ and $[\boldsymbol{\sigma}](t)$ *do not* depend on the whole history $\{\boldsymbol{\alpha}(s),\ 0\le s\le t\}$ of the Markov chain $\boldsymbol{\alpha}$ over the interval $0\le s\le t$, but just on the value of $\boldsymbol{\alpha}(t)$ at the instant t alone, while dependence on the history $\{\mathbf{W}(s),\ 0\le s\le t\}$ of the Brownian motion $\mathbf{W}$ is completely excluded. This is significantly more restrictive than the dependence of the market parameters on both $\mathbf{W}$ and $\boldsymbol{\alpha}$ allowed by Condition 2.7, but this restrictive dependency structure seems to be essential for any approach which addresses the primal problem directly (without use of conjugate duality), such as the dynamic programming approach adopted in Sotomayor and Cadenillas [16] and the completion of squares approach of Zhou and Yin [18].

We next use the market parameters stipulated at Condition 2.7 to postulate *Itô-process price models* for a single risk-free asset and several risky assets in which investment takes place:

Condition 2.10. We postulate a *market model* in terms of the interest-rate process r, the mean rate of return on stocks $\mathbf{b}$, and the volatility process $[\boldsymbol{\sigma}]$ stipulated at Condition 2.7, as well as the Brownian motion $\mathbf{W}$ and finite-state Markov chain $\boldsymbol{\alpha}$ stipulated at Condition 2.1. This market model comprises $N+1$ assets traded continuously over the interval $[0,T]$, one of which is risk-free with price process $\{S_0(t);t\in[0,T]\}$ modelled by

$$\mathrm{d}S_0(t)=r(t)S_0(t)\,\mathrm{d}t\quad\&\quad S_0(0)=1, \tag{2.10}$$

which can be solved explicitly as

$$S_0(t)=exp\left\{\int_0^t r(s)\,\mathrm{d}s\right\}\quad 0\le t\le T. \tag{2.11}$$

The remaining N risky assets, with prices denoted as $\{S_n(t);t\in[0,T]\}$ for $n=\{1,\dots,N\}$ are modelled as *Itô processes* as follows:

$$\mathrm{d}S_n(t)=S_n(t)\left[b_n(s)\,\mathrm{d}t+\sum_{m=1}^{N}\sigma_{nm}(s)\,\mathrm{d}W_m(s)\right]\quad\&\quad S_n(0)=s_n. \tag{2.12}$$

Here $s_n>0$ is stipulated and non-random for all n. Notice that (2.12) can also be solved explicitly as follows:

$$S_n(t)=S_n(0)exp\left\{\int_0^t\sum_{m=1}^{N}\sigma_{nm}(s)\,\mathrm{d}W_m(s)+\int_0^t\left[b_n(s)-\frac{1}{2}\sum_{m=1}^{N}\sigma_{nm}^2(s)\right]\mathrm{d}s\right\}\quad 0\le t\le T. \tag{2.13}$$

Remark 2.11. Observe that the regime switch Markov chain $\boldsymbol{\alpha}$ does not appear explicitly in the equations (2.11) - (2.13), and the role of $\boldsymbol{\alpha}$ in these equations is completely concealed by the notation. In fact the role of the Markov chain $\boldsymbol{\alpha}$ is encapsulated entirely within the filtration $\{\mathcal{F}_t\}$ (see (2.8)), together with the $\mathcal{F}_t$-progressive measurability stipulated by Condition 2.7, for it is here that $\boldsymbol{\alpha}$ determines the market parameters through the Doob measurability theorem. As will become clear when we construct an associated dual problem, regime switching significantly affects the structure of the semimartingales which constitute the dual variables (see Section 4 and in particular Remark 4.12 which follows).

Remark 2.12. We define the $\mathbb{R}^N$-valued *market price of risk process* $\{\boldsymbol{\theta}(t),\ t \in [0,T]\}$ in the usual way, namely

$$\boldsymbol{\theta}(t) := [\boldsymbol{\sigma}(t)^{-1}]\,(\mathbf{b}(t) - r(t)\mathbf{1}) \tag{2.14}$$

where $\mathbf{1} \in \mathbb{R}^N$ is the vector with all unit entries.

Remark 2.13. From Condition 2.7 we see that $\boldsymbol{\theta}(t)$ is uniformly bounded and $\mathcal{F}_t$-progressively measurable on $\Omega \times [0,T]$. From (2.9) and Xu and Shreve [17], there exists a constant $k_1 \in (1,\infty)$ such that

$$\frac{1}{k_1}\max\left\{\left\|([\boldsymbol{\sigma}(\omega,t)^{-1}]\mathbf{z}\right\|, \left\|([\boldsymbol{\sigma}(\omega,t)^{-1}]^T\mathbf{z}\right\|\right\} \le \|\mathbf{z}\| \le k_1\,\min\left\{\left\|([\boldsymbol{\sigma}(\omega,t)^{-1}]z\right\|, \left\|([\boldsymbol{\sigma}(\omega,t)^{-1}]^T z\right\|\right\} \tag{2.15}$$

for all $(\mathbf{z},\omega,t) \in \mathbb{R}^N \times \Omega \times [0,T]$. We will need the bounds (2.15) several times in subsequent chapters.

◁

Notation 2.14. From now on we denote by $\mathcal{F}^\star$ and $\mathcal{P}^\star$ (respectively) the $\mathcal{F}_t$-progressively measurable and $\mathcal{F}_t$-predictable σ-algebra on $\Omega \times [0,T]$. In addition, we write $\phi \in \mathcal{F}^\star$ to indicate that the mapping $\phi : \Omega \times [0,T] \to \mathbb{R}^n$ is $\mathcal{F}_t$-progressively measurable and similarly $\phi \in \mathcal{P}^\star$ if the mapping is $\mathcal{F}_t$-predictable. The qualifier "a.s." always refers to the probability measure $\mathbb{P}$ on $(\Omega,\mathcal{F})$, and "a.e." refers to the product measure $(\mathbb{P} \otimes \lambda)$ on $\mathcal{F} \times \mathcal{B}([0,T])$ where λ is the Lebesgue measure on the Borel σ-algebra of $[0,T]$, denoted by $\mathcal{B}([0,T])$.

We next introduce the *wealth process* $\{X(t); t \in [0,T]\}$ of an investor assuming

Condition 2.15. The initial wealth $X(0) = x_0 \in (0,\infty)$ is given and is non-random, and there are *no constraints* on how the investor can allocate his or her wealth $X(t)$ at each instant t among the assets with prices $S_0, S_1, \ldots S_N$ stipulated at Condition 2.10.

Remark 2.16. We always assume that the investor follows a *self-financing* trading strategy when investing the wealth $X(t)$ at each instant $t \in [0,T]$. We denote by $\pi_n(t)$ the *proportion or fraction* of the total wealth $X(t)$ that the investor allocates to the risky asset with price $S_n(t)$, for each $n = 1,2,\ldots,N$. That is $\pi_n(t)X(t)$ denotes the *dollar amount* invested in the stock with price S_n, $n = 1,2,\ldots,N$. The vector $\boldsymbol{\pi}(t) = [\pi_1(t),\ldots,\pi_N(t)] \in \mathbb{R}^N$ then denotes the *total portfolio* at the instant t, and gives rise to an $\mathbb{R}^N$-valued *portfolio process* $\{\boldsymbol{\pi}(t),\ t \in [0,T]\}$. This process records the complete portfolio invested in the N-risky assets over the whole trading interval $t \in [0,T]$.

Condition 2.17. We assume that the portfolio process $\boldsymbol{\pi}$ is a member of the set Π defined by

$$\Pi := \left\{\boldsymbol{\pi} : \Omega \times [0,T] \to \mathbb{R}^N \mid \boldsymbol{\pi} \in \mathcal{F}^\star \quad \& \quad \int_0^T \|\boldsymbol{\pi}(s)\|^2\,\mathrm{d}t < \infty \quad \text{a.s.}\right\}, \tag{2.16}$$

(recall from Notation 2.14 that $\boldsymbol{\pi} \in \mathcal{F}^\star$ indicates that the process $\boldsymbol{\pi}$ is $\mathcal{F}_t$-progressively measurable).

Remark 2.18. Notice that Condition 2.17 stipulates that the portfolio process $\boldsymbol{\pi}$ is $\mathcal{F}_t$-progressively measurable. This encapsulates the idea that, at each instant t, the portfolio process $\boldsymbol{\pi}(t)$ depends only on the Brownian motion $\{\mathbf{W}(s),\ 0 \le s \le t\}$ and Markov chain $\{\boldsymbol{\alpha}(s),\ 0 \le s \le t\}$ over the interval $0 \le s \le t$. This again underscores the role of regime switching, since dependence on the Markov chain $\boldsymbol{\alpha}$ is now built into the very definition of a portfolio process. The pathwise square-integrability of $\boldsymbol{\pi}$ stipulated at (2.16) is for technical reasons and will ensure that all stochastic integrals are well-defined in what follows.

We next derive a *stochastic differential equation* (SDE) for the wealth process $\{X(t),\ t \in [0,T]\}$ which results from a given portfolio process $\boldsymbol{\pi} \in \Pi$ together with the price equations for S_n, $n = 0, 1, 2, \ldots, N$, at Condition 2.10. This wealth equation is a standard and well known result when the portfolio is denominated in *dollar amounts* but is slightly less well known when the portfolio is denominated in terms of *fractions* of total wealth, which is the denomination that we use in this thesis (see Remark 2.16). We therefore include the derivation of the wealth equation for completeness. From the definition of $X(t)$ and $\pi(t)$, one observes that

$$X(t)[1 - \sum_{n=1}^{N} \pi_i(t)] \tag{2.17}$$

is the *dollar amount* of wealth invested in the risk-free asset with price S_0 (since the investor follows a self-financing trading strategy, see Remark 2.16).
Using (2.10), (2.12), (2.17), and recalling Remark 2.16, it follows that $X(t)$ is given by

$$\begin{aligned}
\mathrm{d}X(t) &= \sum_{n=1}^{N} \pi_n(t) X(t)\, \mathrm{d}S_n(t) + [1 - \sum_{n=1}^{N} \pi_i(t)] X(t)\, \mathrm{d}S_0(t) \\
&= \sum_{n=1}^{N} \pi_n(t) X(t) [b_n(t)\, \mathrm{d}t + \sum_{m=1}^{N} \sigma_{nm}(t)\, \mathrm{d}W_m(t)] + [1 - \sum_{n=1}^{N} \pi_i(t)] X(t) r(t)\, \mathrm{d}t \\
&= X(t)\boldsymbol{\pi}^T(t)[\mathbf{b}(t)\, \mathrm{d}t + [\boldsymbol{\sigma}(t)]\, \mathrm{d}\mathbf{W}(t)] + X(t)r(t)\, \mathrm{d}t - X(t)\boldsymbol{\pi}^T(t) r(t)\mathbf{1}\, \mathrm{d}t \\
&= X(t)[r(t) + \boldsymbol{\pi}^T(t)(\mathbf{b}(t) - \mathbf{1})]\, \mathrm{d}t + X(t)\boldsymbol{\pi}^T(t)[\boldsymbol{\sigma}(t)]\, \mathrm{d}\mathbf{W}(t) \\
&= X(t)[r(t) + \boldsymbol{\pi}^T(t)[\boldsymbol{\sigma}(t)]\boldsymbol{\theta}(t)]\, \mathrm{d}t + X(t)\boldsymbol{\pi}^T(t)[\boldsymbol{\sigma}(t)]\, \mathrm{d}\mathbf{W}(t),
\end{aligned} \tag{2.18}$$

that is, the wealth process X satisfies the SDE

$$\mathrm{d}X(t) = X(t)[r(t) + \boldsymbol{\pi}^T(t)[\boldsymbol{\sigma}(t)]\boldsymbol{\theta}(t)]\, \mathrm{d}t + X(t)\boldsymbol{\pi}^T(t)[\boldsymbol{\sigma}(t)]\, \mathrm{d}\mathbf{W}(t), \tag{2.19}$$

in which the portfolio process $\boldsymbol{\pi}$ appears as a "control" on the right-hand side. For a portfolio process $\boldsymbol{\pi} \in \Pi$ we will denote by $X^{\boldsymbol{\pi}}$ the corresponding wealth process associated with this $\boldsymbol{\pi}$ together with the *initial wealth condition* $X^{\boldsymbol{\pi}}(0) = x_0$ (recall from Remark 2.16 that $\pi_n(t)$ denotes the *fraction* of total wealth, not the actual dollar amount, invested in the asset with price $S_n(t)$). It follows from (2.19) that $X^{\boldsymbol{\pi}}$ is uniquely determined in terms of the portfolio process $\boldsymbol{\pi}$ and initial wealth x_0 as the solution of the so-called "wealth equation"

$$\mathrm{d}X^{\boldsymbol{\pi}}(t) = X^{\boldsymbol{\pi}} \left\{[r(t) + \boldsymbol{\pi}^T(t)[\boldsymbol{\sigma}(t)]\boldsymbol{\theta}(t)]\, \mathrm{d}t + \boldsymbol{\pi}^T(t)[\boldsymbol{\sigma}(t)]\, \mathrm{d}\mathbf{W}(t)\right\} \quad \& \quad X^{\boldsymbol{\pi}}(0) = x_0. \tag{2.20}$$

Remark 2.19. If the portfolio process $\boldsymbol{\pi}$ were such that $\pi_n(t)$ is the *dollar amount* invested in the asset with price $S_n(t)$ (instead of the fraction of total wealth, as in this thesis, recall Remark 2.16) then it is easily seen that the wealth equation becomes

$$\mathrm{d}X^{\boldsymbol{\pi}}(t) = \{r(t)X^{\boldsymbol{\pi}}(t) + \boldsymbol{\pi}^T(t)[\boldsymbol{\sigma}(t)]\boldsymbol{\theta}(t)\}\, \mathrm{d}t + \boldsymbol{\pi}^T(t)[\boldsymbol{\sigma}(t)]\, \mathrm{d}\mathbf{W}(t), \quad \& \quad X^{\boldsymbol{\pi}}(0) = x_0, \tag{2.21}$$

in place of (2.20). This is the case for example in Donnelly [4] and [5] (e.g. eqn. (2.4) in [5]), where it is more appropriate to denominate portfolios in dollar amounts when the goal is quadratic risk minimization. For utility maximization it is frequently advantageous to denominate in terms of fractions of total wealth, as is discussed in the following Remark 2.20.

Remark 2.20. The allocation of portfolios in terms of *fractions* of total wealth, instead of the more commonly used dollar amounts, goes back at least as far as Karatzas, Lehoczky, Shreve and Xu [9], and is now very widely used (see e.g. Cvitanić and Karatzas [3], Cuoco and Liu [2], and Chaps. 5 and 6 of Karatzas and Shreve [10]). This allocation has the advantage that the wealth process $X^{\boldsymbol{\pi}}$ is *strictly positive* for every portfolio $\boldsymbol{\pi} \in \Pi$. To verify this we first define the process $Z(t)$ as

$$\mathrm{d}Z(t) := \left[r(t) + \boldsymbol{\pi}^T(t)[\boldsymbol{\sigma}(t)]\boldsymbol{\theta}(t)\right] \mathrm{d}t + \boldsymbol{\pi}^T(t)[\boldsymbol{\sigma}(t)]\,\mathrm{d}\mathbf{W}(t), \qquad Z(0) = 0. \tag{2.22}$$

By (2.20) and (2.22) gives

$$\mathrm{d}X^{\boldsymbol{\pi}}(t) = X^{\boldsymbol{\pi}}(t)\,\mathrm{d}Z(t) \quad \& \quad X^{\boldsymbol{\pi}}(0) = x_0, \tag{2.23}$$

from which the explicit form of $X^{\boldsymbol{\pi}}$ is equal to

$$\begin{aligned} X^{\boldsymbol{\pi}}(t) &= x_0 exp\left\{Z(t) - \frac{1}{2}\langle Z\rangle(t)\right\} \\ &= x_0 exp\left\{\int_0^t [r(s) + \boldsymbol{\pi}(s)^T[\boldsymbol{\sigma}(s)]\boldsymbol{\theta}(s)]\,\mathrm{d}s + \int_0^t \boldsymbol{\pi}(s)^T[\boldsymbol{\sigma}(s)]\,\mathrm{d}\mathbf{W}(s) - \frac{1}{2}\int_0^t \left\|\boldsymbol{\pi}^T(s)[\boldsymbol{\sigma}(s)]\right\|^2 \mathrm{d}s\right\}, \end{aligned} \tag{2.24}$$

which is strictly positive for all t $\in [0,T]$. Effectively this is saying that, for any allocation of portfolios in terms of *fractions of total wealth*, the corresponding wealth process $X^{\boldsymbol{\pi}}$ is always strictly positive. This turns out to be a decided technical advantage in the duality analysis which follows.

◁

Thus far we have stipulated a market model and obtained the SDE (2.20) relating the wealth process $X^{\boldsymbol{\pi}}$ and the portfolio process $\boldsymbol{\pi}$. To formulate the *optimization problem* or *primal problem* of the investor we next define the *utility* function, which is an essential ingredient for the primal problem. We will denote the utility function by U and stipulate the following properties for the utility function:

Condition 2.21. $U : \mathbb{R} \to [-\infty, \infty)$. For all $x \in \mathbb{R}$, the function U has the following properties:

$$\begin{aligned} &(1) \quad U(x) = -\infty \quad for \quad x \in (-\infty, 0] \\ &(2) \quad U(x) > -\infty \quad for \quad x \in (0, \infty) \\ &(3) \quad U(\cdot) \text{ is strictly increasing on } (0,\infty) \quad i.e.\ \forall\ x_2 > x_1 > 0, \text{ then } U(x_2) > U(x_1) \\ &(4) \quad U(\cdot) \text{ is of class } C_1 \text{ over the interval } (0,\infty) \\ &(5) \quad U(\cdot) \text{ is strictly concave on } (0,\infty) \quad i.e.\ U^{(1)}(x) \text{ is strictly decreasing } \forall\ x \in (0,\infty) \end{aligned} \tag{2.25}$$

◁

Remark 2.22. The stipulation at Condition 2.21(5) that $U(\cdot)$ is of class C_1 over the interval $(0, \infty)$ means that the derivative $U^{(1)}(x)$ exists for every $x \in (0, \infty)$ and $U^{(1)}(\cdot)$ is a continuous function on $(0, \infty)$.

With everything we have introduced so far, we can now define the *value* of the optimization problem:

$$V := \sup_{\boldsymbol{\pi}\in\Pi} E[U(X^{\boldsymbol{\pi}}(T))]. \tag{2.26}$$

To avoid trivialities we asume

Condition 2.23. $-\infty < V < \infty$.

◁

We are now able to formulate the *primal problem* which will be addressed in this thesis as follows:

$$\text{Determine some } \boldsymbol{\pi} \in \Pi \text{ such that } V = E[U(X^{\boldsymbol{\pi}}(T))]. \tag{2.27}$$

It follows from (2.27) and (2.26) that the primal problem involves searching for a portfolio process $\boldsymbol{\pi} \in \Pi$ which *maximizes the expected utility of the wealth at close of trade*, that is maximizes the quantity $E[U(X^{\boldsymbol{\pi}}(T))]$ over all $\boldsymbol{\pi} \in \Pi$.

Remark 2.24. From standard integration theory we know that

$$E[U(X^{\boldsymbol{\pi}}(T))] = E[U^+(X^{\boldsymbol{\pi}}(T))] - E[U^-(X^{\boldsymbol{\pi}}(T))], \tag{2.28}$$

where

$$U^+(x) := \max\{0, U(x)\} \qquad \text{and} \qquad U^-(x) := \max\{0, -U(x)\} \qquad \text{for all } x \in (0, \infty). \tag{2.29}$$

The quantity on the left of (2.28) is well-defined provided that *at least one* of the terms on the right side is finite, that is takes values in $[0, \infty)$. If it were the case that the utility function U were uniformly lower bounded for all $x \in [0, \infty)$, that is, for some constant $c \in \mathbb{R}$ we have

$$U(x) \geq c \qquad \text{for all} \qquad x \in [0, \infty) \tag{2.30}$$

then of course the second term on the right of (2.28) would always be finite. However, we do not wish to impose a lower-bounded condition such as (2.30) because this excludes some important and useful utility functions (such as the logarithmic utility $U(x) = \log(x)$). None of the conditions stipulated above precludes the possibility that we could have

$$E[U^+(X^{\boldsymbol{\pi}}(T))] = E[U^-(X^{\boldsymbol{\pi}}(T))] = \infty \tag{2.31}$$

for some $\pi \in \Pi$, and in this case the left side of (2.28) is undefined. To deal with this we shall adopt the following convention: For any $\mathcal{F}_T$-measurable function $\xi : (\Omega, \mathcal{F}, \mathbb{P}) \to [0, \infty)$, we *define*

$$EU(\xi) := -\infty \quad \text{whenever} \quad EU^-(\xi) = +\infty. \tag{2.32}$$

This means, in particular, that we put $E[U(X^{\boldsymbol{\pi}}(T))] = -\infty$ whenever $\boldsymbol{\pi} \in \Pi$ is such that $E[U^-(X^{\boldsymbol{\pi}}(T))] = \infty$. With this understanding the problem at (2.27) really amounts to maximizing the quantity $E[U(X^{\boldsymbol{\pi}}(T))]$ over all $\boldsymbol{\pi} \in \Pi$ such that $E[U^-(X^{\boldsymbol{\pi}}(T))] < \infty$.

◁

3 Reduction to calculus of variations form

Remark 3.1. Our approach to the primal problem (2.27) is based on the method of *conjugate duality.* In brief the idea is to formulate an associated *dual optimization problem* over a vector space of *dual variables*, which is "dual" to the primal problem (2.27) in a way that we shall shortly make precise. With a dual problem having been defined, we can then establish *optimality relations* between putative solutions of the primal and dual problems in terms of *transversality relations*, *complementary slackness relations* and *feasibility conditions.* These optimality relations are the essential tool for dealing with the primal problem. Neither the vector space of dual variables or the dual functional on the space of dual variables is *a-priori* evident. In this thesis we shall construct both the space of dual variables and the dual functional by an approach which has its origins in a fundamental work of Bismut [1], which was further extended by Labbe *et-al* [11] for quadratic risk minimization problems without regime switching, and by Donnelly *et-al* [4] and [5] for quadratic risk minimization problems which include regime switching. In the present thesis this approach extended to problems of *utility maximization.*

Following Bismut [1], the first step in the construction of an appropriate vector space of dual variables and a dual functional is to remove the portfolio process $\boldsymbol{\pi}$ from the optimization problem (2.27), and to write this optimization problem in "calculus of variation form" as the optimization of a *primal functional* over a vector space of Itô processes which is large enough to include the wealth processes $X^{\boldsymbol{\pi}}$ for each $\boldsymbol{\pi} \in \Pi$ (recall (2.20)). To formulate the vector space of Itô processes define the following spaces of processes (recall Notation 2.14):

$$\mathcal{H}_1 := \left\{ v : \Omega \times [0,T] \to \mathbb{R} \;\middle|\; v \in \mathcal{F}^\star, \quad \int_0^T |v(t)| \, \mathrm{d}t < \infty \quad \text{a.s.} \right\}, \tag{3.1}$$

$$\mathcal{H}_2 := \left\{ \boldsymbol{\xi} : \Omega \times [0,T] \to \mathbb{R}^N \;\middle|\; \boldsymbol{\xi} \in \mathcal{F}^\star, \quad \int_0^T \|\boldsymbol{\xi}(t)\|^2 \, \mathrm{d}t < \infty \quad \text{a.s.} \right\}, \tag{3.2}$$

It is elementary that $\mathcal{H}_1$ and $\mathcal{H}_2$ are vector spaces (with the usual pointwise addition and scalar multiplication of functions on $\Omega \times [0,T]$). Now define the product of sets

$$\mathbb{I} := \mathbb{R} \times \mathcal{H}_1 \times \mathcal{H}_2, \tag{3.3}$$

which, being a product of vector spaces, is itself a vector space. For each $X \equiv (X_0, \dot{X}, \boldsymbol{\Lambda}_X) \in \mathbb{I}$ define the $\mathcal{F}_t$-Itô process $\{X(t), t \in [0,T]\}$ as follows:

$$X(t) = X_0 + \int_0^t \dot{X}(s) \, \mathrm{d}s + \int_0^t \boldsymbol{\Lambda}_X^T(s) \, \mathrm{d}\mathbf{W}(s), \qquad t \in [0,T]. \tag{3.4}$$

It is clear from (3.1) and (3.2) that the $\mathrm{d}s$ and $\mathrm{d}\mathbf{W}(s)$ integrals on the right side of (3.4) are well defined and that

$$\int_0^\cdot \dot{X}(s) \, \mathrm{d}s \in \mathcal{FV}_0^c(\{\mathcal{F}_t\}) \quad \& \quad \int_0^\cdot \boldsymbol{\Lambda}_X^T(s) \, \mathrm{d}\mathbf{W}(s) \in \mathcal{M}_{0,\mathrm{loc}}^c(\{\mathcal{F}_t\}). \tag{3.5}$$

The next result establishes that the integrands $\dot{X}$ and $\boldsymbol{\Lambda}_X$ on the right side of (3.5) are essentially unique. The proof of this result is relegated (along with most proofs in the thesis) to Appendix A:

Lemma 3.2. *Suppose that for some $V_1, V_2 \in \mathcal{H}_1$ and $\boldsymbol{\xi}_1, \boldsymbol{\xi}_2 \in \mathcal{H}_2$, we have*

$$X(t) = X_0 + \int_0^t V_1(s) \, \mathrm{d}s + \int_0^t \boldsymbol{\xi}_1^T(s) \, \mathrm{d}\mathbf{W}(s), \qquad t \in [0,T], \tag{3.6}$$

and

$$X(t) = X_0 + \int_0^t V_2(s)\,\mathrm{d}s + \int_0^t \boldsymbol{\xi}_2^T(s)\,\mathrm{d}\mathbf{W}(s), \qquad t \in [0,T]. \tag{3.7}$$

Then

$$V_1 = V_2 \quad \& \quad \boldsymbol{\xi}_1 = \boldsymbol{\xi}_2 \quad a.e. \qquad \textit{(recall Notation 2.14)}. \tag{3.8}$$

◇

In view of Lemma 3.2 we see that there is a linear bijective relation between the vector space of all Itô processes of the form (3.4) and the vector space $\mathbb{I}$ of triples $(X_0, \dot{X}, \mathbf{\Lambda}_X)$.

Remark 3.3. From (3.4) and the uniqueness established in Lemma 3.2 one sees that, if $X \equiv (X_0, \dot{X}, \mathbf{\Lambda}_X) \in \mathbb{I}$ satisfies the wealth relation (2.19) for some $\boldsymbol{\pi} \in \Pi$, then the following identities must hold:

$$\dot{X}(t) \;:=\; X(t)[r(t) + \boldsymbol{\pi}^T(t)[\boldsymbol{\sigma}(t)]\boldsymbol{\theta}(t)] \tag{3.9}$$

$$\mathbf{\Lambda}_X \;:=\; X(t)[\boldsymbol{\sigma}(t)]^T\boldsymbol{\pi}(t). \tag{3.10}$$

Motivated by (3.9) and (3.10), we shall now restrict out attention to members $X \equiv (X_0, \dot{X}, \mathbf{\Lambda}_X) \in \mathbb{I}$ that are "similar" in structure. To do so, let us first define for all $X \equiv (X_0, \dot{X}, \mathbf{\Lambda}_X) \in \mathbb{I}$ the following:

$$\mathcal{U}(X) := \{\boldsymbol{\pi} \in \Pi | \dot{X}(t) = X(t)[r(t) + \boldsymbol{\pi}^T(t)[\boldsymbol{\sigma}(t)]\boldsymbol{\theta}(t)] \text{ a.e.} \quad \& \quad \mathbf{\Lambda}_X = X(t)[\boldsymbol{\sigma}(t)]^T\boldsymbol{\pi}(t) \text{ a.e.}\}. \tag{3.11}$$

Notice that, if $\mathcal{U}(X) \neq \emptyset$ for some $X \equiv (X_0, \dot{X}, \mathbf{\Lambda}_X) \in \mathbb{I}$, then every member of $\mathcal{U}(X)$ is a portfolio process $\boldsymbol{\pi} \in \Pi$ for which X satisfies the wealth dynamics at (2.19).

◁

From the uniqueness of the solution of (2.20), we have the following lemma:

Lemma 3.4. *For all $X \in \mathbb{I}$ and $\boldsymbol{\pi} \in \Pi$ we have the equivalence*

$$X = X^{\boldsymbol{\pi}} \quad a.e. \qquad \textit{(where } X^{\boldsymbol{\pi}} \textit{ is defined by the SDE (2.20))} \tag{3.12}$$

if and only if

$$X_0 = x_0 \quad \& \quad \boldsymbol{\pi} \in \mathcal{U}(X) \quad a.e.. \tag{3.13}$$

◇

Remark 3.5. Using Lemma 3.4 we can rewrite the primal value V at (2.26) as

$$V = \sup_{\substack{X \in \mathbb{I},\, X_0 = x_0 \\ \mathcal{U}(X) \neq \emptyset}} E[U(X(T))] \tag{3.14}$$

or equivalently, as a (more convenient) minimization problem:

$$-V = \inf_{\substack{X \in \mathbb{I},\, X_0 = x_0 \\ \mathcal{U}(X) \neq \emptyset}} E[-U(X(T))]. \tag{3.15}$$

◁

In view of (3.15) we see that the primal problem (2.27) can be formulated as a search for some $X \in \mathbb{I}$ subject to the constraints that $\mathcal{U}(X) \neq \emptyset$ and $X_0 = x_0$ such that the infimum at (3.15) is attained. The minimizing process $X \in \mathbb{I}$ is then the optimal wealth process. As will become clear, this formulation of the primal problem lends itself particularly well to the use of conjugate duality. Thus, from now on we focus on the minimization problem (3.15) rather the maximization problem (2.27). Since the constraint $\mathcal{U}(X) \neq \emptyset$ is quite difficult to work with, we shall in the following proposition obtain an equivalent condition that is more tractable.

Proposition 3.6. *For each $X \equiv (X_0, \dot{X}, \mathbf{\Lambda}_X) \in \mathbb{I}$, with $X_0 > 0$, we have the following equivalence:*

$$\mathcal{U}(X) \neq \emptyset \tag{3.16}$$

if and only if

$$\inf_{t\in[0,T]} X(t) > 0 \quad a.s., \tag{3.17}$$

and

$$\dot{X}(t) = X(t)r(t) + \mathbf{\Lambda}_X^T(t)\boldsymbol{\theta}(t) \quad a.e.. \tag{3.18}$$

$\diamond$

Because of Proposition 3.6, when $\mathcal{U}(X) \neq \emptyset$ and $X \in \mathbb{I}$ with $X_0 > 0$ we need focus only on a subset $\mathbb{I}_1 \subset \mathbb{I}$ which is defined as follows:

$$\mathbb{I}_1 := \{X \in \mathbb{I} \mid \inf_{t\in[0,T]} X(t) > 0 \quad \text{a.s.}\}. \tag{3.19}$$

Since $x_0 > 0$ (see Condition 2.15) the value at (3.15) can be simplified to

$$-V = \inf_{\substack{X\in\mathbb{I}_1,\, X_0=x_0 \\ \mathcal{U}(X)\neq\emptyset}} E[-U(X(T))]. \tag{3.20}$$

We will now introduce $\{0, \infty\}$-valued penalty functions to account for the constraints $X_0 = x_0$ and $\mathcal{U}(X) \neq \emptyset$.

For each $(\omega, t, x, v, \boldsymbol{\xi}) \in \Omega \times [0,T] \times \mathbb{R} \times \mathbb{R} \times \mathbb{R}^N$ define:

$$l_0(x) := \begin{cases} 0 & \text{if } X_0 = x_0, \\ \infty & \text{otherwise,} \end{cases} \tag{3.21}$$

$$l_T(x) := \begin{cases} -U(x) & \text{if } x \in (0,\infty), \\ \infty & \text{otherwise,} \end{cases} \tag{3.22}$$

$$L(\omega, t, x, v, \boldsymbol{\xi}) := \begin{cases} 0 & \text{if } x \in (0,\infty) \text{ and } v = xr(\omega,t) + \boldsymbol{\xi}^T\boldsymbol{\theta}(\omega,t), \\ \infty & \text{otherwise.} \end{cases} . \tag{3.23}$$

Remark 3.7. Notice that the definition at (3.23) is really motivated by the equivalence given by Proposition 3.6. Using (3.23), for each $X \equiv (X_0, \dot{X}, \mathbf{\Lambda}_X) \in \mathbb{I}_1$ we then have

$$L(\omega, t, X(t), \dot{X}(t), \mathbf{\Lambda}_X(t)) = \begin{cases} 0 & \text{if } X(t) > 0 \text{ and } \dot{X}(t) = X(t)r(t) + \mathbf{\Lambda}_X(t)^T\boldsymbol{\theta}(t), \\ \infty & \text{otherwise.} \end{cases} \tag{3.24}$$

From Proposition 3.6 and (3.24), for each $X \in \mathbb{I}_1$ we have

$$L(\omega, t, X(t), \dot{X}(t), \mathbf{\Lambda}_X(t))) = \begin{cases} 0 & \text{if } \mathcal{U}(X) \neq \emptyset, \\ \infty & \text{otherwise,} \end{cases} \tag{3.25}$$

and therefore

$$E\int_0^T L(\omega, t, X(s), \dot{X}(s), \mathbf{\Lambda}_X(s))\,\mathrm{d}s = \begin{cases} 0 & \text{if } \mathcal{U}(X) \neq \emptyset, \\ \infty & \text{otherwise.} \end{cases} \tag{3.26}$$

$\triangleleft$

Now combining (3.21), (3.22), and (3.26) our primal value at (3.20) can be rewritten as:

$$-V = \inf_{X \in \mathbb{I}_1} \Phi(X) \tag{3.27}$$

where we have defined

$$\Phi(X) := l_0(X_0) + E[l_T(X(T))] + E\int_0^T L(\omega, t, X(s), \dot{X}(s), \mathbf{\Lambda}_X(s))\,\mathrm{d}s, \qquad \text{for all } X \in \mathbb{I}_1. \tag{3.28}$$

Notice that the first term on the right of (3.28) is a $\{0, \infty\}$-valued penalty term which encodes the constraint $X_0 = x_0$, the third term on the right of (3.28) is a $\{0, \infty\}$-valued penalty term which encodes the constraint $\mathcal{U}(X) \neq \emptyset$, while the second term on the right of (3.28) is the objective function which must be minimized. We have therefore reduced the given primal problem (2.27) to the search for an $X \in \mathbb{I}_1$ at which the infimum at (3.27) is attained, and this minimizing X is the optimal wealth process for the primal problem (2.27).

In view of (3.28) we can reduce further the set of candidate $X \in \mathbb{I}$ that we will consider for the primal problem at (3.27), namely define

$$\mathbb{I}_2 := \{X \in \mathbb{I}_1 | X_0 = x_0 \quad \& \quad \mathcal{U}(X) \neq \emptyset\}. \tag{3.29}$$

In view of (3.20) and (3.29) we see that

$$-V = \inf_{X \in \mathbb{I}_2} E[-U(X(T))], \tag{3.30}$$

and by Lemma 3.4, we have that

$$X \in \mathbb{I}_2 \qquad \Leftrightarrow \qquad X = X^{\boldsymbol{\pi}} \quad \text{for some } \boldsymbol{\pi} \in \mathcal{U}(X). \tag{3.31}$$

Remark 3.8. Fix some $X \in \mathbb{I}_2$. Then $\mathcal{U}(X) \neq \emptyset$ (from (3.29)), and in fact $\mathcal{U}(X)$ is a *single-point* set containing *just one* $\boldsymbol{\pi} \in \Pi$. This is dictated by the second condition on the right of (3.11), namely

$$\mathbf{\Lambda}_X = X(t)[\boldsymbol{\sigma}(t)]^T \boldsymbol{\pi}(t) \text{ a.e.}$$

Since $X \in \mathbb{I}_2$ is a member of $\mathbb{I}_1$ and therefore strictly positive, it follows that $\boldsymbol{\pi}$ should necessarily be given by

$$\boldsymbol{\pi}(t) := [\boldsymbol{\sigma}(t)^{-1}]^T \frac{\mathbf{\Lambda}_X(t)}{X(t)}. \tag{3.32}$$

It remains to check that $\boldsymbol{\pi}$ at (3.32) is a member of Π. Clearly $\boldsymbol{\pi}$ is $\mathcal{F}_t$-progressively measurable, and it follows from (3.32) and (2.15) that

$$\|\boldsymbol{\pi}(t)\| \leq k_1 \frac{\|\mathbf{\Lambda}_X(t)\|}{X(t)}, \qquad t \in [0, T].$$

Since $\mathbf{\Lambda}_X \in \mathcal{H}_2$ (see (3.2)) and $\inf_{t \in [0,T]} X(t) > 0$ (see (3.19)) we have that $\int_0^T \|\boldsymbol{\pi}(s)\|^2 \,\mathrm{d}t < \infty$ a.s., so that $\boldsymbol{\pi} \in \Pi$ as required. We therefore conclude the following: if $X \equiv (X_0, \dot{X}, \mathbf{\Lambda}_X) \in \mathbb{I}_2$ then $\mathcal{U}(X)$ is non-empty and a single-point set, the only member of $\mathcal{U}(X)$ being $\boldsymbol{\pi} \in \Pi$ defined by (3.32).

$\triangleleft$

From (3.29) together with (3.21), (3.22), (3.26) (3.30), and Condition 2.23, we obtain

$$\Phi(X) = E[l_T(X(T))] = E[-U(X(T))] \geq -V > -\infty, \qquad X \in \mathbb{I}_2. \tag{3.33}$$

On the other hand, for $X \in \mathbb{I}_1 \setminus \mathbb{I}_2$ we have either $X_0 \neq x_0$ or $\mathcal{U}(X) = \emptyset$ (see (3.29)), and in either case it follows from (3.28), (3.25) and (3.21) that

$$\Phi(X) = \infty \qquad \text{for all } X \in \mathbb{I}_1 \setminus \mathbb{I}_2. \tag{3.34}$$

Using (3.33) and (3.34), we then get the following:

$$\Phi(X) := \begin{cases} -E[U(X(T))] \in (-\infty, +\infty] & \text{when } X \in \mathbb{I}_2, \\ \infty & \text{when } X \in \mathbb{I}_1 \setminus \mathbb{I}_2. \end{cases} \tag{3.35}$$

With (3.35), we can rewrite (3.27) as

$$-V = \inf_{X \in \mathbb{I}_2} \Phi(X). \tag{3.36}$$

Remark 3.9. In view of (3.36), our goal is to obtain

$$-V = \inf_{X' \in \mathbb{I}_2} \Phi(X') = \Phi(X) \quad \text{for some } X \in \mathbb{I}_2. \tag{3.37}$$

Then, from Remark 3.7, we have

$$\boldsymbol{\pi}(t) := [\boldsymbol{\sigma}(t)^{-1}]^T \frac{\Lambda_X(t)}{X(t)} \qquad \text{and} \qquad \boldsymbol{\pi} \in \mathcal{U}(X), \tag{3.38}$$

and by (3.29) and Proposition 3.4

$$X = X^{\boldsymbol{\pi}} \tag{3.39}$$

in the sense of (2.20). Finally, from (3.39) and (3.37) we have that

$$V = E[U(X^{\boldsymbol{\pi}}(T))] \tag{3.40}$$

i.e. $\boldsymbol{\pi}$ defined at (3.38) in terms of $X \in \mathbb{I}_2$ is the optimal portfolio for the primal problem at (2.27), with X being the optimal wealth process.

◁

4 The space of dual process

In this section we begin the construction of an optimization problem which is dual to the primal problem (3.36). The first step in this construction is the definition of an appropriate vector space of *dual variables* on which the functional of the dual optimization problem will be defined in due course. This space of dual variables turns out to be a vector space of $\mathcal{F}_t$-semimartingales of a particular form for the joint filtration defined at (2.8). The Markov chain $\boldsymbol{\alpha}$ will play an essential role in the definition of these semimartingales, and in particular we shall require the so-called *canonical martingales* M of the Markov chain $\boldsymbol{\alpha}$ when we write these semimartingales. We therefore begin with the definition of the *canonical martingales* M of the Markov chain $\boldsymbol{\alpha}$.

Remark 4.1. For $i, j \in S$ with $i \neq j$, define the processes:

$$R_{ij}(t) \quad := \quad \sum_{0<s\leq t} I[\boldsymbol{\alpha}(s-) = i]\, I[\boldsymbol{\alpha}(s) = j], \qquad t \in [0,T], \tag{4.1}$$

$$\tilde{R}_{ij}(t) \quad := \quad \int_0^t g_{ij} I[\boldsymbol{\alpha}(s-) = i]\, \mathrm{d}s, \qquad t \in [0,T], \tag{4.2}$$

where $I[\cdot]$ is the $\{0,1\}$-valued indicator function understood as follows:

$$I[\boldsymbol{\alpha}(s-) = i] := \begin{cases} 1 & \text{if } \boldsymbol{\alpha}(s-) = i, \\ 0 & \text{otherwise.} \end{cases} \tag{4.3}$$

The g_{ij} at (4.2) are the entries of the generator matrix $[\boldsymbol{G}]$ (see Remark 2.2), and $\boldsymbol{\alpha}(s-)$ is defined at Remark B.2 (recall that the paths of $\boldsymbol{\alpha}$ are cádlág, see Condition 2.1). Using (4.1) and (4.2) we will define for each $i, j \in S$, with $i \neq j$, the *canonical martingale* $\{M_{ij}(t), t \in [0,T]\}$ associate with the Markov Chain $\boldsymbol{\alpha}$ as

$$M_{ij}(t) := R_{ij}(t) - \tilde{R}_{ij}(t) \quad t \in [0,T], \tag{4.4}$$

(we shall see at Lemma 4.4(ii) that M_{ij} is indeed a $\mathcal{F}_t$-martingale, which is also square-integrable and purely discontinuous, hence our use of the terminology "canonical martingale"). For the case where $i = j$ we simply put

$$M_{ii}(t) := 0 \quad t \in [0,T]. \tag{4.5}$$

Moreover, define the usual self-filtration $\{\mathcal{F}_t^{\boldsymbol{\alpha}}, t \in [0,T]\}$ of the Markov chain $\boldsymbol{\alpha}$ as follows:

$$\mathcal{F}_t^{\boldsymbol{\alpha}} := \sigma\{\boldsymbol{\alpha}(s), 0 \leq s \leq t\} \vee \mathcal{N}(\mathbb{P}). \tag{4.6}$$

◁

Remark 4.2. Note that, for all $i \in S$, we have that $M_{ii}(t) = 0$ for all $t \in [0,T]$. For this reason, all subsequent calculations will be referred to the case when $i \neq j$ unless specifically indicated.

◁

Remark 4.3. From (4.1) and (4.2), one can see that the following statements hold for each $t \in [0,T]$ and for i, $j \in S$ with $i \neq j$:

(1) $Rij(t)$ counts the number of times $\boldsymbol{\alpha}$ jumps from state i to state j in the time interval $[0,t]$.

(2) $\tilde{R}_{ij}(t)/g_{ij}$ is the amount of time $\boldsymbol{\alpha}$ stays in state i in the time interval $[0,t]$.

(3) R_{ij} is a $\mathcal{F}_t$-adapted, non-decreasing and cádlág process that is null at the origin.

(4) $\tilde{R}_{ij}$ is a $\mathcal{F}_t$-adapted, non-decreasing and continuous process that is null at the origin. Furthermore, the continuity of $\tilde{R}_{ij}$ implies that it is $\mathcal{F}_t$-predictable as well.

(5) Using (3) and (4), we also have that $M_{ij}(t)$ is a a $\mathcal{F}_t$-adapted and cádlág process that is null at the origin.

◁

Next, we state some properties of M_{ij} that will be useful in subsequent sections.

Lemma 4.4. *For all $i, j, a, b \in S, t \in [0,T], n \in \mathbb{N}$ and $k \in \{1, \ldots, N\}$,*

$$\textbf{(i)} \quad E[R_{ij}(t)^n] < \infty \quad \& \quad E[\tilde{R}_{ij}(t)^n] < \infty, \tag{4.7}$$

$$\textbf{(ii)} \quad M_{ij} \in \mathcal{M}_0^2(\{\mathcal{F}_t\}) \textit{ and is purely discontinuous}, \tag{4.8}$$

$$\textbf{(iii)} \quad [M_{ij}, M_{ij}](t) = R_{ij}(t) \quad a.s., \tag{4.9}$$

$$[M_{ij}, M_{ab}](t) = 0 \quad a.s. \quad \textit{if } (i,j) \neq (a,b), \tag{4.10}$$

$$\textbf{(iv)} \quad \langle M_{ij}, M_{ij} \rangle (t) = \tilde{R}_{ij}(t) \quad a.s., \tag{4.11}$$

$$\langle M_{ij}, M_{ab} \rangle (t) = 0 \quad a.s. \quad \textit{if } (i,j) \neq (a,b), \tag{4.12}$$

$$\textbf{(v)} \quad [M_{ij}, W_k](t) = \langle M_{ij}, W_k \rangle (t) = 0 \ a.s., \tag{4.13}$$

where W_k is standard Brownian Motion.

◇

Remark 4.5. For the benefit of the reader the notions of purely discontinuous martingales, as well as the square-bracket process $[M_{ij}, M_{ij}]$ and angle-bracket process $\langle M_{ij}, M_{ij} \rangle$, are summarized in Appendix B (see in particular Sections B.3, B.5 and B.6).

In subsequent sections we will be using stochastic integrals with respect to the purely discontinuous martingales M_{ij}, and we must therefore define an appropriate class of integrands. For this we shall need

Remark 4.6. Let $\mathcal{A}_{loc}^+$ be the set of all processes $A : \Omega \times [0,T] \to \mathbb{R}$ where A is a $\mathcal{F}_t$-adapted process that is increasing, cádlág and locally integrable.

◁

Definition 4.7. For $i, j \in S$ with $i \neq j$, let $L_{loc}(\{\mathcal{F}_t\}, [M_{ij}])$ be the vector space of all $\mathbb{R}$-valued $\mathcal{F}_t$-predictable process H such that the pathwise Lebesgue-Stieltjes integral

$$t \to \left[\int_0^t H^2(s) \, \mathrm{d}[M_{ij}](s) \right]^{\frac{1}{2}}, \qquad t \in [0,T], \tag{4.14}$$

(which is necessarily increasing, cádlág and $\mathcal{F}_t$-adapted) is a member of $\mathcal{A}_{loc}^+$.

◁

Remark 4.8. The significance of the space $L_{loc}(\{\mathcal{F}_t\}, [M_{ij}])$ given by Definition 4.7 is the following: for every $H \in L_{loc}(\{\mathcal{F}_t\}, [M_{ij}])$ we can define the *stochastic integral*

$$(H \bullet M_{ij})(t) := \int_0^t H(s) \, \mathrm{d}M_{ij}(s), \qquad t \in [0,T],$$

with respect to the canonical martingale M_{ij}, and furthermore this integral is a purely discontinuous $\mathcal{F}_t$-local martingale. The exact details of the construction need not concern us, and we shall require only the following simple properties of the integral (given by Theorem 26.13 of Kallenberg [8], as well as the localized version of Theorem 4 on p.102 of Liptser and Shiryayev [12] together with the comments at the top of p.103 of [12]):

Lemma 4.9. *For all $i, j \in S, t \in [0,T]$ and $H \in L_{loc}(\{\mathcal{F}_t\}, [M_{ij}])$, the stochastic integral $H \bullet M_{ij}$ has the following properties:*

$$\textbf{(i)} \quad (H \bullet M_{ij}) \text{ is purely discontinuous } \mathcal{F}_t\text{-local martingale,} \tag{4.15}$$

$$\textbf{(ii)} \quad \Delta(H \bullet M_{ij})(t) = H(t)\Delta M_{ij}(t), \tag{4.16}$$

$$\textbf{(ii)} \quad [H \bullet M_{ij}](t) := \int_0^t H^2(s)\, \mathrm{d}[M_{ij}](s). \tag{4.17}$$

◇

Next we will define a measure on the predictable σ-algebra $\mathcal{P}^\star$ which will be needed (although very occasionally) in the subsequent sections.

Remark 4.10. For $i \neq j \in S$, define the measure $\nu_{[M_{ij}]}$ on the measurable space $(\Omega \times [0,T], \mathcal{P}^\star)$ as follows:

$$\nu_{[M_{ij}]}[A] := \mathrm{E}\left[\int_0^T I_A(\omega, t)\ \mathrm{d}[M_{ij}](t)\right], \quad \forall A \in \mathcal{P}^\star, \tag{4.18}$$

(recall Notation 2.14). This measure is known as the *Doléans measure* of M_{ij}.

◁

With these definitions and results we are almost ready to define the space of dual processes, which is denoted by $\mathbb{J}$. However, we first need to define a class of $D \times D$-matrix valued predictable processes $\mathcal{H}_3$, as follows:

$$\mathcal{H}_3 := \left\{ [\mathbf{\Gamma}] = [\Gamma_{ij}]_{i,j=1}^D \;\middle|\; \Gamma_{ij} : \Omega \times [0,T] \to \mathbb{R} \text{ where } \begin{array}{c} \Gamma_{ii} := 0 \text{ and } \Gamma_{ij} \in L_{loc}(\{\mathcal{F}_t\}, [M_{ij}]) \\ \text{for all } i, j \in S \text{ with } i \neq j, \end{array} \right\} \tag{4.19}$$

The significance of the set $\mathcal{H}_3$ is that, for $[\mathbf{\Gamma}] = [\Gamma_{ij}]_{i,j=1}^D \in \mathcal{H}_3$, the sum of stochastic integrals

$$\sum_{i,j=1}^D \int_0^t \Gamma_{ij}(s)\, \mathrm{d}M_{ij}(s), \qquad t \in [0,T], \tag{4.20}$$

is defined and is a purely discontinuous $\mathcal{F}_t$-local martingale (as follows from Lemma 4.9(i)). Note that the sum at (4.20) is really over all $i, j \in S$ with $i \neq j$ (see (4.5)), and this will be understood in all such sums.

Using (4.19), we can now define the dual process space $\mathbb{J}$ as the product

$$\mathbb{J} := \mathbb{R} \times \mathcal{H}_1 \times \mathcal{H}_2 \times \mathcal{H}_3, \tag{4.21}$$

where $Y \equiv (Y_0, \dot{Y}, \mathbf{\Lambda}_Y, [\mathbf{\Gamma}^Y]) \in \mathbb{J}$ is understood to mean the $\mathbb{R}$-valued and cádlág $\mathcal{F}_t$-semimartingale $\{Y(t),\ t \in [0,T]\}$ with the form:

$$Y(t) := Y_0 + \int_0^t \dot{Y}(s)\, \mathrm{d}s + \int_0^t \mathbf{\Lambda}_Y^T(s)\, \mathrm{d}\mathbf{W}(s) + \sum_{i,j=1}^D \int_0^t \Gamma_{ij}^Y(s)\, \mathrm{d}M_{ij}(s). \tag{4.22}$$

Remark 4.11. Effectively, $\mathbb{J}$ is the vector space of all semimartingales Y with the form given by (4.22) for some quadruple $(Y_0, \dot{Y}, \mathbf{\Lambda}_Y, [\mathbf{\Gamma}^Y]) \in \mathbb{J}$. It follows from Lemma 4.4 that the integrands $\dot{Y}$ and $\mathbf{\Lambda}_Y$ are a.e.-uniquely determined on $\Omega \times [0,T]$, and that Γ_{ij}^Y is $\nu_{[M_{ij}]}$-uniquely determined on $\Omega \times [0,T]$ for each $i, j \in S$.

◁

Remark 4.12. Observe that the dual processes Y at (4.22) are different from the primal processes X at (3.4) by including the sum of $\mathrm{d}M_{ij}(s)$-integrals. The presence of these integrals at (4.22) are entirely a consequence of the regime switching in the market model, in particular the inclusion of $\boldsymbol{\alpha}$ in the basic filtration (2.8). If regime switching were not part of the market model then the basic filtration would be defined instead by

$$\mathcal{F}_t := \sigma\{\mathbf{W}(s), s \in [0,t]\} \vee \mathcal{N}(\mathbb{P}), \tag{4.23}$$

and in this case the dual variables would be semimartingales defined by

$$Y(t) := Y_0 + \int_0^t \dot{Y}(s)\,\mathrm{d}s + \int_0^t \mathbf{\Lambda}_Y^T(s)\,\mathrm{d}\mathbf{W}(s). \tag{4.24}$$

We shall see that the $\mathrm{d}M_{ij}(s)$-integrals at (4.22) are essential for construction of the dual problem in the next section.

◁

Lastly, we state an important proposition that will be essential for the construction of the dual problem in the following section.

Proposition 4.13. *Define* $\{\mathbb{M}(X,Y)(t), t \in [0,T]\}$ *for each* $Y \equiv (Y_0, \dot{Y}, \mathbf{\Lambda}_Y, [\mathbf{\Gamma}^Y]) \in \mathbb{J}$ *and* $X \equiv (X_0, \dot{X}, \mathbf{\Lambda}_X) \in \mathbb{I}_2$ *as follows:*

$$\mathbb{M}(X,Y)(t) := X(t)Y(t) - X_0Y_0 - \int_0^t [X(s)\dot{Y}(s) + \dot{X}(s)Y(s) + (\mathbf{\Lambda}^X(s))^\top \mathbf{\Lambda}_Y(s)]\,\mathrm{d}s. \tag{4.25}$$

Then $\mathbb{M}(X,Y) \in \mathcal{M}_{0,loc}(\{\mathcal{F}_t\})$ *with* $\mathbb{M}(X,Y)(0) = 0$.

◇

5 The dual problem and preliminary optimality relations

In this section, we will construct an optimization problem which is *dual* to the primal problem (2.26) in a sense that will shortly be made precise. As an essential step towards this dual optimization problem, we will first define *convex conjugate transforms* of the functions $l_0(\cdot)$, $l_T(\cdot)$ and $L(\cdot)$ (see (3.21) - (3.23)) as follows:

$$
\begin{aligned}
m_0(y) &:= \sup_{x\in\mathbb{R}} \{xy - l_0(x)\}, && (5.1)\\
m_T(y) &:= \sup_{x\in\mathbb{R}} \{x(-y) - l_T(x)\}, && (5.2)\\
M(\omega,t,y,s,\boldsymbol{\gamma}) &:= \sup_{\substack{x,v\in\mathbb{R}\\ \boldsymbol{\xi}\in\mathbb{R}^N}} \left\{xs + vy + \boldsymbol{\xi}^T\boldsymbol{\gamma} - L(\omega,t,x,v,\boldsymbol{\xi})\right\}. && (5.3)
\end{aligned}
$$

We can explicitly evaluate the conjugate transforms at (5.1) - (5.3). Firstly, from (3.21), we have

$$
m_0(y) = x_0 y, \qquad y\in\mathbb{R}, \tag{5.4}
$$

because $l_0(x) = +\infty$ for all $x \neq x_0$. Next, using (3.22), we get that

$$
m_T(y) = \tilde{U}(y), \qquad y\in\mathbb{R}, \tag{5.5}
$$

where

$$
\tilde{U}(y) := \sup_{x\in\mathbb{R}} \{U(x) - xy\}, \qquad y\in\mathbb{R}. \tag{5.6}
$$

We note from (5.6) and Condition 2.25 that

$$
\tilde{U}(y) = +\infty, \qquad \text{for all } y\in(-\infty,0). \tag{5.7}
$$

Lastly,

$$
\begin{aligned}
M(\omega,t,y,s,\boldsymbol{\gamma}) &= \sup_{\substack{x,v\in\mathbb{R}\\ \boldsymbol{\xi}\in\mathbb{R}^N}} \left\{xs + vy + \boldsymbol{\xi}^T\boldsymbol{\gamma} - L(\omega,t,x,v,\boldsymbol{\xi})\right\}\\
&= \sup_{x>0,\,\boldsymbol{\xi}\in\mathbb{R}^N} \left\{xs + y(xr(\omega,t) + \boldsymbol{\xi}^T\boldsymbol{\theta}(\omega,t) + \boldsymbol{\xi}^T\boldsymbol{\gamma})\right\}\\
&= \sup_{x>0,\,\boldsymbol{\xi}\in\mathbb{R}^N} \left\{x(s + yr(\omega,t)) + \boldsymbol{\xi}^T(y\boldsymbol{\theta}(\omega,t)) + \boldsymbol{\gamma})\right\}\\
&= \begin{cases} 0 & \text{if } s + yr(\omega,t) \le 0 \quad \& \quad y\boldsymbol{\theta}(\omega,t) + \boldsymbol{\gamma} = 0,\\ \infty & \text{otherwise.}\end{cases}
\end{aligned} \tag{5.8}
$$

Remark 5.1. From (5.1) - (5.3), together with (5.5), we immediately obtain the following inequalities:

$$
\begin{aligned}
l_0(x) + m_0(y) &\ge xy, && (5.9)\\
\tilde{U}(y) - U(x) &\ge -xy, && (5.10)\\
L(\omega,t,x,v,\boldsymbol{\xi}) + M(\omega,t,y,s,\boldsymbol{\gamma}) &\ge xs + yv + \boldsymbol{\xi}^T\boldsymbol{\gamma}, && (5.11)
\end{aligned}
$$

for all $(\omega,t)\in\Omega\times[0,T]$, and $(x,v,\boldsymbol{\xi})$, $(y,s,\boldsymbol{\gamma})\in\mathbb{R}\times\mathbb{R}\times\mathbb{R}^N$.

◁

We *tentatively* define the *dual cost function* in terms of the conjugate transforms at (5.1) - (5.3) as follows (recall that $m_T(\cdot) = \tilde{U}(\cdot)$, see (5.5))

$$\Psi(Y) := m_0(Y_0) + E[\tilde{U}(Y(T))] + E\left[\int_0^T M(t, Y(T), \dot{Y}(t), \mathbf{\Lambda}_Y(t)\,\mathrm{d}t\right], \tag{5.12}$$

for all $Y \equiv (Y_0, \dot{Y}, \mathbf{\Lambda}_Y, [\mathbf{\Gamma}^Y]) \in \mathbb{J}$. Our motivation for the definition at (5.12) results from the following *formal* argument: From (5.9) - (5.11), (5.12), (4.25), (3.28) we get the following for each $X \in \mathbb{I}_2$ and $Y \in \mathbb{J}$:

$$\begin{aligned}
&\Phi(X) + \Psi(Y)\\
=&l_0(X_0) + E\left[-U(X(T))\right] + E\left[\int_0^T L(s, X(s), \dot{X}(s), \mathbf{\Lambda}_X(s))\,\mathrm{d}s\right]\\
&\qquad\qquad + m_0(Y_0) + E\left[\tilde{U}(Y(T))\right] + E\left[\int_0^T M(s, Y(s), \dot{Y}(s), \mathbf{\Lambda}_Y(s))\,\mathrm{d}s\right]\\
=&[l_0(X_0) + m_0(Y_0)] + E[\tilde{U}(Y(T)) - U(X(T))]\\
&+ E\left[\int_0^T \{L(s, X(s), \dot{X}(s), \mathbf{\Lambda}_X(s)) + M(s, Y(s), \dot{Y}(s), \mathbf{\Lambda}_Y(s))\}\,\mathrm{d}s\right]\\
\geq& X_0Y_0 - E\left[X(T)Y(T)\right] + E\left[\int_0^T \left\{\dot{X}(s)Y(s) + X(s)\dot{Y}(s) + \mathbf{\Lambda}_X^T(s)\mathbf{\Lambda}_Y(s)\right\}\,\mathrm{d}s\right]\\
=& - E\left[\mathbb{M}(X, Y)(T)\right],
\end{aligned} \tag{5.13}$$

that is, we have *formally* established the inequality

$$\Phi(X) + \Psi(Y) \geq -E\left[\mathbb{M}(X, Y)(T)\right], \qquad (X, Y) \in \mathbb{I}_2 \times \mathbb{J}. \tag{5.14}$$

If $\mathbb{M}(X, Y)$ were actually a *supermartingale* then we would have

$$E\left[\mathbb{M}(X, Y)(T)\right] \leq E\left[\mathbb{M}(X, Y)(0)\right] = 0,$$

so that (5.14) then gives

$$\Phi(X) + \Psi(Y) \geq 0, \qquad (X, Y) \in \mathbb{I}_2 \times \mathbb{J}. \tag{5.15}$$

This inequality, which relates the primal function $\Phi(\cdot)$ and the dual function $\Psi(\cdot)$, is the essential thing needed for conjugate duality. However, at this point we do not actually know that $\mathbb{M}(X, Y)$ is a supermartingale (only that it is a local martingale, by Proposition 4.13(a)). Moreover, even if $\mathbb{M}(X, Y)$ actually were a supermartingale, the analysis at (5.13) would still be deficient, since we do not know that the expectation $E[\tilde{U}(Y(T))]$ on the right of (5.12) even exists for all $Y \in \mathbb{J}$. Indeed, from elementary measure theory we know that $E[\tilde{U}(Y(T))]$ is defined by

$$E[\tilde{U}(Y(T))] = E[\tilde{U}^+(Y(T))] - E[\tilde{U}^-(Y(T))], \tag{5.16}$$

provided that *at least one* of the terms on the right side of (5.16) is finite (i.e. has value in $[0, \infty)$). However, there is nothing in the definition of $\tilde{U}(\cdot)$ or the set of processes $\mathbb{J}$ which precludes the possibility that

$$E[\tilde{U}^+(Y(T))] - E[\tilde{U}^-(Y(T))] = +\infty \qquad \text{for some } Y \in \mathbb{J},$$

in which case $E[\tilde{U}(Y(T))]$ is undefined (here as usual $\tilde{U}^+(y) := \max\{0, \tilde{U}(y)\}$ and $\tilde{U}^-(y) := \max\{0, -\tilde{U}(y)\}$, all $y \in \mathbb{R}$. We are now going to construct a subset $\mathbb{J}_2 \subset \mathbb{J}$ such that $\mathbb{M}(X, Y)$ actually is a supermartingale for each $(X, Y) \in \mathbb{I}_2 \times \mathbb{J}_2$, and $E[\tilde{U}(Y(T))]$ is defined for all $Y \in \mathbb{J}_2$. To motivate the construction of $\mathbb{J}_2$, observe that a weak duality relation of the form (5.15) is useful only for $Y \in \mathbb{J}$ such that $\Psi(Y) < +\infty$, and does not yield any useful information when $\Psi(Y) = +\infty$. Accordingly, we restrict to dual processes $Y \equiv (Y_0, \dot{Y}, \mathbf{\Lambda}_Y, [\mathbf{\Gamma}^Y]) \in \mathbb{J}$ such that

$$M(t, Y(T), \dot{Y}(t), \mathbf{\Lambda}_Y(t)) = 0 \qquad \text{a.e.} \tag{5.17}$$

in order to avoid the third term on the right side of (5.12) taking the value $+\infty$ (recall from (5.8) that, for each $(\omega, t) \in \Omega \times [0, T]$, we have either $M(t, Y(T), \dot{Y}(t), \mathbf{\Lambda}_Y(t)) = +\infty$ or $M(t, Y(T), \dot{Y}(t), \mathbf{\Lambda}_Y(t)) = 0$). In view of (5.8) we should therefore restrict to $Y \equiv (Y_0, \dot{Y}, \mathbf{\Lambda}_Y, [\mathbf{\Gamma}^Y]) \in \mathbb{J}$ such that

$$\dot{Y}(t) + r(t)Y(t) \leq 0 \qquad \text{and} \qquad \mathbf{\Lambda}_Y(t) + Y(t)\boldsymbol{\theta}(t) = 0 \text{ a.e.} \tag{5.18}$$

since we then get $M(t, Y(T), \dot{Y}(t), \mathbf{\Lambda}_Y(t)) = 0$ (from (5.8) we will get $M(t, Y(T), \dot{Y}(t), \mathbf{\Lambda}_Y(t)) = +\infty$ when the conditions at (5.18) fail). Similarly, in view of (5.7), we should restrict to $Y \in \mathbb{J}$ such that

$$Y(T) \geq 0 \qquad \text{a.s.} \tag{5.19}$$

in order to avoid the second term on the right of (5.12) taking the value $+\infty$ (since it follows from (5.7) that $\tilde{U}(Y(T)) = +\infty$ when $Y(T) < 0$). In fact, in order to establish Proposition 5.2 (which follows, and which is the essential tool for applying conjugate duality), we shall actually strengthen the condition at (5.19) to the condition

$$Y(t) \geq 0 \text{ a.e.} \qquad \text{on} \qquad \Omega \times [0, T]. \tag{5.20}$$

In light of (5.20) and (5.18) we therefore restrict to dual processes $Y \in \mathbb{J}_2 \subset \mathbb{J}$ defined by

$$\mathbb{J}_2 := \left\{ Y \in \mathbb{J} \;\middle|\; Y(t) \geq 0 \text{ a.e.}, \; \dot{Y}(t) + r(t)Y(t) \leq 0 \;\&\; \mathbf{\Lambda}_Y(t) + Y(t)\boldsymbol{\theta}(t) = 0 \text{ a.e.} \right\}. \tag{5.21}$$

Having defined $\mathbb{J}_2$, we then have the following essential result:

Proposition 5.2. *For each $X \in \mathbb{I}_2$, $Y \in \mathbb{J}_2$ we have:*

(a) $\mathbb{M}(X, Y)(t) \geq X(t)Y(t) - X_0Y_0$ *a.e. and* $\mathbb{M}(X, Y)$ *is a* $\mathcal{F}_t$*-supermartingale*

(b) $E[\tilde{U}^-(Y(T))] < \infty$ *(here* $\tilde{U}^-(y) := \max\{0, -\tilde{U}(y)\}$*,* $y \in \mathbb{R}$*) is the usual measure-theoretic* negative part *of the function* $\tilde{U}(\cdot)$*).*

◇

It follows from (5.8) that $M(t, Y(T), \dot{Y}(t), \mathbf{\Lambda}_Y(t)) = 0$ a.e. when $Y \in \mathbb{J}_2$, that is

$$E\left[\int_0^T M(t, Y(t), \dot{Y}(t), \mathbf{\Lambda}_Y(t) \, \mathrm{d}t\right] = 0, \qquad \text{for all} \qquad Y \in \mathbb{J}_2. \tag{5.22}$$

Moreover, it follows from Proposition 5.2(b) that $E[\tilde{U}(Y(T))]$ is well defined for each $Y \in \mathbb{J}_2$. In view of this, together with (5.22) and (5.12), we shall now define the dual cost functional as follows:

$$\Psi(Y) := x_0Y_0 + E[\tilde{U}(Y(T))] \qquad \text{for all } \; Y \in \mathbb{J}_2. \tag{5.23}$$

Note that we will never need to have $\Psi(Y)$ defined for $Y \in \mathbb{J} \setminus \mathbb{J}_2$.

Remark 5.3. A useful observation is the following: For each $X \in \mathbb{I}_2$ and $Y \in \mathbb{J}_2$ we have from (5.23), (5.22) and (3.28) that

$$\begin{aligned}
&\Phi(X) + \Psi(Y) \\
=&\left\{l_0(X_0) + E[-U(X(T))] + E\left[\int_0^T L(s, X(s), \dot{X}(s), \mathbf{\Lambda}_X(s))\,\mathrm{d}s\right]\right\} \\
&\qquad + \left\{m_0(Y_0) + E[\tilde{U}(Y(T))]\right\} \\
=&l_0(X_0) + E[-U(X(T))] + E\left[\int_0^T L(s, X(s), \dot{X}(s), \mathbf{\Lambda}_X(s))\,\mathrm{d}s\right] \\
&\qquad + m_0(Y_0) + E[\tilde{U}(Y(T))] + E\left[\int_0^T M(s, Y(s), \dot{Y}(s), \mathbf{\Lambda}_Y(s))\,\mathrm{d}s\right] \\
=&[l_0(X_0) + m_0(Y_0)] + E[\tilde{U}(Y(T)) - U(X(T))] \\
&+ E\left[\int_0^T \{L(s, X(s), \dot{X}(s), \mathbf{\Lambda}_X(s)) + M(s, Y(s), \dot{Y}(s), \mathbf{\Lambda}_Y(s))\}\,\mathrm{d}s\right] \\
\geq& - E[X(T)Y(T)] + X_0Y_0 + E\left[\int_0^T \{\dot{X}(s)Y(s) + X(s)\dot{Y}(s) + \mathbf{\Lambda}_X^T(s)\mathbf{\Lambda}_Y(s)\}\,\mathrm{d}s\right] = -E[\mathbb{M}(X,Y)(T)],
\end{aligned} \tag{5.24}$$

where the inequality follows from (5.9) - (5.11), and the last equality at (5.24) follows from (4.25). We shall make use of (5.24) shortly.

With $\mathbb{J}_2$ defined at (5.21) and $\Psi(\cdot)$ defined at (5.23), we have the following result, which essentially makes rigorous the formal calculation leading to (5.15) (with Y now restricted to $Y \in \mathbb{J}_2$):

Proposition 5.4. *Suppose Condition 2.7, Condition 2.15 and Condition 2.23 hold. Then:*

(a) We have the weak duality principle

$$\Phi(X) + \Psi(Y) \geq 0, \qquad \textit{for all } (X,Y) \in \mathbb{I}_2 \times \mathbb{J}_2. \tag{5.25}$$

(b) If we have

$$\Phi(X) + \Psi(Y) = 0, \tag{5.26}$$

<u>for some</u> $(X,Y) \in \mathbb{I}_2 \times \mathbb{J}_2$ *then*

$$\Phi(X) = \inf_{X' \in \mathbb{I}_2} \Phi(X') = \sup_{Y' \in \mathbb{I}_2} [-\Psi(Y')] = -\Psi(Y), \tag{5.27}$$

that is $X \in \mathbb{I}_2$ *solves the* primal problem *of minimizing* $\Phi(\cdot)$ *over* $\mathbb{I}_2$ *(recall (3.36) and Remark 3.9), and* $Y \in \mathbb{J}_2$ *solves the* dual problem *of minimizing* $\Psi(\cdot)$ *over* $\mathbb{J}_2$*.*

(c) For arbitrary $(X,Y) \in \mathbb{I}_2 \times \mathbb{J}_2$ *we have the equivalence*

$$\Phi(X) + \Psi(Y) = 0 \tag{5.28}$$

if and only if

$$\begin{aligned}
&(1)\ \Psi(Y) \in \mathbb{R}, \\
&(2)\ \tilde{U}(Y(T)) - U(X(T)) = -X(T)Y(T), \\
&(3)\ L(t, X(t), \dot{X}(t), \mathbf{\Lambda}_X(t)) + M(t, Y(t), \dot{Y}(t), \mathbf{\Lambda}_Y(t)), \\
&\qquad\qquad = \dot{X}(t)Y(t) + X(t)\dot{Y}(t) + \mathbf{\Lambda}_X^T(t)\mathbf{\Lambda}_Y(t), \\
&(4)\ \mathbb{M}(X,Y)\ \textit{ is a } \mathcal{F}_t\textit{-martingale and } \mathbb{M}(X,Y)(0) = 0.
\end{aligned} \tag{5.29}$$

◇

Remark 5.5. Proposition 5.4 is the key result for using conjugate duality to address the primal problem (3.36), for it tells us that, if we can somehow construct a pair $(X, Y) \in \mathbb{I}_2 \times \mathbb{J}_2$ such that (5.26) holds, then in particular X solves the primal problem (3.36). Moreover, from the equivalence of (5.28) and (5.29), the construction of such a pair reduces to the construction of some $(X, Y) \in \mathbb{I}_2 \times \mathbb{J}_2$ which satisfies the *optimality relations* (5.29)(1)-(4). In view of the importance of Proposition 5.4 we give the proof here in the main body of the thesis, instead of relegating the proof to Appendix A:

Proof of Proposition 5.4:
(a) The inequality $\Phi(X) + \Psi(Y) \geq 0$ for any $(X, Y) \in \mathbb{I}_2 \times \mathbb{J}_2$ follows from (5.13) and the fact that $\mathbb{M}(X, Y)$ is a $\mathcal{F}_t$-supermartingale null at the origin (established in Proposition 5.2**(a)**).

(b) From the result established at (a) we see that

$$\Phi(X') + \Psi(Y') \geq 0 \qquad \text{for all } X' \in \mathbb{I}_2 \text{ and } Y' \in \mathbb{J}_2, \tag{5.30}$$

and therefore, from (5.30), we have

$$\Phi(X) \geq \inf_{X' \in \mathbb{I}_2} \Phi(X') \geq \sup_{Y' \in \mathbb{J}_2} [-\Psi(Y')] \geq -\Psi(Y) \tag{5.31}$$

for each $(X, Y) \in \mathbb{I}_2 \times \mathbb{J}_2$. When $(X, Y) \in \mathbb{I}_2 \times \mathbb{J}_2$ also satisfies (5.26) then (5.27) follows at once from (5.31).

(c) Suppose that $X \in \mathbb{I}_2$, $Y \in \mathbb{J}_2$ and (5.28) hold. From (3.36) and (5.28),

$$V \geq -\Phi(X) = \Psi(Y). \tag{5.32}$$

Moreover, by basic measure theory,

$$E[\tilde{U}(Y(T))] = E[\tilde{U}^+(Y(T))] - E[\tilde{U}^-(Y(T))], \tag{5.33}$$

and it follows from (5.33) and Proposition 5.2(b) that

$$E[\tilde{U}(Y(T))] > -\infty. \tag{5.34}$$

In view of (5.34) and (5.32) we get (5.29)(1).

To establish (5.29)(2)(3)(4)we are going to use (5.24). In view of Proposition 5.2(a) we have that

$$-E\left[\mathbb{M}(X, Y)(T)\right] \geq -E\left[\mathbb{M}(X, Y)(0)\right] = 0. \tag{5.35}$$

In view of (5.35), (5.24) and (5.26) we obtain

$$0 = \Phi(X) + \Psi(Y) \geq -E\left[\mathbb{M}(X, Y)(T)\right] \geq -E\left[\mathbb{M}(X, Y)(0)\right] = 0, \tag{5.36}$$

which implies

$$\Phi(X) + \Psi(Y) = E\left[\mathbb{M}(X, Y)(T)\right] = E\left[\mathbb{M}(X, Y)(0)\right]. \tag{5.37}$$

From (5.37) and (5.24) we find

$$\begin{aligned}
&[l_0(X_0) + m_0(Y_0)] + E[\tilde{U}(Y(T)) - U(X(T))] \\
&+ E\left[\int_0^T \{L(s, X(s), \dot{X}(s), \mathbf{\Lambda}_X(s)) + M(s, Y(s), \dot{Y}(s), \mathbf{\Lambda}_Y(s))\}\, \mathrm{d}s\right] \\
=& - E[X(T)Y(T)] + X_0 Y_0 + E\left[\int_0^T \{\dot{X}(s)Y(s) + X(s)\dot{Y}(s) + \mathbf{\Lambda}_X^T(s)\mathbf{\Lambda}_Y(s)\}\, \mathrm{d}s\right] = 0,
\end{aligned} \tag{5.38}$$

that is

$$
\begin{aligned}
&[l_0(X_0) + m_0(Y_0) - X_0Y_0] + E[\tilde{U}(Y(T)) - U(X(T)) + X(T)Y(T)] \\
&+ E\left[\int_0^T L(s, X(s), \dot{X}(s), \mathbf{\Lambda}_X(s)) + M(s, Y(s), \dot{Y}(s), \mathbf{\Lambda}_Y(s))\right. \\
&\left. -\dot{X}(s)Y(s) - X(s)\dot{Y}(s) - \mathbf{\Lambda}_X^T(s)\mathbf{\Lambda}_Y(s)\} \,\mathrm{d}s\right] = 0.
\end{aligned} \tag{5.39}
$$

But, from (5.9) - (5.11), we have

$$
l_0(X_0) + m_0(Y_0) - X_0Y_0 \geq 0, \tag{5.40}
$$

$$
\tilde{U}(Y(T)) - U(X(T)) + X(T)Y(T) \geq 0 \qquad \text{a.s.} \tag{5.41}
$$

$$
L(s, X(s), \dot{X}(s), \mathbf{\Lambda}_X(s)) + M(s, Y(s), \dot{Y}(s), \mathbf{\Lambda}_Y(s)) - \dot{X}(s)Y(s) - X(s)\dot{Y}(s) - \mathbf{\Lambda}_X^T(s)\mathbf{\Lambda}_Y(s) \geq 0 \;\; \text{a.e.} \tag{5.42}
$$

In view of (5.39) - (5.42) we get in particular that

$$
\tilde{U}(Y(T)) - U(X(T)) + X(T)Y(T) = 0 \qquad \text{a.s.} \tag{5.43}
$$

and

$$
L(s, X(s), \dot{X}(s), \mathbf{\Lambda}_X(s)) + M(s, Y(s), \dot{Y}(s), \mathbf{\Lambda}_Y(s)) - \dot{X}(s)Y(s) - X(s)\dot{Y}(s) - \mathbf{\Lambda}_X^T(s)\mathbf{\Lambda}_Y(s) = 0 \;\; \text{a.e.} \tag{5.44}
$$

Now (5.43) and (5.44) are (5.29)(2)(3). Moreover, since $E\left[\mathbb{M}(X,Y)(T)\right] = E\left[\mathbb{M}(X,Y)(0)\right]$ (see (5.37)) and $\mathbb{M}(X,Y)$ is a $\mathcal{F}_t$-supermartingale (see Proposition 5.2(a)) it follows that $\mathbb{M}(X,Y)$ is a $\mathcal{F}_t$-martingale, which is (5.29)(4). We have therefore established that (5.28) implies (5.29)(1)(2)(3)(4)for arbitrary $(X,Y) \in \mathbb{I}_2 \times \mathbb{J}_2$.

To show the converse, suppose that (5.29)(1) - (4) hold for some $(X,Y) \in \mathbb{I}_2 \times \mathbb{J}_2$. First note that because $X \in \mathbb{I}_2$ and $Y \in \mathbb{J}_2$, it follows from (5.4) and (3.21) that

$$
l_0(X_0) + m_0(Y_0) = X_0Y_0. \tag{5.45}
$$

By (4.25), (5.29)(2)(3), (5.45) and (5.29)(4), we can rewrite (5.24) as:

$$
\begin{aligned}
\Phi(X) + \Psi(Y) &= -X(t)Y(t) + X_0Y_0 + E\left[\int_0^T \dot{X}(s)Y(s) + X(s)\dot{Y}(s) + \mathbf{\Lambda}_X^T(s)\mathbf{\Lambda}_Y(s)\,\mathrm{d}s\right] \\
&= -E\left[\mathbb{M}(X,Y)(T)\right] \\
&= -E\left[\mathbb{M}(X,Y)(0)\right] \\
&= 0,
\end{aligned} \tag{5.46}
$$

which is exactly (5.28).

□

6 Refined optimality relations

Remark 6.1. From the discussion at Remark 5.5 we know that it is enough to construct a pair $(X, Y) \in \mathbb{I}_2 \times \mathbb{J}_2$ which satisfies the optimality relations (5.29)(1)-(4) of Proposition 5.4, for then the first element $X \in \mathbb{I}_2$ solves the primal problem in the sense that

$$-V = \inf_{X' \in \mathbb{I}_2} \Phi(X') = \Phi(X), \tag{6.1}$$

and we can obtain the optimal portfolio $\boldsymbol{\pi}$ in terms of X from (3.38), that is

$$\boldsymbol{\pi}(t) := [\boldsymbol{\sigma}(t)^{-1}]^T \frac{\Lambda_X(t)}{X(t)}, \qquad t \in [0, T]. \tag{6.2}$$

Unfortunately the optimality relations (5.29)(1)-(4) are not particularly easy to work with, and in the present section our goal is to refine these optimality relations into optimality relations which are *fully equivalent* to (5.29)(1)-(4) but much easier to use. In fact these refined optimality relations will be essential for addressing the concrete examples in Section 7. We start with the following lemma, which will be the tool for refining in particular the optimality relation (5.29)(3)(the proof of Lemma 6.2 is in Appendix A).

Lemma 6.2. *For each* $(\omega, t) \in \Omega \times [0, T]$*, and* $(x, v, \boldsymbol{\xi})$*,* $(y, s, \boldsymbol{\gamma}) \in \mathbb{R} \times \mathbb{R} \times \mathbb{R}^N$ *we have the equivalence*

$$L(\omega, t, x, v, \boldsymbol{\xi}) + M(\omega, t, y, s, \boldsymbol{\gamma}) = xs + yv + \boldsymbol{\xi}^T \boldsymbol{\gamma} \tag{6.3}$$

if and only if

$$\begin{aligned} &(1)\ x > 0, \\ &(2)\ v = xr(t) + \boldsymbol{\xi}^T \boldsymbol{\theta}(\omega, t), \\ &(3)\ s + yr(t) = 0, \\ &(4)\ y\boldsymbol{\theta}(\omega, t) + \boldsymbol{\gamma} = 0. \end{aligned} \tag{6.4}$$

◇

Remark 6.3. Fixing some $(X, Y) \in \mathbb{I}_2 \times \mathbb{J}_2$, from Lemma 6.2 with $(x, v, \boldsymbol{\xi}) := (X(t), \dot{X}(t), \boldsymbol{\Lambda}_X(t))$ and $(y, s, \boldsymbol{\gamma}) := (Y(y), \dot{Y}(t), \boldsymbol{\Lambda}_Y(t))$, one sees that (5.29)(3) i.e.

$$L(t, X(t), \dot{X}(t), \boldsymbol{\Lambda}_X(t)) + M(t, Y(t), \dot{Y}(t), \boldsymbol{\Lambda}_Y(t)) = \dot{X}(t)Y(t) + X(t)\dot{Y}(t) + \boldsymbol{\Lambda}_X^T(t)\boldsymbol{\Lambda}_Y(t) \ \text{ a.e.} \tag{6.5}$$

is equivalent to the following:

$$\begin{aligned} &(1)\ X(t) > 0 \quad \text{a.e.} \\ &(2)\ \dot{X}(t) = X(t)r(t) + \boldsymbol{\Lambda}_X^T(t)\boldsymbol{\theta}(t) \quad \text{a.e.} \\ &(3)\ \dot{Y}(t) + Y(t)r(t) = 0 \quad \text{a.e.} \\ &(4)\ Y(t)\boldsymbol{\theta}(t) + \boldsymbol{\Lambda}_Y(t) = 0 \quad \text{a.e.} \end{aligned} \tag{6.6}$$

◁

In the following remark, we give some properties of $\tilde{U}(\cdot)$ (recall (5.6)) that will be needed in refining (5.29)(2).

Remark 6.4. If we denote by $I : (0, \infty) \to (0, \infty)$ the inverse of $U^{(1)} : (0, \infty) \to (0, \infty)$, then Condition 2.21 ensures that I is continuous and strictly decreasing. Moreover we have

$$\tilde{U}^{(1)}(y) = -I(y), \qquad y \in (0, \infty), \tag{6.7}$$

which implies that $\tilde{U}$ is continuous, convex and strictly decreasing for $y \in (0, \infty)$.

◁

Using the property given by Remark 6.4, we can use the following Lemma to refine (5.29)(2).

Lemma 6.5. *For all* $x,y \in \mathbb{R}$*,*

$$\tilde{U}(y) + U(x) = -xy \tag{6.8}$$

if and only if

$$y > 0 \qquad \& \qquad x = I(y). \tag{6.9}$$

◇

Remark 6.6. It follows from Lemma 6.5 that for some arbitrary $(X, Y) \in \mathbb{I}_2 \times \mathbb{J}_2$, the optimality condition stipulated at (5.29)(2) namely,

$$\tilde{U}(Y(T)) - U(X(T)) = -X(T)Y(T) \tag{6.10}$$

is equivalent to the following two conditions,

$$\begin{aligned} &(1)\ Y(T) > 0 \quad \text{a.s.} \\ &(2)\ X(T) = I(Y(T)) \quad \text{a.s.} \end{aligned} \tag{6.11}$$

◁

Before we can state the refined version of Proposition 5.4, we need to introduce one more lemma, which allows us to refine the relation (5.29)(4).

Lemma 6.7. *Suppose that (5.29)(2)(3) hold for some* $X \in \mathbb{I}_2$ *and* $Y \in \mathbb{J}_2$*. Then for all* $t \in [0, T]$ *we have*

$$\mathbb{M}(X, Y)(t) = X(t)Y(t) - X_0Y_0 \tag{6.12}$$

and

$$\inf_{t\in[0,T]} Y(t) > 0 \quad a.s. \tag{6.13}$$

◇

Remark 6.8. From now on we will only deal with $(X, Y) \in \mathbb{I}_2 \times \mathbb{J}_2$ such that (5.29)(1) - (4) hold. Then, motivated by Lemma 6.2 and Lemma 6.7, we will therefore restriction our attention to $Y \in \mathbb{J}_2$ such that (6.13) and (6.6)(3)hold, that is we restrict to processes $Y \in \mathbb{J}_3 \subseteq \mathbb{J}_2$, where $\mathbb{J}_3$ is defined as follows:

$$\mathbb{J}_3 := \left\{ Y \in \mathbb{J}_2 \;\middle|\; \inf_{t\in[0,T]} Y(t) > 0 \quad \& \quad \dot{Y}(t) + Y(t)r(t) = 0 \quad \text{a.e.} \right\}. \tag{6.14}$$

◁

Using the the results from Lemma 6.2, Lemma 6.5 and Lemma 6.7, we can now give the refined version of the optimality relation in the following proposition.

Proposition 6.9. *For each* $(X, Y) \in \mathbb{I}_2 \times \mathbb{J}_3$*, we have the equivalence*

$$\Phi(X) + \Psi(Y) = 0 \tag{6.15}$$

if and only if

$$\begin{aligned} &(1)\ \Psi(Y) \in \mathbb{R}, \\ &(2)\ X(T) = I(Y(T)) \quad a.s., \\ &(3)\ \dot{Y}(t) + Y(t)r(t) = 0 \quad a.e., \\ &(4)\ Y(t)\boldsymbol{\theta}(t) + \boldsymbol{\Lambda}_Y(t) = 0 \quad a.e., \\ &(5)\ XY \text{ is a } \mathcal{F}_t\text{-martingale.} \end{aligned} \tag{6.16}$$

$\diamond$

We next observe that Conditions (6.16) can be simplified. In fact, when $Y \in \mathbb{J}_3$ then it is immediate from (6.14) and (5.21) that (6.16)(3)(4)already hold, so these conditions do not give any useful information and can be removed. We can then simplify Proposition 6.9 as follows:

Proposition 6.10. *For each* $(X, Y) \in \mathbb{I}_2 \times \mathbb{J}_3$*, we have the equivalence*

$$\Phi(X) + \Psi(Y) = 0 \tag{6.17}$$

if and only if

$$\begin{aligned} &(1)\ \Psi(Y) \in \mathbb{R}, \\ &(2)\ X(T) = I(Y(T)) \quad a.s., \\ &(3)\ XY \text{ is a } \mathcal{F}_t\text{-martingale.} \end{aligned} \tag{6.18}$$

$\diamond$

The optimality relations (6.18) are certainly easier to work with than the relations (5.29). However, the relations (6.18) become even more tractable when the dual process $Y \in \mathbb{J}_3$ (which is strictly positive) is put in *exponential* form. To do this, we first take $\boldsymbol{\theta}$ defined in Remark 2.12, the $\mathbb{R}^N$-valued Brownian and Motion given at (2.1) and define the usual Ito stochastic integral $(\boldsymbol{\theta} \bullet \mathbf{W})$ as:

$$(\boldsymbol{\theta} \bullet \mathbf{W})(t) := \int_0^t \boldsymbol{\theta}^T(s)\, \mathrm{d}\mathbf{W}(s) = \sum_{k=1}^{N} \int_0^t \theta_k(s)\, \mathrm{d}W_k(s), \quad t \in [0, T]. \tag{6.19}$$

Using (6.19), define the process $\{H(t), t \in [0, T]\}$ as:

$$H(t) := exp\left\{-\int_0^t r(s)\, \mathrm{d}s\right\} \mathcal{E}(-\boldsymbol{\theta} \bullet \mathbf{W})(t), \tag{6.20}$$

where $\mathcal{E}(-\boldsymbol{\theta} \bullet \mathbf{W})$ is the stochastic exponential (recall Remark B.44) of $(-\boldsymbol{\theta} \bullet \mathbf{W})$ (the process H is just the familiar *state price density* process).

Moreover for all $[\boldsymbol{\mu}] \in \mathcal{H}_3$ (recall (4.19)), define the $\mathcal{F}_t$-local martingale $([\boldsymbol{\mu}] \bullet M)$ as follows:

$$([\boldsymbol{\mu}] \bullet M)(t) := \sum_{i,j=1}^{D} \mu_{ij} \bullet M_{ij}(t) = \sum_{i,j=1}^{D} \int_0^t \mu_{ij}(s)\, \mathrm{d}M_{ij}(s), \quad t \in [0, T]. \tag{6.21}$$

The processes defined at (6.20) and (6.21) will play a key role in expressing the dual process $Y \in \mathbb{J}_3$ in exponential form.

From Lemma 4.9 **(i)**, the $\mathcal{F}_t$-local martingale $([\boldsymbol{\mu}] \bullet M)$ defined at (6.21) is *purely discontinuous* (see Definition B.17), the stochastic exponential $\mathcal{E}([\boldsymbol{\mu}] \bullet M)$ therefore need not be $\mathbb{P}$-strictly positive. For that reason, we shall from now on only focus on $\mathcal{G}_3 \subseteq \mathcal{H}_3$ where for each $[\boldsymbol{\mu}] \in \mathcal{G}_3$, the stochastic exponential $\mathcal{E}([\boldsymbol{\mu}] \bullet M)$ is strictly positive i.e.

$$\mathcal{G}_3 := \left\{ [\boldsymbol{\mu}] \in \mathcal{H}_3 \,\middle|\, \inf_{t \in [0,T]} \mathcal{E}([\boldsymbol{\mu}] \bullet M)(t) > 0 \quad \text{a.s.} \right\}. \tag{6.22}$$

Finally, using (6.20), (6.21) and (6.22), we have following important proposition which allows us to define Y in exponential form as well as further refine the optimality conditions in Proposition 6.10.

Proposition 6.11. *For each* $(y, [\boldsymbol{\mu}]) \in (0,\infty) \times \mathcal{G}_3$*, the process* $Y := yH\mathcal{E}([\boldsymbol{\mu}] \bullet M)$ *is a member of* $\mathbb{J}$ *(recall (4.21)). More precisely,*

(a) $\mathbb{J}_3 = \{yH\mathcal{E}([\boldsymbol{\mu}] \bullet M) \mid (y, [\boldsymbol{\mu}]) \in (0,\infty) \times \mathcal{G}_3\}$

(b) *If* $Y := yH\mathcal{E}([\boldsymbol{\mu}] \bullet M)$ *for some* $(y, [\boldsymbol{\mu}]) \in (0,\infty) \times \mathcal{G}_3$*, then we have the following identities:*

(i) $Y(t)\boldsymbol{\theta}(t) + \boldsymbol{\Lambda}_Y(t) = 0 \quad t \in [0,T]$

(ii) $\dot{Y}(t) + r(t)Y(t) = 0 \quad t \in [0,T]$

◇

Remark 6.12. Proposition 6.11(a) tells us that the set of dual variables $\mathbb{J}_3$ is identical to the set of all processes $\{yH(t)\mathcal{E}([\boldsymbol{\mu}] \bullet M)(t),\ t \in [0,T]\}$ for all $y \in (0,\infty)$ and $[\boldsymbol{\mu}] \in \mathcal{G}_3$. In view of this, we shall, from now on, regard the set of dual variables as the set of all pairs $(y, [\boldsymbol{\mu}]) \in (0,\infty) \times \mathcal{G}_3$. We can then define the dual cost functional

$$\tilde{\Psi}(t, [\boldsymbol{\mu}]) := \Psi(yH\mathcal{E}([\boldsymbol{\mu}] \bullet M)), \qquad \text{for all } (y, [\boldsymbol{\mu}]) \in (0,\infty) \times \mathcal{G}_3, \tag{6.23}$$

that is (see (5.23))

$$\tilde{\Psi}(y, [\boldsymbol{\mu}]) = x_0 Y_0 + E\left[\tilde{U}(yH(T)\mathcal{E}([\boldsymbol{\mu}] \bullet M)(T))\right], \qquad (y, [\boldsymbol{\mu}]) \in (0,\infty) \times \mathcal{G}_3 \tag{6.24}$$

Using the results from Proposition 6.11 and (6.24), we can rewrite Proposition 6.10 in terms of the dual variables $(y, [\boldsymbol{\mu}]) \in (0,\infty) \times \mathcal{G}_3$:

Proposition 6.13. *For each* $X \in \mathbb{I}_2$ *and* $(y, [\boldsymbol{\mu}]) \in (0,\infty) \times \mathcal{G}_3$ *we have the equivalence*

$$\Phi(X) + \tilde{\Psi}(y, [\boldsymbol{\mu}]) = 0 \tag{6.25}$$

if and only if

$$\begin{aligned} &(1)\ \tilde{\Psi}(y, [\boldsymbol{\mu}]) \in \mathbb{R}, \\ &(2)\ X(T) = I(yH(T)\mathcal{E}([\boldsymbol{\mu}] \bullet M)(T)) \quad a.s., \\ &(3)\ XH\mathcal{E}([\boldsymbol{\mu}] \bullet M) \ \textit{is a } \mathcal{F}_t\textit{-martingale.} \end{aligned} \tag{6.26}$$

◇

Remark 6.14. In view of Proposition 6.13, suppose we constructed some $(X, y, [\boldsymbol{\mu}]) \in \mathbb{I}_2 \times (0,\infty) \times \mathcal{G}_3$ such that (6.26)(1) - (3) hold. If we define Y as:

$$Y := yH\mathcal{E}([\boldsymbol{\mu}] \bullet M), \tag{6.27}$$

then $Y \in \mathbb{J}_3$ by Proposition 6.11 **(a)**. Moreover, Proposition 6.13 also gives:

$$\Phi(X) + \Psi(Y) = 0. \tag{6.28}$$

We then have from the Weak Duality Principle at (6.19) that the optimal solution to the primal problem at (2.27) is given by (recalled Remark 3.9):

$$\boldsymbol{\pi}(t) := [\boldsymbol{\sigma}(t)^{-1}]^T \frac{\boldsymbol{\Lambda}_X(t)}{X(t)}, \qquad t \in [0,T]. \tag{6.29}$$

Moreover, $Y \in \mathbb{J}_3 \subseteq \mathbb{J}_1$ (see (6.14)) implies that Y give by (6.27) minimizes the dual problem of minimizaing $\Psi(\cdot)$ over $\mathbb{J}_1$ (and hence maximizes the the primal problem) i.e.

$$\inf_{Y' \in \mathbb{J}_1} \Psi(Y') = \Psi(Y) = \Psi(yH\mathcal{E}([\boldsymbol{\mu}] \bullet M)). \tag{6.30}$$

◁

7 Example problems

Remark 7.1. In view of Remark 6.14 it is enough to construct some $(X, y, [\boldsymbol{\mu}]) \in \mathbb{I}_2 \times (0, \infty) \times \mathcal{G}_3$ such that the relations (6.26)(1) - (3) of Proposition 6.13 hold, for then the optimal portfolio is given in terms of the primal variable X by (6.29), that is

$$\boldsymbol{\pi}(t) = [\boldsymbol{\sigma}(t)^{-1}]^T \frac{\boldsymbol{\Lambda}_X(t)}{X(t)}, \qquad t \in [0, T]. \tag{7.1}$$

In this section we are going to concentrate on the particular cases of the *power utility* and *logarithmic utility* functions, and in each case we shall obtain the optimal portfolio in *explicit form*, meaning that, at each instant $t \in [0, T]$, the optimal portfolio $\boldsymbol{\pi}(t)$ is explicitly computable in terms of the (known) market parameters $r(t)$, $\mathbf{b}(t)$ and $[\boldsymbol{\sigma}](t)$ at the same instant t. The results of this section will be a significant improvement on the results of Sotomayor and Cadenillas [16], who likewise concentrate on the power utility and logarithmic utility functions (we discuss this further below).

◁

7.1 Power utility function

In this section we shall address problem (2.27) in the case of the *power utility function*, that is, we shall strengthen Condition 2.21 to the following:

Condition 7.2. In Condition 2.21 it is further stipulated that

$$U(x) := \frac{1}{\beta} x^{\beta}, \qquad x \in (0, \infty) \tag{7.2}$$

where $\beta \in (-\infty, 1)\backslash\{0\}$ is a fixed constant.

◁

Even with the specific power utility function at Condition 7.2 it is generally not possible to get an explicit optimal portfolio without adding further conditions to problem (2.27). For example, in the case where regime switching is absent from the market model, it is necessary to suppose that the market parameters are *non-random* or *deterministic* to get an *explicit* optimal portfolio for the case of the power utility function (see e.g. Example 6.6.7 of Karatzas and Shreve [10]). Likewise, with regime switching present in the market model (through the Markov chain $\boldsymbol{\alpha}$ stipulated in Section 2), it will be necessary to add further conditions to get an explicit optimal portfolio, and indeed we shall strengthen Condition 2.7 as follows:

Condition 7.3. Condition 2.7 holds with the additional stipulation that the market parameters r, $\mathbf{b}$ and $[\boldsymbol{\sigma}]$ are $\mathcal{F}_t^{\boldsymbol{\alpha}}$-progressively measurable, where $\mathcal{F}_t^{\boldsymbol{\alpha}}$ is given by (4.6).

◁

Remark 7.4. Condition 7.3 stipulates that, at each instant $t \in [0, T]$, the market parameters $r(t)$, $\mathbf{b}(t)$ and $[\boldsymbol{\sigma}](t)$ are *completely determined* by the paths $\{\boldsymbol{\alpha}(s), 0 \le s \le t\}$ of the Markov chain $\boldsymbol{\alpha}$, and are not at all influenced by the paths $\{\mathbf{W}(s), 0 \le s \le t\}$ of the driving Brownian motion $\mathbf{W}$. This of course is significantly stronger than Condition 2.7, but is nevertheless a good deal weaker than the dependency conditions customarily posited in virtually all works on regime switching, which rely in an essential way on the condition that the market parameters be *Markov modulated* in the sense discussed at Remark 2.9, see in particular Sotomayor and Cadenillas [16] and Zhou and Yin [18] (an exception to this are the results of Donnelly*et-al* [5], in which explicit optimal portfolios are likewise obtained subject to Condition 7.3 in the case of quadratic risk minimization). Our reason for stipulating Condition 7.3 is that it facilitates application of the following *martingale representation theorem* for the Markov chain $\boldsymbol{\alpha}$, which will be indispensable for securing the process $[\boldsymbol{\mu}]$ when we construct some $(X, y, [\boldsymbol{\mu}]) \in \mathbb{I}_2 \times (0, \infty) \times \mathcal{G}_3$ such that the relations (6.26)(1) - (3) of Proposition 6.13 hold.

◁

Theorem 7.5. Suppose that $\{Q(t), t \in [0,T]\}$ is a $\mathcal{F}_t^{\boldsymbol{\alpha}}$-martingale such that $c_1 \leq Q(t) \leq c_2$ for all $t \in [0,T]$ where $0 < c_1 < c_2 < \infty$ are constants. Then there exists a $\nu_{[M_{ij}]}$-unique and $\mathcal{F}_t^{\boldsymbol{\alpha}}$-predictable integrand $[\boldsymbol{\mu}] \in \mathcal{G}_3$ such that $Q(t) = Q(0)\mathcal{E}\left([\boldsymbol{\mu}] \bullet M\right)(t)$ for all $t \in [0,T]$.

◇

Remark 7.6. To summarize, our goal is to address problem (2.27) assuming Condition 2.1, Condition 2.3, Condition 7.3 (which strngthens Condition 2.7), Condition 2.10, Condition 2.15, Condition 2.17, Condition 7.2 (which strengthens Condition 2.21), and Condition 2.23.

◁

Remark 7.7. If we postulate all the conditions of Remark 7.6 but further strengthen Condition 7.3 and suppose that the market parameters r, $\mathbf{b}$ and $[\boldsymbol{\sigma}]$ are *nonrandom* or *deterministic*, then problem (2.27) is solved in Ex 3.6.7 Karatzas and Shreve [10], where it is established that the optimal portfolio process is given by

$$\boldsymbol{\pi}(t) = \frac{[\boldsymbol{\sigma}(t)^{-1}]^T \boldsymbol{\theta}(t)}{1-\beta}, \qquad t \in [0,T]. \tag{7.3}$$

Since the underlying Brownian motion $\mathbf{W}$ is independent of the Markov chain $\boldsymbol{\alpha}$ (recall Condition 2.3), it seems plausible that the portfolio process given by (7.3) should still be optimal even when r, $[\boldsymbol{\sigma}]$ and $\boldsymbol{\theta}$ are random in the sense of Condition 7.3. This intuition was formulated and used by Karatzas and Shreve (see Example 6.7.4 on p.305 of [10]) for a problem of unconstrained utility maximization in which the market parameters depend non-anticipatively on one Brownian motion which is independent of the Brownian motion driving the price model, and goes by the name of "totally unhedgeable coefficients". With Condition 7.3 we have a rather analogous situation, except that the market parameters depend non-anticipatively on the Markov chain $\boldsymbol{\alpha}$ (instead of on a Brownian motion) which is independent of the Brownian motion $\mathbf{W}$ which drives the price model. This makes it quite plausible that, under the conditions of Remark 7.6, the optimal portfolio is indeed given by (7.3). We are now going to use Proposition 6.13 as a *verification tool* to establish that this is the case. Our program is as follows: we shall construct some $X \in \mathbb{I}_2$ from the portfolio $\boldsymbol{\pi}$ at (7.3) and the SDE (2.20) (see (7.4) which follows). We shall then use the martingale representation Theorem 7.5 as a tool to construct a pair of dual variables $(y, [\boldsymbol{\mu}]) \in (0, \infty) \times \mathcal{G}_3$ such that $(X, y, [\boldsymbol{\mu}])$ satisfies the relations the relations (6.26)(1) - (3) of Proposition 6.13, and will establish that the portfolios at (7.3) and (7.1) are identical. This establishes optimality of $\boldsymbol{\pi}$ given by (7.3).

◁

Remark 7.8. Since $\boldsymbol{\theta}$ and $[\boldsymbol{\sigma}]$ are uniformly bounded and $\mathcal{F}_t^{\boldsymbol{\alpha}}$-progressively measurable, hence we also have that $\boldsymbol{\pi}$ given in (7.3) is also uniformly bounded and $\mathcal{F}_t^{\boldsymbol{\alpha}}$-progressively measurable. In particular, this implies that $\boldsymbol{\pi} \in \Pi$ (see (2.16)).

◁

In accordance with Remark 7.7 we define the wealth process X as

$$X := X^{\boldsymbol{\pi}}, \tag{7.4}$$

(recall (2.20)), where $\boldsymbol{\pi}$ is given by (7.3) so that $X \in \mathbb{I}_2$ (recall Lemma 3.4 and (3.19)). Then by (2.20) and (7.3) we have

$$\begin{aligned} \mathrm{d}X(t) &= X\left\{\left[r(t) + \frac{\boldsymbol{\theta}^T(t)[\boldsymbol{\sigma}(t)^{-1}]}{1-\beta}[\boldsymbol{\sigma}(t)]\boldsymbol{\theta}(t)\right]\,\mathrm{d}t + \frac{\boldsymbol{\theta}^T(t)[\boldsymbol{\sigma}(t)^{-1}]}{1-\beta}[\boldsymbol{\sigma}(t)]\,\mathrm{d}\mathbf{W}(t)\right\} \\ &= X\left\{\left[r(t) + \frac{\|\boldsymbol{\theta}(t)\|^2}{1-\beta}\right]\,\mathrm{d}t + \frac{\boldsymbol{\theta}^T(s)}{1-\beta}\,\mathrm{d}\mathbf{W}(t)\right\}. \end{aligned} \tag{7.5}$$

From (2.24) and (7.3) we obtain X explicitly in the form

$$\begin{aligned} X(t) &= x_0 exp\left\{\int_0^t\left[r(s) + \frac{\|\boldsymbol{\theta}(s)\|^2}{1-\beta}\right]\,\mathrm{d}s + \int_0^t \frac{\boldsymbol{\theta}^T(s)}{1-\beta}\,\mathrm{d}\mathbf{W}(s) - \frac{1}{2}\int_0^t \left\|\frac{\boldsymbol{\theta}(s)}{1-\beta}\right\|^2\,\mathrm{d}s\right\} \\ &= x_0 exp\left\{\int_0^t\left[r(s) + \frac{1-2\beta}{2(1-\beta)^2}\|\boldsymbol{\theta}(s)\|^2\right]\,\mathrm{d}s + \frac{1}{1-\beta}\int_0^t \boldsymbol{\theta}^T(s)\,\mathrm{d}\mathbf{W}(s)\right\}. \end{aligned} \tag{7.6}$$

We have now obtained a candidate of $X \in \mathbb{I}_2$ at (7.6) based on the portfolio $\boldsymbol{\pi}$ given at (7.3). We shall now use this X to construct a pair $(y, [\boldsymbol{\mu}]) \in (0,\infty) \times \mathcal{G}_3$ such that $(X, y, [\boldsymbol{\mu}])$ satisfies the relations (6.26)(1)(2)(3) (recall Remark 7.1). In fact, we will construct $(y, [\boldsymbol{\mu}]) \in (0,\infty) \times \mathcal{G}_3$ such that (6.26)(2) is satisfied, and then show that (6.26)(1)(3) are satisfied as well. To do so, first note that, from Condition 7.5, we have

$$U^{(1)}(x) = x^{\beta-1} \in (0,\infty), \qquad x \in (0,\infty). \tag{7.7}$$

From (7.7) we see that $U^{(1)} : (0,\infty) \to (0,\infty)$ has the inverse $I : (0,\infty) \to (0,\infty)$ given by

$$I(s) = \left(U^{(1)}\right)^{-1}(s) = s^{\frac{1}{\beta-1}}, \qquad s \in (0,\infty). \tag{7.8}$$

We are now going to use (6.26)(2) to establish a relation between $y \in (0,\infty)$ and $[\boldsymbol{\mu}] \in \mathcal{G}_3$. To this end fix some arbitrary $(y, [\boldsymbol{\mu}]) \in (0,\infty) \times \mathcal{G}_3$. Then, from (7.8) with $s := yH(T)\mathcal{E}\left([\boldsymbol{\mu}] \bullet M\right)(T)$ we get

$$I(yH(T)\mathcal{E}\left([\boldsymbol{\mu}] \bullet M\right)(T)) = \left[yH(T)\mathcal{E}\left([\boldsymbol{\mu}] \bullet M\right)(T)\right]^{\frac{1}{\beta-1}}. \tag{7.9}$$

From (6.19), (6.20) and Remark B.44, $H(t)$ can explicitly be written as:

$$\begin{aligned} H(t) &= exp\left\{-\int_0^t r(s)\,\mathrm{d}s\right\}\mathcal{E}(-\boldsymbol{\theta} \bullet \mathbf{W})(t) \\ &= exp\left\{-\int_0^t r(s)\,\mathrm{d}s\right\} exp\left\{-\int_0^t \boldsymbol{\theta}^T(s)\,\mathrm{d}\mathbf{W}(s) - \frac{1}{2}\int_0^t \|\boldsymbol{\theta}(s)\|^2\,\mathrm{d}s\right\} \\ &= exp\left\{-\int_0^t\left[r(s) + \frac{1}{2}\|\boldsymbol{\theta}(s)\|^2\right]\,\mathrm{d}s - \int_0^t \boldsymbol{\theta}^T(s)\,\mathrm{d}\mathbf{W}(s)\right\}, \end{aligned} \tag{7.10}$$

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

Using (7.10), we then get

$$\begin{aligned} H(T)^{\frac{1}{\beta-1}} &= \left[exp\left\{-\int_0^T\left[r(s) + \frac{1}{2}\|\boldsymbol{\theta}(s)\|^2\right]\,\mathrm{d}s - \int_0^T \boldsymbol{\theta}^T(s)\,\mathrm{d}\mathbf{W}(s)\right\}\right]^{\frac{1}{\beta-1}} \\ &= exp\left\{\frac{1}{1-\beta}\int_0^T\left[r(s) + \frac{1}{2}\|\boldsymbol{\theta}(s)\|^2\right]\,\mathrm{d}s + \frac{1}{1-\beta}\int_0^T \boldsymbol{\theta}^T(s)\,\mathrm{d}\mathbf{W}(s)\right\}. \end{aligned} \tag{7.11}$$

Using (7.11), we can then rewrite (7.9) as the following identity

$$\begin{aligned} & I(yH(T)\mathcal{E}\left([\boldsymbol{\mu}] \bullet M\right)(T)) \\ &= \left[y\mathcal{E}\left([\boldsymbol{\mu}] \bullet M\right)(T)\right]^{\frac{1}{\beta-1}} exp\left\{\frac{1}{1-\beta}\int_0^T \left[r(s) + \frac{1}{2}\left\|\boldsymbol{\theta}(s)\right\|^2\right] \mathrm{d}s + \frac{1}{1-\beta}\int_0^T \boldsymbol{\theta}^T(s)\,\mathrm{d}\mathbf{W}(s)\right\} \end{aligned} \tag{7.12}$$

which holds for all $y \in (0,\infty)$ and all $[\boldsymbol{\mu}] \in \mathcal{G}_3$. In view of (7.12) and (7.6), we see that (6.26)(2) necessarily dictates

$$\begin{aligned} & x_0 exp\left\{\int_0^T \left[r(s) + \frac{1-2\beta}{2(1-\beta)^2}\left\|\boldsymbol{\theta}(s)\right\|^2\right] \mathrm{d}s + \frac{1}{1-\beta}\int_0^T \boldsymbol{\theta}^T(s)\,\mathrm{d}\mathbf{W}(s)\right\} \\ &= \left[y\mathcal{E}\left([\boldsymbol{\mu}] \bullet M\right)(T)\right]^{\frac{1}{\beta-1}} exp\left\{\frac{1}{1-\beta}\int_0^T \left[r(s) + \frac{1}{2}\left\|\boldsymbol{\theta}(s)\right\|^2\right] \mathrm{d}s + \frac{1}{1-\beta}\int_0^T \boldsymbol{\theta}^T(s)\,\mathrm{d}\mathbf{W}(s)\right\}, \end{aligned} \tag{7.13}$$

that is, upon simplifying (7.13), we further get

$$x_0 = \left[y\mathcal{E}\left([\boldsymbol{\mu}] \bullet M\right)(T)\right]^{\frac{1}{\beta-1}} exp\left\{\frac{\beta}{1-\beta}\int_0^T \left[r(s) + \frac{1}{2(1-\beta)}\left\|\boldsymbol{\theta}(s)\right\|^2\right] \mathrm{d}s\right\}. \tag{7.14}$$

that is

$$y\mathcal{E}\left([\boldsymbol{\mu}] \bullet M\right)(T) = x_0^{\beta-1} exp\left\{\beta\int_0^T \left[r(s) + \frac{1}{2(1-\beta)}\left\|\boldsymbol{\theta}(s)\right\|^2\right] \mathrm{d}s\right\}. \tag{7.15}$$

To summarize: with $X \in \mathbb{I}_2$ given by (7.3) and (7.4), we have established that (6.26)(2) *forces* the relation (7.15) between $y \in (0,\infty)$ and $[\boldsymbol{\mu}] \in \mathcal{G}_3$. We are now going to *work backwards* from (7.15) to construct some $(y,[\boldsymbol{\mu}]) \in (0,\infty)\times\mathcal{G}_3$ such that (6.26)(2) holds. Our tool for this construction will be the martingale representation Theorem 7.5. To this end, define the process $\{Q(t), t \in [0,T]\}$ as follows:

$$Q(t) := E\left[exp\left\{\beta\int_0^T \left[r(s) + \frac{1}{2(1-\beta)}\left\|\boldsymbol{\theta}(s)\right\|^2\right] \mathrm{d}s\right\} \;\Big|\; \mathcal{F}_t^{\boldsymbol{\alpha}}\right], \qquad t \in [0,T]. \tag{7.16}$$

Since r and $\boldsymbol{\theta}$ are uniformly bounded on $\Omega \times [0,T]$ (recall Condition 2.7 and Remark 2.13) there exist constants $0 < a < b < \infty$ such that

$$a < exp\left\{\beta\int_0^T \left[r(s) + \frac{1}{2(1-\beta)}\left\|\boldsymbol{\theta}(s)\right\|^2\right] \mathrm{d}s\right\} < b, \qquad \text{for all } (\omega,t) \in \Omega\times[0,T]. \tag{7.17}$$

Using the bounds at (7.17), we then have

(1) $c_1 < Q(t) < c_2 \quad \forall t \in [0,T]$ for some constants $0 < c_1 < c_2 < \infty$ (7.18)

(2) Q is a $\mathcal{F}_t$-martingale (7.19)

By (7.18), (7.19) and Theorem 7.5 we see that there exists a $\nu_{[M_{ij}]}$-unique and $\mathcal{F}_t^{\boldsymbol{\alpha}}$-predictable integrand $[\boldsymbol{\mu}] \in \mathcal{G}_3$ such that

$$Q(t) = Q(0)\mathcal{E}\left([\boldsymbol{\mu}] \bullet M\right)(t), \qquad \text{for all } t \in [0,T]. \tag{7.20}$$

Additionally, (7.18), (7.19) and (7.20) imply the following three properties of the integrand process $[\boldsymbol{\mu}] \in \mathcal{G}_3$ that has been constructed:

(1) $c_3 < \mathcal{E}\left([\boldsymbol{\mu}] \bullet M\right)(t) < c_4 \quad \forall t \in [0,T]$ for some constant $0 < c_3 < c_4 < \infty$ (7.21)

(2) $\mathcal{E}\left([\boldsymbol{\mu}] \bullet M\right)$ is a $\mathcal{F}_t$-martingale (7.22)

(3) $E[\mathcal{E}\left([\boldsymbol{\mu}] \bullet M\right)(T)] = E[\mathcal{E}\left([\boldsymbol{\mu}] \bullet M\right)(0)] = 1$ and $E[Q(T)] = Q(0)$. (7.23)

We have now constructed some $[\boldsymbol{\mu}] \in \mathcal{G}_3$ such that (7.20) - (7.23) hold, and it remains to construct $y \in (0,\infty)$. With Q defined at (7.16), the motivating relation (7.15) takes the form

$$y\mathcal{E}\left([\boldsymbol{\mu}] \bullet M\right)(T) = x_0^{\beta-1} Q(T). \tag{7.24}$$

Taking expectations on both sides of (7.24) and using (7.23) motivates the following definition

$$y := x_0^{\beta-1} Q(0) \in (0,\infty), \tag{7.25}$$

(recall that $x_0^{\beta-1}$ and $Q(0)$ are strictly positive, from Condition 2.15 and (7.18)).

We now have a candidate $y \in (0,\infty)$ defined at (7.25) as well as a candidate $[\boldsymbol{\mu}] \in \mathcal{G}_3$ constructed with the martingale representation Theorem 7.5 such that (7.20) - (7.23) hold. Since $X \in \mathbb{I}_2$ we therefore have $(X, y, [\boldsymbol{\mu}]) \in \mathbb{I}_2 \times (0,\infty) \times \mathcal{G}_3$. It remains to verify that this $(X, y, [\boldsymbol{\mu}])$ satisfies (6.26)(1)(2)(3). We start by verifying that $(X, y, [\boldsymbol{\mu}])$ satisfies (6.26)(2). From (7.20) we have:

$$\mathcal{E}\left([\boldsymbol{\mu}] \bullet M\right)(T) = Q(T)/Q(0). \tag{7.26}$$

Now if we insert $\mathcal{E}\left([\boldsymbol{\mu}] \bullet M\right)(T)$ at (7.26) and y at (7.25) into the right hand-side of the general identity (7.12) we get

$$\begin{aligned}
&[y\mathcal{E}\left([\boldsymbol{\mu}] \bullet M\right)(T)]^{\frac{1}{\beta-1}} exp\left\{\frac{1}{1-\beta}\int_0^T \left[r(s) + \frac{1}{2}\|\boldsymbol{\theta}(s)\|^2\right] \mathrm{d}s + \frac{1}{1-\beta}\int_0^T \boldsymbol{\theta}^T(s)\,\mathrm{d}\mathbf{W}(s)\right\} \\
=&\left[\{x_0^{\beta-1}Q(0)\}\{Q(T)/Q(0)\}\right]^{\frac{1}{\beta-1}} exp\left\{\frac{1}{1-\beta}\int_0^T \left[r(s) + \frac{1}{2}\|\boldsymbol{\theta}(s)\|^2\right] \mathrm{d}s + \frac{1}{1-\beta}\int_0^T \boldsymbol{\theta}^T(s)\,\mathrm{d}\mathbf{W}(s)\right\} \\
=&\left[x_0^{\beta-1}Q(T)\right]^{\frac{1}{\beta-1}} exp\left\{\frac{1}{1-\beta}\int_0^T \left[r(s) + \frac{1}{2}\|\boldsymbol{\theta}(s)\|^2\right] \mathrm{d}s + \frac{1}{1-\beta}\int_0^T \boldsymbol{\theta}^T(s)\,\mathrm{d}\mathbf{W}(s)\right\} \\
=&x_0 Q(T)^{\frac{1}{\beta-1}} exp\left\{\frac{1}{1-\beta}\int_0^T \left[r(s) + \frac{1}{2}\|\boldsymbol{\theta}(s)\|^2\right] \mathrm{d}s + \frac{1}{1-\beta}\int_0^T \boldsymbol{\theta}^T(s)\,\mathrm{d}\mathbf{W}(s)\right\}.
\end{aligned} \tag{7.27}$$

Putting (7.27) into (7.12) we then get

$$I(yH(T)\mathcal{E}\left([\boldsymbol{\mu}] \bullet M\right)(T)) = x_0 Q(T)^{\frac{1}{\beta-1}} exp\left\{\frac{1}{1-\beta}\int_0^T \left[r(s) + \frac{1}{2}\|\boldsymbol{\theta}(s)\|^2\right] \mathrm{d}s + \frac{1}{1-\beta}\int_0^T \boldsymbol{\theta}^T(s)\,\mathrm{d}\mathbf{W}(s)\right\}. \tag{7.28}$$

Furthermore, from (7.16) we have

$$Q(T) = exp\left\{\beta\int_0^T \left[r(s) + \frac{1}{2(1-\beta)}\|\boldsymbol{\theta}(t)\|^2\right] \mathrm{d}s\right\}, \tag{7.29}$$

and therefore

$$\begin{aligned}
Q(T)^{\frac{1}{\beta-1}} &= \left[exp\left\{\beta\int_0^T \left[r(s) + \frac{1}{2(1-\beta)}\|\boldsymbol{\theta}(t)\|^2\right] \mathrm{d}s\right\}\right]^{\frac{1}{\beta-1}} \\
&= exp\left\{\frac{-\beta}{1-\beta}\int_0^T \left[r(s) + \frac{1}{2(1-\beta)}\|\boldsymbol{\theta}(t)\|^2\right] \mathrm{d}s\right\}.
\end{aligned} \tag{7.30}$$

Replacing $Q(T)^{\frac{1}{\beta-1}}$ in (7.28) with the expression we obtained in (7.30) as well as using the explicit form of $X(T)$ given by (7.6), we obtain

$$
\begin{aligned}
& I(yH(T)\mathcal{E}([\boldsymbol{\mu}] \bullet M)(T)) \\
=& x_0 exp\left\{ \frac{-\beta}{1-\beta} \int_0^T \left[r(s) + \frac{1}{2(1-\beta)} \|\boldsymbol{\theta}(t)\|^2 \right] \mathrm{d}s \right\} \\
& \qquad \times exp\left\{ \frac{1}{1-\beta} \int_0^T \left[r(s) + \frac{1}{2} \|\boldsymbol{\theta}(s)\|^2 \right] \mathrm{d}s + \frac{1}{1-\beta} \int_0^T \boldsymbol{\theta}^T(s)\, \mathrm{d}\mathbf{W}(s) \right\} \\
=& x_0 exp\left\{ \int_0^T \left[r(s) + \frac{1-2\beta}{2(1-\beta)^2} \|\boldsymbol{\theta}(s)\|^2 \right] \mathrm{d}s + \frac{1}{1-\beta} \int_0^T \boldsymbol{\theta}^T(s)\, \mathrm{d}\mathbf{W}(s) \right\} \\
=& X(T),
\end{aligned}
\tag{7.31}
$$

which is exactly (6.26)(2).

To summarize the above, we have established that $(X, y, [\boldsymbol{\mu}]) \in \mathbb{I}_2 \times (0,\infty) \times \mathcal{G}_3$, with $X \in \mathbb{I}_2$ defined by (7.3) and (7.4), $[\boldsymbol{\mu}] \in \mathcal{G}_3$ constructed by Theorem 7.5 to satisfy (7.20) - (7.23) and $y \in (0,\infty)$ defined by (7.25), satisfies the transversality relation (6.26)(2). It remains to show that the triple $(X, y, [\boldsymbol{\mu}]) \in \mathbb{I}_2 \times (0,\infty) \times \mathcal{G}_3$ satisfies the remaining two relations (6.26)(1)(3). We will do so by first showing that (6.26)(3) is satisfied. Note that from (7.10) and (7.6) we have the following:

$$
\begin{aligned}
& X(t)H(t) \\
=& x_0 exp\left\{ \int_0^t [r(s) + \frac{1-2\beta}{2(1-\beta)^2} \|\boldsymbol{\theta}(s)\|^2]\, \mathrm{d}s + \frac{1}{1-\beta} \int_0^t \boldsymbol{\theta}^T(s)\, \mathrm{d}\mathbf{W}(s) \right\} \\
& \qquad \times exp\left\{ -\int_0^t r(s)\, \mathrm{d}s - \int_0^t \boldsymbol{\theta}^T(s)\, \mathrm{d}\mathbf{W}(s) - \frac{1}{2} \int_0^t \|\boldsymbol{\theta}(s)\|^2\, \mathrm{d}s \right\} \\
=& x_0 exp\left\{ \frac{\beta}{1-\beta} \int_0^t \boldsymbol{\theta}^T(s)\, \mathrm{d}\mathbf{W}(s) - \frac{\beta^2}{2(1-\beta)^2} \int_0^t \|\boldsymbol{\theta}(s)\|^2\, \mathrm{d}s \right\} \\
=& x_0 \mathcal{E}\left(\frac{\beta}{1-\beta} \boldsymbol{\theta} \bullet \mathbf{W} \right), \qquad t \in [0,T].
\end{aligned}
\tag{7.32}
$$

To simplify the notation put

$$
\boldsymbol{\lambda}(t) = \frac{\beta}{1-\beta} \boldsymbol{\theta}(t), \qquad t \in [0,T]. \tag{7.33}
$$

Since $\boldsymbol{\theta}$ is uniformly bounded and $\mathcal{F}_t$-progressively measurable (recall Remark 2.13) by Proposition B.46 we have the following:

$$
\mathcal{E}\left(\frac{p}{2} \boldsymbol{\lambda} \bullet \mathbf{W} \right) \quad \text{is a } \textit{square-integrable } \mathcal{F}_t\text{-martingale for each } p \in \mathbb{R}. \tag{7.34}
$$

For each $p \in \mathbb{R}$ we also have the elementary identity

$$
[\mathcal{E}(\boldsymbol{\lambda} \bullet \mathbf{W})(t)]^p = \exp\left\{ \frac{p(p-2)}{4} \int_0^t \|\boldsymbol{\lambda}(s)\|^2\, \mathrm{d}s \right\} [\mathcal{E}(\frac{p}{2} \boldsymbol{\lambda} \bullet \mathbf{W})(t)]^2, \tag{7.35}
$$

From (7.34) together with the Doob L^2-inequality (Theorem B.9), we have

$$
E\left[\max_{t \in [0,T]} [\mathcal{E}(\frac{p}{2} \boldsymbol{\lambda} \bullet \mathbf{W})(t)]^2 \right] < \infty, \qquad \text{for all } p \in \mathbb{R}. \tag{7.36}
$$

Since $\boldsymbol{\lambda}$ is uniformly bounded (from (7.33) and Remark 2.13) it follows from (7.35) and (7.36) that

$$E\left[\max_{t\in[0,T]}[\mathcal{E}(\boldsymbol{\lambda}\bullet\mathbf{W})(t)]^p\right]<\infty,\qquad\text{for all } p\in\mathbb{R}.\tag{7.37}$$

for all $p \in \mathbb{R}$. It now follows from (7.37), (7.33) and (7.32) that

$$E\left[\max_{t\in[0,T]}|X(t)H(t)|^p\right]<\infty,\qquad\text{for all } p\in\mathbb{R}.\tag{7.38}$$

But, from (7.21), we see that

$$|X(t)H(t)\mathcal{E}([\boldsymbol{\mu}]\bullet M)(t)|\le c_4|X(t)H(t)|,\qquad t\in[0,T],\tag{7.39}$$

and then it follows from (7.39) and (7.38) that

$$E\left[\max_{t\in[0,T]}|X(t)H(t)\mathcal{E}([\boldsymbol{\mu}]\bullet M)(t)|\right]<\infty.\tag{7.40}$$

Now since $[\boldsymbol{\mu}]\bullet M$ is a *purely discontinuous* $\mathcal{F}_t$-local martingale, and $\boldsymbol{\theta}\bullet\mathbf{W}$ is a *continuous* $\mathcal{F}_t$-local martingale (in fact a $\mathcal{F}_t$-martingale), from Theorem B.40 we obtain

$$[\boldsymbol{\theta}\bullet\mathbf{W},[\boldsymbol{\mu}]\bullet M](t)=0,\quad\text{for all } t\in[0,T].\tag{7.41}$$

Using (7.41) together with Corollary B.45 yields

$$X(t)H(t)\mathcal{E}\left([\boldsymbol{\mu}]\bullet M\right)(t)=x_0\mathcal{E}\left(\frac{\beta}{1-\beta}\boldsymbol{\theta}\bullet\mathbf{W}\right)\mathcal{E}\left([\boldsymbol{\mu}]\bullet M\right)(t)=x_0\mathcal{E}\left(\frac{\beta}{1-\beta}\boldsymbol{\theta}\bullet\mathbf{W}+[\boldsymbol{\mu}]\bullet M\right)(t).\tag{7.42}$$

Since $\frac{\beta}{1-\beta}\boldsymbol{\theta}\bullet\mathbf{W}+[\boldsymbol{\mu}]\bullet M$ is a $\mathcal{F}_t$-local martingale it follows from (7.42) that

$$XH\mathcal{E}\left([\boldsymbol{\mu}]\bullet M\right)\text{ is a }\mathcal{F}_t\text{-local martingale on }[0,T].\tag{7.43}$$

From (7.43) and (7.40) we obtain

$$XH\mathcal{E}\left([\boldsymbol{\mu}]\bullet M\right)\text{ is a }\mathcal{F}_t\text{-martingale on }[0,T].\tag{7.44}$$

This verifies (6.26)(3), so it remains to verify (6.26)(1). To this end, we fix some $\tau\in(0,\infty)$ and consider the function $f(\varsigma)$ defined as follows:

$$f(\varsigma):=\frac{1}{\beta}\varsigma^\beta-\varsigma\tau,\qquad\varsigma\in(0,\infty).\tag{7.45}$$

In order to maximize f over $\varsigma\in(0,\infty)$, we first differentiate f with respect to ς and we get

$$f^{(1)}(\varsigma)=\varsigma^{\beta-1}-\tau,\qquad\varsigma\in(0,\infty).\tag{7.46}$$

By setting (7.46) equals to 0 we get the critical point ς^* is equal to:

$$\varsigma^*=\tau^{\frac{1}{\beta-1}}\tag{7.47}$$

which is strictly positive as $\tau>0$. Moreover, it corresponds to the maximum because the second derivative of f with respect to ς,

$$f^{(2)}(\varsigma)=(\beta-1)\varsigma^{\beta-2},\tag{7.48}$$

is strictly negative for all $\varsigma > 0$ (recall $\beta \in (-\infty, 1)\backslash\{0\}$ from Condition 7.2).

Now using (7.47) and Condition 7.2 on $\tilde{U}$ defined at (5.5), we have

$$\begin{aligned}\tilde{U}(\tau) &= \sup_{\varsigma>0}\{U(\varsigma) - \varsigma\tau\} \\ &= \frac{1}{\beta}\left(\tau^{\frac{1}{\beta-1}}\right)^{\beta} - \left(\tau^{\frac{1}{\beta-1}}\right)\tau \\ &= \frac{1}{\beta}\tau^{\frac{\beta}{\beta-1}} - \tau^{\frac{\beta}{\beta-1}} \\ &= \frac{1-\beta}{\beta}(\tau)^{\frac{\beta}{\beta-1}}, \qquad \tau \in (0,\infty).\end{aligned} \tag{7.49}$$

That is, $\tilde{U}$ is equal to

$$\tilde{U}(\tau) = \frac{1-\beta}{\beta}(\tau)^{\frac{\beta}{\beta-1}}, \qquad \tau \in (0,\infty). \tag{7.50}$$

If we let $\tau = yH(T)\mathcal{E}([\boldsymbol{\mu}] \bullet M)(T)$ in (7.50) we get,

$$\tilde{U}(yH(T)\mathcal{E}([\boldsymbol{\mu}] \bullet M)(T)) = \frac{1-\beta}{\beta}\left[yH(T)\mathcal{E}([\boldsymbol{\mu}] \bullet M)(T)\right]^{\frac{\beta}{\beta-1}}. \tag{7.51}$$

We will now establish that

$$E\left\{\left[H(T)\mathcal{E}([\boldsymbol{\mu}] \bullet M)(T)\right]^{\frac{\beta}{\beta-1}}\right\} \in \mathbb{R}. \tag{7.52}$$

From (7.21) and (6.20) it follows that

$$\left[H(T)\mathcal{E}([\boldsymbol{\mu}] \bullet M)(T)\right]^{\frac{\beta}{\beta-1}} > 0 \qquad \text{a.s.} \tag{7.53}$$

so the expectation at (7.52) is well defined and must be shown to be finite. From (7.21) there are constants $0 < a_1 < a_2 < \infty$ such that

$$0 < a_1\left[H(T)\right]^{\frac{\beta}{\beta-1}} < \left[H(T)\mathcal{E}([\boldsymbol{\mu}] \bullet M)(T)\right]^{\frac{\beta}{\beta-1}} < a_2\left[H(T)\right]^{\frac{\beta}{\beta-1}}. \tag{7.54}$$

Since $\boldsymbol{\theta}$ is uniformly bounded it follows from (6.20) that

$$E\left[\max_{t\in[0,T]}|H(t)|^p\right] < \infty, \qquad \text{for all } p \in \mathbb{R}, \tag{7.55}$$

(exactly as at (7.38)), in particular

$$E\left[|H(T)|^{\frac{\beta}{\beta-1}}\right] < \infty. \tag{7.56}$$

Now (7.52) is immediate from (7.56) and (7.54), and then (7.52) and (7.51) give

$$E\left[\tilde{U}\left(yH(T)\mathcal{E}([\boldsymbol{\mu}] \bullet M)\,(T)\right)\right] \in \mathbb{R}, \tag{7.57}$$

which implies

$$\tilde{\Psi}(y, [\boldsymbol{\mu}]) = x_0Y_0 + E\left[\tilde{U}(yH(T)\mathcal{E}([\boldsymbol{\mu}] \bullet M)(T))\right] \in \mathbb{R} \tag{7.58}$$

and thus verifying (6.26)(1).

Remark 7.9. We have now constructed a triple $(X, y, [\boldsymbol{\mu}]) \in \mathbb{I}_2 \times (0,\infty) \times \mathcal{G}_3$ where X and y are given by (7.4) and (7.25) respectively and $[\boldsymbol{\mu}]$ is constructed using Theorem 7.5 at (7.20). This triple satisfies (6.26)(1) - (3), as we have just shown. Moreover, $\boldsymbol{\pi}$ given at (7.3) also satisfies the identity at (7.1) i.e.

$$\boldsymbol{\pi}(t) = \frac{[\boldsymbol{\sigma}(t)^{-1}]^T \boldsymbol{\theta}(t)}{1-\beta}, \qquad \text{for all } t \in [0,T]. \tag{7.59}$$

Therefore must be the optimal solution to the problem (2.27) subject to the conditions at Remark 7.6.

◁

Remark 7.10. We have addressed problem (2.27) with the conditions indicated at Remark 7.6 in place. Similar results are obtained by Sotomayor and Cadenillas [16], but subject to a significantly stronger condition on the market parameters than that stipulated at Condition 7.3, namely the market parameters must be *Markov modulated* in the sense discussed at Remark 2.9. This more restrictive dependency structure is completely essential for the approach taken in [16], which is the application of dynamic programming directly to the primal problem, an approach which does not carry over when one has the general non-anticipative dependence on the paths of $\boldsymbol{\alpha}$ that we have postulated at Condition 7.3. More importantly, the duality approach that we have used gives every indication of extending to problems with *convex portfolio constraints* (although such constraints have not been addressed in this thesis). In contrast, direct application of dynamic programming to problems of the form (2.27), but including portfolio constraints, has never been successful, mainly because the constraints ruin the tractability of the Bellman equation. This is particularly true for the rather complicated Bellman equation developed in [16] for regime switching.

◁

7.2 Logarithmic utility function

In this section we shall address problem (2.27) in the case of the *logarithmic utility function*, that is, we shall strengthen Condition 2.21 to the following:

Condition 7.11. In Condition 2.21 it is further stipulated that

$$U(x) := \log(x), \qquad x \in (0,\infty). \tag{7.60}$$

◁

In contrast to the case of the power utility function addressed in Section 7.1, in the present section we shall *not* need the special dependency structure of Condition 7.3, and will be able to get explicit optimal portfolios when only Condition 2.7 holds, that is the market parameters depend non-anticipatively on the paths of both $\boldsymbol{\alpha}$ and $\mathbf{W}$. This is possible because the logarithmic utility function is much more tractable than the power utility function. Indeed, in the classical case where regime switching is absent from the market model (so that the market parameters are progressively measurable with respect to just the filtration of $\mathbf{W}$) one can get explicit optimal portfolios for the logarithmic utility function (see Example 3.6.6 of Karatzas and Shreve [10]). In this section we shall establish that the same thing holds for the logarithmic utility function when regime switching is present in the market model, that is the market parameters are progressively measurable with respect to filtration of $\mathbf{W}$ and $\boldsymbol{\alpha}$, as stipulated at Condition 2.7.

Remark 7.12. To summarize, we are going to address problem (2.27) assuming Condition 2.1, Condition 2.3, Condition 2.7, Condition 2.10, Condition 2.15, Condition 2.17, Condition 7.11 (which strengthens Condition 2.21), and Condition 2.23.

◁

Remark 7.13. If the market parameters r, $[\boldsymbol{\sigma}]$ and $\boldsymbol{\theta}$ are progressively measurable with respect to the filtration of *only the Brownian Motion* $\mathbf{W}$ (i.e. we discard regime switching in the market model), then from Ex 3.6.6 of Karatzas and Shreve [10] the optimal portfolio $\boldsymbol{\pi}$ is given by:

$$\boldsymbol{\pi}(t) = [\boldsymbol{\sigma}(t)^{-1}]^T \boldsymbol{\theta}(t) \tag{7.61}$$

which is equal to (7.3) with $\beta = 0$. Motivated by the principle of *totally unhedgeable coefficients* discussed at Remark 7.7 it is plausible that the portfolio at (7.61) is still optimal even when regime switching is present in the market model, that is the market parameters are $\mathcal{F}_t$-adapted. We are going to use the optimality relations 6.26(1)- (3)of Proposition 6.13 to verify that this is the case, much as we did in Section 7.1 for the power utility function. Howevever, we shall see that this verification is much more easily accomplished for the logarithmic utility function, and in particular will not require the martingale representation Theorem 7.5.

◁

Remark 7.14. From Remark Remark 2.13 we observe that $\boldsymbol{\pi}$ at (7.61) is uniformly bounded and $\mathcal{F}_t$-progressively measurable, and $\boldsymbol{\pi} \in \Pi$ (recall (2.16)).

◁

Following exactly the same approach that was used in Section 7.1 we shall construct a triple $(X, y, [\boldsymbol{\mu}])$ that satisfies (6.26)(1) - (3) inthe case of the power utility function. To do so, we again define the wealth process X as

$$X := X^{\boldsymbol{\pi}} \tag{7.62}$$

where $\boldsymbol{\pi}$ is given by (7.61). This implies that $X \in \mathbb{I}_2$. Next, we will construct a pair $(y, [\boldsymbol{\mu}]) \in (0, \infty) \times \mathcal{G}_3$ such that (6.26)(2) is satisfied.

Using (2.20) (dropping the $\boldsymbol{\pi}$ in the postscript) and letting $\boldsymbol{\pi}$ be given by (7.61) we get

$$\mathrm{d}X = X(t)\left\{\left[r(t) + \|\boldsymbol{\theta}(t)\|^2\right] \, \mathrm{d}t + \boldsymbol{\theta}^T(t)\, \mathrm{d}\mathbf{W}(t)\right\}. \tag{7.63}$$

Exactly as in Remark 2.20 we then get the explicit form of $\{X(t)\,,\, t \in [0, T]\}$ as follows:

$$X(t) = x_0 exp\left\{\int_0^t \left[r(s) + \frac{1}{2}\|\boldsymbol{\theta}(s)\|^2\right] \, \mathrm{d}s + \int_0^t \boldsymbol{\theta}^T(s)\, \mathrm{d}\mathbf{W}(s)\right\}. \tag{7.64}$$

As in Section 7.1, we shall use (7.64) as the basis to construct a pair $(y, [\boldsymbol{\mu}]) \in (0, \infty) \times \mathcal{G}_3$ that satisfies (6.26)(2). From Condition 7.11 we have

$$U^{(1)}(x) = \frac{1}{x} \in (0, \infty), \qquad x \in (0, \infty). \tag{7.65}$$

From (7.65) we see that $U^{(1)} : (0, \infty) \to (0, \infty)$ has the inverse $I : (0, \infty) \to (0, \infty)$ given by

$$I(s) = \left(U^{(1)}\right)^{-1}(s) = \frac{1}{s}, \qquad s \in (0, \infty). \tag{7.66}$$

Again we will use (6.26)(2) to establish the relation that $y \in (0, \infty)$ and $[\boldsymbol{\mu}] \in \mathcal{G}_3$ must necessarily satisfy. To do so, we first fix some arbitrary $(y, [\boldsymbol{\mu}]) \in (0, \infty) \times \mathcal{G}_3$. Then if we let $s := yH(T)\mathcal{E}([\boldsymbol{\mu}] \bullet M)(T)$ in (7.66) where $H(T)$ is given by (6.20), we get

$$I(yH(T)\mathcal{E}([\boldsymbol{\mu}] \bullet M)(T)) = [yH(T)\mathcal{E}([\boldsymbol{\mu}] \bullet M)(T)]^{-1}. \tag{7.67}$$

Using (7.10) we can write out the explicit form of $H^{-1}(t)$, that is

$$\begin{aligned} H^{-1}(t) &= \left[exp\left\{ -\int_0^t \left[r(s) + \frac{1}{2}\|\boldsymbol{\theta}(s)\|^2 \right] \mathrm{d}s - \int_0^t \boldsymbol{\theta}^T(s)\,\mathrm{d}\mathbf{W}(s) \right\} \right]^{-1} \\ &= exp\left\{ \int_0^t \left[r(s) + \frac{1}{2}\|\boldsymbol{\theta}(s)\|^2 \right] \mathrm{d}s + \int_0^t \boldsymbol{\theta}^T(s)\,\mathrm{d}\mathbf{W}(s) \right\}, \qquad t \in [0,T]. \end{aligned} \tag{7.68}$$

Note that from (7.64), (7.68) and the fact that $x_0 > 0$ (recall Condition 2.15), we have the following:

$$H^{-1}(t) = \frac{X(t)}{x_0}, \qquad \text{for all } t \in [0,T]. \tag{7.69}$$

Using (7.69), we see that (7.67) can be rewritten as:

$$I(yH(T)\mathcal{E}([\boldsymbol{\mu}] \bullet M)(T)) = [yH(T)\mathcal{E}([\boldsymbol{\mu}] \bullet M)(T)]^{-1} = [y\mathcal{E}([\boldsymbol{\mu}] \bullet M)(T)]^{-1} \frac{X(T)}{x_0}. \tag{7.70}$$

In view of (7.70), in order for (6.26)(2) to hold, (7.70) must equal to $X(T)$ i.e.

$$X(T) = [y\mathcal{E}([\boldsymbol{\mu}] \bullet M)(T)]^{-1} \frac{X(T)}{x_0}. \tag{7.71}$$

Since X is strictly positive (from (7.64)), we therefore can divide both sides of (7.71) by $X(T)$, which gives

$$1 = [y\mathcal{E}([\boldsymbol{\mu}] \bullet M)(T)]^{-1} x_0^{-1} \tag{7.72}$$

Rewriting (7.72), we get

$$y\mathcal{E}([\boldsymbol{\mu}] \bullet M)(T) = x_0^{-1}, \tag{7.73}$$

which gives the relation that $y \in (0,\infty)$ and $[\boldsymbol{\mu}] \in \mathcal{G}_3$ must necessarily satisfy when (6.26)(2) holds true. It now remains to *construct* some $(y,[\boldsymbol{\mu}]) \in (0,\infty) \times \mathcal{G}_3$ which satisfies (7.73).

Remark 7.15. It is worthwhile to note that the relation analogous to (7.73) in the case of the power utility function is (7.15). Indeed, (7.15) just reduces to (7.73) when $\beta = 0$, and this (in a very formal sense only) corresponds to the case of the logarithmic utility function (see Remark 7.13). Obviously (7.73) is a much simpler relation than (7.15) (this is because the logarithmic utility function is so "nice"), and we are going to see that it is much easier to construct a pair $(y,[\boldsymbol{\mu}])$ which satisfies (7.73) than it was to construct a similar pair to satisfy (7.15). Indeed, in the latter case we had to resort to the martingale representation Theorem 7.5 to construct $[\boldsymbol{\mu}]$, and in order to use this theorem we needed Condition 7.3. None of this will be necessary in the case of the logarithmic utility function.

◁

Since $x_0 > 0$ (see Condition 2.15) and $[\mathbf{0}]$ (the $D \times D$ zero matrix) is a member of $\mathcal{G}_3$ (recall (6.22)), we see that the pair $(y,[\boldsymbol{\mu}]) \in (0,\infty) \times \mathcal{G}_3$ defined by

$$y = x_0^{-1}, \tag{7.74}$$
$$[\boldsymbol{\mu}] = [\mathbf{0}], \qquad \text{i.e. } \mu_{ij} = 0 \quad \text{for all } i,j \in S. \tag{7.75}$$

satisfies (7.73) (since $\mathcal{E}([\boldsymbol{\mu}] \bullet M)(T) = 1$ when $[\boldsymbol{\mu}] := [\mathbf{0}]$).

To recap what we have so far, we now have a candidate for $y \in (0,\infty)$ defined at (7.74) as well as a candidate $[\boldsymbol{\mu}] \in \mathcal{G}_3$ defined at (7.75). Moreover, we also have X defined at (7.61) and (7.62) is a member of $\mathbb{I}_2$. All this implies that the triple $(X, y, [\boldsymbol{\mu}]) \in \mathbb{I}_2 \times (0, \infty) \times \mathcal{G}_3$ and now we will verify that this triple $(X, y, [\boldsymbol{\mu}])$ satisfies (6.26)(1)(2)(3). First we verify it satisfy (6.26)(2). From (7.70) we have

$$I(yH(T)\mathcal{E}([\boldsymbol{\mu}] \bullet M)(T)) = [y\mathcal{E}([\boldsymbol{\mu}] \bullet M)(T)]^{-1} \frac{X(T)}{x_0}. \tag{7.76}$$

If we replace y and $[\boldsymbol{\mu}]$ on the right hand-side of(7.76) by the expressions given at (7.74) and (7.75), respectively and make use of the explicit expression of $X(T)$ given at (7.64), we have

$$\begin{aligned} I(yH(T)\mathcal{E}([\boldsymbol{\mu}] \bullet M)(T)) &= [y\mathcal{E}([\boldsymbol{\mu}] \bullet M)(T)]^{-1} \frac{X(T)}{x_0} \\ &= \left[x_0^{-1}\mathcal{E}([\mathbf{0}] \bullet M)(T)\right]^{-1} \frac{X(T)}{x_0} \\ &= x_0 \frac{X(T)}{x_0} \\ &= X(T). \end{aligned} \tag{7.77}$$

That is, (7.77) implies the following:

$$I(yH(T)\mathcal{E}([\boldsymbol{\mu}] \bullet M)(T)) = X(T), \tag{7.78}$$

which is exactly (6.26)(2).

From (7.78), we see that the the triple $(X, y, [\boldsymbol{\mu}]) \in \mathbb{I}_2 \times (0, \infty) \times \mathcal{G}_3$ where $X \in \mathbb{I}_2$ is given by (7.61) and (7.62), $y \in (0, \infty)$ is given by (7.74) and $[\boldsymbol{\mu}] \in \mathcal{G}_3$ given by (7.75), satisfies (6.26)(2) . Now we are going show that the triple $(X, y, [\boldsymbol{\mu}]) \in \mathbb{I}_2 \times (0, \infty) \times \mathcal{G}_3$ satisfies (6.26)(3). To do so, we define the process $\{Q(t)\ ,\ t \in [0, T]\}$ by

$$Q(t) := X(t)H(t)\mathcal{E}([\boldsymbol{\mu}] \bullet M)(t), \qquad t \in [0, T]. \tag{7.79}$$

Then using the fact that $[\boldsymbol{\mu}] = [\mathbf{0}]$, we can simplify (7.79) to

$$Q(t) = X(t)H(t), \qquad t \in [0, T]. \tag{7.80}$$

Furthermore, by (7.69), $Q(t)$ at (7.80) can be further reduced to:

$$\begin{aligned} Q(t) &= X(t)H(t) \\ &= X(t) \left[\frac{X(t)}{x_0}\right]^{-1} \\ &= x_0 \end{aligned} \tag{7.81}$$

That is $Q(t)$ at defined at (7.79) is equal to

$$Q(t) = x_0, \qquad \text{for all } t \in [0, T]. \tag{7.82}$$

It the follows immediately from (7.79) and (7.82) that

$$X(t)H(t)\mathcal{E}([\boldsymbol{\mu}] \bullet M)(t) = x_0, \qquad \text{for all } t \in [0, T], \tag{7.83}$$

so that $XH\mathcal{E}([\boldsymbol{\mu}] \bullet M)$ is (trivially!) a $\mathcal{F}_t$-martingale i.e.(6.26)(3) is satisfied.

As we have shown at (7.78) and (7.82), the triple $(X, y, [\boldsymbol{\mu}]) \in \mathbb{I}_2 \times (0, \infty) \times \mathcal{G}_3$ we have constructed satisfies (6.26)(2)(3). So now it remains to show that (6.26)(1) also holds true. We shall repeat the steps employed in the the power utility function example. We fix some $\tau \in (0,\infty)$ and consider the function g which is defined as:

$$g(\varsigma) = log(\varsigma) - \varsigma\tau, \qquad \varsigma \in (0,\infty). \tag{7.84}$$

To maximize g over $\varsigma \in (0,\infty)$, we differentiate g with respect to ς that is,

$$g^{(1)}(\varsigma) = \frac{1}{\varsigma} - \tau, \qquad \varsigma \in (0,\infty). \tag{7.85}$$

Solving for the critical point ς^*, we get

$$\varsigma^* = \frac{1}{\tau}, \tag{7.86}$$

which is strictly positive as $\tau \in (0,\infty)$.

Since $g^{(2)}(\varsigma) = -\frac{1}{\varsigma^2} < 0$, for all $\varsigma > 0$ thus ς^* at (7.86) is the maximum to the function g over $(0,\infty)$.

Using (7.86) and Condition 7.11 on $\tilde{U}$ defined at (5.5) we get that for all $\tau > 0$

$$\begin{aligned}\tilde{U}(\tau) &= \sup_{\varsigma>0}\{U(\varsigma) - \varsigma\tau\}\\ &= log\left(\frac{1}{\tau}\right) - \left(\frac{1}{\tau}\right)\tau\\ &= log(\frac{1}{\tau}) - 1.\end{aligned} \tag{7.87}$$

That is,

$$\tilde{U}(\tau) = log(\frac{1}{\tau}) - 1, \qquad \text{where } \tau \in (0,\infty). \tag{7.88}$$

If we let $\tau := yH(T)\mathcal{E}([\boldsymbol{\mu}] \bullet M)(T) = x_0^{-1}H(T)$ (recall (7.74) and (7.75)), then (7.88) becomes:

$$\begin{aligned}\tilde{U}(x_0^{-1}H(T)) &= log\left(\frac{1}{x_0^{-1}H(T)}\right) - 1\\ &= log(H(T)^{-1}) + log(x_0) - 1\end{aligned} \tag{7.89}$$

From (7.89), we see that it suffices to show that

$$E\left[log(H(T)^{-1})\right] \in \mathbb{R} \tag{7.90}$$

First note that from (6.20) we have

$$|H(T)^{-1}| > 0, \qquad \text{a.s.} \tag{7.91}$$

and therefore the expection at (7.90) is well-defined and we shall the establish the finiteness of (7.90) in order to verify (6.26)(1). To this end, we replace $H(T)^{-1}$ using the expression at (7.68) then $E\left[log(H(T)^{-1})\right]$ at (7.90) becomes:

$$E\left[log(H(T)^{-1})\right] = E\left[\int_0^T \left[r(s) + \frac{1}{2}\|\boldsymbol{\theta}(s)\|^2\right] \mathrm{d}s + \int_0^T \boldsymbol{\theta}^T(s)\,\mathrm{d}\mathbf{W}(s)\right] \tag{7.92}$$

Since r and $\boldsymbol{\theta}$ are uniformly bounded (recall Condition 2.7 and Remark 2.13), we have

$$E\left[log(H(T)^{-1})\right] = E\left[\int_0^T \left[r(s) + \frac{1}{2}\|\boldsymbol{\theta}(s)\|^2\right] \mathrm{d}s + \int_0^T \boldsymbol{\theta}^T(s)\, \mathrm{d}\mathbf{W}(s)\right] < \infty. \tag{7.93}$$

It then follows from (7.93) that

$$E\left[\tilde{U}(yH(T)\mathcal{E}([\boldsymbol{\mu}] \bullet M)(T))\right] = E\left[\tilde{U}(x_0^{-1}H(T))\right] \in \mathbb{R}, \tag{7.94}$$

which implies

$$\tilde{\Psi}(y, [\boldsymbol{\mu}]) = x_0 Y_0 + E\left[\tilde{U}(yH(T))\right] \in \mathbb{R}, \tag{7.95}$$

and hence establishing (6.26)(1).

Remark 7.16. We have established that the triple $(x, y, [\boldsymbol{\mu}])$ constructed at (7.62), (7.74) and (7.75) satisfies (6.26)(1)(2)(3). Furthermore, the $\boldsymbol{\pi}$ given by (7.61) also satisfies the identity at (7.61) i.e.

$$\boldsymbol{\pi}(t) = X^{-1}(t)[\boldsymbol{\sigma}(t)^{-1}]^T \boldsymbol{\Lambda}_X(t), \quad \text{for all } t \in [0, T]. \tag{7.96}$$

For that reason, it must be the optimal portfolio.

◁

Remark 7.17. We have addressed problem (2.27) with the conditions indicated at Remark 7.12 in place. Similar results are obtained by Sotomayor and Cadenillas [16] for the logarithmic utility function, but (exactly as for the case of the power utility function, see Remark 7.10) subject to the rather stringent condition that the market parameters be Markov modulated in the sense of Remark 2.9. Again, exactly as at Remark 7.10, this stringent condition it essential to facilitate the dynamic programming approach adopted in [16]. In contrast we need only postulate the much more general dependency structure of Condition 2.7, which allows a full non-anticipative dependence of the market parameters on both of the driving processes $\boldsymbol{\alpha}$ and $\mathbf{W}$ in the market model (see Remark 2.8) in order to get explicit optimal portfolios with the logarithmic utility function. This illustrates fairly clearly the inherent superiority of the conjugate duality approach. This approach gives every indication of extending to problems with *convex portfolio constraints*, with the market parameters again subject to Condition 2.7, which is not the case for the dynamic programming approach of [16].

◁

Appendices

A Proofs

Proofs of several technical results occurring in the main body of the thesis are relegated to this Appendix in order to avoid obscuring the main lines of development. Readers of the thesis will in fact lose very little if they choose not to study the following proofs in detail.

<u>**Proof of Lemma 3.2:**</u> From (3.6) and (3.7), we obtain the following equality

$$\int_0^t [V_1(s) - V_2(s)] \, \mathrm{d}s = \int_0^t [\boldsymbol{\xi}_2(s) - \boldsymbol{\xi}_2(s)] \, \mathrm{d}s. \tag{A.1}$$

Since V_1, $V_2 \in \mathcal{H}_1$ we have

$$\int_0^t [V_1(s) - V_2(s)] \, \mathrm{d}s \in \mathcal{FV}_0^c \cap \mathcal{M}_{0,\mathrm{loc}}^c, \tag{A.2}$$

which implies that

$$\int_0^t [V_1(s) - V_2(s)] \, \mathrm{d}s = 0, \qquad \text{for all } t \in [0, T]. \tag{A.3}$$

It then follows from (A.1) and (A.3) that

$$Q(t) := \int_0^t [\boldsymbol{\xi}_2(s) - \boldsymbol{\xi}_2(s)] \, \mathrm{d}s = 0 \quad \text{for all } t \in [0, T]. \tag{A.4}$$

Using (2.12), we get

$$\langle Q \rangle (t) := \int_0^t |\boldsymbol{\xi}_2(s) - \boldsymbol{\xi}_2(s)|^2 \, \mathrm{d}s = 0 \quad \text{for all } t \in [0, T] \tag{A.5}$$

and (3.8) follows from (A.3) and (A.5).

□

<u>**Proof of Lemma 3.4:**</u> Suppose first that (3.12) holds for some $X \in \mathbb{I}$ and $\boldsymbol{\pi} \in \Pi$. Then from (2.20) and (3.4) we have

$$\begin{aligned} X(t) &= X_0 + \int_0^t \dot{X}(s) \, \mathrm{d}s + \int_0^t \boldsymbol{\Lambda}_X^T(s) \, \mathrm{d}\mathbf{W}(s) \\ &= x_0 + \int_0^t X^{\boldsymbol{\pi}}(s)[r(s) + \boldsymbol{\pi}^T(s)[\boldsymbol{\sigma}(s)]\boldsymbol{\theta}(s)] \, \mathrm{d}s + \int_0^t X^{\boldsymbol{\pi}}(s)\boldsymbol{\pi}^T(s)[\boldsymbol{\sigma}(s)] \, \mathrm{d}\mathbf{W}(s) \quad t \in [0, T]. \end{aligned} \tag{A.6}$$

Then (3.12) and (A.6) implies that

$$X_0 = x_0 \quad \text{a.e.} \tag{A.7}$$

Moreover it follows from (3.12), (A.6) and (3.2) that

$$\dot{X}(t) = X^{\boldsymbol{\pi}}(t)[r(t) + \boldsymbol{\pi}^T(t)[\boldsymbol{\sigma}(t)]\boldsymbol{\theta}(t)] \quad \text{a.e.} \tag{A.8}$$

$$\boldsymbol{\Lambda}_X = X^{\boldsymbol{\pi}}(t)[\boldsymbol{\sigma}(t)]^T\boldsymbol{\pi}(t) \quad \text{a.e.} \tag{A.9}$$

Now from (3.11), (A.8) and (A.9) we have

$$\boldsymbol{\pi} \in \mathcal{U}(X), \tag{A.10}$$

and (3.13) follows from (A.7) and (A.10). This establishes the $\Rightarrow$ of the lemma and to show the converse, suppose that (3.13) holds. Then using the definition of $\mathcal{U}(X)$ (recall Remark 3.3), we have the following

$$\begin{aligned} \dot{X}(t) &= X^{\boldsymbol{\pi}}(t)[r(t) + \boldsymbol{\pi}^T(t)[\boldsymbol{\sigma}(t)]\boldsymbol{\theta}(t)] \quad \text{a.e.} && \text{(A.11)} \\ \boldsymbol{\Lambda}_X &= X^{\boldsymbol{\pi}}(t)[\boldsymbol{\sigma}(t)]^T\boldsymbol{\pi}(t) \quad \text{a.e.} && \text{(A.12)} \end{aligned}$$

By (3.13), (A.11) and (A.12) we get

$$X(t) = x_0 + \int_0^t X^{\boldsymbol{\pi}}(s)[r(s) + \boldsymbol{\pi}^T(s)[\boldsymbol{\sigma}(s)]\boldsymbol{\theta}(s)]\,\mathrm{d}s + \int_0^t X^{\boldsymbol{\pi}}(s)\boldsymbol{\pi}^T(s)[\boldsymbol{\sigma}(s)]\,\mathrm{d}\mathbf{W}(s). \tag{A.13}$$

Since Lemma 3.2 guarantees the uniqueness of the integrands in (A.13), (3.13) follows from (2.9) and (A.13).

□

Proof of Proposition 3.6: For the $\Rightarrow$ of the proposition, let us suppose that (3.15) holds for some $X \in \mathbb{I}$ where $X_0 > 0$. Then this implies that there exists a $\boldsymbol{\pi} \in \Pi$ such that the following are true

$$\begin{aligned} \dot{X}(t) &= X(t)[r(t) + \boldsymbol{\pi}^T(t)[\boldsymbol{\sigma}(t)]\boldsymbol{\theta}(t)] \quad \text{a.e.} && \text{(A.14)} \\ \boldsymbol{\Lambda}_X &= X(t)[\boldsymbol{\sigma}(t)]^T\boldsymbol{\pi}(t) \quad \text{a.e.} && \text{(A.15)} \end{aligned}$$

Inserting (A.15) into (A.14) yields

$$\dot{X}(t) = X(t)r(t) + \boldsymbol{\Lambda}_X^T(t)\boldsymbol{\theta}(t), \tag{A.16}$$

which is exactly (3.18).

To establish (3.17), first note that we have $X_0 > 0$ and $\boldsymbol{\pi} \in \Pi$. Moreover, using (A.14) and (A.15) we get

$$\mathrm{d}X(t) = X(t)\left\{[r(t) + \boldsymbol{\pi}^T(t)[\boldsymbol{\sigma}(t)]\boldsymbol{\theta}(t)]\,\mathrm{d}t + \boldsymbol{\pi}^T(t)[\boldsymbol{\sigma}(t)]^T\,\mathrm{d}\mathbf{W}(t)\right\}. \tag{A.17}$$

By Remark 2.20, we see that X given by (A.17) is $\mathbb{P}$-strictly positive , which implies that (3.16) is true.

From (A.16) and (A.17), we have established the $\Rightarrow$ of the proposition. Now to show the converse, we assume that (3.17) and (3.18) are true and define the following

$$\begin{aligned} \boldsymbol{\pi}(t) &:= [\boldsymbol{\sigma}(t)^{-1}]^T\frac{\boldsymbol{\Lambda}_X(t)}{X(t)}, && \text{(A.18)} \\ \Phi(\omega) &:= \inf_{t\in[0,T]} X(t,\omega) > 0. && \text{(A.19)} \end{aligned}$$

It then follow from (A.18) and (A.19) that

$$\boldsymbol{\pi}(t) \leq [\boldsymbol{\sigma}(t)^{-1}]^T\frac{\boldsymbol{\Lambda}_X(t)}{\Phi(\omega)}. \tag{A.20}$$

Using (A.20) and the first inequality of (2.15) (with $z := \boldsymbol{\Lambda}_X$) we get

$$\frac{1}{k_1}\left\|[\boldsymbol{\sigma}(t)^{-1}]^T\boldsymbol{\Lambda}_X(t)\right\| \leq \|\boldsymbol{\Lambda}_X(t)\| \tag{A.21}$$

or equivalently,

$$\left\| [\boldsymbol{\sigma}(t)^{-1}]^T \boldsymbol{\Lambda}_X(t) \right\| \le k_1 \left\| \boldsymbol{\Lambda}_X(t) \right\|. \tag{A.22}$$

By (A.20) and (A.22), we have

$$\|\boldsymbol{\pi}(t)\| \le \frac{k_1}{\Phi(\omega)} \left\| \boldsymbol{\Lambda}_X(t) \right\|. \tag{A.23}$$

Now since $\boldsymbol{\Lambda}_X \in \mathcal{H}_2$ (recall (3.2)), (A.23) gives

$$\int_0^T \|\boldsymbol{\pi}(s)\|^2 \, \mathrm{d}s < \infty \quad \text{a.s.} \tag{A.24}$$

which implies that $\boldsymbol{\pi} \in \Pi$ (recall (2.16)). This together with (3.18) establishes (3.16).

□

Proof of Proposition 4.13: Suppose we fix some arbitrary $X \in \mathbb{I}$ and $Y \in \mathbb{J}$ (recall (3.3) and (4.21)). Then X and Y are $\mathcal{F}_t$-semimartingale. Moreover, by Itô's formula we have that

$$X(t)Y(t) = X_0 Y_0 + \int_0^t X(s)\,\mathrm{d}Y(s) + \int_0^t Y(s)\,\mathrm{d}X(s) + [X,Y](t), \quad t \in [0,T]. \tag{A.25}$$

From (3.4) and (4.22) we have the following

$$\begin{aligned} \mathrm{d}X(s) &= \dot{X}(s)\,\mathrm{d}s + \boldsymbol{\Lambda}_X^T(s)\,\mathrm{d}\mathbf{W}(s), && \text{(A.26)} \\ \mathrm{d}Y(s) &= \dot{Y}(s)\,\mathrm{d}s + \boldsymbol{\Lambda}_Y^T(s)\,\mathrm{d}\mathbf{W}(s) + \sum_{i,j=1}^{D} \Gamma_{ij}^Y(s)\,\mathrm{d}M_{ij}(s). && \text{(A.27)} \end{aligned}$$

Using (A.26) and (A.27), we can rewrite (A.25) as

$$\begin{aligned} & X(t)Y(t) \\ =& X_0 Y_0 + \int_0^t X(s) \left[\dot{Y}(s)\,\mathrm{d}s + \boldsymbol{\Lambda}_Y^T(s)\,\mathrm{d}\mathbf{W}(s) + \sum_{i,j=1}^{D} \Gamma_{ij}^Y(s)\,\mathrm{d}M_{ij}(s) \right] \\ & + \int_0^t Y(s) \left[\dot{X}(s)\,\mathrm{d}s + \boldsymbol{\Lambda}_X^T(s)\,\mathrm{d}\mathbf{W}(s)) + \int_0^t \boldsymbol{\Lambda}_X^T(s)\boldsymbol{\Lambda}_Y(s)\,\mathrm{d}s \right] \\ =& X_0 Y_0 + \int_0^t \left[X(s)\dot{Y}(s) + Y(s)\dot{X}(s) + \boldsymbol{\Lambda}_X^T(s)\boldsymbol{\Lambda}_Y(s) \right] \mathrm{d}s \\ & + \int_0^t X(s)\boldsymbol{\Lambda}_Y^T(s)\,\mathrm{d}\mathbf{W}(s) + \int_0^t Y(s)\boldsymbol{\Lambda}_X^T(s)\,\mathrm{d}\mathbf{W}(s) + \sum_{i,j=1}^{D} \int_0^t X(s)\Gamma_{ij}^Y(s)\,\mathrm{d}M_{ij}(s). \end{aligned} \tag{A.28}$$

Then (A.28) implies that

$$\mathbb{M}(X,Y)(t) = \int_0^t X(s)\boldsymbol{\Lambda}_Y^T(s)\,\mathrm{d}\mathbf{W}(s) + \int_0^t Y(s)\boldsymbol{\Lambda}_X^T(s)\,\mathrm{d}\mathbf{W}(s) + \sum_{i,j=1}^{D} \int_0^t X(s)\Gamma_{ij}^Y(s)\,\mathrm{d}M_{ij}(s). \tag{A.29}$$

Since every term on the right hand side of (A.29) is a $\mathcal{F}_t$-local martingale, we also have $\mathbb{M}(X,Y) \in \mathcal{M}_{0,\mathrm{loc}}(\{\mathcal{F}_t\})$.

□

Proof of Proposition 5.2 **(a)** First note that for each $(X,Y) \in \mathbb{I}_2 \times \mathbb{J}_2$(recalled (3.29) and (5.21)), it follows from (3.24) and (5.8) that:

$$L(t, X(t), \dot{X}(t), \mathbf{\Lambda}_X(t)) + M(t, Y(t), \dot{Y}(t), \mathbf{\Lambda}_Y(t)) = 0. \tag{A.30}$$

Then (A.30) together with (5.11) with $(x, v, \boldsymbol{\xi}) := (X_0, \dot{X}, \Gamma^X) \in \mathbb{I}_2$ and $(y, s, \boldsymbol{\gamma}) := (Y_0, \dot{Y}, \mathbf{\Lambda}_Y, [\mathbf{\Gamma}^Y]) \in \mathbb{J}_2$ imply the following inequality,

$$0 \geq \dot{X}(t)Y(t) + X(t)\dot{Y}(t) + \mathbf{\Lambda}_X^T(t)\mathbf{\Lambda}_Y(t). \tag{A.31}$$

Now from (4.25) and (A.31), one can see that for each $(X,Y) \in \mathbb{I}_2 \times \mathbb{J}_2$

$$\mathbb{M}(X,Y)(t) \geq X(t)Y(t) - X_0Y_0, \quad \text{for all } t \in [0,T]. \tag{A.32}$$

Since

$$X(t) > 0 \quad \text{a.e.} \quad \& \quad Y(t) \geq 0 \quad \text{a.e.} \tag{A.33}$$

therefore (A.32) implies

$$\mathbb{M}(X,Y)(t) \geq -X_0Y_0, \quad \text{for all } t \in [0,T]. \tag{A.34}$$

From (A.34) we see that $\mathbb{M}(X,Y) \in \mathcal{M}_{0,\text{loc}}(\{\mathcal{F}_t\})$ is uniformly bounded from below, hence $\mathbb{M}(X,Y)$ is a $\mathcal{F}_t$-supermartingale for all $X \in \mathbb{I}_2$ and $Y \in \mathbb{J}_2$.

□

(b) Suppose we fix some aribitrary $X \in \mathbb{I}_2$ and $Y \in \mathbb{J}_2$. Then from **(a)** we have that $\mathbb{M}(X,Y)$ is a $\mathcal{F}_t$-supermartingale that's null at the origin. This together with the fact that it is uniformly bounded from below by $-X_0Y_0$ yields the following inequalities:

$$\begin{aligned} -X_0Y_0 \leq E\mathbb{M}(X,Y)(T) \leq 0 \\ 0 \leq E\mathbb{M}(X,Y)(T) + X_0Y_0 \leq X_0Y_0. \end{aligned} \tag{A.35}$$

Using (4.25), (A.31), (A.35) and the fact that $X(t)Y(t) \geq 0$ for all $t \in [0,T]$, we then have

$$\begin{aligned} 0 \leq EX(T)Y(T) &\leq EX(T)Y(T) - E\int_0^T [\dot{X}(s)Y(s) + X(s)\dot{Y}(s) + \mathbf{\Lambda}_X^T(s)\mathbf{\Lambda}_Y(s)]\,\mathrm{d}s \\ &= E\mathbb{M}(X,Y)(T) + X_0Y_0 \leq X_0Y_0. \end{aligned} \tag{A.36}$$

Now from (5.10), for all $\omega \in \Omega$, $x > 0$ and $y \geq 0$ we have

$$-\tilde{U}(\omega, y) \leq -U(\omega, x) + xy. \tag{A.37}$$

Moreover, (A.37) implies

$$-\tilde{U}(\omega, y) \leq max\{0, -U(\omega, x)\} + xy. \tag{A.38}$$

Since both terms on the right hand side of (A.38) are non-negative, we also have the following inequality

$$max\{0, -\tilde{U}(\omega, y)\} \leq max\{0, -U(\omega, x)\} + xy. \tag{A.39}$$

Using (2.29), we can rewrite (A.39) as

$$\tilde{U}^-(\omega, y) \leq U^-(\omega, x) + xy. \tag{A.40}$$

With (A.40), we have that for all $(X,Y) \in \mathbb{I}_2 \times \mathbb{J}_2$,

$$E\tilde{U}^-(Y(T)) \leq EU^-(X(T)) + X(T)Y(T). \tag{A.41}$$

Using (A.36), (A.41) and Condition 2.21(2) we get

$$E\tilde{U}^-(Y(T)) \leq EU^-(X(T)) + X_0Y_0 < \infty, \tag{A.42}$$

which establishes **(b)**.

$\square$

Proof of Lemma 6.2 Fixing some $(\omega, t) \in \Omega \times [0, T]$, and $(x, v, \boldsymbol{\xi})$, $(y, s, \boldsymbol{\gamma}) \in \mathbb{R} \times \mathbb{R} \times \mathbb{R}^N$ and suppose that (6.3) holds. Then (3.23) and (5.8) give the following equalities

$$\begin{aligned} &(1)\ x > 0, \\ &(2)\ v = xr(t) + \boldsymbol{\xi}^T \boldsymbol{\theta}(t), \\ &(3)\ s + yr(t) \leq 0, \\ &(4)\ y\boldsymbol{\theta}(t) + \boldsymbol{\gamma} = 0, \\ &(5)\ xs + yv + \boldsymbol{\xi}^T \boldsymbol{\gamma} = 0. \end{aligned} \tag{A.43}$$

First we note that (6.4)(1)(2)(4) immediately follows from (A.43)(1)(2)(4). Moreover, to show (6.4)(3) holds, consider the following:

$$\begin{aligned} x(s + yr(t)) &= xs + y[xr(t)] \\ &= xs + y[v - \boldsymbol{\xi}^T \boldsymbol{\theta}(t)] \quad \text{by (A.43)(2)} \\ &= xs + yv - \boldsymbol{\xi}^T [y\boldsymbol{\theta}(t)] \\ &= xs + yv + \boldsymbol{\xi}^T \boldsymbol{\gamma} \quad \text{by (A.43)(4)} \\ &= 0 \quad \text{by (A.43)(5).} \end{aligned} \tag{A.44}$$

From (A.44) we see that (6.4)(3) holds. By (A.43)(1)(2)(4) and (A.44) we see that the $\Rightarrow$ of the lemma have been established. Now to establish the converse, suppose (6.4)(1) - (4) hold for some $(\omega, t) \in \Omega \times [0, T]$, and $(x, v, \boldsymbol{\xi})$, $(y, s, \boldsymbol{\gamma}) \in \mathbb{R} \times \mathbb{R} \times \mathbb{R}^N$. Then using (A.44) again we have

$$0 = xs + yv + \boldsymbol{\xi}^T \boldsymbol{\gamma}. \tag{A.45}$$

Furthermore (6.4)(1) - (4) gives

$$L(\omega, t, x, v, \boldsymbol{\xi}) + M(\omega, t, y, s, \boldsymbol{\gamma}) = 0, \tag{A.46}$$

and (6.3) follows immediately from (A.45) and (A.46).

$\square$

Proof of Lemma 6.5 First we fix some x,$y \in \mathbb{R}$ and suppose that (6.8) holds. Since $xy \in \mathbb{R}$, we have

$$\tilde{U}(y) \in \mathbb{R} \qquad \text{recall (5.6).} \tag{A.47}$$

From (A.47), we see that the following is also true

$$y \in (0, \infty) \qquad \text{by (5.7),} \tag{A.48}$$

which is exactly the first condition in (6.9). Now to show the second condition in (6.9), namely $x = I(y)$ we differentiate both sides of (6.8) with respect to y and we get

$$\tilde{U}^{(1)}(y) = -x. \tag{A.49}$$

Using Remark 6.4, we have

$$x = I(y), \tag{A.50}$$

which is exactly the second conditio (6.9). Now for the $\Leftarrow$ of the lemma, suppose both conditions in (6.9) holds for all x,$y \in \mathbb{R}$. By Remark 6.4, we have

$$\tilde{U}^{(1)}(y) = -x \quad \& \quad x > 0. \tag{A.51}$$

Integrating the first equality in (A.51) with respect to y will give us

$$k + \tilde{U}(y) = -xy, \tag{A.52}$$

for some constant $k \in \mathbb{R}$. We could then let $k = -U(x)$, which is permitted since $x > 0$ implies $U(x) \in \mathbb{R}$. Moreover, $y > 0 \Rightarrow \tilde{U}(y) \in \mathbb{R}$, so we have

$$\tilde{U}(y) + U(x) = -xy, \tag{A.53}$$

which gives (6.8).

□

Proof of Lemma 6.7 Fixing some $(X, Y) \in \mathbb{I}_2 \times \mathbb{J}_2$ such that (5.29)(2)(3) hold. Then (5.8), (3.25) and Lemma 6.2 imply the following

$$\dot{X}(t)Y(t) + X(t)\dot{Y}(t) + \mathbf{\Lambda}_X^T(t)\mathbf{\Lambda}_Y(t) = 0. \tag{A.54}$$

Using (A.54) together with (4.25), we see that (6.12) follows immediately.

Now to establish (6.13), let $\{Z(t), t \in [0, T]\}$ be defined as follows:

$$Z(t) := X(t)Y(t), \quad t \in [0, T]. \tag{A.55}$$

Then

$$Z(t) \geq 0 \quad \text{a.s.} \quad \text{for all } t \in [0, T]. \tag{A.56}$$

Now (A.56) is true because:

$$X(t) > 0 \quad \text{a.s.} \quad \& \quad Y(t) \geq 0 \quad \text{a.s.,} \quad \text{for all } t \in [0, T], \tag{A.57}$$

which follows from the fact that $(X, Y) \in \mathbb{I}_2 \times \mathbb{J}_2$ (recall (3.29) and (5.21)). Moreover, by (6.12) and the fact that $\mathbb{M}(X, Y)$ is a $\mathcal{F}_t$-supermartingale (from Proposition 5.2 **(a)**) we have that Z is also a $\mathcal{F}_t$-supermartingale as they are only differ by a constant.

Since we also have

$$Y(T) > 0 \quad \text{a.s.} \tag{A.58}$$

then (A.58) together with

$$\inf_{t \in [0,T]} X(t) > 0 \tag{A.59}$$

yields the following:

$$Z(T) > 0 \quad \text{a.s.} \tag{A.60}$$

Lastly, combining (A.56), (A.60) and the fact that Z is a $\mathcal{F}_t$-supermartingale, we get that Z must also be $\mathbb{P}$-strictly positive i.e.

$$\inf_{t \in [0,T]} Z(t) > 0, \tag{A.61}$$

From (A.55) and (A.61), we then have

$$Y(t) > 0 \quad \text{a.e.} \tag{A.62}$$

□

Proof of Proposition 6.9 Since both Proposition 5.4 and Proposition 6.9 involve *if and only if* statements, it suffices to show that for some $(X,Y) \in \mathbb{I}_2 \times \mathbb{J}_3$, if (6.15) is true then (6.16) and (5.29) are equivalent. First we fix some $(X,Y) \in \mathbb{I}_2 \times \mathbb{J}_3$ such that (6.15) holds. Then (5.29)(1) and (6.16)(1) are trivially equivalent.

Next, since $X \in \mathbb{I}_2 \subseteq \mathbb{I}_1$ (recall (3.19) and (3.29)) , we have:

$$\inf_{t\in[0,T]} X(t) > 0 \quad \text{a.s.} \tag{A.63}$$

$$\dot{X}(t) = X(t)r(t) + \boldsymbol{\Lambda}_X^T(t)\boldsymbol{\theta}(t) \quad \text{a.e.} \tag{A.64}$$

Then (A.63) and (A.64) together with (6.16)(3)(4) are equivalent to (5.29)(3) by Lemma 6.2 and Remark 6.3.

To show that (6.16)(2) is equivalent to (5.29)(2) we first note that $Y \in \mathbb{J}_3$ (recall 5.14.2) implies

$$Y(T) > 0 \ \text{ a.s..} \tag{A.65}$$

Using (A.65), Lemma 6.5 and Remark 6.6., we then have the equivalence between (6.16)(2) and (5.29)(2). We have therefore established the following: For every $(X,Y) \in \mathbb{I}_2 \times \mathbb{J}_3$ such that (6.15) holds the following equivalences are true:

- (5.29)(1) holds *iff* (6.16)(1) holds
- (5.29)(2) holds *iff* (6.16)(2) holds
- (5.29)(3) holds *iff* (6.16)(3)(4) hold

It remains to prove the equivalence between (6.16)(5) and (5.29)(4)when $(X,Y) \in \mathbb{I}_2 \times \mathbb{J}_3$ satisfies (6.15). But then, from Proposition 5.4, we see in particular that (X,Y) satisfies (5.29)(2)(3), so we can use Lemma 6.7 to assert that

$$\mathbb{M}(X,Y)(t) = X(t)Y(t) - X_0Y_0. \tag{A.66}$$

In view of (A.66) it is immediate that

- (5.29)(4) holds *iff* (6.16)(5) holds.

□

Proof of Proposition 6.11 (a) We first show the $\Leftarrow$ of the claim by supposing that $Y := yH\mathcal{E}\left([\boldsymbol{\mu}] \bullet M\right)$ for some fixed $(y,[\boldsymbol{\mu}]) \in (0,\infty) \times \mathcal{G}_3$. Then we have

$$\inf_{t\in[0,T]} Y(t) > 0 \text{ because } y, H \text{ and } \mathcal{E}\left([\boldsymbol{\mu}] \bullet M\right) > 0 \quad \text{a.s.} \tag{A.67}$$

Next, using Itô's formula on Y and we obtain

$$\begin{aligned} Y(t) = \quad & yH(t)\mathcal{E}\left([\boldsymbol{\mu}] \bullet M\right) \quad (t) && \text{(A.68)} \\ = \quad & yH(0)\mathcal{E}\left([\boldsymbol{\mu}] \bullet M\right) \quad (0) + \int_0^t H(s)\,\mathrm{d}\mathcal{E}\left([\boldsymbol{\mu}] \bullet M\right)(s) && \text{(A.69)} \\ & + y\int_0^t \mathcal{E}\left([\boldsymbol{\mu}] \bullet M\right)(s_-)\,\mathrm{d}H(s) + y[H, \mathcal{E}\left([\boldsymbol{\mu}] \bullet M\right)](t). && \text{(A.70)} \end{aligned}$$

We shall examine each terms in (A.69) and (A.70) individually. First of all, because of (6.20) and (6.21), the first term in (A.69) becomes

$$Y(0) = yH(0)\mathcal{E}\left([\boldsymbol{\mu}] \bullet M\right)(0) = y. \tag{A.71}$$

Secondly, since $[\boldsymbol{\mu}] \bullet M \in \mathcal{M}_{0,\text{loc}}(\{\mathcal{F}_t\})$, the stochastic exponential of $[\boldsymbol{\mu}] \bullet M$ namely, $\mathcal{E}\left([\boldsymbol{\mu}] \bullet M\right)$ must satisfy the identity

$$\mathcal{E}\left([\boldsymbol{\mu}] \bullet M\right)(t) = 1 + \int_0^t \mathcal{E}\left([\boldsymbol{\mu}] \bullet M\right)(s_-)\,\mathrm{d}[\boldsymbol{\mu}] \bullet M(s) \tag{A.72}$$

or equivalently

$$\begin{aligned} \mathrm{d}\mathcal{E}\left([\boldsymbol{\mu}] \bullet M\right)(s) &= \mathcal{E}\left([\boldsymbol{\mu}] \bullet M\right)(s_-)\,\mathrm{d}[\boldsymbol{\mu}] \bullet M(s) \\ &= \sum_{i,j=1}^{D} \mathcal{E}\left([\boldsymbol{\mu}] \bullet M\right)(s_-)\mu_{ij}(s)\,\mathrm{d}M_{ij}(s). \end{aligned} \tag{A.73}$$

Next, by observing the first term in (A.70) if we define the processes $\{A(t), t \in [0,T]\}$ and $\{B(t), t \in [0,T]\}$ as

$$A(t) := exp\left\{-\int_0^t r(s)\,\mathrm{d}s\right\}, \tag{A.74}$$

$$B(t) := \mathcal{E}(-\boldsymbol{\theta} \bullet \mathbf{W})(t). \tag{A.75}$$

Then

$$H(t) = A(t)B(t), \quad \text{for all } t \in [0,T]. \tag{A.76}$$

Moreover, since $A \in \mathcal{FV}_0^c(\{\mathcal{F}_t\})$, we have

$$[A, B](t) = 0, \quad \text{for all } t \in [0,T]. \tag{A.77}$$

Using (A.76), (A.77) and Itô's formula, we can reweire $H(t)$ as

$$H(t) = 1 + \int_0^t A(s)\,\mathrm{d}B(s) + \int_0^t B(s)\,\mathrm{d}A(s) \tag{A.78}$$

where

$$\mathrm{d}B(s) = \mathrm{d}\mathcal{E}(-\boldsymbol{\theta} \bullet \mathbf{W})(s) = -\mathcal{E}(-\boldsymbol{\theta} \bullet \mathbf{W})(s)\,\mathrm{d}\boldsymbol{\theta} \bullet \mathbf{W}(s) = B(s)\boldsymbol{\theta}^T(s)\,\mathrm{d}\mathbf{W}(s), \tag{A.79}$$

$$\mathrm{d}A(s) = -r(s)exp\left\{-\int_0^t r(s)\,\mathrm{d}s\right\}\mathrm{d}s = -r(s)A(s)\,\mathrm{d}s. \tag{A.80}$$

If substituting (A.79) and (A.80) into (A.78), we obtain

$$\begin{aligned} H(t) &= 1 - \int_0^t A(s)B(s)\boldsymbol{\theta}^T(s)\,\mathrm{d}\mathbf{W}(s) - \int_0^t B(s)r(s)A(s)\,\mathrm{d}s \\ &= 1 - \int_0^t H(s)\boldsymbol{\theta}^T(s)\,\mathrm{d}\mathbf{W}(s) - \int_0^t r(s)H(s)\,\mathrm{d}s. \end{aligned} \tag{A.81}$$

From (A.81) we have

$$\mathrm{d}H(s) = -H(s)\boldsymbol{\theta}^T(s)\,\mathrm{d}\mathbf{W}(s) - r(s)H(s)\,\mathrm{d}s. \tag{A.82}$$

Lastly, since H is a continuous $\mathcal{F}_t$-semimartingale (from Theorem B.43) and $\mathcal{E}\left([\boldsymbol{\mu}] \bullet M\right)$ is a purely discontinuous $\mathcal{F}_t$-semimartingale (from $[\boldsymbol{\mu}] \bullet M$ being a purely discontinuous semimartingale and Theorem B.43), it follows from Theorem B.40 that the second term in (A.70) can be simplified to

$$y[H, \mathcal{E}\left([\boldsymbol{\mu}] \bullet M\right)](t) = \sum_{0 \le s \le t} \Delta H(s) \Delta \mathcal{E}\left([\boldsymbol{\mu}] \bullet M\right) = 0, \qquad \text{for all } t \in [0,T]. \tag{A.83}$$

Finally, using (A.71), (A.73), (A.82) and (A.83) on (A.68) we get

$$\begin{aligned} Y(t) = y + y \sum_{i,j=1}^{D} \int_0^t H(s) \mathcal{E}\left([\boldsymbol{\mu}] \bullet M\right)(s_-) \mu_{ij}(s) \,\mathrm{d}M_{ij}(s) + \\ y \int_0^t H(s) \mathcal{E}\left([\boldsymbol{\mu}] \bullet M\right)(s_-) [\boldsymbol{\theta}^T(s) \,\mathrm{d}\mathbf{W}(s) - r(s) \,\mathrm{d}s] \\ = y + \sum_{i,j=1}^{D} \int_0^t Y(s_-) \mu_{ij}(s) \,\mathrm{d}M_{ij}(s) + \int_0^t \boldsymbol{\theta}^T(s) \,\mathrm{d}\mathbf{W}(s) - \int_0^t Y(s_-) r(s) \,\mathrm{d}s. \end{aligned} \tag{A.84}$$

Let us define the following:

$$Y_0 := y, \tag{A.85}$$

$$\dot{Y}(t) := -r(t) Y(t), \tag{A.86}$$

$$\boldsymbol{\Lambda}_Y(t) := -Y(t) \boldsymbol{\theta}(t), \tag{A.87}$$

$$[\boldsymbol{\Gamma}^Y](t) := Y(t) [\boldsymbol{\mu}](t), \quad (\textbf{i.e. } \Gamma_{ij}^Y(t) := Y(t) \mu_{ij}(t)). \tag{A.88}$$

Using (A.85) - (A.88), we can rewrite (A.84) as

$$Y(t) = Y_0 + \int_0^t \dot{Y}(s) \,\mathrm{d}s + \int_0^t \boldsymbol{\Lambda}_Y^T(s) \,\mathrm{d}\mathbf{W}(s) + \sum_{i,j=1}^{D} \int_0^t \Gamma_{ij}^Y(s) \,\mathrm{d}M_{ij}(s) \tag{A.89}$$

i.e. $Y \equiv (Y_0, \dot{Y}, \boldsymbol{\Lambda}_Y, [\boldsymbol{\Gamma}^Y])$. We also have

$$Y \equiv (Y_0, \dot{Y}, \boldsymbol{\Lambda}_Y, [\boldsymbol{\Gamma}^Y]) \in \mathbb{J}. \tag{A.90}$$

Additionally, we have $y \in (0, \infty)$ (by initial assumption), H (from (6.20)) and $\mathcal{E}\left([\boldsymbol{\mu}] \bullet M\right)$ (by $[\boldsymbol{\mu}] \in \mathcal{G}_3$) are strictly positive we thus have

$$\inf_{t \in [0,T]} Y(t) > 0 \quad \text{a.s..} \tag{A.91}$$

By (A.90) and (A.91) we have $Y \in \mathbb{J}_3$ and thus establishes $\Leftarrow$ of the proposition. Now for the converse, suppose that $Y \equiv (Y_0, \dot{Y}, \boldsymbol{\Lambda}_Y, [\boldsymbol{\Gamma}^Y]) \in \mathbb{J}_3$, then we have

$$Y(t) = Y_0 + \int_0^t \dot{Y}(s) \,\mathrm{d}s + \int_0^t \boldsymbol{\Lambda}_Y^T(s) \,\mathrm{d}\mathbf{W}(s) + \sum_{i,j=1}^{D} \int_0^t \Gamma_{ij}^Y(s) \,\mathrm{d}M_{ij}(s). \tag{A.92}$$

Since $Y \in \mathbb{J}_3 \subseteq \mathbb{J}_1$, by (6.14) and (5.21) we have the following two identities:

$$\dot{Y}(t) = -r(t) Y(t), \tag{A.93}$$

$$\boldsymbol{\Lambda}_Y(t) = -Y(t) \boldsymbol{\theta}(t). \tag{A.94}$$

Using (A.93), (A.94) and (A.91), (A.92) becomes

$$Y(t) = Y_0 + \int_0^t Y(s)[-r(s)\,\mathrm{d}s - \boldsymbol{\theta}^T(s)\,\mathrm{d}\mathbf{W}(s) + \sum_{i,j=1}^{D} Y^{-1}(s)\Gamma_{ij}^{Y}(s)\,\mathrm{d}M_{ij}(s)] \tag{A.95}$$

or

$$\mathrm{d}Y(s) = Y(s)[-r(s)\,\mathrm{d}s - \boldsymbol{\theta}^T(s)\,\mathrm{d}\mathbf{W}(s) + \sum_{i,j=1}^{D} Y^{-1}(s)\Gamma_{ij}^{Y}(s)\,\mathrm{d}M_{ij}(s)]. \tag{A.96}$$

Now since $\inf_{t\in[0,T]} Y(t) > 0$ a.s. we define process $\{\mu_{ij}(t), t \in [0,T]\}$ as

$$\mu_{ij}(t) := Y^{-1}(s)\Gamma_{ij}^{Y}, \quad \text{for all } t \in [0,T].can \tag{A.97}$$

Using (A.96) and (A.97) define a new process $\{\eta(t), t \in [0,T]\}$ as

$$\eta(t) := -\int_0^t r(s)\,\mathrm{d}s - \int_0^t \boldsymbol{\theta}^T(s)\,\mathrm{d}\mathbf{W}(s) + \sum_{i,j=1}^{D}\int_0^t \mu_{ij}(t)\,\mathrm{d}M_{ij}(s). \tag{A.98}$$

From (A.98), we then have

$$\mathrm{d}\eta(t) = r(s)\,\mathrm{d}s - \boldsymbol{\theta}^T(s)\,\mathrm{d}\mathbf{W}(s) + \sum_{i,j=1}^{D} \mu_{ij}(t)\,\mathrm{d}M_{ij}(s). \tag{A.99}$$

Inserting (A.99) into (A.96), we get

$$\mathrm{d}Y(s) = Y(s)\,\mathrm{d}\eta(s). \tag{A.100}$$

Since η is a $\mathcal{F}_t$-semimartingale that is null at the origin, Y is the stochastic exponential of η and has the explicit form

$$Y(t) = Y_0 exp\left\{\eta(t) - \frac{1}{2}\left\langle \eta^c, \eta^c\right\rangle(t)\right\}, \qquad t \in [0,T]. \tag{A.101}$$

Moreover, we have

$$\left\langle \eta^c, \eta^c\right\rangle(t) = \left\langle \boldsymbol{\theta} \bullet \mathbf{W}(s), \boldsymbol{\theta} \bullet \mathbf{W}(s)\right\rangle(t) = \int_0^t \|\boldsymbol{\theta}(s)\|^2\,\mathrm{d}s, \qquad \text{for all } t \in [0,T]. \tag{A.102}$$

Using (A.102), we can write (A.101) as

$$Y(t) = Y_0 exp\left\{-\int_0^t r(s)\,\mathrm{d}s - \boldsymbol{\theta} \bullet \mathbf{W}(t) + [\boldsymbol{\mu}] \bullet M(t) - \frac{1}{2}\int_0^t \|\boldsymbol{\theta}(s)\|^2\,\mathrm{d}s\right\}, \quad t \in [0,T]. \tag{A.103}$$

We can rewrite (A.103) as

$$Y(t) = Y_0 exp\left\{-\int_0^t r(s)\,\mathrm{d}s\right\} exp\left\{-\boldsymbol{\theta} \bullet \mathbf{W}(t) - \frac{1}{2}\int_0^t \|\boldsymbol{\theta}(s)\|^2\,\mathrm{d}s\right\} exp\left\{[\boldsymbol{\mu}] \bullet M(t)\right\}, \quad t \in [0,T]. \tag{A.104}$$

Now because $(\boldsymbol{\theta} \bullet \mathbf{W})$ is a continuous $\mathcal{F}_t$-semimartingale, thus by Theorem B.43 we have

$$\mathcal{E}(-\boldsymbol{\theta} \bullet \mathbf{W})(t) = exp\left\{-\boldsymbol{\theta} \bullet \mathbf{W}(t) - \frac{1}{2}\int_0^t \|\boldsymbol{\theta}(s)\|^2\,\mathrm{d}s\right\} \quad \text{a.s.} \quad \text{for all } t \in [0,T]. \tag{A.105}$$

Next, since $[\boldsymbol{\mu}] \bullet M$ is a purely discontinuous semimartingale, from Definition B.38 we have for all $i, j \in S$,

$$\langle ([\boldsymbol{\mu}] \bullet M)^c, ([\boldsymbol{\mu}] \bullet M)^c \rangle (t) = 0, \qquad \text{for all } t \in [0, T]. \tag{A.106}$$

Furthermore $[\boldsymbol{\mu}] \in \mathcal{G}_3$ implies that

$$\inf_{t \in [0,T]} \mathcal{E}\left([\boldsymbol{\mu}] \bullet M\right)(t) > 0 \quad \text{a.s.} \tag{A.107}$$

Then using (A.106), (A.107) and Theorem B.43 we have that

$$\mathcal{E}\left([\boldsymbol{\mu}] \bullet M\right)(t) = exp\left\{[\boldsymbol{\mu}] \bullet M(t)\right\} \;\; \text{a.s.} \quad \text{for all } t \in [0, T]. \tag{A.108}$$

Now by (A.108) and (A.105), we can rewrite (A.104) as

$$Y(t) = Y_0 exp\left\{-\int_0^t r(s)\,\mathrm{d}s\right\}\mathcal{E}(-\boldsymbol{\theta} \bullet \mathbf{W})(t)\mathcal{E}\left([\boldsymbol{\mu}] \bullet M\right)(t), \quad t \in [0, T]. \tag{A.109}$$

If we define $H(t)$ as

$$H(t) := exp\left\{-\int_0^t r(s)\,\mathrm{d}s\right\}\mathcal{E}(-\boldsymbol{\theta} \bullet \mathbf{W})(t), \quad t \in [0, T]. \tag{A.110}$$

Then $Y(t)$ in (A.109) becomes

$$Y(t) = Y_0 H(t)\mathcal{E}\left([\boldsymbol{\mu}] \bullet M\right)(t), \quad t \in [0, T]. \tag{A.111}$$

Since Y is $\mathbb{P}$-strictly positive and so are Y_0 and H, we must then have that $\mathcal{E}(-\boldsymbol{\theta} \bullet \mathbf{W})$ is strictly positive as well. This implies that:

$$y := Y_0 \in (0, \infty), \tag{A.112}$$

$$[\boldsymbol{\mu}] \in \mathcal{G}_3. \tag{A.113}$$

Lastly, by (A.112) and (A.113), we have

$$Y \in \left\{ yH\mathcal{E}\left([\boldsymbol{\mu}] \bullet M\right) \;\middle|\; y \in (0, \infty) \quad \& \quad [\boldsymbol{\mu}] \in \mathcal{G}_3 \right\} \tag{A.114}$$

which proves the $\Rightarrow$ of the proposition.

□

(b) **(i)** and **(i)** follow from **(a)**, (5.21) and (6.14).

□

Proof of Proposition 6.13: Just like in the proof of Proposition 6.9, it suffices to show that (6.18) is equivalent to (6.26) when (6.25) is true for some $X \in \mathbb{I}_2$, $y \in (0, \infty)$ and $[\boldsymbol{\mu}] \in \mathcal{G}_3$. But before that, we need to first show that if (6.25) holds for some $X \in \mathbb{I}_2$, $y \in (0, \infty)$ and $[\boldsymbol{\mu}] \in \mathcal{G}_3$, then (6.15) is also true. To this end we assume that (6.25) holds i.e.

$$\Phi(X) + \Psi(yH\mathcal{E}\left([\boldsymbol{\mu}] \bullet M\right)) = 0, \tag{A.115}$$

for some $X \in \mathbb{I}_2$ $y \in (0, \infty)$ and $[\boldsymbol{\mu}] \in \mathcal{G}_3$. It then follows from Proposition 6.11(a) that $Y \in \mathbb{J}_3$ if we define Y as:

$$Y := yH\mathcal{E}\left([\boldsymbol{\mu}] \bullet M\right). \tag{A.116}$$

Now using (A.116) we can then rewrite (A.115) as

$$\Phi(X) + \Psi(Y) = 0, \tag{A.117}$$

for some $X \in \mathbb{I}_2$, $Y \in \mathbb{J}_3$, which is exactly (6.15). This proves that (6.15) holds if (6.25) is true for some $X \in \mathbb{I}_2$, $y \in (0, \infty)$ and $[\boldsymbol{\mu}] \in \mathcal{G}_3$.

Now we are going to show that (6.18)(1)(2) and (6.26)(1)(2) are equivalent when (6.25) holds for some $X \in \mathbb{I}_2$, $y \in (0, \infty)$ and $[\boldsymbol{\mu}] \in \mathcal{G}_3$. First we suppose for some arbitrary $X \in \mathbb{I}_2$, $y \in (0, \infty)$ and $[\boldsymbol{\mu}] \in \mathcal{G}_3$ such that (6.25) holds i.e.

$$\Phi(X) + \Psi(yH\mathcal{E}([\boldsymbol{\mu}] \bullet M)) = 0. \tag{A.118}$$

Again we note that by Proposition 6.11(a), we have $Y \in \mathbb{J}_3$ if Y is defined as:

$$Y := yH\mathcal{E}([\boldsymbol{\mu}] \bullet M). \tag{A.119}$$

Then using (A.119), we have that (6.26)(1)(2) are trivially equivalent to (6.18)(1)(2) i.e.

$$\Psi(yH\mathcal{E}([\boldsymbol{\mu}] \bullet M)) = \Psi(Y) \in \mathbb{R}, \tag{A.120}$$

$$X(T) = I(yH(T)\mathcal{E}([\boldsymbol{\mu}] \bullet M)(T)) = I(Y(T)) \quad \text{a.s.}, \tag{A.121}$$

where $X \in \mathbb{I}_2$, $y \in (0, \infty)$, $[\boldsymbol{\mu}] \in \mathcal{G}_3$ and $Y \in \mathbb{J}_3$.

Lastly, to establish the equivalence between (6.18)(3) and (6.26)(3) we suppose that (6.18)(1) - (3) hold for some $(X, Y) \in \mathbb{I}_2 \times \mathbb{J}_3$. Then (6.18)(3) implies that XY is a $\mathcal{F}_t$-martingale. Moreover, since $Y \in \mathbb{J}_3$ we can, by Proposition 6.11(a) write Y as:

$$Y := yH\mathcal{E}([\boldsymbol{\mu}] \bullet M), \tag{A.122}$$

for some fixed $y \in (0, \infty)$ and $[\boldsymbol{\mu}] \in \mathcal{G}_3$. Using (6.18)(3) and (A.122) we have that $XyH\mathcal{E}([\boldsymbol{\mu}] \bullet M)$ is a $\mathcal{F}_t$-martingale, which implies that (6.26)(3) is true as they are only differ by a fixed constant, namely $y \in (0, \infty)$.

For the converse, we posit (6.26)(1) - (3) for some $X \in \mathbb{I}_2$, $y \in (0, \infty)$ and $[\boldsymbol{\mu}] \in \mathcal{G}_3$ Under (6.26)(3), we have that $XH\mathcal{E}([\boldsymbol{\mu}] \bullet M)$ is a $\mathcal{F}_t$-martingale. This implies that $XyH\mathcal{E}([\boldsymbol{\mu}] \bullet M)$ is also a $\mathcal{F}_t$-martingale. Then using this together with Proposition 6.11(a), which gives

$$Y := yH\mathcal{E}([\boldsymbol{\mu}] \bullet M) \in \mathbb{J}_3 \tag{A.123}$$

imply that (6.18)(3) is also true i.e. XY is a $\mathcal{F}_t$-martingale for some $(X, Y) \in \mathbb{I}_2 \times \mathbb{J}_3$.

□

Proof of Theorem 7.5 From the martingale representation theorem for general semimartingales (see III.4.39 of [7]), one easily sees that there exists a $\mathcal{F}_t^{\boldsymbol{\alpha}}$-predictable integrand $\Gamma \in \mathcal{H}_3$ such that

$$Q(t) = Q(0) + \sum_{i,j=1}^{D} \int_0^t \Gamma_{ij}(s)\, \mathrm{d}M_{ij}(s), \qquad t \in [0, T]. \tag{A.124}$$

Since we have $Q(s) \geq c_1 > 0$, for all $s \in [0, T]$, we shall define a $\mathcal{F}_t^{\boldsymbol{\alpha}}$-predictable process $\{\mu_{ij}(s), s \in [0, T]\}$ as follow

$$\mu_{ij}(s) := \Gamma_{ij}(s)/Q(s-), \quad i, j \in S.. \tag{A.125}$$

Then it follows form (A.124) and (A.125) that

$$Q(t) = Q(0) + \sum_{i,j=1}^{D} \int_0^t Q(s-)\mu_{ij}(s)\,\mathrm{d}M_{ij}(s) \tag{A.126}$$

or equivalently (from (6.21))

$$Q(t) = Q(0)\mathcal{E}([\boldsymbol{\mu}] \bullet M)(t), \qquad t \in [0, T]. \tag{A.127}$$

Lastly, the $\nu_{[M_{ij}]}$-uniqueness of each $[\boldsymbol{\mu}]_{ij}$ follows from Lemma 4.4 **(iv)** i.e.

$$[M_{ij}, M_{ab}](t) = 0, \qquad \text{for all } t \in [0, T]. \tag{A.128}$$

when $(i, j) \neq (a, b)$.

□

B Standard definitions and results

B.1 cádlág stochastic processes

Definition B.1. A process $X = \{X(t), t \in [0,T]\}$ is *cádlág* if the sample function $t \to X(t,\omega) : [0,T] \to \mathbb{R}$ is right continuous with finite left-hand limits for each and every ω.

$\diamond$

Remark B.2. If X is cádlág , then we define

$$X(0_-) := X(0), \tag{B.1}$$

$$X(t_-) := \lim_{s \uparrow t,\, s<t} X(s), \qquad t \in (0,T] \tag{B.2}$$

Furthermore, we define the *jump process* $\{\Delta X(t), t \in [0,T]\}$ as

$$\Delta X(t) := X(t) - X(t_-), \qquad t \in (0,T]. \tag{B.3}$$

$\triangleleft$

Definition B.3. A *filtered probability space* is a pair $((\Omega, \mathcal{F}, \mathbb{P}), \{\mathcal{F}_t\})$ consisting of a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and a filtration $\{\mathcal{F}_t\,,\, t \in [0,T]\}$ on $\mathcal{F}$. A *standard filtered probability space* is a filtered probability space $((\Omega, \mathcal{F}, \mathbb{P}), \{\mathcal{F}_t\})$ with the following additional properties:

- $(\Omega, \mathcal{F}, \mathbb{P})$ is a complete probability space
- $\mathcal{F}_0$ includes all $\mathbb{P}$-null events in $\mathcal{F}$
- the filtration $\{\mathcal{F}_t\}$ is right-continuous, i.e. $\mathcal{F}_t = \bigcap_{u>t} \mathcal{F}_u$ for all $t \in [0,T]$

$\diamond$

Remark B.4. We shall denote by E the expectation with respect to the measure $\mathbb{P}$.

$\triangleleft$

Definition B.5. A process $\{X(t), t \in [0,T]\}$ on a filtered probability space $((\Omega, \mathcal{F}, \mathbb{P}), \{\mathcal{F}_t\})$ is *non-decreasing* if the mappings $t \to X(t,\omega)$ are non-decreasing on $[0,T]$ for all $\omega \in \Omega$.

$\diamond$

B.2 Spaces of martingales

In this section we formulate the definitions of martingales and local martingales, together with some results on these that are needed in the thesis. All of our definitions and results are restricted to the finite closed trading interval $[0,T]$ because all processes throughout the thesis are on this interval.

Definition B.6. A $\mathbb{R}$-valued, $\mathcal{F}_t$-adapted processes $\{M(t), t \in [0,T]\}$ on $(\Omega, \mathcal{F}, \mathbb{P})$ is called a $\{\mathcal{F}_t\}$-*martingale* when

- $E|M(s)| < \infty, \qquad$ for all $s \in [0,T]$
- $E\left(M(t)|\mathcal{F}_s\right) = M(s), \quad \mathbb{P}$-a.s for all $0 \le s \le t \le T$.

We shall use $\mathcal{M}(\{\mathcal{F}_t\})$ to denote the set of all $\mathcal{F}_t$-martingales on $(\Omega, \mathcal{F}, \mathbb{P})$.

◇

Remark B.7. There are no other probability measures besides $\mathbb{P}$ in the thesis, and in particular no measure changes, and therefore measure $\mathbb{P}$ will be understood and not specifically indicated in the notation at Remark B.6, Remark B.10 and elsewhere. In particular, in order to avoid cumbersome notation, we always use E to denote expectation with respect to the probability measure $\mathbb{P}$ (instead of the more complete $E^{\mathbb{P}}$), and we denote the set of $\mathcal{F}_t$-martingales on $(\Omega, \mathcal{F}, \mathbb{P})$ by $\mathcal{M}(\{\mathcal{F}_t\})$ (instead of the more complete $\mathcal{M}(\{\mathcal{F}_t\}, \mathbb{P})$ or even more detailed $\mathcal{M}(\{\mathcal{F}_t\}, (\Omega, \mathcal{F}, \mathbb{P}))$ etc).

Definition B.8. Given a constant $p \in (1, \infty]$, a martingale M is called L_p-bounded when $E|M(T)|^p < \infty$ (we then have $\sup_{t\in[0,T]} E|M(t)|^p < \infty$, as follows from Jensen's inequality). An L_2-bounded martingale is also called *square integrable.*

◇

Next we will state a very useful theorem of L_p-bounded martingales

Theorem B.9. Doob's L_p-Inequality
Let $p \in (1, \infty)$. Let M be a cádlág martingale with respect to the filtered probability space $((\Omega, \mathcal{F}, \mathbb{P}), \{\mathcal{F}_t\})$ which is L_p-bounded. Then

$$E\left(\sup_{t\in[0,T]} |M(t)|^p\right) \leq \left(\frac{p}{p-1}\right)^p E|M(T)|^p. \tag{B.4}$$

◇

Remark B.10. We introduce some notations for classes of martingales that appear throughout this thesis (see Remark B.7):

(a) $\mathcal{M}_0(\{\mathcal{F}_t\})$ denotes the set of $M \in \mathcal{M}(\{\mathcal{F}_t\}))$ which are $\mathbb{P}$-a.s. null at the origin i.e. $M(0) = 0$ $\mathbb{P}$-a.s.

(b) $\mathcal{M}^c(\{\mathcal{F}_t\})$denotes the set of $M \in \mathcal{M}(\{\mathcal{F}_t\})$ which are sample path continuous.

(c) $\mathcal{M}_0^c(\{\mathcal{F}_t\}) := \mathcal{M}_0(\{\mathcal{F}_t\}) \cap \mathcal{M}^c(\{\mathcal{F}_t\})$ i.e. the set of martingales that are null at the origin $\mathbb{P}$-a.s and have continuous sample path.

(d) $\mathcal{M}^2(\{\mathcal{F}_t\})$ denotes the set of $M \in \mathcal{M}(\{\mathcal{F}_t\})$ which are square-integrable.

(e) $\mathcal{M}_0^2(\{\mathcal{F}_t\}) := \mathcal{M}_0(\{\mathcal{F}_t\}) \cap \mathcal{M}^2(\{\mathcal{F}_t\})$ i.e. the set of martingales that are null at the origin $\mathbb{P}$-a.s and square-integrable.

◁

B.3 Spaces of local martingales

Definition B.11. For a sequence $\mathcal{F}_t$-stopping time, $\{T^m\, , \, m \in \mathbb{N}\}$, we write

$$T^m \Uparrow T \tag{B.5}$$

to indicate the following

1. $0 \leq T^m(\omega) \leq T^{m+1}(\omega) \leq T$ for all $\omega \in \Omega$ and for all $m \in \mathbb{N}$
2. there exists $M(\omega) \in \mathbb{N}$ such that $T^m(\omega) = T$, for all $m \geq M(\omega)$ and for all $\omega \in \Omega$

◇

Remark B.12. Definition B.11 ensures that the right end-point T of the interval $[0, T]$ is included in the localization sequence. It also rules out the possibility that the T^m are all *strictly* less than T (i.e. $T^m < T$) for all $m = 1, 2, ...$

◁

Definition B.13. A real-valued process $\{M(t), t \in [0, T]\}$ on $(\Omega, \mathcal{F}, \mathbb{P})$ is called an $\mathcal{F}_t$-local martingale if there is a sequence $\{T^m, m \in \mathbb{N}\}$ of $\mathcal{F}_t$-stopping times such that

1. $T^m \Uparrow T \qquad \mathbb{P}$-a.s.
2. $\{M(t \wedge T^m), t \in [0, T]\} \in \mathcal{M}(\{\mathcal{F}_t\}))$ for each $m \in \mathbb{N}$

We shall use $\mathcal{M}_{\text{loc}}(\{\mathcal{F}_t\})$ to denote the set of $\mathcal{F}_t$-local martingales on $(\Omega, \mathcal{F}, \mathbb{P})$.

◇

Remark B.14. The sequence $\{T^m, m \in \mathbb{N}\}$ of $\mathcal{F}_t$-stopping times in Definitions B.13 is called a *localizing sequence* for M.

◁

Remark B.15. We will introduce some notations for some classes of local martingales that will appear in this thesis,

(a) $\mathcal{M}_{0,\text{loc}}(\{\mathcal{F}_t\})$ denotes the set of $M \in \mathcal{M}_{\text{loc}}(\{\mathcal{F}_t\})$ which are $\mathbb{P}$-a.s. null at the origin.

(b) $\mathcal{M}^c_{\text{loc}}(\{\mathcal{F}_t\})$ denotes the set of $M \in \mathcal{M}_{\text{loc}}(\{\mathcal{F}_t\})$ which are sample-path continuous.

(c) $\mathcal{M}^c_{0,\text{loc}}(\{\mathcal{F}_t\}) := \mathcal{M}_{0,\text{loc}}(\{\mathcal{F}_t\}) \cap \mathcal{M}^c_{\text{loc}}(\{\mathcal{F}_t\})$ i.e. the set of local martingales that are null at the origin $\mathbb{P}$-a.s and have continuous sample path.

(d) $\mathcal{M}^2_{0,\text{loc}}(\{\mathcal{F}_t\}) := \mathcal{M}_{\text{loc}}(\{\mathcal{F}_t\}) \cap \mathcal{M}^2_{\text{loc}}$. i.e. the set of local martingales that are null at the origin $\mathbb{P}$-a.s and square-integrable.

◁

From Jacod and Shiryayev [7], Definition I.4.11, we have the following definitions.

Definition B.16. Two local martingales N and M are *orthogonal* if their product $L = MN$ is a local martingale.

◇

Definition B.17. A local martingale M is called a *purely discontinuous local martingale* if

1. $M(0) = 0$
2. It is orthogonal to all *continuous* local martingales i.e. $MN \in \mathcal{M}_{\text{loc}}(\{\mathcal{F}_t\})$ for each $N \in \mathcal{M}^c_{\text{loc}}(\{\mathcal{F}_t\})$.

◇

B.4 Finite variation processes

Definition B.18. A process $\{A(t) : t \in [0,T]\}$ is of *finite variation* if it is an $\mathcal{F}_t$-adapted and cádlág process such that each path $t \to A(\omega, t)$ is of finite variation on $[0,T]$ i.e. for all $\omega \in \Omega$

$$V_A(\omega, t) < \infty \tag{B.6}$$

for the process $t \to A(\omega, t)$, $t \in [0,T]$ where the *variation process* $V_A(\omega, t)$is defined as

$$V_A(\omega, t) := \int_{(0,t]} |\,\mathrm{d}A(\omega, s)| = \sup \sum_{i=1}^{n} |A(\omega, s_i) - A(\omega, s_{i-1})|, \qquad t \in [0,T] \tag{B.7}$$

(the supremum is taken over all *finite* partitions $0 = s_0 < s_1 < .... < s_n = t$ of $[0,T]$).

◇

Remark B.19. We will introduce some notations for some classes of finite variation processes that will appear in this thesis,

(a) $\mathcal{FV}(\{\mathcal{F}_t\})$ denotes the set of all $\mathbb{R}$-valued, $\mathcal{F}_t$-adapted, cádlág processes on the filtered probability space $((\Omega, \mathcal{F}, \mathbb{P}), \{\mathcal{F}_t\})$ which are of finite variation.

(b) $\mathcal{FV}_0(\{\mathcal{F}_t\})$ denotes the set of $A \in \mathcal{FV}(\{\mathcal{F}_t\})$ which are $\mathbb{P}$-a.s. null at the origin.

(c) $\mathcal{FV}^+(\{\mathcal{F}_t\})$ denotes the set of $A \in \mathcal{FV}(\{\mathcal{F}_t\})$ which are non-decreasing.

(d) $\mathcal{FV}_0^+(\{\mathcal{F}_t\})$ denotes the set of $A \in \mathcal{FV}^+(\{\mathcal{F}_t\})$ which are $\mathbb{P}$-a.s. null at the origin.

◁

Definition B.20. A process $\{A(t) : t \in [0,T]\}$ is of *integrable variation* if it is of finite variation and

$$E\left[V_A(\omega, T)\right] < \infty \tag{B.8}$$

for V_A given by (B.7).

◇

Remark B.21. We will introduce some notations for some classes of integrable variation processes that will appear in this thesis,

(a) $\mathcal{IV}(\{\mathcal{F}_t\})$ denotes the set of $A \in \mathcal{FV}(\{\mathcal{F}_t\})$ which are of integrable variation.

(b) $\mathcal{IV}_0(\{\mathcal{F}_t\}))$ denotes the set of $A \in \mathcal{IV}(\{\mathcal{F}_t\})$ which are $\mathbb{P}$-a.s. null at the origin.

(c) $\mathcal{IV}^+(\{\mathcal{F}_t\})$ denotes the set of $A \in \mathcal{FV}^+(\{\mathcal{F}_t\})$ which are integrable, that is $E\left[A(T)\right] < \infty$.

(d) $\mathcal{IV}_0^+(\{\mathcal{F}_t\})$ denotes the set of $A \in \mathcal{IV}^+(\{\mathcal{F}_t\})$ which are $\mathbb{P}$-a.s. null at the origin.

◁

B.5 Angle bracket processes for locally square integrable martingales

The definitions of the angle bracket quadratic variation and co-quadratic variation processes is motivated by the following theorem from Jacod and Shiryaev [7], Theorem I.4.2:

Theorem B.22. For each pair $N, M \in \mathcal{M}_{loc,2}(\{\mathcal{F}_t\})$, there exists a real-valued, cádlág $\mathcal{F}_t$-adapted and 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 predictable
3. $NM - \langle N, M\rangle \in \mathcal{M}_{\text{loc}}(\{\mathcal{F}_t\})$

and the following identity is true,

$$\langle N, M\rangle = \frac{1}{4}(\langle N+M, N+M\rangle - \langle N-M, N-M\rangle). \tag{B.9}$$

Moreover, if $N, M \in \mathcal{M}^2(\{\mathcal{F}_t\})$ then we also have $\langle N, M\rangle \in \mathcal{IV}(\{\mathcal{F}_t\})$ and $NM - \langle N, M\rangle$ is a *uniformly integrable* martingale. Additionally, $\langle N, M\rangle$ is non-decreasing when $N = M$

◇

Remark B.23. From Theorem B.22, we see that the process $\langle N, M\rangle$ is uniquely determined and we called it the angle-bracket quadratic co-variation process of N and M. Furthermore, if $N = M \in \mathcal{M}^2_{\text{loc}}(\{\mathcal{F}_t\})$, then $\langle M, M\rangle$ (or $\langle M\rangle$ for short) is called the angle-bracket quadratic variation process of M.

◁

Remark B.24. A continuous local martingale is locally bounded, and therefore of course locally square-integrable, that is we have

$$\mathcal{M}^c_{\text{loc}}(\{\mathcal{F}_t\}) \subset \mathcal{M}^2_{\text{loc}}(\{\mathcal{F}_t\}). \tag{B.10}$$

It follows that the angle-bracket quadratic covariation $\langle N, M\rangle$ of the continuous local martingales $N, M \in \mathcal{M}^c_{\text{loc}}(\{\mathcal{F}_t\})$ is also given by Theorem B.22.

◁

B.6 Square bracket processes for local martingales

Remark B.25. In Theorem B.22 it is essential that M and N be *locally square integrable* martingales, that is $N, M \in \mathcal{M}^2_{\text{loc}}(\{\mathcal{F}_t\})$. If we simply assume that $N, M \in \mathcal{M}_{\text{loc}}(\{\mathcal{F}_t\})$, then there generally will not exist an angle-bracket quadratic co-variation process $\langle M, N\rangle$ with the properties stated in Theorem B.22. In this section therefore, we will formulate the *square-bracket* processes which exist even when the local martingales are not necessarily square integrable.

◁

From Jacod and Shiryayev [7], equation I.4.46 and Proposition I.4.50 and Rogers and Williams [15], Theorem VI.36.6 and Theorem VI.37.8, we have the following theorem.

Theorem B.26. For each pair $N, M \in \mathcal{M}_{\text{loc}}(\{\mathcal{F}_t\})$, there exists a real-valued, cádlág $\mathcal{F}_t$-adapted and finite variation process $[N, M]$, which is unique up to indistinguishability, such that

1. $[N, M](0) = 0$ a.s.
2. $\Delta[N, M](t) = \Delta N(t)\Delta M(t)$ for all $t > 0$
3. $NM - [N, M] \in \mathcal{M}_{\text{loc}}(\{\mathcal{F}_t\})$

and the following identity is true,

$$[N, M] = \frac{1}{4}([N + M, N + M] - [N - M, N - M]). \tag{B.11}$$

Moreover, if $N, M \in \mathcal{M}^2(\{\mathcal{F}_t\})$ then $[N, M] \in \mathcal{IV}_0(\{\mathcal{F}_t\})$ and $NM - [N, M]$ is a *uniformly integrable* martingale. Additionally, $[N, M]$ is non-decreasing when $N = M$, in which case we will denote $[N, M]$ by $[M]$.

◇

Remark B.27. We call $[N, M]$ the square bracket quadratic co-variation process of N and M. Furthermore, if $N = M \in \mathcal{M}_{\text{loc}}(\{\mathcal{F}_t\})$, then $[M, M]$ (or $[M]$ for short) is called the square bracket quadratic variation process of the local martingale M.

◁

From Protter [13], Chapter II, Section 6, Corollary 3, page 73, we have the the following corollary that gives conditions for a local martingale over $[0, T]$ to be a square-integrable martingale.

Corollary B.28. *If $M \in \mathcal{M}_{loc}(\{\mathcal{F}_t\})$, then we have $E|M(t)|^2 < \infty$ for all $t \geq 0$ (i.e. $M \in \mathcal{M}^2(\{\mathcal{F}_t\})$) if and only if $E[M, M](t) < \infty$ for all $t \geq 0$ Moreover, if $E[M, M](t) < \infty$ for all $t \geq 0$, then*

$$E|M(t)|^2 = E[M, M](t), \qquad t \geq 0. \tag{B.12}$$

◇

Next, we shall state a modified version of result of Protter [13], Section 6, Corollary 4, page 74, which relates a local martingale and a L^2-bounded martingale.

Corollary B.29. *Let $\{M(t), t \in [0, T]\} \in \mathcal{M}_{loc}(\{\mathcal{F}_t\})$, if $E[M, M](T) < \infty$, then M is a L^2-bounded $\mathcal{F}_t$-martingale i.e.*

$$\sup_{t\in[0,T]} E|M(t)|^2 = E|M(T)|^2 < \infty \tag{B.13}$$

and (B.12) continue to hold for all $t \in [0, T]$.

◇

B.7 Semimartingales and their decomposition

Definition B.30. A $\mathbb{R}$-valued $\mathcal{F}_t$-adapted process $\{X(t), t \in [0, T]\}$ on $(\Omega, \mathcal{F}, \mathbb{P})$ is called a *semimartingale* if it can be written in the form

$$X(t) = X(0) + M(t) + A(t), \qquad t \in [0, T] \tag{B.14}$$

for some $M \in \mathcal{M}_{0,\text{loc}}(\{\mathcal{F}_t\})$ and $A \in \mathcal{FV}_0(\{\mathcal{F}_t\})$.

◇

Remark B.31. In (B.14), M is called the *local martingale part* of X whereas A is the *finite variation part*. It is worth noting however, that the decomposition on the right side of (B.14) is generally not unique, this means that we could have

$$X = X(0) + \tilde{M} + \tilde{A} \tag{B.15}$$

for some $\tilde{M} \in \mathcal{M}_{0,\text{loc}}(\{\mathcal{F}_t\})$ and $\tilde{A} \in \mathcal{FV}_0(\{\mathcal{F}_t\})$ that are distinct from the M and A in (B.14). Furthermore, We will use the notation $\mathcal{SM}(\{\mathcal{F}_t\})$ to denote the set of all $\mathcal{F}_t$-adapted semimartingales on $(\Omega, \mathcal{F}, \mathbb{P})$.

◁

As we noted in the previous remark, the decomposition of X at (B.14) is generally not unique. However, from Jacod and Shiryaev [7] Theorem I.4.18 one can see that the decomposition of M, the local martingale part of X is unique up to indistinguishability, as stated in the following theorem:

Theorem B.32. Any local martingale $M \in \mathcal{M}_{\text{loc}}(\{\mathcal{F}_t\})$ admits a unique (up to indistinguishability) decomposition

$$M = M(0) + M^c + M^d, \tag{B.16}$$

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 B.33. We call M^c the *continuous part* and M^d the *purely discontinuous part* of the local martingale M.

◁

From Theorem B.32, we have the following result proposition (see C. Donnelly [4], Appendix B, Corollary B.2.39, page 120):

Proposition B.34. *Let $X \in \mathcal{SM}(\{\mathcal{F}_t\})$ and fix any two arbitrary decompositions*

$$X = X(0) + M + A \tag{B.17}$$

$$X = X(0) + \tilde{M} + \tilde{A} \tag{B.18}$$

for some $M, \tilde{M} \in \mathcal{M}_{0,loc}(\{\mathcal{F}_t\})$ and some $A, \tilde{A} \in \mathcal{FV}_0(\{\mathcal{F}_t\})$. Applying Theorem B.32, we can decompose the local martingales M and $\tilde{M}$ as follow,

$$M = M(0) + M^c + M^d \tag{B.19}$$

$$\tilde{M} = \tilde{M}(0) + \tilde{M}^c + \tilde{M}^d \tag{B.20}$$

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. Then M^c and $\tilde{M}^c$ are indistinguishable.

◇

We will now define the *square bracket process* of a pair of semimartingales, but first we will need the following result (from Section 2.1 of Liptser and Shiryaev [12]).

Proposition B.35. *If $X, Y \in \mathcal{SM}(\{\mathcal{F}_t\})$ then for all $t \geq 0$, the following is true*

$$\sum_{0\leq s\leq t} |\Delta X(s)\Delta Y(s)| < \infty \qquad a.s. \tag{B.21}$$

$\diamond$

Using Proposition B.35 we can define the square bracket process for X, $Y \in \mathcal{SM}(\{\mathcal{F}_t\})$ in the following definition.

Definition B.36. Given any $X, Y \in \mathcal{SM}(\{\mathcal{F}_t\})$ define

$$[X, Y](t) := \langle X^c, Y^c\rangle (t) + \sum_{0\leq s\leq t} \Delta X(s)\Delta Y(s) \qquad t \geq 0 \tag{B.22}$$

Here X^c and Y^c are the continuous local martingale parts X and Y respectively, which are uniquely by Theorem B.32.

$\diamond$

Remark B.37. Since $X^c, Y^c \in \mathcal{M}^c_{0,\text{loc}}(\{\mathcal{F}_t\})$, the first term on the right of (B.22) is given by Theorem B.22. Furthermore, Proposition B.35 ensures that the second term on the right of (B.22) is absolutely convergent $\mathbb{P}$-a.s.

$\triangleleft$

From Protter [13], Chapter II, Section 6, page 71 we have the following definition and theorem.

Definition B.38. Let X be a semimartingale and let X^c denote its continuous local martingale part. Then X is called a *purely discontinuous semimartingale* if $\langle X^c, X^c\rangle = 0$.

$\diamond$

Theorem B.39. If a semimartingale X is adapted, cádlág with paths of finite variation then X is a purely discontinuous semimartingale.

$\diamond$

From a special case of Jacod and Shiryaev [7], Theorem I.4.52, we have the following theorem.

Theorem B.40. Let X be a purely discontinuous member of $\mathcal{SM}(\{\mathcal{F}_t\})$. Then for any $Y \in \mathcal{SM}(\{\mathcal{F}_t\})$ we have

$$[X, Y](t) = \sum_{0\leq s\leq t} \Delta X(s)\Delta Y(s). \tag{B.23}$$

$\diamond$

B.8 Itô's formula for general semimartingales

From Rogers and Williams [15], Theorem VI.38.3, we have the following *Itô integration by parts formula* for semimartingales:

Theorem B.41. Let $X, Y \in \mathcal{SM}(\{\mathcal{F}_t\})$. Then

$$X(t)Y(t) = X(0)Y(0) + \int_0^t X(s_-)\,\mathrm{d}Y(s) + \int_0^t Y(s_-)\,\mathrm{d}X(s) + [X, Y](t). \tag{B.24}$$

◇

From Rogers and Williams [15], Theorem VI.39.1, we have a more general version of Itô's formula for semimartingales.

Theorem B.42. Let $f : \mathbb{R}^N \to \mathbb{R}$ be a function of class C_2. Suppose $\mathbf{X} = (X_1, ..., X_N)$ is a semimartingale in $\mathbb{R}^N$, that is $X_i \in \mathcal{SM}(\{\mathcal{F}_t\})$ for all $i = 1, 2, ..., N$. Then

$$\begin{aligned} &f(\mathbf{X}(t)) - f(\mathbf{X}(0)) \\ =& \sum_{i=0}^{N} \int_0^t \frac{\partial f}{\partial X_i} \mathbf{X}(s_-) \, \mathrm{d}\mathbf{X}(s) + \frac{1}{2} \sum_{i=0}^{N} \sum_{j=0}^{N} \int_0^t \frac{\partial^2 f}{\partial X_i X_j} \mathbf{X}(s_-) \, \mathrm{d}[X_i^c, X_j^c](s) \\ &\qquad + \sum_{0 \le s \le t} \left(f(\mathbf{X}(s)) - f(\mathbf{X}(s_-) - \sum_{i=1}^{N} \frac{\partial f}{\partial X_i} \mathbf{X}(s_-) \Delta X_i(s) \right) \end{aligned} \tag{B.25}$$

where X_i^c is the continuous local martingale part of the semimartingale X_i for $i = 1, 2, ..., N$.

◇

B.9 Doléans-Dade exponential results

From Elliott [6], Chapter 13, Theorem 13.5 and Remark 13.6, we have the following theorem.

Theorem B.43. Suppose $X \in \mathcal{SM}_0(\{\mathcal{F}_t\})$. Let X^c denote its continuous local martingale part, then there is an unique $Z \in \mathcal{SM}(\{\mathcal{F}_t\})$ such that

$$Z(t) = 1 + \int_0^t Z(s_-) \, \mathrm{d}X(s). \tag{B.26}$$

In addition, $Z(s)$ is given by the expression

$$Z(t) = exp\left\{ X(t) - \frac{1}{2} \langle X^c, X^c \rangle \right\} \prod_{s \in [0,t]} (1 + \Delta X(s)) exp \left\{ -\Delta X(s) \right\} \qquad t \ge 0 \tag{B.27}$$

where the infinite product is absolutely convergent $\mathbb{P}$-a.s.

◇

Remark B.44. One can see that from (B.27) Z is strictly positive if and only if

$$\Delta X(t) > -1 \quad \text{a.s.} \quad \text{for all } t \ge 0. \tag{B.28}$$

In particular, if X is continuous then $[X, X] = \langle X, X \rangle$ and (B.27) becomes

$$Z(t) = exp\left\{ X(t) - \frac{1}{2}[X, X](t) \right\} \quad \text{a.s.} \quad t \ge 0. \tag{B.29}$$

We will use the notation $\mathcal{E}(X)(t)$ to represent $Z(t)$ i.e. $Z(t) = \mathcal{E}(X)(t)$. Moreover, we call $\mathcal{E}(X)$ the *Doléans-Dade exponential* of the semimartingale X.

◁

From Elliott [6], Chapter 13, Corollary 13.58, we also have the following result.

Corollary B.45. *If X and Y are semimartingales, then*

$$\mathcal{E}(X)\mathcal{E}(Y) = \mathcal{E}(X + Y + [X, Y]). \tag{B.30}$$

In particular, if $[X, Y] = 0$ then

$$\mathcal{E}(X)\mathcal{E}(Y) = \mathcal{E}(X + Y). \tag{B.31}$$

◇

From Donnelly ([4], Section 4, Proposition 4.6.3), we have the following proposition.

Proposition B.46. *Suppose $\{\boldsymbol{\lambda}(t)\,,\, t \in [0, T]\}$ is an uniformly bounded and $\mathcal{F}_t$-progressively measurable process on $\Omega \times [0, T]$. Then*

$$\mathcal{E}(p\,\boldsymbol{\lambda} \bullet \mathbf{W}) \text{ is a square integrable } \mathcal{F}_t\text{-martingale for any } p \in \mathbb{R}. \tag{B.32}$$

◇

B.10 Compensator Results

Before we introduce some results of the *compensator*, we will need the following new notations for sets of integrable variation process.

Definition B.47. A process $\xi : \Omega \times [0, T] \to \mathbb{R}$ is called *locally integrable* if there exists a sequence of $\mathcal{F}_t$-stopping times $\{T^m\}$ such that

$$\textbf{(i)} \quad T^m < T^{m+1} \qquad \forall n \geq 0 \tag{B.33}$$

$$\textbf{(ii)} \quad \lim_{n\to\infty} T^m = \infty \tag{B.34}$$

$$\textbf{(iiii)} \quad E[A(T^m)] < \infty \qquad \forall n \geq 0 \tag{B.35}$$

◇

Definition B.48. Recalling Remark B.19 we denote by $\mathcal{IV}_{\text{loc}}(\{\mathcal{F}_t\})$ the set of all $A \in \mathcal{FV}_0(\{\mathcal{F}_t\})$ such that there exists a sequence of A dependent $\mathcal{F}_t$-stopping times $\{T^M\}$ where

1. $T^m \Uparrow T$
2. $A[0, T^m] \in \mathcal{IV}_0(\{\mathcal{F}_t\})$ for each $A \in \mathcal{FV}_0(\{\mathcal{F}_t\})$

We call members of $\mathcal{IV}_{\text{loc}}(\{\mathcal{F}_t\})$ *locally integrable variation* processes.

◇

Remark B.49. It's immediate from Definition B.48 that

$$\mathcal{IV}(\{\mathcal{F}_t\}) \subset \mathcal{IV}_{\text{loc}}(\{\mathcal{F}_t\}). \tag{B.36}$$

Moreover, we will denote by $\mathcal{IV}^+_{0,\text{loc}}(\{\mathcal{F}_t\})$ the set of all process with sample paths that are null at the origin, non-decreasing and of locally integrable variation i.e.

$$\mathcal{IV}^+_{0,\text{loc}}(\{\mathcal{F}_t\}) = \mathcal{FV}^+_0(\{\mathcal{F}_t\}) \cap \mathcal{IV}_{\text{loc}}(\{\mathcal{F}_t\}). \tag{B.37}$$

◁

From Jacod and Shiryaev [7], Theorem I.3.17, we have the result.

Theorem B.50. Let $A \in \mathcal{IV}^+_{0,\mathrm{loc}}(\{\mathcal{F}_t\})$. Then there exists a *predictable* process $A^p \in \mathcal{IV}^+_{0,\mathrm{loc}}(\{\mathcal{F}_t\})$ such that

$$A - A^p \in \mathcal{M}_{0,\mathrm{loc}}(\{\mathcal{F}_t\}). \tag{B.38}$$

Moreover, A^p is unique in the sense that, if for any predictable process $\tilde{A} \in \mathcal{IV}^+_{0,\mathrm{loc}}$ we have

$$A - \tilde{A} \in \mathcal{M}_{0,\mathrm{loc}}(\{\mathcal{F}_t\}), \tag{B.39}$$

then A^p and $\hat{A}$ are indistinguishable.

◇

Remark B.51. The process A^p in (B.38) is called the *compensator* (or *dual predictable projection*) of the given process A, and it is unique to within indistinguishability.

◁

Theorem B.52. Let $A \in \mathcal{IV}^+_{0,\mathrm{loc}}(\{\mathcal{F}_t\})$ and any predictable process $\bar{A} \in \mathcal{IV}^+_{0,\mathrm{loc}}(\{\mathcal{F}_t\})$, the following are equivalent:

1. $\bar{A}$ is the compensator of A i.e. $\bar{A} = A^p$
2. for all nonnegative predictable processes H one has

$$E\int_0^\infty H(s)\,\mathrm{d}A(s) = E\int_0^\infty H(s)\,\mathrm{d}\bar{A}(s) \tag{B.40}$$

◇

By Theorem B.26 and given a local martingale M we have that $[M] \in \mathcal{FV}^+_0(\{\mathcal{F}_t\})$. From Rogers and Williams [15], Theorem VI.34.2 we have the following theorem giving conditions on M which ensure that $[M] \in \mathcal{IV}^+_{0,\mathrm{loc}}(\{\mathcal{F}_t\})$ and shows that the compensator of $[M]$ and $\langle M\rangle$ are identical.

Theorem B.53. Let $M \in \mathcal{M}_{0,\mathrm{loc}}(\{\mathcal{F}_t\})$, the following statements are equivalent

1. $M \in \mathcal{M}^2_{0,\mathrm{loc}}(\{\mathcal{F}_t\})$
2. $[M] \in \mathcal{IV}^+_{0,\mathrm{loc}}(\{\mathcal{F}_t\})$ i.e. $[M]$ is locally integrable

Furthermore, under these equivalent conditions we also have the following

$$\langle M\rangle = [M]^p, \tag{B.41}$$

that is $\langle M\rangle$ is the compensator of $[M]$.

◇

B.11 Convex analysis

In this section we will summarize some basic definitions and results from convex analysis that will be utilized in this thesis.

Definition B.54. Let $f : \mathbb{R}^m \to \mathbb{R} \cup \{+\infty\}$ be a function on $\mathbb{R}^m$. Then the *convex conjugate* of f is a function, denoted by $f^* : \mathbb{R}^m \to \mathbb{R} \cup \{+\infty\}$ defined as follows:

$$f^*(\mathbf{y}) := \sup_{\mathbf{x} \in \mathbb{R}^m} \left\{ \mathbf{x}^T \mathbf{y} - f(\mathbf{x}) \right\}. \tag{B.42}$$

◇

Definition B.55. Let V be a real vector space and let $f : V \to \mathbb{R} \cup \{+\infty\}$ be a function on V. Then f is said to be *convex* if for every x, $y \in V$,

$$f(\epsilon x + (1-\epsilon)y) \leq \epsilon f(x) + (1-\epsilon) f(y) \qquad \text{for all } \epsilon \in [0,1]. \tag{B.43}$$

◇

Glossary

$[M]$, $[M, M]$ square-bracket quadratic variation process of M, page 62
$[N, M]$ square-bracket quadratic co-variation process of N and M, page 61
$\langle M\rangle$, $\langle M, M\langle$ angle-bracket quadratic variation process of M, page 61
$\langle N, M\rangle$ angle-bracket quadratic co-variation process of N and M, page 61
$\boldsymbol{\alpha}$ markov chain, page 4
$\mathbf{b}$ mean rate of return process, page 5
$\mathcal{E}$ Doléans-Dade exponential, page 65
$\mathcal{F}_t$ standard filtration generated by $\mathbf{W}$ and $\boldsymbol{\alpha}$ with the $\mathbb{P}$-null sets. 4
$\mathcal{F}_t^{\boldsymbol{\alpha}}$ standard filtration generated by $\boldsymbol{\alpha}$ with the $\mathbb{P}$-null sets. 31
$\mathcal{FV}$ the set of processes of finite variation, page 60
$\mathcal{FV}_0$ the set of processes of finite variation that are null at the origin, page 60
$\mathcal{FV}^+$ the set of processes that are non-decreasing, page 60
$\mathcal{FV}_0^+$ the set of processes that are non-decreasing and null at the origin, page 60
H State price density, page 28
$\mathcal{IV}$ the set of processes in $\mathcal{FV}$ of integrable variation, page 60
$\mathcal{IV}_0$ the set of processes in $\mathcal{FV}_0$ of integrable variation, page 60
$\mathcal{IV}^+$ the set of processes in $\mathcal{FV}^+$ which are integrable, page 60
$\mathcal{IV}_0^+$ the set of processes in $\mathcal{FV}^+$ which are null at the origin, page60
$\mathcal{IV}_{\text{loc}}$ processes of locally integrable variation, 66
$\mathcal{IV}_{0,\text{loc}}^+$ the set of processes in $\mathcal{FV}_0^+$ which are of locally integrable variation, 66
$\mathcal{M}$ the set of martingales, page 57
$\mathcal{M}_0$ the set of martingales null at the origin, page 58
$\mathcal{M}^c$ the set of continuous martingales, page 58
$\mathcal{M}_0^c$ the set of continuous martingales null at the origin, page 58
$\mathcal{M}^2$ the set of square-integrable martingales, page 58
$\mathcal{M}_0^2$ the set of square-integrable martingales null at the origin, page 58
$T^m \Uparrow T$ increasing stopping times, page 58
$\mathcal{M}_{\text{loc}}$ the set of local martingales, page 59
$\mathcal{M}_{0,\text{loc}}$ the set of local martingales null at the origin, page 59
$\mathcal{M}_{\text{loc}}^c$ the set of continuous local martingales, page 59
$\mathcal{M}_{0,\text{loc}}^c$ the set of continuous local martingales null at the origin, page 59
$\mathcal{M}_{\text{loc}}^2$ the set of locally square-integrable martingales, 59
M_{ij} the canonical martingale associated with Markov chain $\boldsymbol{\alpha}$, page 16
$\nu_{[M_{ij}]}$ Doléans measure of M_{ij}, page 18
$\mathcal{P}^\star$ the predictable σ-algebra on $\Omega \times [0, T]$, page 7
$\mathcal{F}^\star$ the progressively measurable σ-algebra on $\Omega \times [0, T]$, page 7
$\boldsymbol{\pi}$ portfolio process, page 7
Π portfolio process space, page 7
g_{ij} the $(i, j)^{th}$ entry of the generator of the Markov chain, page 4
$\boldsymbol{\theta}$ the market price of risk, page 7
$[\boldsymbol{\sigma}]$ the volatility process, page 5
g the generator of the Markov chain, page 4
$\mathbf{W}$ a N-dimensional Brownian Motion, page 4
$X^{\boldsymbol{\pi}}$ solution to the wealth equation for $\boldsymbol{\pi}$, page 7

References

[1] J.M. Bismut *Conjugate convex functions in optimal stochastic control*, Journal of Math. Analysis and Applications, pp.384-404, v.44, 1973.

[2] D. Cuoco and H. Liu, A martingale characterization of consumption choices and hedging costs with margin requirements, *Mathematical Finance*, pp.355–385, v.10, (2000).

[3] J. Cvitanić and I. Karatzas, Convex duality in constrained portfolio optimization, *Annals Appl. Probability*, pp.767–818, v.2, (1992).

[4] C. Donnelly. *Convex Duality in Constrained Mean-Variance Portfolio Optimization Under a Regime-Switching Model*, PhD thesis, University of Waterloo, Canada, 2008.

[5] C. Donnelly and A.J. Heunis, Quadratic risk minimization in a regime-switching model with portfolio constraints, *SIAM Journal on Control and Optimization*, pp.2431–2461, v.50, (2012).

[6] R. J. Elliott. *Stochastic Calculus and Applications*, Springer, New York, 1982.

[7] J. Jacod and A. N. Shiryaev. *Limit Theorems for Stochastic Processes*, Springer, Berlin, 1987.

[8] O. Kallenberg, *Foundations of Modern Probability, 2nd Ed.*, Springer, (2002).

[9] I. Karatzas, J.P. Lehoczky, S.E. Shreve, and G.L. Xu, Martingale and duality methods for utility maximization in an incomplete market, *SIAM J. Control and Optimization*, pp.702–730, v.29, (1991).

[10] I. Karatzas and S.E. Shreve. *Methods of Mathematical Finance*, Springer, New York, 1998.

[11] C. Labbé and A.J. Heunis, Convex duality in constrained mean-variance portfolio optimization, *Adv. Appl. Probab.*, pp.77–104, v.39, (2007).

[12] R. Sh. Liptser and A.N. Shiryayev. *Theory of Martingales*, Kluwer Academic, Dordrecht, 1989.

[13] P. E. Protter. *Stochastic Integration and Differential Equations, 2nd Ed.*, Springer, New York, 2005.

[14] P. Ramchandani. *Quadratic loss minimization in a regime switching model with control and state constraints*, PhD thesis, University of Waterloo, Canada, 2015.

[15] L.C.G. Rogers and D. Williams. *Diffusions, Markov Processes and Martingales, Volume II Itô Calculus*, Cambridge University Press, 2000.

[16] L. R. Sotomayor and A. Cadenillas, Explicit solutions of consumption-investment problems in financial markets with regime switching, *Mathematical Finance*, pp.251–279, v.19, (2009).

[17] G.L. Xu and S.E. Shreve. *A Duality Method for Optimal Consumption and Investment Under Short-Selling Prohibition. I. General Market Coefficients*, Annals of Applied Probability, pp.87-112, v.2, 1992.

[18] 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*, pp.1466–1482, v.42, (2003).