Quadratic Loss Minimization in a Regime Switching Model with Control and State Constraints

by

Pradeep Ramchandani

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

Waterloo, Ontario, Canada, 2015

Author's Declaration

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

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

Abstract

In this thesis, we address a convex stochastic optimal control problem in mathematical finance, with the goal of minimizing a general quadratic loss function of the wealth at close of trade. We study this problem in the setting of an Ito process market model, in which the underlying filtration to which the market parameters are adapted is the joint filtration of the driving Brownian motion for the market model, together with the filtration of an independent finite-state Markov chain which models occasional changes in "regime states", that is our model allows for "regime switching" among a finite number of regime states. Other aspects of the problem that we address in this thesis are:

(1) The portfolio vector of holdings in the risky assets is confined to a given closed and convex constraint set;

(2) There is a "state constraint" in the form of a stipulated almost-sure lower bound on the wealth at close of trade.

The combination of constraints represented by (1) and (2) makes the optimization problem quite challenging. The powerful and effective method of *auxiliary markets*, of Cvitanic and Karatzas [Ann. Appl. Prob., v.2, 767-818, 1992] for dealing with convex portfolio constraints, does not appear to extend to problems with regime-switching, while the more recent approach of Donnelly and Heunis [SIAM Jour. Control Optimiz., v.50, 2431-2461, 2012], which deals with both regime-switching and the convex portfolio constraints (1), is nevertheless confounded when one adds state constraints of the form (2) to the problem. The reason for this is clear: state constraints of the form (2) typically involve "singular" Lagrange multipliers which fall well outside the scope of the "well-behaved" Lagrange multipliers, manifested either as random variables or stochastic processes, which suffice when one is dealing only with portfolio constraints such as (1) above. In these circumstances we resort to an "abstract" duality approach of Rockafellar and Moreau, which has been applied with considerable success to finite-dimensional problems of stochastic mathematical programming in which singular Lagrange multipliers also naturally arise. The main goal of this thesis is to adapt and extend the Rockafellar-Moreau approach to the stochastic optimal control problem summarized above. We find that this is indeed possible, although some considerable effort is required in view of the infinite dimensionality of the problem. We construct an appropriate space of Lagrange multipliers, synthesize a dual optimization problem, establish optimality relations which give necessary and sufficient conditions for the given optimization problem and its dual to each have a solution with zero duality gap, and use the optimality relations to synthesize an optimal portfolio in terms of the Lagrange multipliers.

Acknowledgements

First and foremost, I would like to thank my supervisor, Andrew Heunis for his continued patient guidance and assistance throughout my degree. Due to his vast knowledge of mathematics and emphasis on giving precise mathematical proofs, I have learnt a lot from all the weekly meetings.

Secondly, I would like to thank my committee members, Patrick Mitran, Shojaeddin Chenouri, Takis Konstantopoulos and Ravi Mazumdar for their time. I would also like to thank Kalikinkar, Rajat, Kaushik and Sharad for their friendship and support.

Contents

Chapter 1

Introduction, Background and Motivation

1.1 Introduction

In this thesis we study a stochastic optimal control problem which arises from the allocation of an agent's wealth among a variety of assets according to a widely used notion of "quadratic loss optimality". The goal is to continuously trade in a designated set of risky assets (i.e. shares) and a single bond (effectively a money-market account) in such a way as to minimize the expected value of a general *quadratic loss function* of the wealth at close of trade. By appropriately specializing the quadratic loss function one can model genuine trading scenarios such as quadratic hedging of contingent claims and mean-variance hedging. For example, in quadratic hedging of a contingent claim, an agent is responsible for "paying out" at some future fixed date T a sum of money (i.e. the contingent claim) represented by a square-integrable random variable γ which is "measurable" with respect to all the trading information available in the market up to the closing date T. The agent must then use the increasing information available in the market as time evolves to trade among the available shares and money-market account over the trading interval $0 \leq t \leq T$, in order to generate a wealth-stream $\{X(t),\ 0 \leq t \leq T\}$ in such a way that the *mean-square discrepancy*

$$E|X(T) - \gamma|^2 \tag{1.1}$$

between the contingent claim γ for which the agent is responsible, and the actual wealth $X(T)$ which the agent will have generated by the date T in order to meet this obligation, is *minimized*. If one regards the allocation of wealth among shares over the interval $0 \leq t \leq T$ as a "control" and the wealth-process $\{X(t),\ 0 \leq t \leq T\}$ which is being thus "controlled" as the "state", then we essentially have a stochastic optimal control problem the goal of which is to minimize the quantity $E|X(T)-\gamma|^2$. Such trading scenarios based on quadratic minimization were first envisaged (for a simple "one-period") model in the Nobel prize-winning work of Markowitz [28], and are now quite widely used in practice.

To further enhance the applicability of the stochastic optimal control model, one can postulate convex *portfolio constraints*, that is insist, as part of the definition of the optimization problem, that the vector of allocations of wealth in various stocks always takes values in a designated convex *constraint set*. For example, a regulatory body might specify that traders not short-sell a designated set of stocks, and this can be represented by means of a convex *portfolio constraint* (simple examples of this are illustrated later in the thesis). We should mention that portfolio constraints are effectively control constraints in a stochastic control setting, and as such constitute a definite challenge. The definitive work on portfolio constraints (for utility maximization rather than quadratic minimization) is that of Cvitanic and Karatzas [5]. As we shall see in the discussion that follows, the particular aspects of the problem we intend to address unfortunately rules out the elegant methodology of [5], and we shall have to establish a substitute framework for dealing with portfolio constraints.

Despite the widespread use of the quadratic loss criterion, there is nevertheless one distinct drawback associated with this criterion, namely even when the discrepancy $E|X(T) - \gamma|^2$ is minimized one can still have $P\{X(T) < 0\} > 0$, that is strictly negative wealth (or bankruptcy) with positive probability at close of trade. Obviously this is a very undesirable outcome from the point of view of the agent, and in practice this problem is typically taken care of by a variety of heuristic and *ad-hoc* devices, none of which are completely satisfactory. In fact, the only honest way of dealing with this drawback is to include as a specific *constraint* in the quadratic minimization problem the requirement that $X(T) \geq 0$ a.s., so that the goal of the stochastic control problem is now to minimize the quadratic loss function subject to this additional constraint, together with the convex portfolio constraint of the previous paragraph. A major challenge in implementing this idea is that the constraint $X(T) \geq 0$ a.s. is an almost-sure "state constraint", which must be dealt with in addition to the afore-mentioned portfolio constraint or "control constraint", so that we are effectively dealing with a stochastic control problem which exhibits both a *control constraint* and an a.s. *state constraint*. We indicated earlier that control constraints by themselves constitute a significant challenge. It is well known that this challenge is very substantially compounded by the addition of a state constraint, even in the case of *deterministic* optimal control, and of course these challenges become even greater when one is dealing with a *stochastic* optimal control problem. Indeed, the archetypical works on *deterministic* optimal control problems with state and control constraints (i.e. without any randomness) by Dubovitskii and Mil'yutin [8], Rockafellar [35] and Makowski and Neustadt [27] make abundantly clear the challenges posed by a combination of control and state constraints. Among the most serious of these challenges is that the "Lagrange multiplier" which enforces the combination of control and state constraints often turns out to be "strange" or "degenerate" or "singular" in some sense (such as being only a *finitely additive* measure instead of a *countably additive* measure). This is in direct contrast to the case of problems which only have control constraints and no state constraints; these are certainly challenging as we have already noted, calling upon Lagrange multipliers which are "infinite dimensional" and which can be quite exotic (such as countably additive measures or stochastic processes), but at least these Lagrange multipliers are regular and well behaved mathematical objects without the degeneracies which are all too characteristic of the multipliers which arise when one adds in a state constraint as well. The fact that the enforcement of a combination of control and state constraints calls upon such singular Lagrange multipliers is really a reflection of just how challenging such combined constraints actually are. In the case of *stochastic* optimal control problems with control and almost-sure state constraints the challenges already inherent in the simpler deterministic case are of course considerably multiplied by the stochastic character of the problem, and in fact there are few if any general results on such problems.

Remark 1.1.1 We conclude this introductory discussion with the following remark: since our goal is to minimize the quantity (1.1) would it not be preferable simply to use the standard Black-Scholes approach to *replicate* the contingent claim γ in the sense of determining a trading strategy such that

$$X(T) = \gamma, \tag{1.2}$$

for in this case we certainly minimize the quantity at (1.1), and we seem to avoid the rather thorny issues indicated above as well. We shall see in Chapter 2 that the conditions of the problem addressed in this thesis do not, in any way, match the conditions under which one can apply a Black-Scholes approach and that effectively we *must* use the stochastic control approach that has been outlined in the preceding discussion.

1.2 Background and Motivation

Despite the difficulties and challenges noted in Section 1.1, there is a pioneering work of Bielecki, Pliska, Jin and Zhou [2] which addresses a special case of the trading problem of quadratic minimization in the previous section with the state constraint $X(T) \geq 0$ a.s. included but with *unconstrained* portfolios. The approach of [2] exploits the fact that, for certain models of the share-price dynamics, the wealth process $\{X(t),\ 0 \leq t \leq T\}$ is determined by a stochastic differential equation which is particularly simple, so simple in fact that the Lagrange multipliers for the state constraint turn out to be scalar variables which can be determined by application of the theorem of separating hyperplanes in finite-dimensional space. In particular, with the simple market dynamics adopted in [2], the Lagrange multipliers do not exhibit any of the singularities seen in Dubovitskii and Mil'yutin [8] and Makowski and Neustadt [27], which address rather general (although deterministic) system dynamics. We do not intend to enter into a detailed discussion of the approach of [2], and wish only to emphasize that the portfolios in that work are assumed to be *unconstrained* and the share-price dynamics are modeled by classical Itô stochastic differential equations driven by an underlying Brownian motion which effectively models the continuous microscopic random effects at play in the market. The market parameters (that is, the prevailing interest rate process, and the processes which model the mean return-rate and volatility of the shares) are assumed to be random and adapted to the filtration of the underlying Brownian motion. The crucial thing in this market model is that the posited underlying Brownian motion is the *one and only* source of randomness in the model, and since the portfolios are assumed to be unconstrained this means that one can hedge completely against this randomness (in the parlance of mathematical finance theory the market model adopted in [2] is *complete*). It is really this property of the market model in [2] which leads to the very simple structure of the Lagrange multipliers for the state constraint $X(T) \geq 0$ a.s.

Despite its simplicity, the market model adopted by Bielecki, Pliska, Jin and Zhou [2] is of definite interest when markets are functioning under "normal" or "regular" conditions and there are no externally mandated portfolio constraints. However, as we have already noted, it may well be the case that there are externally imposed portfolio constraints; in the parlance of mathematical finance portfolio constraints render the market model *incomplete*, and this effectively rules out the approach of [2] which does rely crucially on completeness of the market model arising from the absence of portfolio constraints. Furthermore, empirical evidence suggests that a Brownian motion alone does not provide sufficient randomness to model all aspects of observed market behavior. In particular, markets are observed to make the occasional *discontinuous* and random change in "regime state" (e.g. from a "bullish" to a "bearish" state), and these random changes cannot be modeled by a Brownian motion alone, since Brownian motion is really best suited to modeling cumulative changes resulting from highly persistent "small-scale" *continuous* random perturbations. Despite the fact that the discontinuous changes in regime state are only rather occasional they are nevertheless large-scale changes with an important impact on the market, so it is important to attempt to model these changes accurately. One of the most successful models for such changes is a *finite-state Markov chain*, and models in which the basic source of randomness is not just a Brownian motion but is instead a Brownian motion *together with* a finite-state Markov chain are known as *regime switching market models*. Thus, in these models the Brownian motion effectively accounts for persistent small-scale continuous perturbations whereas the finite-state Markov chain accounts for occasional large-scale discontinuous perturbations. Such regime switching market models are now being increasingly encountered in practice and are seen to be considerably more realistic than the simpler models in which all randomness is due only to a Brownian motion, such as the market model adopted in [2]. On the other hand, the theoretical analysis of regime switching models is considerably more challenging than is the case for the simpler complete market model of [2]. We do not intend to discuss in detail the rather technical reasons for this, and note only that when randomness is modeled by the combination of a Brownian motion and a finite-state Markov chain then (much as for portfolio constraints) the resulting market model is *incomplete*, and this again rules out generalizing the approach of Bielecki *et-al*, which relies on a complete market model. Despite the challenges presented by market models which include regime switching, the many advantages of this model make it a very worthwhile object of study, and it has recently received significant attention in the literature on mathematical finance. Particularly significant are the following works:

(I) Zhou and Yin [47]: This addresses the problem of minimizing a general quadratic loss function of the wealth at close of trade in a market model incorporating a simplified form of regime switching, in which the market parameters at every instant are determined *entirely* by the regime switch Markov chain at that same instant (the market parameters are then said to be *Markov-modulated*), and there are *no constraints* either on the portfolio (so that the market model is incomplete by virtue of regime switching only) or on the wealth at close of trade i.e. there is no constraint on the wealth at close of trade analogous to $X(T) \geq 0$. This absence of any constraints, together with the simple structure of the Markov-modulated market model, makes it possible to proceed by direct analysis of the primal problem through an elegant and simple, but nevertheless very problem-specific, completion-of-squares approach. In particular, this approach does not extend to problems which involve either portfolio or state constraints, nor to problems in which the market parameters are not just Markov-modulated but depend non-trivially on both the driving Brownian motion and the finite-state Markov chain.

(II) Sotomayor and Cadenillas [42]: This addresses the problem of *utility maximization* (as opposed to the quadratic minimization addressed by [47]) but in much the same setting as [47], that is regime switching is included in the market model in the form of simple Markov-modulated market parameters, the portfolio is unconstrained, and there is no a.s. state constraint on the wealth at close of trade. Since the objective function involves (non-quadratic) utility maximization, it is not possible to proceed by the completion-of-squares approach of [47]. Instead, Sotomayor and Cadenillas [42] exploit the special features of their problem (absence of any constraints, Markov-modulated market parameters) to analyze the primal problem directly by means of dynamic programming and the Hamilton-Jacobi equation.

(III) Donnelly and Heunis [7]: The work [7] generalizes the quadratic minimization problem of Zhou and Yin [47] in two ways, namely it allows for convex constraints on the portfolio, and the market parameters are determined not just by the instantaneous value of the regime switch Markov chain (as in the Markov-modulated case), but instead depend *non-anticipatively* on both the regime switch Markov chain *and* the Brownian motion. This is a much more general market model than the simple Markov-modulated models of [47] and [42], and this fact, together with the portfolio constraints, completely rules out application of either the completion-of-squares-approach of [47] or the dynamic programming approach of [42]. A possible natural approach to this problem would involve extending the *auxiliary market model* approach of Cvitanic and Karatzas [5] from the domain of non-regime switching models in which the only source of randomness is a Brownian motion (for which the approach of [5] works extremely well) to the domain of regime switching models in which the underlying source of randomness is both a finite state Markov chain and a Brownian motion. Unfortunately it is far from clear how to accomplish this extension. Instead, in [7], a conjugate duality framework complementary to that of Cvitanic and Karatzas [5], but capable of dealing with regime switching models, is established. This framework actually exploits an extraordinarily flexible stochastic calculus of variations problem introduced in a very classic work of Bismut [3], and furnishes the means to construct an optimal portfolio in terms of the solution of an associated *dual problem* by means of *Kuhn-Tucker* optimality relations. In particular, it is seen in [7] that the Lagrange multiplier which enforces the portfolio constraint is a square-integrable Ito process driven by the *joint filtration* of the underlying Brownian motion and finite state Markov chain.

With the preceding discussion in mind, we can finally formulate, at least in general non-quantitative terms, the main goal of this thesis. We intend to address a quadratic minimization problem of the general kind studied by Bielecki, Pliska, Jin and Zhou [2], Zhou and Yin [47], and Donnelly *et-al* [7], with market parameters which depend non-anticipatively on both a driving regime switch Markov chain and Brownian motion (exactly as in [7]), with convex portfolio constraints (again exactly as in [7]), together with the a.s. state constraint $X(T) \geq 0$ a.s. This problem is clearly substantially more general than that addressed in [47], since it includes a general market model together with both portfolio constraints and a.s. state constraints (none of which are present in [47]). In addition, our problem significantly generalizes the problem addressed by Bielecki, Pliska, Jin and Zhou [2], since it includes both portfolio constraints and regime switching, two elements which are not present in [2] and which rule out application of the elegant (but problem-specific) separating hyperplanes approach of [2]. On the other hand, the problem that we intend to study looks rather close to the problem addressed by Donnelly *el-al* [7], in fact being different from this problem only in that it also includes the state constraint $X(T) \geq 0$ a.s., which is not present in [7]. One might therefore imagine that the conjugate duality approach of [7] could somehow be "fine-tuned" or adapted or extended to work for the problem addressed here. Unfortunately, the marvelous flexibility of the calculus-of-variations formulation of Bismut [3], which is central to [7], does not go quite far enough to allow this, for it is predicated on a *complete absence* of any state constraints whatsoever. In fact, as we shall see in later chapters, addition of the a.s. state constraint significantly changes the entire structure of the Lagrange multiplier which now becomes a *pair*, comprising a square-integrable Ito process (to enforce the portfolio constraint, exactly as in [7]) paired with a finitely additive measure (to enforce the a.s. state constraint). This "compound" Lagrange multiplier unfortunately falls outside the scope of the elegant and powerful duality theory of Bismut [3], which *a-priori* posits Lagrange multipliers that are Ito processes only, and this means that the approach of [7], based as it is on the duality theory of Bismut [3], again does not extend to the problem of this thesis.

All of this discussion leads inevitably to the conclusion that stochastic optimal control with portfolio constraints *together* with state constraints in a regime switching market model presents its own quite special challenges which demand an approach very different from that in any of the prior works [2], [7], and [47]. In particular, with this model the Lagrange multipliers which enforce the state and portfolio constraints are no longer simple scalar variables (as is the case in [2]) or square integrable Ito processes (as is the case in [7]) but are "singular" in the sense of also involving finitely additive measures. That is, we essentially encounter the same sort of singular Lagrange multipliers that are found in the works of Dubovitskii and Mil'yutin [8] and Makowski and Neustadt [27]. These nonstandard multipliers, together with the effects of regime switching, feature largely in the challenges posed by this problem.

1.3 General Approach

Although the problem we intend to study does present definite challenges as outlined in the previous section, it nevertheless has one very nice feature that will be the cornerstone of our entire approach, namely it is a *convex* optimization problem. The importance of this fact cannot be over-stated, for it will allow us to use the general method of *convex duality.* The essence of this method is to associate with the given convex optimization problem (usually called the *primal problem*) an associated *dual optimization problem.* It is usually the case that a dual problem is much better behaved than the primal problem, and in fact one can readily set up conditions which ensure existence of a solution of the dual problem, whereas it is often very difficult to do the same thing for the given primal problem. Solutions of the dual problem are very important, for these are just the Lagrange multipliers which enforce the constraints in the primal problem, and one can typically *construct* a solution of the given primal problem in terms of these Lagrange multipliers. However, a particular challenge in implementing the method of convex duality, particularly for the sort of *dynamic problems* encountered in optimal control (either deterministic or stochastic), is that the structure of an appropriate dual problem is usually far from clear *a-priori.* Indeed, for the sort of stochastic control problems encountered in mathematical finance it is usually the case that the dual problem is arrived at by a lengthy process of guesswork and trial and error, and then subsequently verified to work (this is not unlike trying to guess the solution of a complicated differential equation and then verifying by substitution that the guessed solution actually works). For the stochastic optimal control problem that we address here it is very difficult to follow this "trial and error" approach, since the structure of appropriate *dual variables* (i.e. the space of variables over which the dual functional is defined) is itself not even *a-priori* clear when one has a combination of almost-sure state constraints and portfolio constraints. Moreover, it must be said that resort to a "trial and error" approach is also not very satisfying! For abstract problems of convex optimization a very flexible and powerful approach for constructing appropriate dual variables and dual problems (without any guesswork) was worked out as long ago as the 1970's by Rockafellar and Moreau (this approach is comprehensively summarized in the monographs of Rockafellar [36] and Ekeland and Témam [9]). In fact, regarding the Rockafellar-Moreau approach, Ekeland and Témam state (see p.xii in [9]) "This very flexible abstract theory can be adapted to a wide variety of situations", and then demonstrate this flexibility in [9] on a number of challenging problems concerned with the calculus of variations of partial differential equations. To the best of our knowledge there are, however, almost no applications of the Rockafellar-Moreau approach to problems of stochastic control which exhibit the combination of constraints outlined above as well as regime switching in the market model. One exception is the recent work [14] on a simplified version of the problem of this thesis, which includes portfolio constraints as well as an almost-sure inequality constraint on the wealth at close of trade, but which does not include regime switching, the addition of which makes the problem of this thesis significantly more challenging than the problem addressed in [14]. In this thesis we are going to see that the basic idea of Rockafellar and Moreau is so flexible that it can actually be extended (albeit with significant effort) to our stochastic optimal control problem which features state constraints, control constraints and regime switching in the dynamics, and in particular provides the means for *synthetically constructing* the appropriate vector space of dual variables, dual functional, and optimality relations, without any trial and error guesswork. Although it will not be demonstrated in this thesis, the resulting approach is also powerful enough to constitute a *unified method* for obtaining all of the prior results of [2], [7], and [47].

In addition to the Rockafellar-Moreau approach noted above, this thesis is also indebted to a work of Rockafellar and Wets [37] which addresses *static* problems of stochastic convex optimization. This is in contrast to the present thesis which addresses a convex stochastic control problem i.e. a *dynamic* problem of stochastic convex optimization. Although dynamic optimization problems are considerably more challenging than static problems, it is nevertheless the case that some of the insights from [37] carry over directly to this thesis. In particular, Rockafellar and Wets were the first to understand that almost-sure inequality constraints unavoidably involve Lagrange multipliers which are not "normal" or "regular" (in the sense of being constants, or random variables or stochastic processes) but are "singular" or "pathological" objects (in much the same way that the well-known impulse function or Dirac delta-function of system theory is singular or pathological). We shall find that the almost-sure inequality constraints in the dynamic problem of this thesis similarly demand singular Lagrange multipliers which are precise analogs of the singular Lagrange multipliers occurring in the static problems of Rockafellar and Wets [37].

1.4 Organization of the Thesis

In Chapter 2, we define the market model and quadratic loss minimization (QLM) problem that is the main focus of this work in precise mathematical terms. We also reformulate this problem into an abstract form to which the Rockafellar Moreau approach mentioned in Section 1.3 is suited. Chapter 3 presents a self-contained summary of the Rockafellar-Moreau approach, which is the main technical tool of this whole work, and Chapter 4 is concerned with the solution of the QLM problem using the Rockafellar-Moreau approach. Following Chapter 5 are several appendices. Proofs of several technical results occurring in the main body of the thesis are relegated to Appendix A 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 proofs in Appendix A in detail. The remaining Appendices B - 5 for the most part do not contain any new work as such, and have been included only to provide some of the technical background necessary to read this thesis. We suggest that readers consult these appendices for reference only (when needed) rather than read the appendices in their entirety.

Chapter 2

Market Model and Quadratic Loss Minimization

The goals of this chapter are twofold. We first precisely define the main mathematical elements of the *regime switching* market model that will be assumed throughout. This model is quite widely used and is certainly not novel to this thesis, but we give the model in complete detail in order for the thesis to be readable. What is novel is the quadratic loss minimization (QLM) problem in a market model with regime-switching and including a combination of portfolio constraints and state constraints that we also formulate later in this chapter.

2.1 The Market Model

As we noted previously our regime switching market model is quite standard and widely used. Here we borrow in its entirety the presentation due to Donnelly [6] who gives a particularly clear and complete formulation of this model.

The market consists of a single bank account (or money market account) and a number of stocks in which an agent is allowed to trade. We assume that market is subject to regime-switches from time to time. For example the market could be in a "bullish" phase, with stock prices generally rising, and this phase can be considered a regime. Suddenly there is a stock market crash, and the market enters a "bearish" phase, in which stock prices are mostly falling. The "bearish" phase is another regime, so this is an example of a regime-switch; the market makes a random switch from a "bullish" regime to a "bearish" regime. Of course, the "bearish" regime will eventually give way to another "bullish" regime. In this case we have just two regime states ("bullish" and "bearish") and the switching between these states is modeled by a two-state Markov chain. More generally, as will be seen later in this section, we are going to allow for D regime states (D being a fixed finite integer).

We make the assumption that each time t in the trading interval $[0, T]$, an investor in the market will know everything that has occurred up to time t. This assumption is expressed mathematically using a *filtration*. This is a structure which contains all the events which could have occurred up to each time t. We construct a filtration from a Brownian motion, which drives the stock prices, and a Markov chain, which models the regime-switching.

Before defining the market model itself, it is necessary to define the regime-switching Markov chain and the Brownian motion which provide the source of randomness in the market model. We also need to specify the probability space on which all processes in the market model are defined. Once we have this probability space, we can define the basic filtration which contains the information available to an investor. We can then use the filtration to define a measurability property required of the stochastic integrands and portfolio investment processes.

Condition 2.1.1 All investment activity takes place over a finite time interval $[0, T]$, where $0 < T < \infty$ is non-random and fixed in advance.

The probability space

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a complete probability space. By complete we mean that all the $\mathbb{P}$-negligible subsets of Ω are $\mathcal{F}$-measurable.

Modeling the regime-switching

Regime switching is modeled using a continuous time Markov chain α defined on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$. We assume that there are only finitely many possible regime states $I = \{1, 2, \ldots, D\}$, so the Markov chain takes values in the finite state space I. We assume that the regime state Markov chain $\{\alpha(t) \,|\, t \in [0, T]\}$ is time homogeneous in the sense that

$$\begin{aligned} p_{ij}(t) &:= P(\alpha(t) = j \mid \alpha(s) = i) \quad 0 \le s \le t \le T,\ i, j = 1, 2, \ldots, D \qquad (2.1) \\ &= P(\alpha(t - s) = j \mid \alpha(0) = i). \qquad (2.2) \end{aligned}$$

We denote by Q the generator of the Markov chain α, that is Q is the $D \times D$ matrix $Q = (q_{ij})_{i,j=1}^{D}$, with the following properties.

$$q_{ij} \geq 0 \quad \forall\, i \neq j \quad \text{and} \quad -\, q_{ii} = \sum_{j \neq i} q_{ij}, \tag{2.3}$$

$$\text{and} \qquad p_{ij}(t) = (e^{Qt})_{ij}, \quad t \in [0, \infty), \quad i, j = 1, 2, \ldots, D. \tag{2.4}$$

Appendix (G.1) is a compendium of useful standard background on finite state Markov chains, included for easy reference in the thesis.

Brownian Motion for the Market Model

We assume that the stock prices are driven by a standard, N-dimensional Brownian motion $\mathbf{W} \equiv \{\mathbf{W}(t) = (W_1(t), W_2(t), \ldots, W_N(t))^\top \,|\, t \in [0, T]\}$, defined on the complete probability space $(\Omega, \mathcal{F}, \mathbb{P})$. By a standard, N-dimensional Brownian motion, we mean the following.

Definition 2.1.2 A standard, N-dimensional Brownian motion in an $\mathbb{R}^N$-valued process $\mathbf{W} \equiv \{\mathbf{W}(t) = (W_1(t), \ldots, W_N(t))^\top | \, t \in [0, T]\}$ such that

1. $\mathbf{W} \equiv \{\mathbf{W}(t) \,|\, t \in [0, T]\}$ is null at the origin;
2. the sample paths $t \to \mathbf{W}(\omega, t)$ are continuous for each $\omega \in \Omega$; and
3. For each $s < t$ such that $s, t \in [0, \infty)$, the $\mathbb{R}^N$-valued increment $\mathbf{W}(t) - \mathbf{W}(s)$ is distributed according to $N(0, (t-s)I_N)$ and is independent of the filtration $\mathcal{F}_s^{\circ, \mathbf{W}} := \sigma\{\mathbf{W}(u) : \; u \in [0, s]\}$, where I_N is the $N \times N$ identity matrix.

Independence Assumption

The following independence assumption is very natural from a modeling viewpoint. The idea is that the Brownian motion models the movement of the prices of individual stocks due to micro-economic effects which occur over very short time periods, and the Markov chain models the movement of the prices due to occasional macro-economic effects which occur over much longer time periods.

Condition 2.1.3 The Brownian motion process $\{\mathbf{W}(t) \,|\, t \in [0, T]\}$ is independent of the Markov Chain process $\{\alpha(t) \,|\, t \in [0, T]\}$, in the sense that

$$\mathbb{P}[A \cap B] = \mathbb{P}[A]\mathbb{P}[B], \quad \forall A \in \mathcal{F}_T^{\circ, \alpha} \quad \forall B \in \mathcal{F}_T^{\circ, W}. \tag{2.5}$$

Here
$$\mathcal{F}_T^{\circ, \alpha} \;:=\; \sigma\{\alpha(t) : \mathrm{t} \in [0, T]\} \qquad \text{and} \tag{2.6}$$
$$\mathcal{F}_T^{\circ, \mathbf{W}} \;:=\; \sigma\{\mathbf{W}(t) : \mathrm{t} \in [0, T]\}. \tag{2.7}$$

Generating the filtration and defining predictability

Having defined the Brownian motion $\mathbf{W}$ and the regime-switching Markov chain α on the common probability space $(\Omega, \mathcal{F}, \mathbb{P})$, we next construct the filtration with respect to which we shall define the problem addressed in this thesis.

The raw filtration $\{\mathcal{F}_t^\circ : t \in [0,T]\}$ generated by $\mathbf{W}$ and α is defined in the standard way as

$$\mathcal{F}_t^\circ := \sigma\{\mathbf{W}(s), \alpha(s) : s \in [0,t]\} \quad \forall t \in [0,T]. \tag{2.8}$$

In order to use the results of stochastic calculus we want a filtration which has the usual regularity properties, namely it must contain all the $\mathbb{P}$-null sets in the σ-algebra $\mathcal{F}$ and be right-continuous. The filtration at (2.8) does not unfortunately have these properties. Nevertheless, it is well known that the Brownian motion $\mathbf{W}$ is a *Feller process* with state space $\mathbb{R}^N$, and, likewise, the finite state Markov chain α is also a Feller process with finite state space I (see III(2.22) of Revuz and Yor [31]). From Condition 2.1.3 the Feller processes α and $\mathbf{W}$ are *independent*, and therefore the *joint* process $(\mathbf{W}, \alpha)$, with state space $\mathbb{R}^N \times I$ is again a Feller process (as follows from Kallenberg [19], Ex. 10, Chap.19, p.389). Now it follows from III(2.10) of Revuz and Yor [31] that the filtration

$$\mathcal{F}_t := \mathcal{F}_t^\circ \vee \mathcal{N}(\mathbb{P}), \quad \forall t \in [0,T], \tag{2.9}$$

is right-continuous and includes the $\mathbb{P}$-null events in $\mathcal{F}$ as required.

Remark 2.1.4 It is essential to note the role of the Markov chain α in determining that structure of the raw filtration at (2.8), and therefore of the filtration at (2.9), for the regime switching in the market model formulated in this section resides completely in the presence of the Markov chain α at (2.8). The significance of the filtration at (2.8)-(2.9) is that it represents the *total information* available to the investor, that is at every instant $t \in [0,T]$ the investor knows the paths of the processes $\{\mathbf{W}(s), s \in [0,t]\}$ and $\{\alpha(s), s \in [0,t]\}$. This is the information used by the investor in formulating the portfolio process which should therefore be adapted to (more precisely predictable with respect to) the filtration $\{\mathcal{F}_t,\ t \in [0,T]\}$ (see Definition 2.2.1 which follows). In market models which do *not* account for regime switching one uses in place of (2.8) - (2.9) a filtration which depends *only* on the driving Brownian motion $\mathbf{W}$, that is the filtration defined by

$$\mathcal{F}_t^{\mathbf{W}} := \sigma\{\mathbf{W}(s) : s \in [0,t]\} \vee \mathcal{N}(\mathbb{P}), \quad \forall t \in [0,T]. \tag{2.10}$$

Although the role of the regime switch Markov chain α at (2.8) may look innocuous, it nevertheless does make the formulation of optimal portfolios significantly more challenging than in the case of the simpler filtration at (2.10).

Remark 2.1.5 In the terminology of Definition B.2.3 we therefore see that the pair $((\Omega, \mathcal{F}, \mathbb{P}), \{\mathcal{F}_t\})$, for the filtration at (2.9), is a *standard filtered probability space*. Consequently, we can use the tools of stochastic calculus for integration against general (i.e. discontinuous) semimartingales relative to the filtration $\{\mathcal{F}_t\}$.

We use throughout the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and we use the qualifier almost surely ("a.s.").

To ensure that the stochastic integrals we deal with are properly defined, it is necessary that the stochastic integrands have a type of measurability called predictable measurability (or simply predictability). We define this next:

Definition 2.1.6 Let $\mathcal{P}^*$ denote the minimal σ-algebra on $\Omega \times [0, T]$ generated by the set of all continuous $\mathbb{R}$-valued $\{\mathcal{F}_t\}$-adapted processes. Then a process $\{X(t),\ t \in [0, T]\}$ on $(\Omega, \mathcal{F}, \mathbb{P})$ is said to be $\mathcal{F}_t$-predictably measurable when it is $\mathcal{P}^*$-measurable when regarded as a mapping on $\Omega \times [0, T]$.

The canonical martingales of the Markov chain

Associated with the Markov chain α are a set of canonical martingales $\{M_{ij} : i, j \in I,\ i \neq j\}$. Their construction and properties are detailed in Appendix (G). We summarize the most important items here.

For each $i, j = 1, \ldots, D$ and for all $t \in [0, T]$, set

$$M_{ij}(t) := \begin{cases} \sum_{0<s\le t} \chi[\alpha(s_-) = i]\chi[\alpha(s) = j] - q_{ij} \int_0^t \chi[\alpha(s) = i]\,\mathrm{d}s & \text{if } i \neq j, \\ 0 & \text{if } i = j, \end{cases} \tag{2.11}$$

where the indicator function χ is such that for each $i = 1, \ldots, D$,

$$\chi[\alpha(s) = i] = \begin{cases} 1 & \text{if } \alpha(s) = i, \\ 0 & \text{otherwise.} \end{cases} \tag{2.12}$$

Remark 2.1.7 We define M_{ii} for notational convenience as $M_{ii} := 0$ to make it clear that the set $\{M_{ii} : i \in I\}$ is not part of the set of canonical martingales of the Markov chain, which is based only on *distinct* states i and j. This is discussed further in Appendix (G).

We have $M_{ij} \in \mathcal{M}_{0,2}(\{\mathcal{F}_t\}, \mathbb{P})$ and M_{ij} is a finite variation process (see Remark G.2.15 and Lemma G.2.16). The set of martingales $\{M_{ij} : i, j \in I, i \neq j\}$ are the canonical martingales of the Markov chain α. We define the $D \times D$ "matrix of martingales"

$$M := (M_{ij})_{i,j=1}^{D}. \tag{2.13}$$

We may loosely refer to this matrix M as the set of canonical martingales of the Markov chain α. However, this should be understood as excluding the diagonal elements, all of which are zero (see Remark 2.1.7).

Since $M_{ij} \in \mathcal{M}_{0,2}(\{\mathcal{F}_t\}, \mathbb{P})$, Theorem B.2.29 ensures existence of the associated square-bracket quadratic variation process $[M_{ij}]$ (see Remark B.2.30) of M_{ij}. Moreover, Lemma G.2.19 establishes that this is explicitly given by

$$[M_{ij}](t) := \begin{cases} \sum_{0<s\leq t} \chi[\alpha(s_-) = i]\chi[\alpha(s) = j], & \text{if } i \neq j, \\ 0, & \text{if } i = j, \end{cases} \tag{2.14}$$

$\forall\, \mathrm{t} \in [0, T]$.

Since $M_{ij} \in \mathcal{M}_{0,2}(\{\mathcal{F}_t\}, \mathbb{P})$, Theorem B.2.24 ensures existence of the associated angle-bracket quadratic variation process $\langle M_{ij} \rangle$ (see Remark B.2.26) of M_{ij}, and Lemma G.2.21 establishes that it is given explicitly by

$$\langle M_{ij} \rangle(t) := \begin{cases} \int_0^t q_{ij}\chi[\alpha(s_-) = i]\, \mathrm{d}s, & \text{if } i \neq j, \\ 0, & \text{if } i = j, \end{cases} \tag{2.15}$$

$\forall\, \mathrm{t} \in [0, T]$.

From (2.14) and (2.15), the set of canonical martingales $\{M_{ij} : i, j \in I,\, i \neq j\}$ can be written as the difference of its square and angle bracket quadratic variation process.

Almost everywhere

Throughout this thesis we shall encounter several natural measures on the measure space $(\Omega \times [0, T], \mathcal{P}^*)$. One such measure is the product measure $(\mathbb{P} \otimes Leb)$, where Leb represents Lebesgue measure on the Borel σ-algebra on $[0, T]$. This measure will be particularly important for stochastic integrals with respect to the Brownian motion $\mathbf{W}$.

Another measure on the measurable space $(\Omega \times [0, T], \mathcal{P}^*)$ is $\nu_{[M_{ij}]}$, defined by

$$\nu_{[M_{ij}]}[A] := E \int_0^T \chi_A(\omega, t)\, \mathrm{d}[M_{ij}](t), \quad \forall A \in \mathcal{P}^*, \tag{2.16}$$

for each $i, j = 1, \ldots, D$, $i \neq j$. This is the *Doléans measure* generated by the finite-variation process $[M_{ij}]$, and will be particularly important for dealing with stochastic integrals with respect to the canonical martingales M_{ij}. In all cases we *specifically indicate* which of these measures is intended for relations which hold almost everywhere ("a.e.") on $\Omega \times [0, T]$.

Notation 2.1.8 Frequently we can subsume statements involving equality $\nu_{[\mathbf{M}]}$ - a.e. into a single statement as follows: The notation

$$\mathbf{G} = \mathbf{H} \quad \nu_{[\mathbf{M}]} - a.e. \tag{2.17}$$

for $\mathbb{R}^{D\times D}$-mappings $\mathbf{G} := (G_{ij})_{i,j=1}^{D}$, $\mathbf{H} := (H_{ij})_{i,j=1}^{D}$ on the set $\Omega \times [0, T]$, will indicate that

$$G_{ij} \quad = \quad H_{ij} \quad \nu_{[M_{ij}]} - \text{a.e.}, \qquad \forall i, j \in I,\, i \neq j. \tag{2.18}$$

Market Model, Stocks and Bonds

The market model that we adopt in this thesis comprises of $N + 1$ assets traded continuously over the interval [0, T], namely a single bond, with price $S_0(t)$ and N stocks with prices $\{S_n(t)\}$, $n = 1, 2, \ldots, N$.

We begin by modeling the bond. The price at time t of the bond will be denoted by $S_0(t)$, with the convention that $S_0(0) = 1$. The price process of the bond is modeled by equation

$$\mathrm{d}S_0(t) = r(t)S_0(t)\,\mathrm{d}t, \quad 0 \leq t \leq T, \tag{2.19}$$

in which $r(t)$ is a *given* process, called the *risk-free interest rate process* at time t, and subject to

Condition 2.1.9 The risk-free rate of return $\{r(t)\}$ is a uniformly bounded, non-negative, $\{\mathcal{F}_t\}$-predictable, $\mathbb{R}$-valued stochastic process on the set $\Omega \times [0, T]$.

We next define the stock price processes for our model. The price of one unit holding in the n^{th} stock at time t will be denoted by $S_n(t)$, with the convention that $S_n(0)$ is some positive constant, for each $n = 1, \ldots, N$. The price process of the n^{th} stock satisfies for each $n = 1, \ldots, N$, and for all $t \in [0, T]$,

$$\mathrm{d}S_n(t) = S_n(t)\left[b_n(t)\,\mathrm{d}t + \sum_{m=1}^{N} \sigma_{nm}(t)\,\mathrm{d}W_m(t)\right]. \tag{2.20}$$

$b_n(t)$ is called the *mean rate of return process* of the n^{th} stock at time t, and $\sigma_{nm}(t)$ is the $(n,m)^{th}$ entry of the $N \times N$ matrix *volatility process* $\boldsymbol{\sigma}(t)$ for $n, m = 1, \ldots, N$.

Condition 2.1.10 The entries of the mean rate of return process $\boldsymbol{b}(t) = \{b_n(t)\}_{n=1}^N$ and the entries of the volatility process $\boldsymbol{\sigma}(t) = \{\sigma_{nm}(t)\}_{n,m=1}^N$ are uniformly bounded and $\{\mathcal{F}_t\}$-predictable, $\mathbb{R}$-valued processes on the set $\Omega \times [0, T]$.

Definition 2.1.11 The processes $\{r(t)\}$, $\{\mathbf{b}(t)\}$, $\{\boldsymbol{\sigma}(t)\}$ are called the *market coefficients* or *market parameters* of the market model.

Remark 2.1.12 We observe from the Doob measurability theorem that Condition 2.1.9 and Condition 2.1.10 formulate a very general predictable dependence of the market parameters on the Markov chain α and the Brownian motion $\mathbf{W}$. This very general dependency structure was introduced by Donnelly [6], and is considerably richer and more flexible than the dependency structures adopted in [47] and [42], in which the market parameters are posited to depend at each instant t only on the "value" $\alpha(t-)$ of the regime state Markov chain immediately prior to instant t and nothing else (this rather special dependence is known as "Markov modulation" and is completely essential to the approaches of [47] and [42], which cannot be generalized to the more general dependency structure introduced by Donnelly [6] which we also posit here).

Remark 2.1.13 In the preceding market model the role of the regime state Markov chain α seems almost invisible, since it does not appear as an explicit element in the modeling equations (2.19) and (2.20). It is as a building block of the filtration $\{\mathcal{F}_t\}$ (see (2.9)) and the $\{\mathcal{F}_t\}$-predictability posited by Condition 2.1.9 and Condition 2.1.10 that α plays an essential role by determining the market parameters through the Doob measurability theorem. As we shall see the contribution of the Markov chain α to the filtration $\{\mathcal{F}_t\}$ vastly affects the structure of the dual variables when we construct an associated dual problem.

Further conditions and market price of risk process

In addition to Condition 2.1.10 we shall also postulate

Condition 2.1.14 There exists a constant $\kappa \in (0, \infty)$ such that

$$\mathbf{z}'\boldsymbol{\sigma}(t,\omega)\boldsymbol{\sigma}'(t,\omega)\mathbf{z} \geq \kappa \parallel \mathbf{z} \parallel^2 \quad \forall(\mathbf{z}, t, \omega) \in \mathbb{R}^N \times [0, \mathrm{T}] \times \Omega. \tag{2.21}$$

where we use $\parallel \mathbf{z} \parallel$ to denote the usual Euclidean length of a vector $\mathbf{z} \in \mathbb{R}^N$.

Remark 2.1.15 Condition 2.1.14 is a technical condition which is nevertheless a very standard condition in portfolio optimization theory, and in particular implies that the matrix $\boldsymbol{\sigma}(t,\omega)$ is non-singular for each $(t,\omega) \in [0,T] \times \Omega$. This technical condition is rather consistent with observed market data, and greatly facilitates the use of stochastic calculus tools.

Remark 2.1.16 We will collectively call Conditions 2.1.3, 2.1.9, 2.1.10, 2.1.14 the *market conditions.*

In view of Condition 2.1.10, Condition 2.1.14 and Karatzas, Lehoczky, and Shreve [20], there exists a constant $\kappa \in (0,\infty)$ such that

$$\max\left\{\| (\boldsymbol{\sigma}^T(t,\omega))^{-1}z \|, \| (\boldsymbol{\sigma}(t,\omega))^{-1})z \|\right\} \leq \frac{1}{\sqrt{\kappa}} \| z \| \qquad \forall (z,t,\omega) \in \mathbb{R}^N \times [0,\mathrm{T}] \times \Omega. \tag{2.22}$$

From (2.22) and Condition 2.1.10, there exists a constant $\kappa_{\boldsymbol{\sigma}} \in (0,\infty)$ such that

$$\max\left\{\| (\boldsymbol{\sigma}^T(t,\omega))z \|, \| (\boldsymbol{\sigma}(t,\omega))z \|, \| \boldsymbol{\sigma}^T(t,\omega))^{-1}z \|, \| (\boldsymbol{\sigma}(t,\omega))^{-1})z \|\right\} \leq \kappa_{\boldsymbol{\sigma}} \| z \|, \tag{2.23}$$

for all $(z,t,\omega) \in \mathbb{R}^N \times [0,\mathrm{T}] \times \Omega$. The simple bounds expressed at (2.22) and (2.23) will be used numerous times in this thesis.

Definition 2.1.17 The *market price of risk* is the mapping $\boldsymbol{\theta} : \Omega \times [0,\mathrm{T}] \to \mathbb{R}^N$ given by

$$\boldsymbol{\theta}(\omega,t) := \boldsymbol{\sigma}^{-1}(\omega,t)(\boldsymbol{b}(\omega,t) - r(\omega,t)\mathbf{1}), \tag{2.24}$$

where $\mathbf{1} \in \mathbb{R}^N$ has all unit entries.

Remark 2.1.18 From Condition 2.1.9 and (2.23) it follows that $\mathbf{R}^N$-valued process $\{\theta(t)\}$ is $\{\mathcal{F}_t\}$-predictable and uniformly bounded on $\Omega \times [0,\mathrm{T}]$. That is, there exists a constant $\kappa_{\boldsymbol{\theta}} \in (0,\infty)$ such that

$$\| \boldsymbol{\theta}(\omega,t) \| \leq \kappa_{\boldsymbol{\theta}}, \qquad \forall\, (\omega,t) \in \Omega \times [0,\mathrm{T}]. \tag{2.25}$$

Remark 2.1.19 At each time t, we know the values of the market coefficients $r(\omega,t)$, $\mathbf{b}(\omega,t)$, $\boldsymbol{\sigma}(\omega,t)$, $\boldsymbol{\theta}(\omega,t)$. The goal of portfolio optimization is to characterize, and, if possible, compute the optimal portfolio in terms of these known quantities.

2.2 The quadratic loss minimization problem

In the present section we formulate the problem of interest in this thesis, namely a stochastic optimal control problem with the goal of minimizing a quadratic loss function of the wealth at close of trade subject to constraints. For convenience we will refer to this as the *quadratic loss minimization* (or QLM) problem. It turns out that the quadratic loss minimization problem models a variety of hedging and mean-variance optimization problems that are of genuine practical importance. In order for the formulation of the QLM problem to be meaningful we first need to introduce several prior notions, in particular the portfolio process, the wealth equation, and appropriate spaces of stochastic integrands. All of these will be needed in order to formulate the problem.

The Investor

We consider an investor with some given initial wealth $x_0 > 0$. The total wealth of an investor in the market at time $t \in [0,T]$ is denoted by $X^{\boldsymbol{\pi}}(t)$. The reason for the superscript $\boldsymbol{\pi}$ will be apparent in the next few paragraphs. We assume that the investor consumes nothing and that there are no transaction costs.

We denote by $\pi_0(t)$ the dollar amount of wealth that the investor holds in the bond at time t. We denote by $\pi_n(t)$ the dollar amount of wealth that the investor holds in stock n at time t, for each $n = 1, \ldots, N$. Defining the vector

$$\boldsymbol{\pi}(t) := (\pi_1(t), \ldots, \pi_N(t))^\top, \tag{2.26}$$

we can express the total wealth $X^{\boldsymbol{\pi}}(t)$ of the investor at time t in terms of asset holdings $(\pi_0(t), \boldsymbol{\pi}^\top(t))$ at time t as

$$X^{\boldsymbol{\pi}}(t) = \pi_0(t) + \boldsymbol{\pi}^\top(t)\mathbf{1}, \qquad \forall t \in [0,T], \tag{2.27}$$

where $\mathbf{1} \in \mathbb{R}^N$ has all unit entries.

Note that the value of $\pi_0(t)$ can be retrieved from (2.27) if we know the values of $X^{\boldsymbol{\pi}}(t)$ and $\boldsymbol{\pi}(t)$. As a result, we will define a portfolio process as $\boldsymbol{\pi} = \{\boldsymbol{\pi}(t) : t \in [0,T]\}$, the investor's dollar holdings in the stocks only.

It is now clear that the superscript $\boldsymbol{\pi}$ of $X^{\boldsymbol{\pi}}(t)$ alludes to the investor's portfolio holdings $\boldsymbol{\pi}(t)$ in the N stocks.

We formally define a portfolio process as follows.

Definition 2.2.1 A portfolio process $\{\boldsymbol{\pi}(t) : t \in [0,T]\}$ for the market model is a $\{\mathcal{F}_t\}$-predictable process $\boldsymbol{\pi} : \Omega \times [0,T] \to \mathbb{R}^N$ such that $\int_0^T \parallel \boldsymbol{\pi}(t) \parallel^2 \, \mathrm{d}t < \infty$ a.s.

From Karatzas and Shreve, [22], chapter 1, it is known that the wealth process in terms of $\boldsymbol{\pi}$, $\boldsymbol{\theta}$ and $\boldsymbol{\sigma}$ is given by following linear stochastic differential equation in which the portfolio process $\boldsymbol{\pi}(t)$, reflecting the dollar amounts invested in each stock at instant t, is essentially an "input" or "driving" process and $X^{\boldsymbol{\pi}}(t)$ is the resulting investor wealth:

$$\mathrm{d}X^{\boldsymbol{\pi}}(t) = [r(t)X^{\boldsymbol{\pi}}(t) + \boldsymbol{\pi}^\top(t)\boldsymbol{\sigma}(t)\boldsymbol{\theta}(t)]\mathrm{d}t + \boldsymbol{\pi}^\top(t)\boldsymbol{\sigma}(t)\,\mathrm{d}\mathbf{W}(t),\ X^{\boldsymbol{\pi}}(0) = x_0. \quad (2.28)$$

We call (2.28) the *wealth equation.*

The *wealth process* $X^{\boldsymbol{\pi}} = \{X^{\boldsymbol{\pi}}(t) : \mathrm{t} \in [0, \mathrm{T}]\}$ is the unique (up to indistinguishability) solution of the wealth equation (2.28). It follows from the elementary theory of linear stochastic differential equations applied to (2.28) that the process $X^{\boldsymbol{\pi}}$ is the continuous, $\{\mathcal{F}_t\}$-adapted, $\mathbb{R}$-valued process given by the explicit formula

$$X^{\boldsymbol{\pi}}(t) = S_0(t)\left\{x_0 + \int_0^t S_0^{-1}(\tau)\boldsymbol{\pi}^\top(\tau)\boldsymbol{\sigma}(\tau)\boldsymbol{\theta}(\tau)\,\mathrm{d}\tau + \int_0^t S_0^{-1}(\tau)\boldsymbol{\pi}^\top(\tau)\boldsymbol{\sigma}(\tau)\mathrm{d}\mathbf{W}(\tau)\right\}. \quad (2.29)$$

We refer to $X^{\boldsymbol{\pi}}$ as the solution to the wealth equation (2.28) for the portfolio process $\boldsymbol{\pi}$.

Examining the right-hand side of (2.29), we see that the only parameter which is under the sole control of the investor is the portfolio process $\boldsymbol{\pi}$. The initial wealth x_0 is fixed and all the other parameters are market-determined parameters, known to the investor but not controlled by the investor. Notice how (2.29) displays the wealth as a very simple relation in terms of the portfolio $\boldsymbol{\pi}$.

Spaces of integrands

We will see in the next section that in order to formulate the QLM problem alluded to above, we must have $E|X^{\boldsymbol{\pi}}(T)|^2 < \infty$. This implies that the wealth processes $X^{\boldsymbol{\pi}}$ which we wish to consider as potential solutions must be square-integrable. With this necessity in mind, we define a space $\mathbb{B}$ consisting of right-continuous, square-integrable $\mathcal{F}_t$ semimartingales.

The wealth process which solves the wealth equation (2.28) is of course a *continuous* $\mathcal{F}_t$ semimartingale, and for this reason, we also define a subspace $\mathbb{A}$ of $\mathbb{B}$ whose members are continuous processes, in addition to being square-integrable.

Potential solutions to the QLM problem will turn out to be wealth processes in the smaller space $\mathbb{A}$. The larger space $\mathbb{B}$ of right-continuous semimartingales will play an essential role later on, for it will be seen that the Lagrange multipliers for the constraints in the QLM problem will be members of this space.

We start by defining some appropriate spaces of integrands.

$$L_{21} := \left\{ \Upsilon : \Omega \times [0, \mathrm{T}] \to \mathbb{R} | \Upsilon \in \mathcal{P}^* \;\;\&\;\; E\left(\int_0^T |\Upsilon(t)| \,\mathrm{d}t \right)^2 < \infty \right\}. \tag{2.30}$$

$$L_2(\mathbf{W}) := \left\{ \boldsymbol{\xi} : \Omega \times [0, \mathrm{T}] \to \mathbb{R}^N \,|\, \boldsymbol{\xi} \in \mathcal{P}^* \;\&\; E\int_0^{\mathrm{T}} \parallel \boldsymbol{\xi}(t) \parallel^2 \mathrm{d}t < \infty \right\}. \tag{2.31}$$

$$L_2(\mathbf{M}) := \Big\{ \mathbf{\Gamma} = \{\Gamma_{ij}\}_{i,j=1}^D : \Omega \times [0, \mathrm{T}] \to \mathbb{R}^{D\times D} | \Gamma_{ii} = 0, \, (\mathbb{P} \otimes Leb) - a.e., \forall i \in I,$$
$$\Gamma_{ij} \in \mathcal{P}^* \;\;\&\;\; E\int_0^T |\Gamma_{ij}(t)|^2 \,\mathrm{d}[M_{ij}](t) < \infty, \, \forall i,j \in \mathbf{I}, i \neq j \Big\}. \tag{2.32}$$

The preceding are all vector spaces of predictable processes satisfying various integrability conditions. Now define the product vector space $\mathbb{B}$ as follows:

$$\mathbb{B} := \mathbb{R} \times L_{21} \times L_2(\mathbf{W}) \times L_2(\mathbf{M}). \tag{2.33}$$

We then write $Y \in \mathbb{B}$ to indicate that $Y \equiv \{Y(t) : t \in [0, T]\}$ is a right-continuous semimartingale of the form

$$Y(t) := Y_0 + \int_0^t \Upsilon^Y(\tau) \,\mathrm{d}\tau + \sum_{n=1}^N \int_0^t \xi_n^Y(\tau) \,\mathrm{d}W_n(\tau) + \sum_{i,j=1}^D \int_0^t \Gamma_{i,j}^Y(\tau) \,\mathrm{d}M_{ij}(\tau), \tag{2.34}$$

for some $Y_0 \in \mathbb{R}$, $\Upsilon^Y \in L_{21}$, $\boldsymbol{\xi} := (\xi_1^Y, \ldots, \xi_N^Y)^\top \in L_2(\mathbf{W})$, and $\mathbf{\Gamma}^Y := (\Gamma_{ij}^Y)_{i,j=1}^D \in L_2(\mathbf{M})$. Alternatively (and more explicitly) we shall write

$$Y \equiv (Y_0, \Upsilon^Y, \boldsymbol{\xi}^Y, \mathbf{\Gamma}^Y) \in \mathbb{B}, \tag{2.35}$$

to indicate that (2.34) holds for $Y \in \mathbb{B}$, and we call the quadruple $(Y_0, \Upsilon^Y, \boldsymbol{\xi}^Y, \mathbf{\Gamma}^Y)$ the *components* of Y.

We also define a subspace $\mathbb{A}$ of $\mathbb{B}$ as

$$\mathbb{A} := \{X \equiv (X_0, \Upsilon^X, \boldsymbol{\xi}^X, \mathbf{\Gamma}^X) \in \mathbb{B} \,|\, \mathbf{\Gamma}^X = 0, \, \nu_{[\mathbf{M}]} - \text{a.e.}\}. \tag{2.36}$$

Remark 2.2.2 Effectively a process X is a member of $\mathbb{A}$ when it is given by

$$X(t) := X_0 + \int_0^t \Upsilon^X(\tau)\,\mathrm{d}\tau + \sum_{n=1}^{N} \int_0^t \xi_n^X(\tau)\,\mathrm{d}W_n(\tau) \tag{2.37}$$

for some $X_0 \in \mathbb{R}$, $\Upsilon^X \in L_{21}$, and $\boldsymbol{\xi} := (\xi_1^X, \ldots, \xi_N^X)^\top \in L_2(\mathbf{W})$ (as follows from (2.34) with the $\mathrm{d}M_{ij}$-integrals removed). The subspace $\mathbb{A}$ therefore consists of all *continuous processes* in the space $\mathbb{B}$.

The next result just establishes that the integrands in the representation (2.34) are uniquely determined. For completeness the rather standard proof of this result is placed in Appendix A.

Proposition 2.2.3 Suppose we have $X \equiv (X_0, \Upsilon^X, \boldsymbol{\xi}^X, \boldsymbol{\Gamma}^X) \in \mathbb{B}$ and $Y \equiv (Y_0, \Upsilon^Y, \boldsymbol{\xi}^Y, \boldsymbol{\Gamma}^Y) \in \mathbb{B}$. If for all $t \in [0, T]$,

$$\begin{aligned} &X_0 + \int_0^t \Upsilon^X(\tau)\,\mathrm{d}\tau + \sum_{n=1}^{N} \int_0^t \boldsymbol{\xi}_n^X(\tau)\,\mathrm{d}W_n(\tau) + \sum_{i,j=1}^{D} \int_0^t \Gamma_{i,j}^X(\tau)\,\mathrm{d}M_{ij}(\tau) \\ &= Y_0 + \int_0^t \Upsilon^Y(\tau)\,\mathrm{d}\tau + \sum_{n=1}^{N} \int_0^t \boldsymbol{\xi}_n^Y(\tau)\,\mathrm{d}W_n(\tau) + \sum_{i,j=1}^{D} \int_0^t \Gamma_{i,j}^Y(\tau)\,\mathrm{d}M_{ij}(\tau) \end{aligned} \tag{2.38}$$

then $X_0 = Y_0, \Upsilon^X = \Upsilon^Y, \boldsymbol{\xi}^X = \boldsymbol{\xi}^Y, \boldsymbol{\Gamma}^X = \boldsymbol{\Gamma}^Y$.

The following lemma shows that the members of $\mathbb{B}$ are square integrable. The proof of this result is again placed in Appendix A.

Lemma 2.2.4 For all $Y \equiv (Y_0, \Upsilon^Y, \boldsymbol{\xi}^Y, \boldsymbol{\Gamma}^Y) \in \mathbb{B}$, we have

$$E\left(\sup_{t\in[0,T]} |Y(t)|^2\right) < \infty. \tag{2.39}$$

Remark 2.2.5 Observe that the wealth process $X^{\boldsymbol{\pi}}$ which solves the wealth equation (2.28) for a portfolio process $\boldsymbol{\pi}$ is a continuous process. The next proposition, the elementary proof of which is located in Appendix A, establishes that $X^{\boldsymbol{\pi}} \in \mathbb{A}$ (i.e. $X^{\boldsymbol{\pi}}$ is square integrable) when the portfolio process $\boldsymbol{\pi}$ is in $L_2(\mathbf{W})$.

Proposition 2.2.6 For $X^{\boldsymbol{\pi}}$ which solves the wealth equation (2.28) for a portfolio process $\boldsymbol{\pi}$, we have

$$X^{\boldsymbol{\pi}} \in \mathbb{A} \quad \text{if and only if} \quad \boldsymbol{\pi} \in L_2(\mathbf{W}). \tag{2.40}$$

Remark 2.2.7 Suppose that $\boldsymbol{\pi} \in L_2(\mathbf{W})$. Then from the wealth equation (2.28) and Proposition 2.2.6, we see that

$$X^{\boldsymbol{\pi}} \equiv (x_0, rX^{\boldsymbol{\pi}} + \boldsymbol{\pi}^\top \boldsymbol{\sigma}\boldsymbol{\theta}, \boldsymbol{\sigma}^\top \boldsymbol{\pi}, 0) \in \mathbb{A}. \tag{2.41}$$

2.2.1 The Investor's Problem

Having formulated the wealth equation (see (2.28)) and spaces of integrands, we are finally able to define precisely the QLM problem that will be the focus of this thesis. We first define the so-called *quadratic loss function*:

Definition 2.2.8 Define a quadratic *loss function* J of the form

$$J(x, \omega) := \frac{1}{2}[a(\omega)x^2 + 2c(\omega)x] + q, \quad (x, \omega) \in \mathbb{R} \times \Omega, \tag{2.42}$$

subject to the following condition:

Condition 2.2.9 *a and c are $\mathcal{F}_T$-measurable, square-integrable random variables on $(\Omega, \mathcal{F}, \mathbb{P})$, $q \in \mathbb{R}$ is a constant and a satisfies*

$$0 < \inf_{\omega \in \Omega}\{a(\omega)\} \leq \sup_{\omega \in \Omega}\{a(\omega)\} < \infty. \tag{2.43}$$

From Proposition 2.2.6, the portfolio processes we are interested in must lie in the space $L_2(\mathbf{W})$. This ensures that the corresponding wealth process $X^{\boldsymbol{\pi}}$ is in $\mathbb{A}$ and therefore square-integrable. Thus, we require that the *admissible portfolios* be members of $L_2(\mathbf{W})$. We now introduce a *portfolio constraint*, namely we shall insist that the portfolio vector $\boldsymbol{\pi}$ always remain in a given *constraint set* $K \subset \mathbb{R}^N$:

Condition 2.2.10 *$K \subset \mathbb{R}^N$ is non-empty, closed and convex with $\mathbf{0} \in K$.*

Example 2.2.11 Several important examples of convex constraints sets K are given by Karatzas and Shreve [22], in particular:

1. Unconstrained case : $K = \mathbb{R}^N$.
2. Prohibition of short selling : $K = [0, \infty)^N$. This requires that a *nonnegative* dollar amount always be invested in each and every stock.

3. Incomplete market : $K = \{\boldsymbol{\pi} \in \mathbb{R}^N; \pi_{M+1} = \cdots = \pi_N = 0\}$, for some $M \in \{1, \ldots, N-1\}$. This constraint effectively prohibits investment in the stocks with prices $S_{M+1}, \ldots, S_N$.

Definition 2.2.12 Define the set $\mathcal{A}$ of *admissible portfolios* as

$$\mathcal{A} := \{\boldsymbol{\pi} \in L_2(\mathbf{W}) | \boldsymbol{\pi}(t) \in K \ (\mathbb{P} \otimes Leb) - \text{a.e.}\}. \tag{2.44}$$

We next formulate a *generalization* of the terminal wealth constraint $X^{\boldsymbol{\pi}}(T) \geq 0$ a.s. which was discussed in Chapter 1. To this end we shall suppose

Condition 2.2.13 We are given a random variable B which is $\mathbb{P}$-essentially bounded.

In the forthcoming definition of the QLM problem addressed in this thesis (see Problem 2.2.16) we are going to insist that portfolios $\boldsymbol{\pi}$ always satisfy two constraints, namely the *portfolio constraint* that $\boldsymbol{\pi} \in \mathcal{A}$ (see Definition 2.2.12) and the *terminal wealth constraint* that $X^{\boldsymbol{\pi}}(T) \geq B$ a.s. It is this combination of constraints which makes the problem particularly challenging.

Definition 2.2.14 The *primal value*, denoted by η, is defined as

$$\eta := \inf_{\substack{\boldsymbol{\pi} \in \mathcal{A} \\ X^{\boldsymbol{\pi}}(T) \geq B \, a.s.}} E[J(X^{\boldsymbol{\pi}}(T))], \tag{2.45}$$

where $X^{\boldsymbol{\pi}}$ is the solution to the wealth equation (2.28) corresponding to $\boldsymbol{\pi}$. The set of admissible portfolios $\mathcal{A}$ is given by Definition 2.2.12 and the loss function is given by Definition 2.2.8.

The elementary proof of the following Lemma is given in Appendix A

Lemma 2.2.15 The *primal value* given by (2.45) is such that $-\infty < \eta < \infty$.

We are now ready to define the problem that we are going to address in this thesis:

Problem 2.2.16 The *quadratic loss minimization* (QLM) problem is to determine the existence of, and to characterize, a portfolio process $\bar{\boldsymbol{\pi}} \in \mathcal{A}$ such that $X^{\bar{\boldsymbol{\pi}}}(T) \geq B$-a.s. and $\eta = E[J(X^{\bar{\boldsymbol{\pi}}}(T))]$. This portfolio $\bar{\boldsymbol{\pi}}$ satisfies the constraints of the problem and minimizes the quadratic loss function among all portfolios which satisfy the problem constraints.

By existence and characterization, we mean demonstrating the existence of $\bar{\boldsymbol{\pi}}$ and characterizing its dependence on the market coefficient $r(t)$, $\mathbf{b}(t)$ and $\boldsymbol{\sigma}(t)$ and the filtration $\{\mathcal{F}_t\}$.

Remark 2.2.17 In Problem 2.2.16 recall that the wealth process $X^{\boldsymbol{\pi}}$ is given in terms of the portfolio process $\boldsymbol{\pi} \in L_2(\mathbf{W})$ by the stochastic differential equation (2.28), in which the initial value $x_0 \in (0, \infty)$ is the given initial wealth of the investor.

Remark 2.2.18 The random variable B in the constraint $X^{\boldsymbol{\pi}}(T) \geq B$ a.s. at (2.45) reduces to the no-bankruptcy condition $X^{\boldsymbol{\pi}}(T) \geq 0$ that we discussed in Chapter 1 if we simply put $B := 0$. However, being able to choose this random variable results in a more general problem. For example, fixing $B = c$ for a constant $c > 0$ *guarantees* the investor a wealth in excess of c at close of trade.

Remark 2.2.19 In this remark we indicate more precisely how the QLM Problem 2.2.16 relates to, and generalizes, the problems addressed by Bielecki, Pliska, Jin and Zhou [2], Zhou and Yin [47], and Donnelly [7] that we discussed as motivation in Section 1.2 of Chapter 1.

(a) Take $K := \mathbb{R}^N$ at Definition 2.2.12 (that is, the portfolio is *unconstrained* i.e. $\mathcal{A} = L_2(\mathbf{W})$), put $B := 0$, and remove regime switching from the market model by assuming that the market parameters are adapted *only* to the filtration of the Brownian motion process $\mathbf{W}$, without any dependency on a regime switch Markov chain, that is adapted to the filtration $\{\mathcal{F}_t^{\mathbf{W}},\ t \in [0,T]\}$ defined at (2.10) (compare (2.8), (2.9) and Condition 2.1.10). With these simplifications Problem 2.2.16 reduces to the problem addressed in [2]. As we noted in Chapter 1, these special features make it possible to construct an optimal portfolio by an application of the separating hyperplanes theorem. However, this elegant approach is completely ruled out by either the presence of regime switching or portfolio constraints, and certainly by the combination of these that we have in Problem 2.2.16.

(b) Again take $K := \mathbb{R}^N$ at Definition 2.2.12 (i.e. unconstrained portfolios), remove the constraint $X^{\boldsymbol{\pi}}(T) \geq B$ (see (2.45) and Problem 2.2.16), and strengthen Condition 2.1.9 and Condition 2.1.10 by requiring that the market parameters be *Markov-modulated*. Essentially this means that at every instant $t \in [0,T]$ the market parameters $r(t)$, $b_n(t)$ and $\sigma_{nm}(t)$ are $\sigma\{\alpha(t)\}$-measurable, that is $r(t)$, $b_n(t)$ and $\sigma_{nm}(t)$ are *determined completely* by the random variable $\alpha(t)$. Obviously this is a much simpler type of dependency than the predictability with respect to $\{\mathcal{F}_t\}$ stipulated at Condition 2.1.9 and Condition 2.1.10, since dependence on the Brownian motion $\mathbf{W}$ is excluded completely and dependence on the Markov chain α is limited only to the instant t instead of the history of α on the interval $[0, t]$. With these simplifications Problem 2.2.16 reduces to the problem addressed in [47]. It is these very special features which allow the application of results from classical LQ-control in [47]. However, the presence of either the portfolio constraint (i.e. $\boldsymbol{\pi} \in \mathcal{A}$), or the state constraint $X^{\boldsymbol{\pi}}(T) \geq B$, or the general form of regime switching at Condition 2.1.10, all of which are built into Problem 2.2.16, completely rules out the approach of [47] based on classical LQ-control.

(c) If we remove the constraint $X^{\boldsymbol{\pi}}(T) \geq B$ (see (2.45) and Problem 2.2.16), then Problem 2.2.16 reduces to the problem addressed by Donnelly [7] on the basis of the general conjugate duality theory of Bismut [3]. As we noted in Chapter 1 the presence of this constraint effectively rules out the approach of [7], since it brings into play a Lagrange multiplier which is *singular* (in the sense of being only a finitely additive, rather than a countably additive measure), and which falls outside the scope of the conjugate duality theory of [3].

Remark 2.2.20 The most important special case of the quadratic loss function formulated at Definition 2.2.8 arises as follows: a *contingent claim* is a specified $\mathcal{F}_T$-measurable random variable γ. An agent (e.g. a pension fund) is responsible for paying out this contingent claim at T (the date at which the contingent claim matures). The challenge of the agent is then to trade in the specified bond and stocks so as to generate a wealth $X^{\boldsymbol{\pi}}(T)$ which most closely "approximates" the contingent claim γ. The sense of "approximation" must of course be made precise, and Markowitz [28] suggested minimization of the mean square discrepancy $E[|X^{\boldsymbol{\pi}}(T) - \gamma|^2]$, a criterion which is now widely used in practice. If we take $a(\omega) := 2$, $c(\omega) := -2\gamma(\omega)$, $q := E[\gamma^2]$ in Definition 2.2.8 then we see that

$$E[J(X^{\boldsymbol{\pi}}(T))] = E[|X^{\boldsymbol{\pi}}(T) - \gamma|^2],$$

as required.

Remark 2.2.21 At the very end of Section 1.1, in Remark 1.1.1, we briefly drew attention to the possibility (or otherwise) of a Black-Scholes hedging approach to our Problem 2.2.16. Having formulated the problem in some detail we can now argue why it is definitely *not* possible to use hedging to address Problem 2.2.16. The essence of a hedging problem (for European options) is the following: one is given a contingent claim γ, and the goal is to determine both an *initial wealth* x_0 and a *trading strategy* π over the trading interval $0 \leq t \leq T$ such that the wealth at instant $t = T$ resulting from the initial investment x_0 and the trading strategy π is *exactly* equal to the contingent claim γ, that is

$$X^{x_0,\pi}(T) = \gamma, \tag{2.46}$$

in which $X^{x_0,\pi}(t)$ denotes the quantity on the right hand side of (2.29). The essential idea is this: if one starts with an *appropriate initial wealth* x_0 (the so-called hedging price of the contingent claim), then there is a trading strategy π which *exactly replicates* the contingent claim at $t = T$ in the sense that (2.46) holds. The basic assumptions on which the Black-Scholes theory rests are two-fold namely

(i) Determination of an initial wealth x_0 which makes the perfect hedging at (2.46) possible for some portfolio π is *part of the hedging problem.* This initial wealth is the *hedging price* of the contingent claim, and is given taking the expectation of the discounted contingent claim with respect to a "risk neutral measure";

(ii) The market model must be *complete*, which, in the present setting, excludes any constraints whatsoever on either the portfolio or wealth processes, as well as excluding regime switching in the market model. Indeed, with constraints and/or regime switching in the market model, one cannot in general expect to replicate a contingent claim γ in the sense of finding some x_0 together with a (constrained) portfolio π such that the (constrained) wealth process $X^{x_0,\pi}$ satisfies (2.46).

In view of (ii) the presence of constraints and regime switching in Problem 2.2.16 immediately rules out any approach based on Black-Scholes hedging. Would Black-Scholes hedging become a viable possibility if one considered a simplified version of Problem 2.2.16 in which all constraints are discarded and regime switching is removed? Again the answer is in the negative, this time because assumption (i) is not satisfied. In Problem 2.2.16 the initial wealth x_0 is arbitrarily assigned, being the initial amount of wealth with which the investor must live, and if this x_0 is less than the hedging price of the contingent claim then there generally does not exist any portfolio π (even an unconstrained portfolio) for which (2.46) holds. It is for this very reason that Zhou and Yin [47], who addressed the unconstrained version of Problem 2.2.16, formulated this problem in stochastic control terms rather than as a hedging problem. We see from the preceding discussion that both assumptions (i) and (ii) fail in the case of Problem 2.2.16, and so we adopt a stochastic control approach as well.

For the Problem 2.2.16 to make sense we must have $X^{\boldsymbol{\pi}}(T) \geq B$-a.s. for some $\boldsymbol{\pi} \in \mathcal{A}$, since, if this failed to hold, then there would not exist any portfolio $\boldsymbol{\pi}$ which satisfies the constraints of the problem. In fact, as will be seen in Remark 4.2.4, we shall need to strengthen this to the following very mild *Slater type* condition in order to secure existence of Lagrange multipliers:

Condition 2.2.22 There is some $\hat{\boldsymbol{\pi}} \in \mathcal{A}$ and constant $\epsilon \in (0, \infty)$ such that $X^{\hat{\boldsymbol{\pi}}}(T) \geq B + \epsilon$ a.s.

Remark 2.2.23 If there failed to exist some $\boldsymbol{\pi} \in \mathcal{A}$ such that $X^{\boldsymbol{\pi}}(T) \geq B$-a.s. then of course Problem 2.2.16 would be ill-defined, since the portfolio constraint $\boldsymbol{\pi} \in \mathcal{A}$ on $\boldsymbol{\pi}$ would be incompatible with the terminal wealth constraint $X^{\boldsymbol{\pi}}(T) \geq B$ on $\boldsymbol{\pi}$. Effectively, this means that one has been "unrealistic" in the choice of the random variable B when specifying Problem 2.2.16, since this level of wealth can never be weakly exceeded at close of trade with the portfolio constraint in place. The assumption that $X^{\boldsymbol{\pi}}(T) \geq B$-a.s. for some $\boldsymbol{\pi} \in \mathcal{A}$ is therefore absolutely essential for Problem 2.2.16 just to be well defined. Condition 2.2.22 is a very mild strengthening of this assumption, and really just compels one to make a "reasonable" choice of the random variable B, in the sense that there should exist some admissible portfolio $\hat{\boldsymbol{\pi}} \in \mathcal{A}$ which not only exceeds the stipulated terminal wealth B but does so with a small "margin" ϵ to spare. An example of such a choice of B is $B = \alpha x_0 S_0(T)$ (for some constant $\alpha \in [0, 1)$); since $\{r(t)\}$ is uniformly bounded (see condition 2.1.9) we see that Condition 2.2.22 holds with $\hat{\boldsymbol{\pi}} = 0$. This choice of B constitutes portfolio insurance for an amount corresponding to investing a fraction α of the initial wealth in a money-market account, and is a very common form of portfolio insurance. Conditions similar to Condition 2.2.22, in the sense of requiring satisfaction of the stipulated problem constraints with a small "margin" ϵ to spare, were first introduced by Slater [41] many decades ago for finite dimensional nonlinear optimization problems, and are completely essential for securing existence of Lagrange multipliers for the constraints (this is treated extensively in the context of convex finite dimensional optimization problems in Part VI of Rockafellar [34]). Condition 2.2.22 will similarly be indispensable for securing existence of Lagrange multipliers for the rather complex constraints in Problem 2.2.16 (this will be seen at Remark 4.2.2, Proposition 4.2.3 and Remark 4.2.4 in the next chapter). Our formulation of the Slater type Condition 2.2.22 is motivated by Rockafellar and Wets [37], who use Slater type conditions to obtain existence of Lagrange multipliers for *static* problems of stochastic convex programming. In the present thesis we extend the use of this type of condition to the *dynamic* Problem 2.2.16.

2.2.2 Problem Reformulation

The discussion of Remark 2.2.19 indicates that the QLM Problem 2.2.16 demands an approach quite different from the prior and motivating works [2], [7] and [47]. Indeed, Problem 2.2.16 really amounts to a stochastic optimal control problem with both a *control constraint* (in the form of the portfolio constraint $\boldsymbol{\pi} \in \mathcal{A}$) and a *state constraint* (in the form of the constraint $X^{\boldsymbol{\pi}}(T) \geq B$ a.s.). There are very few available results on such problems, but the prior works of Dubovitskii and Milyutin [8] and Makowski and Neustadt [27], in the considerably simpler setting of *deterministic* optimal control, warn us at the very least to expect rather troublesome "singular" Lagrange multipliers. There is, however, one very nice property possessed by Problem 2.2.16, which is not a feature of the problems studied in [8] and [27], this being that the problem is *convex*. This convexity is an immediate consequence of the simple linear dynamics of the wealth equation (2.28) and the fact that the portfolio constraints set K is assumed convex (see Condition 2.2.10), and its importance simply cannot be overstated. Indeed, we shall exploit this convexity as the key to applying the abstract, and very powerful, conjugate duality theory of Rockafellar and Moreau [36] to Problem 2.2.16. This theory has hitherto been applied mainly to *static* optimization problems, and of course Problem 2.2.16 is a control problem, that is *dynamic*. Nevertheless, we shall see that this general theory, by virtue of its abstractness, can nevertheless be made to apply to Problem 2.2.16, although (again by virtue of its abstractness) this application is far from routine or mechanical. In fact, one of the major challenges of this thesis is to "re-shape" the Rockafellar-Moreau theory and make it apply to the dynamic Problem 2.2.16. In particular, we shall see that the Rockafellar-Moreau approach introduces the "singular" Lagrange multipliers mentioned above in a very simple and natural way.

In order to apply the Rockafellar-Moreau approach we must first "remove" the portfolio $\boldsymbol{\pi}$ as the underlying problem variable with respect to which we optimize, and reformulate Problem 2.2.16 as the minimization of a convex function f over the linear space of Itô processes $\mathbb{A}$. We devote the remainder of the present chapter to this reformulation (which is very simple indeed). With this reformulation in place we shall then be able to apply the general approach of Rockafellar and Moreau to Problem 2.2.16 in the following Chapter. As the first step in this reduction we introduce

Definition 2.2.24 Define the set $\mathbb{D}$ of *admissible wealth processes* as

$$\mathbb{D} := \{X^{\boldsymbol{\pi}} \mid \boldsymbol{\pi} \in \mathcal{A}\}. \tag{2.47}$$

Remark 2.2.25 From Proposition 2.2.6 it follows that $\mathbb{D} \subset \mathbb{A}$.

Notice that a wealth process X is admissible provided that it corresponds to an admissible portfolio $\boldsymbol{\pi} \in \mathcal{A}$, and that an admissible wealth process X need not necessarily satisfy the all-important constraint $X(T) \geq B$ a.s. We take care of this latter constraint by building it into the so-called *primal function* defined as follows:

Definition 2.2.26 Define the *primal function* $f : \mathbb{A} \to (-\infty, \infty]$ as follows:

$$f(X) := \begin{cases} EJ(X(T)), & \text{when}\, X \in \mathbb{D} \text{ and}\, X(T) \geq B\, \text{a.s.}, \\ +\infty, & \text{otherwise.} \end{cases} \tag{2.48}$$

Remark 2.2.27 From Lemma 2.2.4, Condition 2.2.9 and (2.42) it follows that $J(X(T))$ is P-integrable for each $X \in \mathbb{A}$, while Condition 2.2.22 ensures that the value of η in Problem 2.2.16 is finite, i.e. $\eta \in \mathbb{R}$. From (2.48), (2.47) and (2.45), we then have

$$\eta = \inf_{\substack{\boldsymbol{\pi} \in \mathcal{A} \\ X^{\boldsymbol{\pi}}(T) \geq B\, a.s.}} E[J(X^{\boldsymbol{\pi}}(T))] = \inf_{X \in \mathbb{A}} f(X). \tag{2.49}$$

Remark 2.2.28 We have reduced Problem 2.45, and therefore the QLM Problem 2.2.16, to one of minimizing the convex function $f(\cdot)$ over the linear space $\mathbb{A}$ (see (2.49)). Our goal is therefore to construct some $\bar{X} \in \mathbb{A}$ such that

$$\eta = f(\bar{X}). \tag{2.50}$$

Since $\eta \in \mathbb{R}$ it follows from (2.48) that $\bar{X} \in \mathbb{D}$ and $\bar{X}(T) \geq B$ a.s. In view of (2.47) we then see that $\bar{X} = X^{\bar{\boldsymbol{\pi}}}$ for some $\bar{\boldsymbol{\pi}} \in \mathcal{A}$ which is then the optimal portfolio.

Chapter 3

Rockafellar-Moreau Approach

The QLM Problem 2.2.16 is a *convex* optimization problem in the sense that the objective function to be minimized is convex (in fact quadratic) and all constraints in the problem are convex. This convexity is a huge asset and is effectively the key to dealing with the otherwise rather intractable aspects of Problem 2.2.16. Convex optimization problems have the very nice property that one can typically associate with them a so-called *dual optimization* problem which is often much more tractable than the given (or *primal*) problem. If a solution of the dual optimization problem exists then it represents the *Lagrange multipliers* which enforce the constraints in the problem, and the solution of the given primal problem can be determined from the Lagrange multipliers by means of *optimality relations* (or Kuhn-Tucker relations) which relate solutions of the primal and dual problems. The challenge in applying this *convex duality* approach to a given convex problem is that one often has little *a-priori* knowledge concerning the appropriate space of *dual variables* over which the dual functional must be defined, as well as the dual functional and optimality relations, which are themselves far from clear as well. Over the years a collection of "standard" convex optimization problems has been built up, for which the space of dual variables, dual functional and optimality relations have really been determined by a process of trial and error. Particularly important members of this collection are linear programming and quadratic programming problems. However, it is not very difficult to come up with new convex problems which are not in this "library" of standard problems, the QLM Problem 2.2.16 being a case in point, and for these one has to go through an arduous process of guesswork, experimentation, and trial and error to ascertain the structure of the dual variables, dual functional and optimality relations. The Rockafellar-Moreau approach that we summarize in this chapter vastly reduces the amount of trial and error work that is needed to construct an appropriate dual problem for a given convex primal problem. The works of Rockafellar [36] and Ekeland and Temam [9] show that essentially all of the standard convex optimization problems mentioned above fall within the scope of the Rockafellar-Moreau approach. Much more important however, is the fact that Rockafellar-Moreau approach provides a very systematic way of dealing with convex optimization problems for which nothing about the corresponding duality structure is *a-priori* known (this is certainly the case for Problem 2.2.16). In fact, the book of Ekeland and Temam [9] illustrates several examples from continuuum mechanics, calculus of variations for partial differential equations, and filtering, all of which involve convex optimization problems for which nothing about an associated duality structure is known in advance, and demonstrates how one can *synthesize* a vector space of dual variables, dual functional and optimality relations by means of the Rockafellar-Moreau approach. Concerning this approach Ekeland and Temam ([9], p.xii) remark "This very flexible abstract theory can be adapted to a wide variety of situations". One of the main goals of this thesis is to show how this "very flexible abstract theory" can be adapted to our Problem 2.2.16. Of course this adaptation is far from routine (being very different from the cases illustrated in [36] and [9]) and involves some definite challenges. Addressing these challenges constitutes much of the technical core of the thesis.

In this chapter we are going to set out the main elements of the Rockafellar-Moreau approach. The elements of this approach are actually set forth in the work [36], but in a rather scattered form which is difficult to access, as well as in [9]. The self-contained summary in this chapter is completely adequate for reading this thesis. As a supplement to the present chapter we include in Section E.1 of Appendix E a simple "tutorial example" (suggested by Rockafellar and Wets [37]) in which the Rockafellar-Moreau approach is used on a static problem of convex stochastic optimization. Despite its simplicity this example illustrates "in miniature" the much more sophisticated application of the Rockafellar-Moreau approach to the QLM Problem 2.2.16 defined in Chapter 2, and also makes clear how "singular" Lagrange multipliers can unavoidably occur in even the very simplest problems of convex stochastic optimization with almost-sure inequality constraints. Among other things, this simple example illustrates the essential role of the *Yosida-Hewitt* decomposition of the adjoint space L_∞^* which is summarized in Appendix D, for it is the Yosida-Hewitt decomposition theorem which in fact gives the singular Lagrange multipliers.

3.1 Main Steps of the R-M Approach

Suppose that $\mathbb{X}$ is a given real linear space, and the primal problem is to minimize a given convex function

$$f : \mathbb{X} \to [-\infty, \infty] \tag{3.1}$$

on $\mathbb{X}$. Notice that this very simple formulation does not involve any loss of generality, since the objective function f can always be defined to have the value $+\infty$ at all points of $\mathbb{X}$ excluded by possible constraints in the problem (exactly as (2.49)). The Rockafellar Moreau approach gives a systematic method for constructing an appropriate vector space of dual variables, a Lagrangian, and a dual function, and is implemented in the following two steps:

Step I. Fix some real linear space of perturbations $\mathbb{U}$, and some *convex perturbation function*

$$F : \mathbb{X} \times \mathbb{U} \to [-\infty, \infty] \tag{3.2}$$

which must be consistent with the objective function f at (3.1) in that

$$F(x, 0) = f(x), \quad x \in \mathbb{X}, \tag{3.3}$$

that is when $u = 0$ (meaning zero perturbation) the function $F(x, u)$ reproduces the objective function f. Notice the complete freedom of choice that we have in selecting a space of perturbations $\mathbb{U}$ and a perturbation F; the only requirement is that the consistency relation (3.3) holds.

Step II. Fix a real linear space $\mathbb{Y}$ of so-called *dual variables* and some bilinear form $\langle \cdot, \cdot \rangle$ on $\mathbb{U} \times \mathbb{Y}$, that is, $(\mathbb{U}, \mathbb{Y}, \langle \cdot, \cdot \rangle)$ is a dual system in the sense of Remark B.4.7. Having fixed the vector spaces $\mathbb{U}$ and $\mathbb{Y}$, together with the bilinear form $\langle \cdot, \cdot \rangle$ on $\mathbb{U} \times \mathbb{Y}$, we can formulate a *Lagrangian function* and a *dual function* as follows:

Definition 3.1.1 Define the *Lagrangian* $K : \mathbb{X} \times \mathbb{Y} \to [-\infty, \infty]$ by the concave conjugate

$$K(x, y) := \inf_{u \in \mathbb{U}} [\langle u, y \rangle + F(x, u)], \quad (x, y) \in \mathbb{X} \times \mathbb{Y}. \tag{3.4}$$

Definition 3.1.2 Define the dual function $g : \mathbb{Y} \to [-\infty, \infty]$ by

$$g(y) := \inf_{x \in \mathbb{X}} K(x, y) = \inf_{(x,u) \in \mathbb{X} \times \mathbb{U}} [\langle u, y \rangle + F(x, u)], \quad y \in \mathbb{Y}. \tag{3.5}$$

It is immediate from (3.3) to (3.5) that $g(\cdot)$ is concave on $\mathbb{Y}$ (being the point-wise infimum of a collection of affine functionals on $\mathbb{Y}$).

Remark 3.1.3 Taking $u = 0$ at (3.4) we get

$$f(x) \geq K(x, y) \geq g(y), \quad (x, y) \in \mathbb{X} \times \mathbb{Y}. \tag{3.6}$$

From (3.6) we obtain the basic inequality

$$\inf_{x \in \mathbb{X}} f(x) \geq \sup_{y \in \mathbb{Y}} g(y). \tag{3.7}$$

The quantity on the left of (3.7) is called the *primal value* while the quantity on the right of (3.7) is called that *dual value*. If the primal value is *strictly greater* than the dual value then the difference between these quantities constitutes the so-called *duality gap*. We can never get anything useful out of the convex duality approach if the duality gap is non-zero (i.e. strictly positive). It is therefore essential to choose the space of perturbations $\mathbb{U}$, the space of dual variables $\mathbb{Y}$ and the bilinear form $\langle \cdot, \cdot \rangle$ in such a way that the the left and right sides of (3.7) are equal i.e. the duality gap is zero. Theorem 3.1.4 which follows next is an essential result which, among other things, establishes conditions which ensure that the duality gap is zero, and furthermore guarantees existence of a solution of the following *dual optimization problem*:

$$\text{maximize the dual function } g : \mathbb{Y} \to [-\infty, \infty]. \tag{3.8}$$

Theorem 3.1.4 (Rockafellar-Moreau) Suppose that $\langle \mathbb{U}, \mathbb{Y} \rangle$ is a given dual system and $\mathcal{U}$ is a $\langle \mathbb{U}, \mathbb{Y} \rangle$-compatible topology $\mathcal{U}$ on $\mathbb{U}$ (recall Definition C.1.10). If there exists some $x_1 \in \mathbb{X}$ and some $\mathcal{U}$-neighborhood G of the origin $0 \in \mathbb{U}$ such that

$$\sup_{u \in G} F(x_1, u) < \infty, \tag{3.9}$$

then

$$\inf_{x \in \mathbb{X}} f(x) = \sup_{y \in \mathbb{Y}} g(y) = g(\bar{y}) \quad \text{for some } \bar{y} \in \mathbb{Y}. \tag{3.10}$$

Remark 3.1.5 The vector spaces $\mathbb{U}$ and $\mathbb{Y}$, the convex function $F(\cdot, \cdot)$ on $\mathbb{X} \times \mathbb{U}$ and the bilinear form $\langle \cdot, \cdot \rangle$ on $\mathbb{U} \times \mathbb{Y}$ are at our discretion, although subject to (3.3); different choices of these yield different spaces of dual variables $\mathbb{Y}$ as well as different Lagrangian and dual functions. These items must be chosen with the following considerations in mind:

1. The conditions of Theorem 3.1.4 should hold;

2. The dual function at (3.5) should have a reasonably tractable form, so that we can obtain useful necessary conditions resulting from the optimality of $\bar{y}$ given by Theorem 3.1.4.

3. It should be possible to write the condition $f(x) = g(y)$(for an arbitrary pair $(x, y) \in \mathbb{X} \times \mathbb{Y}$) as reasonably explicit *Kuhn-Tucker* optimality relations, involving in particular transversality conditions, complementary slackness conditions, and feasibility conditions. These relations, together with the necessary conditions from (2), should furthermore be useful for constructing $\bar{x} \in \mathbb{X}$ in terms of the maximizer $\bar{y} \in \mathbb{Y}$ given by Theorem 3.1.4, such that $f(\bar{x}) = g(\bar{y})$.

Remark 3.1.6 In verifying the conditions of the Rockafellar-Moreau Theorem 3.1.4 one should really choose the $\langle \mathbb{U}, \mathbb{Y} \rangle$-compatible topology $\mathcal{U}$ to be the Mackey topology $\tau(\mathbb{U}, \mathbb{Y})$ since, according to the Mackey-Arens Theorem C.1.12, this is the $\langle \mathbb{U}, \mathbb{Y} \rangle$-compatible topology with the *largest collection* of open sets, making it easier to find a $\mathcal{U}$-neighborhood G of the origin $0 \in \mathbb{U}$ such that (3.9) holds for some $x_1 \in \mathbb{X}$. Unfortunately the Mackey topology is not very easy to deal with, since the abstract characterization of this topology given by Remark C.1.11 is of little use in concrete applications. As a practical matter, in most applications of Theorem 3.1.4 (including the applications in this thesis) the space of perturbations $\mathbb{U}$ is a *normed* vector space, and the space of dual variables $\mathbb{Y}$ is typically chosen to be the norm-dual of $\mathbb{U}$, that is

$$\mathbb{Y} = \mathbb{U}^{\star}, \tag{3.11}$$

with the canonical dual pair $\langle \mathbb{U}, \mathbb{Y} \rangle$ defined by

$$\langle u, y \rangle = y(u), \qquad u \in \mathbb{U}, \;\; y \in \mathbb{Y}. \tag{3.12}$$

It then follows from the Mackey Theorem C.1.13 that we can take $\mathcal{U}$ to be the norm-topology on $\mathbb{U}$ when applying Theorem 3.1.4, that is we need to show

$$\sup_{\substack{u \in \mathbb{U} \\ ||u|| < \epsilon}} F(x_1, u) < \infty, \tag{3.13}$$

for some $x_1 \in \mathbb{X}$ and some (small) $\epsilon \in (0, \infty)$ in order to verify (3.9) (here $|| \cdot ||$ denotes the norm on $\mathbb{U}$). This significantly simplifies the application of Theorem 3.1.4.

Chapter 4

Application of the Rockafellar-Moreau Approach to the QLM Problem 2.2.16

With the formulation (2.49) of the QLM Problem 2.2.16 in place we are in a position to apply the general approach of Rockafellar and Moreau summarized in Chapter 3 for synthesizing a vector space of *dual variables*, a *dual optimization problem*, and associated *optimality relations* which are the natural dual partner of our primal QLM Problem 2.2.16. Also needed for application of the Rockafellar-Moreau approach is a capital result in functional analysis due to Yosida and Hewitt [46] which decomposes the adjoint L_∞^* of the space L_∞ into 'regular" and "singular" parts, since the Lagrange multipliers associated with the state constraint $X^{\boldsymbol{\pi}}(T) \geq B$ a.s. in Problem 2.2.16 turn out to be members of this adjoint space. We summarize the necessary background on the Yosida-Hewitt decomposition in Appendix D.

4.1 Perturbation, Lagrangian and the Dual Problem

As indicated in Chapter 3, the Rockafellar Moreau approach gives a systematic method for constructing an appropriate vector space of dual variables, a Lagrangian, and a dual function. The first step is to fix some real linear space of perturbations $\mathbb{U}$, and some convex perturbation function.

Notation 4.1.1 From now on write L_p for the spaces $L_p(\Omega, \mathcal{F}_T, \mathbb{P})$ for all $p \in [1, \infty]$, where $\mathcal{F}_T$ is defined by (2.9). Similarly write L_∞^* for the norm-dual $L_\infty^*(\Omega, \mathcal{F}_T, \mathbb{P})$ of the real linear space $L_\infty(\Omega, \mathcal{F}_T, \mathbb{P})$.

We shall now implement the perturbational Rockafellar-Moreau approach summarized in Chapter 3 with the goal of synthesizing a vector space of dual variables, together with a Lagrangian, a dual function and optimality relations. According to the summary in Chapter 3 the first step in this approach is to define a linear space of perturbations and a perturbation function (see Step I in Chapter 3). Here we borrow an ingenious perturbation from Rockafellar and Wets [37] involving *essentially bounded* perturbations. It will be seen that this is exactly what is needed in order to associate an appropriate dual variable with the almost-sure inequality constraint $X^\pi(T) \geq B$ in QLM Problem 2.2.16. Our definition of the space of perturbations is therefore

$$\mathbb{U} := L_2 \times L_\infty. \tag{4.1}$$

Definition 4.1.2 Define the perturbation function $F : \mathbb{A} \times \mathbb{U} \to (-\infty, \infty]$ as

$$F(X, u) := \begin{cases} EJ(X(T) - u_1), & \text{when } X \in \mathbb{D} \text{ and } X(T) \geq B + u_2 \, a.s., \\ +\infty, & \text{otherwise,} \end{cases} \tag{4.2}$$

for all $X \in \mathbb{A}$, $u = (u_1, u_2) \in \mathbb{U}$.

Remark 4.1.3 Convexity of $F(\cdot, \cdot)$ on $\mathbb{A} \times \mathbb{U}$ follows from convexity of $J(\cdot)$ (see (2.42)). We also have the consistency relation that is emphasized in Step I of Chapter 3, namely

$$F(X, 0) = f(X), \qquad X \in \mathbb{A}, \tag{4.3}$$

as follows from (4.2) and (2.48).

In the search for dual solutions we are going to restrict attention to the vector subspace $\mathbb{B}_1 \subset \mathbb{B}$ given by

$$\mathbb{B}_1 := \{Y \equiv (Y_0, \Upsilon^Y, \boldsymbol{\xi}^Y, \boldsymbol{\Gamma}^Y) \in \mathbb{B} \,|\, \Upsilon^Y(t) = -r(t)Y(t), \, (\mathbb{P} \otimes Leb) - a.e.\}. \tag{4.4}$$

Define the $\mathbb{R}$-valued processes $\{\beta(t); t \in [0,T]\}$ and $\{\Xi(y,\boldsymbol{\lambda},\boldsymbol{\gamma})(t); t \in [0,T]\}$ (for $(y,\boldsymbol{\lambda},\boldsymbol{\gamma}) \in \mathbb{R} \times L_2(\mathbf{W}) \times L_2(\mathbf{M})$) as follows:

$$\beta(t) := \exp\left\{-\int_0^t r(\tau)\,\mathrm{d}\tau\right\}, \tag{4.5}$$

$$\Xi(y,\boldsymbol{\lambda},\boldsymbol{\gamma})(t) := \beta(t)\left\{y + \int_0^t \beta^{-1}(\tau)\boldsymbol{\lambda}^\top(\tau)\,\mathrm{d}\mathbf{W}(\tau) + \sum_{i,j=1}^{D}\int_0^t \beta^{-1}(\tau)\gamma_{i,j}(\tau)\,\mathrm{d}M_{ij}(\tau)\right\}. \tag{4.6}$$

The vector subspace $\mathbb{B}_1$ at (4.4), and the functions β at (4.5) and Ξ at (4.6), are borrowed from Donnelly [7], which addresses a simplified version of the QLM Problem 2.2.16 that includes the portfolio constraint $\pi \in \mathcal{A}$ but does not include the almost-sure state constraint $X^\pi(T) \geq B$. As was emphasized in Chapter 1 the presence of this state constraint effectively rules out the approach of [7] for the QLM Problem 2.2.16. It is nevertheless the case some of the technical machinery developed in [7] can be used and adapted in this thesis. We shall in fact see, although by an approach quite different from that followed in [7], that $\mathbb{B}_1$ defined by (4.4) is actually still the vector space of dual variables for the portfolio constraint $\pi \in \mathcal{A}$, exactly as it was in [7], and one of the major challenges of the present chapter is to come up with an appropriate vector space of dual variables for the state constraint $X^\pi(T) \geq B$ in the QLM Problem 2.2.16.

Elementary properties of the set $\mathbb{B}_1$ and the mapping Ξ are summarized in the next proposition, for later use in this chapter. The proof is a simple application of Itô's formula and is given in Appendix A

Lemma 4.1.4 Define the vector space $\mathbf{S} := \mathbb{R} \times L_2(\mathbf{W}) \times L_2(\mathbf{M})$ (where $L_2(\mathbf{W})$ and $L_2(\mathbf{M})$ are defined by (2.31) and (2.32)), and recall the representation (2.34) of general elements Y of $\mathbb{B}$. Then

1. $\mathbb{B}_1$ is a real linear space;
2. If $Y := \Xi(y,\boldsymbol{\lambda},\boldsymbol{\gamma})$ for some $(y,\boldsymbol{\lambda},\boldsymbol{\gamma}) \in \mathbf{S}$, we get

$$Y \equiv (y, -rY_-, \boldsymbol{\lambda}, \boldsymbol{\gamma}) \in \mathbb{R} \times L_{21} \times L_2(\mathbf{W}) \times L_2(\mathbf{M}). \tag{4.7}$$

   Moreover, recalling (4.4), we have

$$Y \in \mathbb{B}_1. \tag{4.8}$$

3. $\Xi : \mathbf{S} \to \mathbb{B}_1$ is a linear bijection.

We shall next implement Step II in the summary of the Rockafellar-Moreau method given in Chapter 3. This involves defining a vector space $\mathbb{Y}$ of dual variables together with a duality pairing of the vector spaces $\mathbb{Y}$ and $\mathbb{U}$. Define the real linear space $\mathbb{Y}$ as

$$\mathbb{Y} := \mathbb{B}_1 \times L_\infty^* \tag{4.9}$$

and bilinear form (see Definition B.4.5) $\langle \cdot, \cdot \rangle$ on $\mathbb{U} \times \mathbb{Y}$ as follows:

$$\langle (u_1, u_2), (Y, Z) \rangle := E[u_1 Y(T)] + Z(u_2), \quad (u_1, u_2) \in \mathbb{U}, \quad (Y, Z) \in \mathbb{Y}. \tag{4.10}$$

The motivation for the definition of $\mathbb{Y}$ at (4.9) is to be found in Remark 3.1.6 in Chapter 3. We want $\mathbb{Y}$ to be the norm-dual of the space of perturbations $\mathbb{U}$. Of course L_2 is the norm-dual of L_2, and we will see from Lemma 4.2.1 that $\mathbb{B}_1$ is isomorphic to L_2, hence we can take $\mathbb{B}_1$ to be the norm-dual of L_2 as we do in the first factor space at (4.9). As for the second factor space L_∞^* at (4.9), this is of course the norm-dual of the second factor space L_∞ at (4.1).

Remark 4.1.5 For $(Y, Z) \in \mathbb{Y}$ we have $Y \in \mathbb{B}_1$, thus $Y(T) \in L_2$ (from Lemma 2.2.4 and (4.4)), i.e. the right side of (4.10) is well-defined.

Motivated by (3.4) we define the *Lagrangian*

$$K(X, (Y, Z)) := \inf_{(u_1, u_2) \in \mathbb{U}} [\langle (u_1, u_2), (Y, Z) \rangle + F(X, u_1, u_2)], \quad X \in \mathbb{A},\ (Y, Z) \in \mathbb{Y}. \tag{4.11}$$

We next evaluate the right side of (4.11) explicitly. To this end, define sets $\mathbb{A}_1$ and $\mathbb{D}_1$ of Itô processes, and the convex conjugate $J^*(y, \omega)$ as

$$\begin{aligned} \mathbb{A}_1 &:= \{X \in \mathbb{A} \,|\, \text{there is some}\, \alpha \in \mathbb{R}\, \text{s.t.}\, X(T) \geq B + \alpha\, \text{a.s.}\}, && (4.12) \\ \mathbb{D}_1 &:= \mathbb{D} \cap \mathbb{A}_1 \neq \emptyset, && (4.13) \\ J^*(y, \omega) &:= \sup_{x \in \mathbb{R}} [xy - J(x, \omega)] \quad (y, \omega) \in \mathbb{R} \times \Omega. && (4.14) \end{aligned}$$

Observe that the non-emptyness of $\mathbb{D}_1$ asserted at (4.13) follows from Condition 2.2.22. The set $\mathbb{A}_1$ comprises all square-integrable Itô processes $X \in \mathbb{A}$ with property that $X(T) - B$ is *uniformly lower bounded* by a constant $\alpha \in \mathbb{R}$, while $\mathbb{D}_1$ is the set of all admissible wealth processes $X \in \mathbb{D}$ with the same property. The set $\mathbb{D}_1$ is therefore a sort of "reduced" set of admissible wealth processes. The reason for defining this set is that it arises very naturally when we evaluate the Lagrangian K in explicit form (see Proposition 4.1.7 which follows). In order to evaluate the Lagrangian we first explicitly calculate the convex conjugate J^* of the quadratic loss function J. The calculation is elementary but we work out all the details in Appendix A for completeness:

Lemma 4.1.6 For all $(y,\omega) \in \mathbb{R} \times \Omega$ we have

$$J^*(y,\omega) = \frac{(y-c(\omega))^2}{2a(\omega)} - q, \tag{4.15}$$

for the convex conjugate given by (4.14).

In Proposition 4.1.7 we explicitly calculate the Lagrangian function defined at (4.11). The proof of this proposition is in Appendix A:

Proposition 4.1.7 For each $X \in \mathbb{A}$ and each $(Y,Z) \in \mathbb{Y}$, the *Lagrangian* $K(X,(Y,Z))$ defined by (4.11) is explicitly given by

$$K(X,(Y,Z)) = \begin{cases} E[X(T)Y(T)] - E[J^*(Y(T))] + \inf\limits_{\substack{u_2 \in L_\infty \\ u_2 \leq X(T)-B}} Z(u_2), & \text{if } X \in \mathbb{D}_1 \ \& \ Z \leq 0, \\ -\infty, & \text{if } X \in \mathbb{D}_1 \ \& \ Z \nleq 0, \\ +\infty, & \text{if } X \notin \mathbb{D}_1. \end{cases} \tag{4.16}$$

Remark 4.1.8 From Equation (4.15), Lemma 2.2.4 and Condition 2.2.9 it follows that $X(T)Y(T)$ and $J^*(Y(T))$ are P-integrable so that the expectations in (4.16) are defined for each $(X,Y) \in \mathbb{A} \times \mathbb{B}_1$.

Remark 4.1.9 The infimum on the right of (4.16) (for $X \in \mathbb{D}_1$ and $Z \leq 0$) is just the usual Lagrange "weighting" of the constraint $X(T) \geq B$ by the "multiplier" $Z \in L_\infty^*$. This simplifies to the familiar form $Z(X(T)-B)$ when $X(T) \in L_\infty$, for then $X(T) - B \in L_\infty$, so that $Z(X(T)-B)$ is defined for each $Z \in L_\infty^*$. However, in general $X(T)$ is only square-integrable (recall that $X(T) \in L_2$), therefore this simplification is not available to us, and we must express the Lagrange weighting indirectly in terms of the infimum on the right of (4.16) (when $X \in \mathbb{D}_1$ and $Z \leq 0$).

Motivated by (3.5) and recalling $\mathbb{Y}$ defined at (4.9) we define *dual function*.

Definition 4.1.10 The *dual function* is the mapping $g : \mathbb{Y} \to [-\infty, \infty]$ given by

$$g(Y,Z) := \inf_{X \in \mathbb{A}} K(X,(Y,Z)), \quad (Y,Z) \in \mathbb{Y}, \tag{4.17}$$

where K is the Lagrangian given by (4.11).

In order to expand the dual function at Definition 4.1.10 put

$$\varkappa(Y,Z) := \sup_{X\in\mathbb{D}_1}\left\{-E[X(T)Y(T)] - \inf_{\substack{u_2\in L_\infty\\ u_2\leq X(T)-B}} Z(u_2)\right\} \qquad (Y,Z)\in\mathbb{Y}. \quad (4.18)$$

Then, for each $(Y,Z)\in\mathbb{Y}$ the dual function is given by

$$g(Y,Z) = \begin{cases} -\varkappa(Y,Z) - E[J^*(Y(T))], & Z\leq 0,\\ -\infty, & \text{otherwise,}\end{cases} \quad (4.19)$$

as follows immediately from (4.17), (4.18) and Proposition 4.1.7.

Remark 4.1.11 From Definition 2.2.26 of the primal function f, Definition 4.1.10 of the dual function g, and the definition of the Lagrangian K (see (4.11)) we immediately have the *weak duality relation*

$$f(X)\geq K(X,(Y,Z))\geq g(Y,Z), \quad X\in\mathbb{A}, \quad (Y,Z)\in\mathbb{Y}, \quad (4.20)$$

(exactly as at Remark 3.1.3).

4.2 Optimality Relations

We have defined the convex perturbation function $F(\cdot,\cdot)$ on $\mathbb{A}\times\mathbb{U}$ (see Definition 4.1.2), the space of dual variables (see (4.9)), the bilinear form $\langle\cdot,\cdot\rangle$ on $\mathbb{U}\times\mathbb{Y}$ (see (4.10)), constructed the Lagrangian K (see Proposition 4.1.7), and defined the dual function g (see Definition 4.1.10). With Remark 3.1.5 in mind, we shall now show in Proposition 4.2.3 that there exists a maximizer of the dual function with zero duality gap (see (E.20)), so that, in particular, *there exists a solution of the dual problem of maximizing the dual function g on the space of dual variables* $\mathbb{Y}$. We shall see that the Rockafellar-Moreau Theorem 3.1.4 is essential for establishing Proposition 4.2.3. We proceed further to establish Kuhn-Tucker optimality relations (see Remark 3.1.5) which will be used in the next section to construct an optimal wealth process $\bar{X}\in\mathbb{A}$ (see (2.50)) in terms of the maximizer of the dual function.

To this end we need the next result, which is a simple consequence of the martingale representation theorem. The elementary proof is relegated to Appendix A:

Lemma 4.2.1 Suppose Condition 2.1.9 and (2.1.10) and recall (4.4). Then for each $\zeta\in L_2$ (recall Notation 4.1.1), there is a unique $Y\in\mathbb{B}_1$, such that $Y(T)=\zeta$ a.s.

Remark 4.2.2 Proposition 4.2.3 which follows establishes that the duality gap is zero (again recall Remark 3.1.3) and that there exists a maximizer of the dual function g. This maximizer is the Lagrange multiplier which "enforces" the constraints in the QLM Problem 2.2.16, which is the central problem addressed in this thesis. The proposition is therefore central to the entire conjugate-duality approach of the thesis, since this relies completely on existence of Lagrange multipliers and a duality gap of zero. As such, Proposition 4.2.3 is the most important single result in the thesis. In view of this we give the proof of the proposition here in the main body of the present chapter, rather than relegating it to an appendix as we have done with most other proofs. As will be seen (at Remark 4.2.4 which follows) the Rockafellar-Moreau Theorem 3.1.4 and the Slater-type Condition 2.2.22 are the central elements of the proof:

Proposition 4.2.3 Suppose Conditions 2.1.9, 2.1.10 and 2.1.14. Then there exists a maximizer of the dual function $g : \mathbb{Y} \to [-\infty, \infty]$ (see Definition 4.1.10) with zero duality gap. That is, there exists some $(\bar{Y}, \bar{Z}) \in \mathbb{Y}$ such that

$$\inf_{X \in \mathbb{A}} f(X) = \sup_{(Y,Z)\in\mathbb{Y}} g(Y, Z) = g(\bar{Y}, \bar{Z}) \in \mathbb{R}. \tag{4.21}$$

Here f is the primal function given by Definition 2.2.26.

Proof. Define the "maximum" norm $||\cdot||_{\mathbb{U}}$ for $(u_1, u_2) \in \mathbb{U}$ as

$$||(u_1, u_2)||_{\mathbb{U}} := \max\{||u_1||_{L_2}, ||u_2||_{L_\infty}\}. \tag{4.22}$$

Let $\mathcal{U}$ be the norm-topology (topology generated by $||\cdot||_{\mathbb{U}}$) on $\mathbb{U}$. For $(Y, Z) \in \mathbb{Y}$, Lemma 2.2.4 implies that $Y(T) \in L_2$. Moreover, Z is a norm-continuous linear functional on L_∞. Hence the mapping $(u_1, u_2) \to \langle (u_1, u_2), (Y, Z)\rangle$ is clearly $\mathcal{U}$-continuous on $\mathbb{U}$ for each $(Y, Z) \in \mathbb{Y}$ (recall (4.10)).

Now fix some $\mathcal{U}$-continuous linear functional ϕ on $\mathbb{U}$. From (4.22) it is clear that $u_2 \to \phi(0, u_2)$ is a continuous linear functional on L_∞. By (D.14), we get

$$\phi(0, u_2) = Z(u_2), \tag{4.23}$$

for all $u_2 \in L_\infty$, for some $Z \in L_\infty^*$.
Also, again from (4.22), the mapping $u_1 \to \phi(u_1, 0)$ is norm-continuous on L_2, thus the Riesz representation theorem D.1.2 gives some $\zeta \in L_2$ such that

$$\phi(u_1, 0) = E[u_1\zeta], \tag{4.24}$$

for all $u_1 \in L_2$.
In view of Lemma 4.2.1, there is some $Y \in \mathbb{B}_1$ such that $Y(T) = \zeta$ almost surely. Hence we get

$$\phi(u_1, 0) = E[u_1 Y(T)], \tag{4.25}$$

for all $u_1 \in L_2$.
We thus get that

$$\begin{aligned}\phi(u_1, u_2) &= \phi(u_1, 0) + \phi(0, u_2) \\ &= E[u_1 Y(T)] + Z(u_2) \\ &= \langle (u_1, u_2), (Y, Z) \rangle \end{aligned} \tag{4.26}$$

for all $(u_1, u_2) \in \mathbb{U}$, as required to see that the topology $\mathcal{U}$ on $\mathbb{U}$ is $\langle \mathbb{U}, \mathbb{Y} \rangle$-compatible (see Definition C.1.10).

We now proceed to verify the remaining conditions of Theorem 3.1.4. Let $\tilde{X} := X^{\tilde{\boldsymbol{\pi}}}$, for $\tilde{\boldsymbol{\pi}} \in \mathcal{A}$ given by Condition 2.2.22. Then

$$\tilde{X} \in \mathbb{D}, \tag{4.27}$$

$$\tilde{X}(T) \geq B + u_2, \quad \forall\ u_2 \in L_\infty \,\text{s.t.} ||u_2||_{L_\infty} < \epsilon. \tag{4.28}$$

From Lemma 2.2.4 it follows that $\tilde{X}(T) \in L_2$. Then, from Condition 2.2.9 and Definition 2.2.8, one sees that the mapping

$$u_1 \to E[J(\tilde{X}(T) - u_1)] : L_2(\Omega, \mathcal{F}_T, \mathbb{P}) \to \mathbb{R} \tag{4.29}$$

is norm-continuous. From the continuity of this map, (4.27), (4.28) and the definition of perturbation function (4.1.2) it is immediate that

$$\sup_{\substack{(u_1,u_2)\in\mathbb{U} \\ ||(u_1,u_2)||_{\mathbb{U}}<\alpha}} F(\tilde{X}, (u_1, u_2)) < \infty \tag{4.30}$$

for some $\alpha \in (0, \infty)$. Now put

$$G := \{(u_1, u_2) \in \mathbb{U} \,|\, ||(u_1, u_1)||_{\mathbb{U}} < \alpha\}. \tag{4.31}$$

Then G is a $\mathcal{U}$-neighborhood G of the origin $(0, 0) \in \mathbb{U}$, and from (4.30) we have

$$\sup_{(u_1,u_2)\in G} F(\tilde{X}, (u_1, u_2)) < \infty. \tag{4.32}$$

Recall that $\tilde{X} \in \mathbb{A}$, that $\mathcal{U}$ is a $\langle \mathbb{U}, \mathbb{Y} \rangle$-compatible topology on $\mathbb{U}$, and $F(\cdot)$ is convex on $\mathbb{A} \times \mathbb{U}$ and satisfies the consistency relation stated in Remark 4.1.3. It follows that all conditions of Theorem 3.1.4 have been verified, and therefore

$$\inf_{X \in \mathbb{A}} f(X) = \sup_{(Y,Z) \in \mathbb{Y}} g(Y, Z) = g(\bar{Y}, \bar{Z}) \tag{4.33}$$

for some $(\bar{Y}, \bar{Z}) \in \mathbb{Y}$. The common value is finite, as follows from Remark 2.2.27.

Remark 4.2.4 Proposition 4.2.3 establishes existence of a Lagrange multiplier $(\bar{Y}, \bar{Z})$ for (2.49), together with zero duality gap. Notice that Proposition 4.2.3 is a consequence of Theorem 3.1.4, the use of which relies on showing (4.32), and notice the crucial role played by the Slater-type Condition 2.2.22 in establishing Proposition 4.2.3. In fact, one sees from the proof that it is the choice of essentially bounded perturbations u_2 at (4.2), together with the choice of the norm-topology for $\mathcal{U}$ on $\mathbb{U}$ (recall (4.1)), which secures (4.30) from Condition 2.2.22. This effectively mandates the definition of $\mathbb{Y}$ at (4.9) with L^*_∞ for the second factor, since, with this definition of $\mathbb{Y}$ and the bilinear form (4.10), the norm-topology $\mathcal{U}$ is $\langle \mathbb{U}, \mathbb{Y} \rangle$-compatible, as required to use Theorem 3.1.4.

Remark 4.2.5 The Lagrange multiplier $(\bar{Y}, \bar{Z})$ given by Proposition 4.2.3 is a pair in the vector space $\mathbb{Y} := \mathbb{B}_1 \times L^*_\infty$ (see (4.9)). In this pair the first member $\bar{Y} \in \mathbb{B}_1$ is the Lagrange multiplier which accounts for the *portfolio constraint* $\pi \in \mathcal{A}$, while the second member $\bar{Z} \in L^*_\infty$ is the Lagrange multiplier for the a.s. state constraint $X^\pi(T) \geq B$ in the QLM Problem 2.2.16. As we noted at Remark 2.2.19(c), removal of the constraint $X^\pi(T) \geq B$ from Problem 2.2.16 results in the problem addressed by Donnelly [7], which involves only the portfolio constraint $\pi \in \mathcal{A}$. Accordingly, the Lagrange multiplier in [7] should just be a process $\bar{Y} \in \mathbb{B}_1$ needed to account for the portfolio constraint. That this is the case is clear (see Proposition 4.8 and (4.23) in [7]).

Remark 4.2.6 For later use we make the following simple observations. Remark 4.1.8 and Proposition 4.2.3 ensure that the integral $EJ^*(\bar{Y}(T))$ exists in $\mathbb{R}$ and that $g(\bar{Y}, \bar{Z}) \in \mathbb{R}$. It then follows from (4.19) that $\bar{Z} \leq 0$ and $\varkappa(\bar{Y}, \bar{Z}) \in \mathbb{R}$.

Notation 4.2.7 From now on the notation $\partial J^*(y, \omega)$ indicates the gradient of the mapping $y \to J^*(y, \omega)$ given by Equation (A.36) for ω fixed, that is

$$\partial J^*(y, \omega) = (y - c(\omega))/a(\omega), \qquad (y, \omega) \in \mathbb{R} \times \Omega. \tag{4.34}$$

The next Proposition 4.2.8 gives *Kuhn-Tucker* optimality relations (see Remark 3.1.5) which are equivalent to the equality $f(X) = g(Y, Z)$ for arbitrary $X \in \mathbb{A}$ and arbitrary $(Y, Z) \in \mathbb{Y}$. The proof is placed in Appendix A.

Proposition 4.2.8 (Kuhn-Tucker optimality relations) Suppose Conditions 2.1.9, 2.1.10 and 2.2.9. For each $(X, (Y, Z)) \in \mathbb{A} \times \mathbb{Y}$, we have the equivalence

$$f(X) = g(Y, Z) \tag{4.35}$$

iff

$$X(T) - B \geq 0, X \in \mathbb{D}_1, Z \leq 0, \tag{4.36}$$

$$\inf_{\substack{u_2 \in L_\infty \\ u_2 \leq X(T) - B}} Z(u_2) = 0, \tag{4.37}$$

$$E[X(T)Y(T)] + \varkappa(Y, Z) = 0, \tag{4.38}$$

$$X(T) = (\partial J^*)(Y(T)). \tag{4.39}$$

Remark 4.2.9 Equations (4.36)-(4.39) constitute the *Kuhn-Tucker optimality relations* for the QLM Problem 2.2.16.. In particular, (4.36) gives *feasibility conditions* on the primal variable X and the dual variable Z. On the other hand (4.37) and (4.38) are *complementary slackness relations* for the constraints in the problem, that is relations between the primal variable X, the dual variable (Y, Z), and the *problem constraints*. In particular, (4.37) is the complementary slackness relation specific to the constraint $X(T) \geq B$ in the primal cost functional (2.48), while (4.38) is a further complementary slackness relation specific to the portfolio constraint $X \in \mathbb{D}$ in the primal cost functional (2.48). Finally, (4.39) is a *transversality relation* which (as is always the case) relates the primal variable X, the dual variable (X, Y), and the *problem cost functional* J at the right boundary point $t = T$. In Section 4.3 which follows we shall *construct* a process $\bar{X} \in \mathbb{A}$ in terms of the solution $(\bar{Y}, \bar{Z}) \in \mathbb{Y}$ of the dual problem (the existence of which is established by Proposition 4.2.3) such that $(\bar{X}, (\bar{Y}, \bar{Z})) \in \mathbb{A} \times \mathbb{Y}$ verifies the Kuhn-Tucker optimality relations (4.36) - (4.39). It then follows from the equivalence given by Proposition 4.2.8 that

$$f(\bar{X}) = g(\bar{Y}, \bar{Z}). \tag{4.40}$$

Upon combining (4.40) and (4.21) we find

$$f(\bar{X}) = \inf_{X \in \mathbb{A}} f(X), \tag{4.41}$$

as required to establish (2.50). It follows from Remark 2.2.28 that $\bar{X}$ is the optimal wealth process for the QLM problem 2.2.16.

4.3 Candidate Optimal Wealth and Portfolio Processes

Remark 4.3.1 In Proposition 4.2.3 we have established existence of a maximizer $(\bar{Y}, \bar{Z}) \in \mathbb{Y}$ of the dual function $g : \mathbb{Y} \to [-\infty, \infty]$. As noted at Remark 4.2.9, our goal in the present section is to construct an Ito process $\bar{X} \in \mathbb{A}$ (recall (2.36)) such that the triple $(\bar{X}, (\bar{Y}, \bar{Z}))$ satisfies the optimality relations (4.36) - (4.39) of Proposition 4.2.8. It then follows, as noted at Remark 4.2.9, that $\bar{X}$ is the optimal wealth process for the QLM problem 2.2.16. As a technical tool in the construction of the process $\bar{X}$ we shall need the following on semimartingales, which is an immediate consequence of Proposition I-1 of Bismut [3]:

Proposition 4.3.2 For any $(X_0, \Upsilon^X, \boldsymbol{\xi}^X, \boldsymbol{\Gamma}^X) \in \mathbb{B}$ and $(Y_0, \Upsilon^Y, \boldsymbol{\xi}^Y, \boldsymbol{\Gamma}^Y) \in \mathbb{B}$ (recall (2.33) - (2.34)), the process $\{\mathbb{M}(X,Y)(t) : t \in [0,T]\}$ defined by

$$\begin{aligned} \mathbb{M}(X,Y)(t) \quad &:= \quad X(t)Y(t) - X_0 Y_0 - \int_0^t (\Upsilon^X(\tau)Y(\tau) + X(\tau)\Upsilon^Y(\tau))\mathrm{d}\tau \\ &\quad - \sum_{n=1}^{N} \int_0^t \xi_n^X(\tau)\xi_n^Y(\tau)\,\mathrm{d}\tau - \sum_{i,j=1}^{D} \int_0^t \Gamma_{i,j}^X(\tau)\Gamma_{i,j}^Y(\tau)\,\mathrm{d}[M_{ij}](\tau) \end{aligned} \tag{4.42}$$

is a martingale, more precisely $\mathbb{M}(X,Y) \in \mathcal{M}_0(\{\mathcal{F}_t\}, \mathbb{P})$.

Remark 4.3.3 Proposition 4.2.3 ensures that $g(\bar{Y}, \bar{Z}) \in \mathbb{R}$ and Remark 4.1.8 ensures that $E[J^*(\bar{Y}(T))]$ exists in $\mathbb{R}$. Now it follows from (4.19) that $\bar{Z} \leq 0$ and $\varkappa(\bar{Y}, \bar{Z}) \in \mathbb{R}$. These observations will be used several times in what follows.

We shall need several useful correspondences between sets and functions. The simplest associates with each set C in a linear space $\mathbb{X}$ the characteristic function $\bar{\delta}_{\mathbb{X}}(\cdot|C)$ of $C \subset \mathbb{X}$ which is defined as follows:

Definition 4.3.4 The characteristic function $\bar{\delta}(x|C)$ of a set C in a linear space $\mathbb{X}$ is defined as

$$\bar{\delta}_{\mathbb{X}}(x|C) = \begin{cases} 0, & \text{if } x \in C \\ +\infty, & \text{if } x \notin C. \end{cases} \tag{4.43}$$

Remark 4.3.5 C is a convex set if and only if $\bar{\delta}_{\mathbb{X}}(\cdot|C)$ is a convex function on $\mathbb{X}$. Characteristic functions play a fundamental role in convex analysis, and should not be confused with the familiar indicator functions of real analysis (see (D.15)). We shall always refer to (4.43) as a *characteristic function* and (D.15) as an *indicator function.*

We shall next establish several *necessary conditions* resulting from the optimality of $(\bar{Y}, \bar{Z}) \in \mathbb{Y}$ given by Proposition 4.2.3, that is, we shall obtain several consequences of the fact that

$$g(\bar{Y}, \bar{Z}) \geq g(Y, Z) \qquad \text{for all } (Y, Z) \in \mathbb{Y}. \tag{4.44}$$

To this end we shall also need the following technical result which effectively gives the *support function* (see Lemma B.1.1) of the set of admissible wealth processes $\mathbb{D}$ (see (2.47)). The proof of Proposition 4.3.6 is located in Appendix A.

Proposition 4.3.6 Suppose Conditions 2.1.9, 2.1.10, 2.1.14 and recall (B.1). For each $Y \in \mathbb{B}_1$ we have

$$\sup_{X \in \mathbb{D}} E[-X(T)Y(T)] \;\; = \;\; -x_0 Y_0 + E \int_0^T \delta(-\boldsymbol{\sigma}(\tau)[\boldsymbol{\theta}(\tau) Y(\tau) + \boldsymbol{\xi}^Y(\tau)] | K) \mathrm{d}\tau. \tag{4.45}$$

Here $x_0 \in (0, \infty)$ is the given nonrandom initial wealth (see (2.28) and Remark 2.2.17).

We shall now establish the first of several *necessary conditions* resulting from the optimality of $(\bar{Y}, \bar{Z})$ for the dual problem. We include the proof here (instead of relegating it to Appendix A) because the argument illustrates the essential idea of exploiting the all important optimality of $(\bar{Y}, \bar{Z})$ at (4.44).

Proposition 4.3.7 Suppose Conditions 2.1.9, 2.1.10, 2.1.14 and recall (B.1) and Proposition 4.2.3. For each $Y \in \mathbb{B}_1$ we have

$$E[\partial J^*(\bar{Y}(T)) Y(T)] - x_0 Y_0 + E \int_0^T \delta(-\boldsymbol{\sigma}(\tau)[\boldsymbol{\theta}(\tau) Y(\tau) + \boldsymbol{\xi}^Y(\tau)] | K) \mathrm{d}\tau \geq 0, \tag{4.46}$$

where $x_0 \in (0, \infty)$ is again the nonrandom initial wealth.

Proof. From Proposition 4.2.3 (i.e. from (4.44)), it follows that, for every $\epsilon \in (0, \infty)$ and $Y \in \mathbb{B}_1$, we have

$$g(\bar{Y} + \epsilon Y, \bar{Z}) \leq g(\bar{Y}, \bar{Z}) \in \mathbb{R}, \tag{4.47}$$

and then, from (4.47) and (4.19), we get

$$-\varkappa(\bar{Y} + \epsilon Y, \bar{Z}) - E[J^*(\bar{Y}(T) + \epsilon Y(T))] \leq -\varkappa(\bar{Y}, \bar{Z}) - E[J^*(\bar{Y}(T))]. \tag{4.48}$$

Since $\varkappa(\bar{Y}, \bar{Z}) \in \mathbb{R}$ (see Remark 4.3.3), we then obtain from (4.48)

$$\varkappa(\bar{Y} + \epsilon Y, \bar{Z}) - \varkappa(\bar{Y}, \bar{Z}) + E[J^*(\bar{Y}(T) + \epsilon Y(T))] - E[J^*(\bar{Y}(T))] \geq 0, \tag{4.49}$$

(the simple arithmetical operation at (4.49) is inadmissible if $\varkappa(\bar{Y}, \bar{Z}) = +\infty$). Again using $\varkappa(\bar{Y}, \bar{Z}) \in \mathbb{R}$ (by Remark 4.3.3), from (4.18) we have

$$\varkappa(\bar{Y} + \epsilon Y, \bar{Z}) - \varkappa(\bar{Y}, \bar{Z}) \leq \epsilon \sup_{X \in \mathbb{D}_1} E[-X(T)Y(T)]. \tag{4.50}$$

for all $\epsilon \in (0, \infty)$ and $Y \in \mathbb{B}_1$. Since $\mathbb{D}_1 \subset \mathbb{D}$ (recall (4.13)), from (4.49) and (4.50) we obtain

$$\sup_{X \in \mathbb{D}} E[-X(T)Y(T)] + E[\frac{J^*(\bar{Y}(T) + \epsilon Y(T)) - J^*(\bar{Y}(T))}{\epsilon}] \geq 0, \tag{4.51}$$

for all $\epsilon \in (0, \infty)$ and $Y \in \mathbb{B}_1$.

In view of Lemma 4.1.6 and the dominated convergence theorem, we can take $\epsilon \to 0$ at (4.51) to get

$$\sup_{X \in \mathbb{D}} E[-X(T)Y(T)] + E[\partial J^*(\bar{Y}(T))Y(T)] \geq 0, \tag{4.52}$$

for each $Y \in \mathbb{B}_1$. Now (4.46) follows from (4.45) and (4.52).

We will later use the inequality (4.46) as tool for constructing the optimal wealth process $\bar{X}$. For this construction we shall also need the *state price density* H defined in terms of the market price of risk $\boldsymbol{\theta}$ (see Definition 2.1.17) and bond price S_0 (see (2.19)) as follows:

Definition 4.3.8 The *state price density* is the process $H : \Omega \times [0, T] \to \mathbb{R}$ given by

$$H(t) = [S_0(t)]^{-1} \exp\left(- \int_0^t \boldsymbol{\theta}^\top(\tau) \, \mathrm{d}\mathbf{W}(\tau) - \tfrac{1}{2} \int_0^t \| \boldsymbol{\theta}(\tau) \|^2 \, \mathrm{d}\tau \right), \ t \in [0, T]. \tag{4.53}$$

The following proposition establishes that the state-price density H is strongly bounded. The proof, which just involves a routine boundedness computation, is in Appendix A:

Proposition 4.3.9 For every $p \in \mathbb{R}$ we have

$$E\left(\sup_{t\in[0,T]} |H(t)|^p\right) < \infty. \tag{4.54}$$

The following result is again elementary, giving the Ito process components of the state-price density H (recall Remark 2.2.2). The proof is relegated to Appendix A.

Proposition 4.3.10 Recalling (2.36), we have $H \equiv (1, -rH, -H\boldsymbol{\theta}, 0) \in \mathbb{A}$. That is

$$H_0 = 1,\ \Upsilon^H = -rH \in L_{21},\ \boldsymbol{\xi}^H = -H\boldsymbol{\theta} \in L_2(\mathbf{W}) \text{ and } \boldsymbol{\Gamma}^H = \mathbf{0} \in L_2(\mathbf{M}). \tag{4.55}$$

The next result is another tool that we shall need for constructing an optimal portfolio $\bar{X}$ in terms of the Lagrange multipliers (i.e. the maximizer $(\bar{Y}, \bar{Z}) \in \mathbb{Y}$ given by Proposition 4.2.3). The proof is rather lengthy but uses only elementary ideas of stochastic calculus, and is placed in Appendix A:

Proposition 4.3.11 Fix some $\zeta \in L_2$ (recall Notation 4.1.1), and define the $\mathbb{R}$-valued process $\{X(t)\}$ in terms of ζ and the state price density H (see Definition 4.3.8) as follows:

$$X(t) := H(t)^{-1}E[\zeta H(T)|\mathcal{F}_t],\ t \in [0, T]. \tag{4.56}$$

Then the following hold:

(a) X is square integrable, that is

$$E\left(\sup_{t\in[0,T]} |X(t)|^2\right) < \infty. \tag{4.57}$$

(b) For X defined as in (4.56) and the state price density H(see (4.53)), the product process XH is both a martingale and a locally square integrable martingale. More precisely (recall Definition B.2.6, Definition B.2.14, Remark B.2.7 and Remark B.2.16) we have

$$XH \in \mathcal{M}_2^{loc}(\{\mathcal{F}_t\}, \mathbb{P}) \cap \mathcal{M}(\{\mathcal{F}_t\}, \mathbb{P}), \tag{4.58}$$

(note that one does not have $\mathcal{M}_2^{loc}(\{\mathcal{F}_t\}, \mathbb{P}) \subset \mathcal{M}(\{\mathcal{F}_t\}, \mathbb{P})$ although it is always the case that $\mathcal{M}_2(\{\mathcal{F}_t\}, \mathbb{P}) \subset \mathcal{M}(\{\mathcal{F}_t\}, \mathbb{P})$).

(c) XH can be expanded as a sum of stochastic integrals, in particular there exist integrand processes

$$\tilde{\boldsymbol{\xi}} = (\tilde{\xi}_1, \dots, \tilde{\xi}_N)^\top \in L_2^{loc}(\mathbf{W}) \tag{4.59}$$

$$\text{and} \qquad \tilde{\boldsymbol{\Gamma}} = (\tilde{\Gamma}_{ij})_{i,j=1}^D \in L_2^{loc}(\mathbf{M}) \tag{4.60}$$

such that

$$X(t)H(t) = X_0 + \sum_{n=1}^N \int_0^t \tilde{\xi}_n(\tau)\, \mathrm{d}W_n(\tau) + \sum_{i,j=1}^D \int_0^t \tilde{\Gamma}_{i,j}(\tau)\, \mathrm{d}M_{ij}(\tau). \tag{4.61}$$

(d) The $\mathbb{R}^N$-valued process $\{\tilde{\boldsymbol{\pi}}(t),\ t \in [0,T]\}$ defined by

$$\tilde{\boldsymbol{\pi}}(t) := [\boldsymbol{\sigma}^\top]^{-1}\{H^{-1}(t)\tilde{\boldsymbol{\xi}}(t) + X(t_-)\boldsymbol{\theta}(t)\}, \quad t \in [0,T], \tag{4.62}$$

pathwise square integrable, that is

$$\int_0^T \parallel \tilde{\boldsymbol{\pi}}(t) \parallel^2 \ \mathrm{d}t < \infty \quad \text{a.s.} \tag{4.63}$$

(e) The process X given by (4.56) is a semimartingale of the form

$$\begin{aligned} X(t) \quad = \quad & X_0 + \int_0^t \{r(\tau)X(t) + \tilde{\boldsymbol{\pi}}^\top(\tau)\boldsymbol{\sigma}(\tau)\boldsymbol{\theta}(\tau)\}\mathrm{d}\tau + \int_0^t \tilde{\boldsymbol{\pi}}^\top(\tau)\boldsymbol{\sigma}(\tau)\, \mathrm{d}\mathbf{W}(\tau) \\ & + \sum_{i,j=1}^D \int_0^t H(\tau)^{-1}\tilde{\Gamma}_{ij}(\tau)\, \mathrm{d}M_{ij}(\tau). \end{aligned} \tag{4.64}$$

(f) The process $\tilde{\boldsymbol{\pi}}$ (see (4.62)) and $\boldsymbol{\Gamma}$ (recall (4.60)) are such that

$$\boldsymbol{\sigma}^\top \tilde{\boldsymbol{\pi}} \in L_2(\mathbf{W}) \quad \text{and} \quad H^{-1}\tilde{\boldsymbol{\Gamma}} \in L_2(\mathbf{M}). \tag{4.65}$$

Corollary 4.3.12 Equation (4.63) can be strengthened to

$$\tilde{\boldsymbol{\pi}} \in L_2(\mathbf{W}). \tag{4.66}$$

Proof. Follows from (4.65) and the uniform boundedness of $\boldsymbol{\sigma}$.

Corollary 4.3.13 Recalling (2.33), we have

$$X \equiv (X_0, rX_- + \tilde{\boldsymbol{\pi}}\boldsymbol{\sigma}\boldsymbol{\theta}, \boldsymbol{\sigma}^\top \tilde{\boldsymbol{\pi}}, H^{-1}\tilde{\boldsymbol{\Gamma}}) \in \mathbb{B} \tag{4.67}$$

Proof. From (4.65), we have $\boldsymbol{\sigma}^\top \tilde{\boldsymbol{\pi}} \in L_2(\mathbf{W})$ and $H^{-1}\tilde{\boldsymbol{\Gamma}} \in L_2(\mathbf{M})$. So we need only show that $rX_- + \tilde{\boldsymbol{\pi}}^\top \boldsymbol{\sigma}\boldsymbol{\theta} \in L_{21}$. However, from (4.57), we have $E\left(\sup_{t\in[0,T]} |X(t)|^2\right) < \infty$ and from Corollary 4.3.12, $E\int_0^T \parallel \tilde{\boldsymbol{\pi}}(t) \parallel^2 \,\mathrm{d}t < \infty$. From these two facts and uniform boundedness of $r, \boldsymbol{\sigma}$, and $\boldsymbol{\theta}$, we have $rX_- + \tilde{\boldsymbol{\pi}}\boldsymbol{\sigma}\boldsymbol{\theta} \in L_{21}$.

Remark 4.3.14 It is evident from (4.56) that the process X is *completely determined* by the choice of the random variable $\zeta \in L_2$. The integrand processes $\tilde{\boldsymbol{\xi}}$ and $\tilde{\boldsymbol{\Gamma}}$ at (4.59) - (4.60) are obtained from application of the martingale representation theorem (see Theorem G.3.5) to the locally square-integrable martingale XH (see (4.58)); since X, and therefore XH, is completely determined by the random variable $\zeta \in L_2$, it follows that the integrand processes $\tilde{\boldsymbol{\xi}}$ and $\tilde{\boldsymbol{\Gamma}}$ are themselves *completely determined* by the random variable $\zeta \in L_2$. It then follows from (4.62) that the process $\tilde{\boldsymbol{\pi}}$ is also *completely determined* by the random variable $\zeta \in L_2$. Later (see (4.69)) we are going to use the component $\bar{Y}$ of the dual optimal solution $(\bar{Y}, \bar{Z})$ given by Proposition 4.2.3 to appropriately construct the random variable $\zeta \in L_2$ such that X defined at (4.56) is actually the *optimal wealth process* and $\tilde{\boldsymbol{\pi}}$ defined at (4.62) is the corresponding *optimal portfolio process* for the QLM Problem 2.2.16. In order to carry out this program we shall need the following result on a particular type of linear stochastic integral equation. This result is established in Donnelly *et-al* (see Lemma A.3 of [7]). For completeness the proof is also given in Appendix A.

Lemma 4.3.15 For each $\boldsymbol{\rho} \in L_2(\mathbf{W})$ and $\boldsymbol{\gamma} = (\gamma)_{i,j=1}^D \in L_2(\mathbf{M})$, there exists $\boldsymbol{\lambda} \in L_2(\mathbf{W})$ such that for all $t \in [0,T]$ we have a.s.

$$\boldsymbol{\lambda}(t) + \boldsymbol{\theta}(t) \int_0^t \boldsymbol{\lambda}^\top(\tau)\,\mathrm{d}\mathbf{W}(\tau) = \boldsymbol{\rho}(t) - \boldsymbol{\theta}(t) \sum_{i,j=1}^D \int_0^t \gamma_{ij}(\tau)\,\mathrm{d}M_{ij}(\tau), \tag{4.68}$$

where $\boldsymbol{\lambda}$ is unique in the sense that if there exists $\bar{\boldsymbol{\lambda}} \in L_2(\mathbf{W})$ such that (4.68) holds with $\boldsymbol{\lambda}$ replaced by $\bar{\boldsymbol{\lambda}}$, then $\boldsymbol{\lambda} = \bar{\boldsymbol{\lambda}}$ ($\mathbb{P} \otimes Leb$)-a.e.

Remark 4.3.16 In the present section we have accumulated several technical tools, in particular Propositions 4.3.2, 4.3.7, and 4.3.11, Corollaries 4.3.12 and 4.3.13, and Lemma 4.3.15. With these tools we are now finally ready to begin construction of the optimal wealth process $\bar{X}$ (recall Remark 4.3.1). Recalling $\bar{Y} \in \mathbb{B}_1$ given by Proposition 4.2.3 define

$$\zeta := \partial J^*(\bar{Y}). \tag{4.69}$$

We then have

$$\zeta \in L_2, \tag{4.70}$$

as follows immediately from Notation 4.2.7, Lemma 2.2.4 and Condition 2.2.9. All results in Proposition 4.3.11 therefore hold in particular when ζ is defined at (4.69). We now define a *candidate optimal wealth process* $\bar{X}$ in accordance with (4.56) with ζ defined at (4.69) in terms of the component $\bar{Y}$ of the dual solution $(\bar{Y}, \bar{Z})$ (see (4.44)), that is

$$\bar{X}(t) := H(t)^{-1} E[\partial J^*(\bar{Y}(T)) H(T) | \mathcal{F}_t], \quad t \in [0, T]. \tag{4.71}$$

The remainder of this section is devoted to showing that $\bar{X} \in \mathbb{A}$, and that the triplet $(\bar{X}, (\bar{Y}, \bar{Z})) \in \mathbb{A} \times \mathbb{Y}$ (recall Proposition 4.2.3) satisfies the optimality relations (4.36) - (4.39) of Proposition 4.2.8. It then follows that $\bar{X}$ is the optimal wealth process for the QLM problem 2.2.16 (see Remark 4.3.1).

In view of Proposition 4.3.11(c) we have

$$\bar{X}(t)H(t) = \bar{X}_0 + \sum_{n=1}^{N} \int_0^t \bar{\xi}_n(\tau)\, \mathrm{d}W_n(\tau) + \sum_{i,j=1}^{D} \int_0^t \bar{\Gamma}_{i,j}(\tau)\, \mathrm{d}M_{ij}(\tau), \tag{4.72}$$

for some integrand processes

$$\bar{\boldsymbol{\xi}} = (\bar{\xi}_1, \ldots, \bar{\xi}_N)^\top \in L_2^{loc}(\mathbf{W}) \tag{4.73}$$

$$\text{and} \qquad \bar{\boldsymbol{\Gamma}} = \left(\bar{\Gamma}_{ij}\right)_{i,j=1}^{D} \in L_2^{loc}(\mathbf{M}). \tag{4.74}$$

The integrand processes $\bar{\boldsymbol{\xi}}$ and $\bar{\boldsymbol{\Gamma}}$ are just the integrand processes $\tilde{\boldsymbol{\xi}}$ and $\tilde{\boldsymbol{\Gamma}}$ at Proposition 4.3.11(c) for the particular case of ζ given by (4.69). Motivated by Proposition 4.3.11(d) we also define the *candidate optimal portfolio process* in terms of $\bar{X}$ and the integrand $\bar{\boldsymbol{\xi}}$ at (4.73) in accordance with (4.62), that is

$$\bar{\boldsymbol{\pi}}(t) := [\boldsymbol{\sigma}^\top]^{-1} \{ H^{-1}(t) \bar{\boldsymbol{\xi}}(t) + \bar{X}(t_-) \boldsymbol{\theta}(t) \}, \quad t \in [0, T]. \tag{4.75}$$

Remark 4.3.17 It is most important to understand that the process $\bar{X}$ defined at (4.71), and the process $\bar{\boldsymbol{\pi}}$ defined at (4.75), are each *completely determined* by the component $\bar{Y}$ of the optimal dual solution $(\bar{Y}, \bar{Z})$ given by Proposition 4.2.3. This is immediate from the dependency on general $\zeta \in L_2$ noted at Remark 4.3.14 and the definition of ζ in terms of $\bar{Y}$ (see (4.69) and (4.70)). It will be seen that the *optimality* of $(\bar{Y}, \bar{Z})$ (see (4.44)) is the essential thing for establishing that $\bar{X}$ is the optimal wealth process and $\bar{\boldsymbol{\pi}}$ is the corresponding optimal portfolio for the QLM Problem 2.2.16.

Remark 4.3.18 In the remainder of this section we are going to show that

$$\bar{\boldsymbol{\pi}} \in \mathcal{A}, \tag{4.76}$$

that is $\bar{\boldsymbol{\pi}}$ defined by (4.75) is an *admissible* portfolio process (recall Definition 2.2.12) and that

$$\bar{X} \in \mathbb{D} \qquad \text{with} \qquad \bar{X} = X^{\bar{\boldsymbol{\pi}}}, \tag{4.77}$$

that is the process $\bar{X}$ defined at (4.71) is an admissible wealth process (recall (2.47)) and is actually the wealth process corresponding to the admissible portfolio $\bar{\boldsymbol{\pi}}$ (recall (2.28)). We shall then verify that the pair $(\bar{X}, (\bar{Y}, \bar{Z}))$ satisfies the optimality relations (4.36)-(4.39) (recall Remark 4.3.16). Since all the preceding constructions in the present section have been quite extensive we shall first briefly recall and highlight the main results which have been established so far: From Proposition 4.3.11(a) and Remark 4.3.16 it follows that $\bar{X}$ defined at 4.71 is bounded in the following sense:

$$E\left(\sup_{t\in[0,T]} |\bar{X}(t)|^2\right) < \infty. \tag{4.78}$$

Moreover, from Corollary 4.3.12, Remark 4.3.16 and (4.75), we have

$$\bar{\boldsymbol{\pi}} \in L_2(\mathbf{W}). \tag{4.79}$$

Finally, recalling $\bar{\boldsymbol{\xi}}$ and $\bar{\boldsymbol{\Gamma}}$ as defined in (4.72), Corollary 4.3.13 and Remark 4.3.16 we get

$$\bar{X} \equiv (\bar{X}_0, r\bar{X}_- + \bar{\boldsymbol{\pi}}\boldsymbol{\sigma}\boldsymbol{\theta}, \boldsymbol{\sigma}^\top\bar{\boldsymbol{\pi}}, H^{-1}\bar{\boldsymbol{\Gamma}}) \in \mathbb{B}. \tag{4.80}$$

Recalling the equivalence of (2.34) and (2.35), we can write (4.80) in expanded form to see that $\bar{X}$ defined at (4.71) satisfies the integral relation

$$\begin{aligned}\bar{X}(t) = \bar{X}_0 \quad &+ \quad \int_0^t [r(\tau)\bar{X}(\tau_-) + \bar{\boldsymbol{\pi}}^\top(\tau)\boldsymbol{\sigma}(\tau)\boldsymbol{\theta}(\tau)]\mathrm{d}\tau \\ &+ \quad \int_0^t \bar{\boldsymbol{\pi}}^\top(\tau)\boldsymbol{\sigma}(\tau)\,\mathrm{d}W(\tau) + \sum_{i,j=1}^{D}\int_0^t H^{-1}(\tau)\bar{\Gamma}_{i,j}\,\mathrm{d}M_{ij}(\tau). \quad (4.81)\end{aligned}$$

We now proceed to establish that (4.76) and (4.77) hold. For this we will need the following result, which is just a slight refinement of Proposition 4.3.7:

Lemma 4.3.19 Suppose Conditions 2.1.9,2.1.10,2.1.14,2.2.9 and 2.2.22 and recall $\bar{X}$ and $\bar{\boldsymbol{\pi}}$ defined at (4.71) and (4.75) respectively in terms of $\bar{Y} \in \mathbb{B}_1$ given by Proposition 4.2.3. Then for each $Y \in \mathbb{B}_1$, we have

$$\begin{aligned}(\bar{X}_0 - x_0)Y_0 + E\int_0^t \bar{\boldsymbol{\pi}}^\top(\tau)\boldsymbol{\sigma}(\tau)[\boldsymbol{\theta}(\tau)Y(\tau) + \boldsymbol{\xi}^Y(\tau)]\,\mathrm{d}\tau \\ +E\sum_{i,j=1}^{D}\int_0^t H(\tau)^{-1}\bar{\Gamma}_{ij}(\tau)\Gamma_{ij}^Y(\tau)\,\mathrm{d}[M_{ij}](\tau) \\ +E\int_0^T \delta(\boldsymbol{\sigma}(\tau)[\boldsymbol{\theta}(\tau)Y(\tau) + \boldsymbol{\xi}^Y(\tau)]|K)\mathrm{d}\tau \geq 0. \quad (4.82)\end{aligned}$$

The elementary proof of Lemma 4.3.19 is placed in Appendix A.

Lemma 4.3.19 and Lemma 4.3.15 are now used as technical tools for establishing Proposition 4.3.20 which follows. From this result it will be clear that (4.76) and (4.77) hold for $\bar{\boldsymbol{\pi}}$ defined at (4.75) and $\bar{X}$ defined at (4.71) (see the discussion in Remark 4.3.21 which follows). The proof of Proposition 4.3.20 is somewhat lengthy and technical and is placed in Appendix A.

Proposition 4.3.20 Suppose Conditions 2.1.9,2.1.10,2.1.14,2.2.9 and 2.2.22 and recall $\bar{X}$ and $\bar{\boldsymbol{\pi}}$ defined at (4.71) and (4.75) respectively in terms of $\bar{Y} \in \mathbb{B}_1$ given by Proposition 4.2.3. Then

(a) x_0 being the given initial wealth of investor (recall (2.28) and Remark 2.2.17), we have

$$\bar{X}(0) \equiv \bar{X}_0 = x_0, \quad (4.83)$$

that is, the value at $t = 0$ of the semimartingale $\bar{X}$ defined at (4.71) is equal to the initial wealth x_0.

(b) For the candidate portfolio process $\bar{\boldsymbol{\pi}}$ given by (4.75), we have

$$\bar{\boldsymbol{\pi}}(t) \in K \ \ (\mathbb{P} \otimes Leb) - \text{a.e.}, \qquad \text{and therefore} \qquad \bar{\boldsymbol{\pi}} \in \mathcal{A}, \tag{4.84}$$

that is, the candidate portfolio process $\bar{\boldsymbol{\pi}}$ satisfies the stipulated portfolio constraints, and is therefore an admissible portfolio process for the QLM Problem (recall Definition 2.2.12 and Problem 2.2.16).

(c) For $\bar{\boldsymbol{\Gamma}}$ at (4.81), we have

$$\bar{\boldsymbol{\Gamma}} = 0 \quad \nu_{[\mathbf{M}]} - \text{a.e.} \tag{4.85}$$

that is, the $\mathrm{d}M_{ij}$-stochastic integral on the right side of (4.81) is identically zero.

Remark 4.3.21 From Proposition 4.3.20(b) we know that $\bar{\boldsymbol{\pi}}$ is an admissible portfolio process. Moreover, from Proposition 4.3.20(a)(c) together with (4.81), we have

$$\begin{aligned} \bar{X}(t) = x_0 &+ \int_0^t [r(\tau)\bar{X}(\tau) + \bar{\boldsymbol{\pi}}^\top(\tau)\boldsymbol{\sigma}(\tau)\boldsymbol{\theta}(\tau)] \, \mathrm{d}\tau \\ &+ \int_0^t \bar{\boldsymbol{\pi}}^\top(\tau)\boldsymbol{\sigma}(\tau) \, \mathrm{d}W(\tau), \end{aligned} \tag{4.86}$$

and in particular

$$\bar{X} \in \mathbb{A}, \tag{4.87}$$

(see (2.36)). Upon comparing (4.86) with the general relation (2.28) we see that $\bar{X} = X^{\bar{\boldsymbol{\pi}}}$. That is, the semimartingale $\bar{X}$ defined at (4.71) in terms of the component $\bar{Y}$ of the optimal dual solution $(\bar{Y}, \bar{Z})$ given by Proposition 4.2.3, is actually the wealth process corresponding to the admissible portfolio $\bar{\boldsymbol{\pi}} \in \mathcal{A}$, that is $\bar{X} \in \mathbb{D}$ (see (2.47)). We have therefore established (4.76) and (4.77).

Continuing the program summarized in Remark 4.3.18 we next verify that $(\bar{X}, (\bar{Y}, \bar{Z}))$ satisfies the optimality relations (4.36)-(4.39). In order to accomplish this we need to extract yet another necessary condition resulting from the optimality of $(\bar{Y}, \bar{Z})$ obtained in Proposition 4.2.3 (see also (4.44)). Recall that we have already obtained one such necessary condition in Proposition 4.3.7, and that this necessary condition was the essential thing needed to obtain (4.76) and (4.77). In the course of establishing Proposition 4.3.7 we varied *only* the first argument of the dual function $g(\cdot,\cdot)$ while keeping the second argument fixed at $\bar{Z}$ (see (4.47)). In Proposition 4.3.22 which follows we shall obtain yet another necessary condition from the optimality of $(\bar{Y},\bar{Z})$ (which complements the necessary condition at Proposition 4.3.7), by varying *both* arguments (Y, Z) in a particular way (see (4.90) which follows). The necessary condition of Proposition 4.3.22 will itself be essential for establishing that $\bar{X}$ satisfies the terminal wealth constraint, that is $\bar{X}(T) \geq B$, a.s. (i.e. that $\bar{X}$ satisfies the first relation of (4.36)). Since Proposition 4.3.22 relies on the all-important optimality of $(\bar{Y},\bar{Z})$ (see (4.44)) we give the proof here instead of relegating it to Appendix A.

Proposition 4.3.22 Suppose Conditions 2.1.9,2.1.10,2.1.14,2.2.9 and 2.2.22 and recall (4.13), Notation 4.1.1 and $\bar{X}$ defined at (4.71). Then for each $\zeta \in L_2$, we have

$$\sup_{\substack{X\in\mathbb{D}_1, u_2\in L_\infty \\ u_2\leq X(T)-B}} E[\zeta(u_2 - (X(T) - B))] + \bar{\delta}_{L^*_\infty}(-\zeta|\mathbb{G}) + E[(\bar{X}(T) - B)\zeta] \geq 0. \quad (4.88)$$

Proof. Fix some $\zeta \in L_2$; from Proposition 4.2.1 we have

$$Y(T) = \zeta, \quad (4.89)$$

for some $Y \in \mathbb{B}_1$. Then Proposition 4.2.3 gives

$$g(\bar{Y} + \epsilon Y, \bar{Z} - \epsilon\zeta) \leq g(\bar{Y},\bar{Z}) \quad \text{for all}\, \epsilon \in (0,\infty). \quad (4.90)$$

Hence from (4.90) and (4.19),

$$\begin{aligned} &\varkappa(\bar{Y} + \epsilon Y, \bar{Z} - \epsilon\zeta) + EJ^*(\bar{Y}(T) + \epsilon\zeta) + \bar{\delta}_{L^*_\infty}(\bar{Z} - \epsilon\zeta|\mathbb{G}) \\ &\quad \geq \varkappa(\bar{Y},\bar{Z}) + EJ^*(\bar{Y}(T)) + \bar{\delta}_{L^*_\infty}(\bar{Z}|\mathbb{G}), \quad \epsilon \in (0,\infty) \end{aligned} \quad (4.91)$$

Since $\bar{Z} \in \mathbb{G}$ (by Remark 4.3.3), we have

$$\begin{aligned} \bar{\delta}(\bar{Z}|\mathbb{G}) &= 0 && (4.92)\\ \text{and} \qquad \bar{\delta}_{L^*_\infty}(\bar{Z} - \epsilon\zeta|\mathbb{G}) &\leq \epsilon\bar{\delta}_{L^*_\infty}(-\zeta|\mathbb{G}), \quad \epsilon \in (0,\infty), && (4.93) \end{aligned}$$

since $\bar{Z} - \epsilon\zeta \in \mathbb{G}$ when $-\zeta \in \mathbb{G}$. From (4.18) and (4.89), we get

$$\varkappa(\bar{Y} + \epsilon Y, \bar{Z} - \epsilon\zeta) = \sup_{X \in \mathbb{D}_1} \left\{ -E[X(T)(\bar{Y} + \epsilon Y)(T)] + \sup_{\substack{u_2 \in L_\infty \\ u_2 \leq X(T) - B}} (\epsilon\zeta - \bar{Z})(u_2) \right\}$$

$$\leq \sup_{X \in \mathbb{D}_1} \left\{ -E[X(T)\bar{Y}(T)] + \sup_{\substack{u_2 \in L_\infty \\ u_2 \leq X(T) - B}} (-\bar{Z})(u_2) + \epsilon \sup_{\substack{u_2 \in L_\infty \\ u_2 \leq X(T) - B}} \{E[\zeta u_2] - E[\zeta X(T)]\} \right\}$$

$$\leq \varkappa(\bar{Y}, \bar{Z}) + \epsilon \sup_{X \in \mathbb{D}_1} \left\{ \sup_{\substack{u_2 \in L_\infty \\ u_2 \leq X(T) - B}} E[\zeta(u_2 - X(T))] \right\}, \quad \text{for each } \epsilon \in (0, \infty). \tag{4.94}$$

Since $\varkappa(\bar{Y}, \bar{Z}) \in \mathbb{R}$ (see Remark 4.2.6), we can combine (4.91),(4.92),(4.93) and (4.94) to get

$$\sup_{\substack{X \in \mathbb{D}_1, u_2 \in L_\infty \\ u_2 \leq X(T) - B}} E[\zeta(u_2 - X(T))] + \bar{\delta}_{L^*_\infty}(-\zeta|\mathbb{G}) + E\left[\frac{J^*(\bar{Y}(T) + \epsilon\zeta) - J^*(\bar{Y}(T))}{\epsilon}\right] \geq 0, \tag{4.95}$$

for all $\epsilon \in (0, \infty)$. Since $\bar{X}(T) = \partial J^*(\bar{Y}(T))$, it follows from Condition 2.2.9 and dominated convergence, that the third term on left of (4.95) converges to $E[\bar{X}(T)\zeta]$ as $\epsilon \to 0$. Since (4.95) holds for arbitrarily chosen $\zeta \in L_2$ we obtain (4.88).

With Proposition 4.3.22 it is very easy to establish that the wealth process $\bar{X}$ satisfies the terminal wealth constraint in the QLM Problem 2.2.16:

Proposition 4.3.23 Suppose Conditions 2.1.9,2.1.10,2.1.14,2.2.9 and 2.2.22. Recall Notation 4.1.1 and $\bar{X}$ defined at (4.71) in terms of $\bar{Y} \in \mathbb{B}_1$, given by Proposition 4.2.3. Then $\bar{X}(T) \geq B$, a.s.

Proof. Put

$$\zeta := I_{\bar{X}(T) < B}, \tag{4.96}$$

where I is the indicator function given by Definition D.1.12. Since ζ is an indicator function with values in the two-point set $\{0, 1\}$ it is clear that $\zeta \in L_2$ and $\zeta \geq 0$. We therefore get

$$\bar{\delta}_{L^*_\infty}(-\zeta|\mathbb{G}) = 0. \tag{4.97}$$

Now suppose the contrary of what we wish to establish, namely $P[\bar{X}(T) < B] > 0$. Then from (4.96)

$$E[(\bar{X}(T) - B)\zeta] < 0. \tag{4.98}$$

Moreover, for each $X \in \mathbb{D}_1$ and $u_2 \in L_\infty$ with $u_2 \leq X(T) - B$, we have

$$E[\zeta(u_2 - (X(T) - B))] \leq 0. \tag{4.99}$$

It follows from (4.97), (4.98) and (4.99) that the quantity on the left side of (4.88) is strictly negative when we suppose that $P[\bar{X}(T) < B] > 0$. Since this contradicts the result of Proposition 4.3.22, we get $\bar{X}(T) \geq B$, a.s.

Remark 4.3.24 Recall from Remark 4.2.9 that our goal is to construct $\bar{X} \in \mathbb{A}$ in terms of the maximizer $(\bar{Y}, \bar{Z}) \in \mathbb{Y}$ of the dual function given by Proposition 4.2.3 such that $(\bar{X}, (\bar{Y}, \bar{Z})) \in \mathbb{A} \times \mathbb{Y}$ satisfies the optimality relations (4.36)-(4.39). We defined $\bar{X}$ in terms of $(\bar{Y}, \bar{Z})$ (see (4.71)), and have established that $\bar{X} \in \mathbb{A}$ (see (4.87)). We have also accomplished most of the work needed for verifying that $(\bar{X}, (\bar{Y}, \bar{Z}))$ satisfies (4.36)-(4.39). In fact, from Proposition 4.3.23 it follows that

$$\bar{X}(T) - B \geq 0, \tag{4.100}$$

so that clearly

$$\bar{X} \in \mathbb{A}_1, \tag{4.101}$$

(see (4.12)). Since we have already shown that $\bar{X} \in \mathbb{D}$ (see Remark 4.3.21 and (4.77)) we then obtain from (4.101)

$$\bar{X} \in \mathbb{D}_1, \tag{4.102}$$

(see (4.13)). From Remark 4.3.3, it follows that

$$\bar{Z} \leq 0, \tag{4.103}$$

and in view of (4.100), (4.102) and (4.103), the pair $(\bar{X}, \bar{Z})$ satisfies Equation (4.36) of Proposition 4.2.8. Furthermore, it is immediate from (4.71) that

$$\bar{X}(T) = (\partial J^*)(\bar{Y}(T)), \tag{4.104}$$

hence (4.39) of Proposition 4.2.8 is verified as well. It therefore only remains to verify that the complementary slackness relations (4.37) and (4.38) also hold. This is the goal of the following proposition which again exploits the optimality of $(\bar{Y}, \bar{Z})$ established at Proposition 4.2.3 (see also (4.44)):

Proposition 4.3.25 Suppose Conditions 2.1.9,2.1.10,2.1.14,2.2.9 and 2.2.22. Recall Notation 4.1.1, $(\bar{Y},\bar{Z})$ given by Proposition 4.2.3 and $\bar{X}$ defined at (4.71). Then

$$\varkappa(\bar{Y},\bar{Z}) + E[\bar{X}(T)\bar{Y}(T)] = \inf_{\substack{u_2\in L_\infty \\ u_2\leq \bar{X}(T)-B}} \bar{Z}(u_2) = 0. \tag{4.105}$$

Proof. The proof of this proposition again relies on the optimality of the pair $(\bar{Y},\bar{Z})$ established at Proposition 4.2.3. In fact, from Proposition 4.2.3, we get

$$g(\bar{Y}-\epsilon\bar{Y},\bar{Z}-\epsilon\bar{Z}) \leq g(\bar{Y},\bar{Z}) \tag{4.106}$$

and Remark 4.3.3 implies

$$\bar{Z}-\epsilon\bar{Z} \leq 0, \tag{4.107}$$

for all $\epsilon\in[0,1)$.

It therefore follows from (4.19), (4.106) and (4.107) that

$$\varkappa(\bar{Y}-\epsilon\bar{Y},\bar{Z}-\epsilon\bar{Z}) + EJ^*(\bar{Y}(T)-\epsilon\bar{Y}(T)) \geq \varkappa(\bar{Y},\bar{Z}) + EJ^*(\bar{Y}(T)), \tag{4.108}$$

for all $\epsilon\in(0,1)$.

From (4.18) one sees that

$$\varkappa(\bar{Y}-\epsilon\bar{Y},\bar{Z}-\epsilon\bar{Z}) = (1-\epsilon)\varkappa(\bar{Y},\bar{Z}), \tag{4.109}$$

for all $\epsilon\in[0,1)$. Using this in (4.108), with $\varkappa(\bar{Y},\bar{Z})\in\mathbb{R}$ (see Remark 4.3.3) then gives that

$$-\varkappa(\bar{Y},\bar{Z}) + \frac{1}{\epsilon}E[J^*((1-\epsilon)\bar{Y}(T)) - J^*(\bar{Y}(T))] \geq 0 \qquad \text{for all}\,\epsilon\in(0,1). \tag{4.110}$$

We now let $\epsilon\to 0$ at (4.110): from Lemma 4.1.6 and the dominated convergence theorem we easily obtain

$$\lim_{\epsilon\to 0}\frac{1}{\epsilon}E[J^*((1-\epsilon)\bar{Y}(T)) - J^*(\bar{Y}(T))] = -E[\bar{Y}(T)\partial J^*(\bar{Y}(T))], \tag{4.111}$$

and upon combining (4.111) with (4.110) and (4.104) we find

$$\varkappa(\bar{Y},\bar{Z}) + E[\bar{X}(T)\bar{Y}(T)] \leq 0. \tag{4.112}$$

Moreover, since $\bar{Z}(0) = 0$, it is immediate from Proposition 4.3.23 that

$$\inf_{\substack{u_2 \in L_\infty \\ u_2 \leq \bar{X}(T) - B}} \bar{Z}(u_2) \leq 0, \tag{4.113}$$

and, since $\bar{X} \in \mathbb{D}_1$ (see Remark 4.3.24), it follows from (4.18) that

$$\varkappa(\bar{Y}, \bar{Z}) + E[\bar{X}(T)\bar{Y}(T)] + \inf_{\substack{u_2 \in L_\infty \\ u_2 \leq \bar{X}(T) - B}} \bar{Z}(u_2) \geq 0. \tag{4.114}$$

The proposition is now immediate from (4.112), (4.113) and (4.114).

Remark 4.3.26 It is immediate from Proposition 4.3.25 that $(\bar{X}, (\bar{Y}, \bar{Z})) \in \mathbb{A} \times \mathbb{Y}$ satisfies the complementary slackness relations (4.37) and (4.38). In view of Remark 4.3.24 we have established that $(\bar{X}, (\bar{Y}, \bar{Z}))$ satisfies the entire set of Kuhn-Tucker optimality relations (4.36)-(4.39). Proposition 4.2.8 then gives

$$f(\bar{X}) = g(\bar{Y}, \bar{Z}). \tag{4.115}$$

However, from Remark 4.1.11 we also have

$$f(X) \geq g(Y, Z) \text{ for each } X \in \mathbb{A} \text{ and } (Y, Z) \in \mathbb{Y}, \tag{4.116}$$

and from (4.115) and (4.116) we find

$$\bar{X} \in \mathbb{A}, \qquad \text{and} \qquad f(\bar{X}) = \inf_{X \in \mathbb{A}} f(X), \tag{4.117}$$

that is $\bar{X}$ is the minimizer of the primal function f on the set $\mathbb{A}$ of admissible wealth processes. From Remark 4.3.24 we have $\bar{X} = X^{\bar{\boldsymbol{\pi}}} \in \mathbb{D}$ and $\bar{X}(T) \geq B$ a.s. From this, together with (2.49), we finally obtain

$$\eta = E[J(X^{\bar{\boldsymbol{\pi}}}(T))] = f(X^{\bar{\boldsymbol{\pi}}}) = f(\bar{X}), \tag{4.118}$$

that is $\bar{\boldsymbol{\pi}}$ is the optimal portfolio for the QLM Problem 2.2.16.

The results of this chapter are summarized in the following Proposition.

Proposition 4.3.27 Suppose Conditions 2.1.9,2.1.10,2.1.14,2.2.9 and 2.2.22. Then there exists some $(\bar{Y}, \bar{Z}) \in \mathbb{Y} := \mathbb{B}_1 \times L_\infty^*$ which maximizes the dual function $g(Y, Z)$ on $\mathbb{Y}$ (recall Definition 4.1.10 and Proposition 4.2.3). Define the $\mathbb{R}$-valued process

$\bar{X}$ in terms of $\bar{Y}$ and state price density H (see Definition 4.3.8) and the $\mathbb{R}^N$-valued process $\bar{\boldsymbol{\pi}} \in \mathcal{P}^*$(see Definition 2.1.6) by

$$\begin{aligned}\bar{X}(t) &:= H(t)^{-1}E[\partial J^*(\bar{Y}(T))H(T)|\mathcal{F}_t], \quad t \in [0,T],\\ \bar{\boldsymbol{\pi}}(t) &:= [\boldsymbol{\sigma}^\top]^{-1}\{H^{-1}(t)\bar{\boldsymbol{\xi}}(t) + \bar{X}(t_-)\boldsymbol{\theta}(t)\}, \quad t \in [0,T],\end{aligned}$$

where $\bar{\boldsymbol{\xi}} \in L_2^{loc}(\mathbf{W})$ is the d$\mathbf{W}$-integrand given by the martingale representation theorem (see (4.72) - (4.73)). Then

(a) $\bar{X} \in \mathbb{A}$ (recall (2.36)), and in particular $\bar{\boldsymbol{\Gamma}} = 0$, $\nu_{[\mathbf{M}]} = 0$-a.e. (see Proposition 4.3.20).

(b) $\bar{\boldsymbol{\pi}} \in \mathcal{A}$ (recall (4.79) and (4.84)), $\bar{X} = X^{\bar{\boldsymbol{\pi}}}$ (recall (2.28)), $X^{\bar{\boldsymbol{\pi}}}(T) \geq B$ a.s. and

$$\begin{aligned}E\left[J(X^{\bar{\boldsymbol{\pi}}}(T))\right] &= \inf_{\substack{\boldsymbol{\pi}\in\mathcal{A}\\ X^{\boldsymbol{\pi}}(T)\geq B\, a.s.}} E[J(X^{\boldsymbol{\pi}}(T))]\\ &= \sup_{(Y,Z)\in\mathbb{Y}} g(Y,Z) = g(\bar{Y},\bar{Z}).\end{aligned} \tag{4.119}$$

In particular,

$$E\left[J(X^{\bar{\boldsymbol{\pi}}}(T))\right] = \eta, \tag{4.120}$$

(recall (2.49)) and $\bar{\boldsymbol{\pi}}$ is the optimal portfolio for the QLM Problem 2.2.16.

Chapter 5

Conclusion and suggestions for Future Work

In this thesis we have addressed a problem of quadratic loss minimization in the presence of portfolio (or control) constraints and terminal wealth (or state) constraints with regime switching included in the market model. This combination leads to a fairly challenging convex stochastic control problem the key to which is to exploit convexity in order to introduce a dual problem. The solution of the dual problem gives the associated Lagrange multipliers, in terms of which one constructs the optimal portfolio (or optimal control). A method of Rockafellar and Moreau, hitherto mainly used for finite dimensional problems of convex optimization, is adapted to the stochastic control problem of this thesis, which is an *infinite dimensional* problem of convex optimization.

The present work can clearly be taken further in a variety of possible directions, of which the following seem to be the most interesting:

(I) The state constraint in this thesis is only on the wealth *at close of trade* T, namely $X^{\pi}(T) \geq B$ a.s. (see Definition 2.2.14), and as such is a *European-type* state constraint. A considerably more challenging problem results when this is replaced with an *American-type* state constraint, in which a continuous $\mathcal{F}_t$-adapted process $\{B(t),\ t \in [0,T]\}$ stipulates an acceptable "floor level" of wealth over the entire trading interval, so that the problem is to minimize the quadratic loss function $E[J(X^{\boldsymbol{\pi}}(T))]$ subject to the portfolio constraint $\pi \in \mathcal{A}$ (exactly as in the QLM Problem 2.2.16) but with the state constraint $X^{\pi}(t) \geq B(t)$ a.s. now over the entire trading interval $t \in [0,T]$ (instead of just at close-of-trade T, as in the

QLM Problem 2.2.16). The primal value of this problem is then given by

$$\eta := \inf_{\substack{\pi \in \mathcal{A} \\ X^{\pi}(t) \geq B(t)\ a.s. \\ t \in [0,T]}} E[J(X^{\pi}(T))], \tag{5.1}$$

(c.f. (2.45) for Problem 2.2.16). In place of the Slater Condition 2.2.22 one postulates that there is some $\hat{\pi} \in \mathcal{A}$ and constant $\epsilon \in (0, \infty)$ such that $X^{\hat{\pi}}(t) \geq B(t) + \epsilon$ for all $t \in [0,T]$ a.s., and the perturbation by $u_2 \in L_\infty$ at (4.2) must be replaced with the perturbation $X(t) \geq B(t) + u_2(t)$, $t \in [0,T]$ a.s. for each $u_2 \in L_\infty((\Omega, \{\mathcal{F}_t\}, P); C[0,T])$ (this denotes the Banach space of all real-valued, continuous, $\mathcal{F}_t$-adapted and P-essentially bounded processes $\{u_2(t);\ t \in [0,T]\}$ with norm $||u_2|| := P - \text{ess-sup}\{\max_{t \in [0,T]} |u_2(t)|\})$. With this perturbation, the natural analogue of the "second" factor space L_∞^* at (4.9) is then the norm dual of $L_\infty((\Omega, \{\mathcal{F}_t\}, P); C[0,T])$, a characterization of which is essentially provided by Ioffe and Levin (Theorem 3 on page 57 of [16]), a result which generalizes the Yosida-Hewitt decomposition (recall Theorem D.1.11) to essentially bounded measurable functions taking values in a separable Banach space. The Rockafellar-Moreau approach provides a rational method for dealing with this constraint, but definite technical challenges going well beyond those of the present thesis must nevertheless still be resolved.

(II) If the a.s. state constraint $X^{\pi}(T) \geq B$ is removed from the QLM Problem 2.2.16 then this problem simplifies to the problem addressed by Donnelly [7], namely minimize $E[J(X^{\pi}(T))]$ subject to the portfolio constraint $\pi \in \mathcal{A}$. In this case it is possible to get *explicit* optimal portfolios in feedback form when the convex portfolio constraint set K (see (2.44)) is "nice" (e.g. a cone) and the market parameters are adapted to the filtration of the Markov chain α that is

$$\mathcal{F}_t^{\alpha} := \sigma\{\alpha(s) : s \in [0,t]\} \vee \mathcal{N}(\mathbb{P}), \quad \forall t \in [0,T], \tag{5.2}$$

(see Section 5 of [7]). It is therefore natural to ask if one can also get comparably explicit optimal portfolios in feedback form for the QLM Problem 2.2.16, which is effectively the problem addressed in [7] with the a.s. constraint $X^{\pi}(T) \geq B$ included (recall Remark 2.2.19(c)). The calculation of explicit optimal portfolios in Section 5 of [7] relies on the fact that the Lagrange multiplier is just a simple Ito process $\bar{Y} \in \mathbb{B}_1$. In the case of Problem 2.2.16 the Lagrange multiplier is a pair $(\bar{Y}, \bar{Z}) \in \mathbb{B}_1 \times L_\infty^*$, in which the second member $\bar{Z} \in L_\infty^*$ is the Lagrange multiplier for the state constraint $X^{\pi}(T) \geq B$ (see Remark 4.2.5), and it is this second Lagrange multiplier $\bar{Z}$ which ruins the possibility of getting explicit optimal portfolios in feedback form along the lines of [7]. This is because $\bar{Z} \in L_\infty^*$ typically involves a *singular* element $\bar{Z}^\circ \in \mathcal{Q}(\Omega, \mathcal{F}_T, \mathbb{P})$ and there is as yet no adequate calculus for handling such singular elements. Despite this, Proposition 4.3.27 is still an essential starting point for the computation of optimal portfolios, since the duality characterization of optimality that it furnishes is indispensable for implementing any of the powerful duality-based primal-dual Lagrangian algorithms (motivated by Chap.4 of Ito and Kunisch [17]). In particular, one sees from Proposition 4.3.27 that the optimal portfolio $\bar{\pi}$ depends explicitly on *just the first element* $\bar{Y}$ of the solution $(\bar{Y}, \bar{Z})$ of the dual problem, and consequently we need approximate the second element $\bar{Z}$ in only a rather weak sense. Now the space $L_2(\Omega, \mathcal{F}_T, \mathbb{P})$ is a $\sigma\{L_\infty^*, L_\infty\}$-dense subspace of L_∞^* (as follows from Theorem 6.24(3) in Aliprantis and Border [1]), that is $\bar{Z} \in L_\infty^*$ can be approximated arbitrarily closely in the $\sigma\{L_\infty^*, L_\infty\}$-topology by (much more tractable) elements of $L_2(\Omega, \mathcal{F}_T, \mathbb{P})$. With such an approximation at hand it becomes possible to extend the approach in Section 5 of [7] to get at least approximate optimal portfolios in feedback form. This suggests the goal of implementing a primal-dual Lagrangian algorithm in function space for approximating a triple $(\bar{X}, (\bar{Y}, \bar{Z})) \in \mathbb{D} \times (\mathbb{B}_1 \times L_\infty^*)$ such that $f(\bar{X}) = g(\bar{Y}, \bar{Z})$ in which $\bar{Z}$ is approximated by elements of $L_2(\Omega, \mathcal{F}_T, \mathbb{P})$ in the (weak*) topology $\sigma(L_\infty^*, L_\infty)$.

Appendix A

Supplementary Results and Proofs

The proofs of several of the technical results in the main body of the thesis are collected in the present Appendix. The reason for this is that these proofs, which are sometimes quite lengthy and technical, are usually of subsidiary interest, and can obscure the main lines of development if placed in the main part of the thesis, and are accordingly best relegated to an appendix. In fact, we have retained in the main body of the thesis only those proofs which illustrate ideas of genuinely "intrinsic" interest, and have placed all other proofs in this Appendix. Readers of the thesis will actually lose very little if they choose not to study the proofs which follow.

Proof of Proposition 2.2.3: Set $t = 0$ in (2.38) to obtain immediately $X_0 = Y_0$. Then we have a.s.

$$\int_0^t (\boldsymbol{\xi}^X - \boldsymbol{\xi}^Y)^\top(\tau)\,\mathrm{d}\mathbf{W}(\tau) = \int_0^t \Upsilon^Y(\tau) - \Upsilon^X(\tau)\,\mathrm{d}\tau + \sum_{i,j=1}^{D} \int_0^t (\Gamma^Y_{i,j} - \Gamma^X_{i,j})(\tau)\,\mathrm{d}M_{ij}(\tau) \tag{A.1}$$

for all $t \in [0, T]$. The right-hand side of (A.1) is the sum of continuous finite variation process and a finite-variation local martingale, both null at the origin. So the right-hand side is a finite-variation semimartingale. The left-hand side of (A.1) is clearly a continuous local martingale, null at the origin. However, the left-hand side must also have paths of finite-variation, since the right-hand side had paths of finite-variation.

If a continuous local martingale N has paths of finite variation then $N(t) = 0$ a.s. for all $t \in [0, T]$ (see Rogers and Williams [32], Theorem IV.30.4). Thus from (A.1) we obtain

$$\int_0^t (\boldsymbol{\xi}^X - \boldsymbol{\xi}^Y)^\top(\tau)\,\mathrm{d}\mathbf{W}(\tau) = 0 \ a.s., \qquad \forall\, t \in [0,T] \tag{A.2}$$

Evaluating at time $t = T$, squaring, taking expectations and using the Itô isometry, we get from (A.2),

$$0 = E\left(\int_0^t (\boldsymbol{\xi}^X - \boldsymbol{\xi}^Y)^\top(\tau)\,\mathrm{d}\mathbf{W}(\tau)\right)^2 = E\int_0^t \parallel (\boldsymbol{\xi}^X - \boldsymbol{\xi}^Y)(\tau) \parallel^2 \ \mathrm{d}\tau \tag{A.3}$$

This implies that $\boldsymbol{\xi}^X = \boldsymbol{\xi}^X (\mathbb{P} \otimes Leb)$-a.e. From (A.1) and (A.2) we then obtain

$$\int_0^t \Upsilon^X(\tau) - \Upsilon^Y(\tau)\,\mathrm{d}\tau = \sum_{i,j=1}^{D} \int_0^t (\Gamma^Y_{i,j} - \Gamma^X_{i,j})(\tau)\,\mathrm{d}M_{ij}(\tau) \quad \text{a.s.} \tag{A.4}$$

The left-hand side of (A.4) is a continuous process, null at the origin. This implies that the finite-variation local martingale on the right-hand side of the equation is continuous. As noted above, for any continuous local martingale N which has paths of finite variation, we have $N(t) = 0$-a.s. for all $t \in [0,T]$. Thus

$$\sum_{i,j=1}^{D} \int_0^t (\Gamma^Y_{i,j} - \Gamma^X_{i,j})(\tau)\,\mathrm{d}M_{ij}(\tau) = 0 \quad \text{a.s.}, \qquad \forall\, t \in [0,T] \tag{A.5}$$

Evaluating at time $t = T$, squaring, taking expectations and using the it$\hat{o}$ isometry, we get from (A.5),

$$0 = E\left(\sum_{i,j=1}^{D} \int_0^t (\Gamma^Y_{i,j} - \Gamma^X_{i,j})(\tau)\,\mathrm{d}M_{ij}(\tau)\right)^2 = \sum_{i,j=1}^{D} E\int_0^T |(\Gamma^Y_{i,j} - \Gamma^X_{i,j})(\tau)|^2\,\mathrm{d}[M_{ij}](\tau) \tag{A.6}$$

This implies that $\Gamma^Y_{i,j} = \Gamma^X_{i,j}$ $(\nu_{[M_{ij}]})$-a.e. for each $i,j = 1,\ldots,D$, $i \neq j$. By virtue of $\mathbf{\Gamma}^X, \mathbf{\Gamma}^Y \in L_2(M)$, we have $\Gamma^X_{i,i} = \Gamma^Y_{i,i} = 0\,(\mathbb{P} \otimes Leb)$-a.e. for each $i = 1,\ldots,D$. Hence, $\mathbf{\Gamma}^X = \mathbf{\Gamma}^Y$ $(\nu_{[\mathbf{M}]} - a.e.)$

Finally from (A.4) and (A.5), we are left with

$$\int_0^t \Upsilon^X(\tau) - \Upsilon^Y(\tau)\,\mathrm{d}\tau = 0 \tag{A.7}$$

So we must have $\Upsilon^X(\tau) = \Upsilon^Y(\tau)$ for all $t \in [0,T]$.

<u>Proof of Lemma 2.2.4:</u> From (2.34), we have a.s.

$$Y(t) := Y_0 + \int_0^t \Upsilon(\tau)\,\mathrm{d}\tau + \sum_{n=1}^{N} \int_0^t \boldsymbol{\xi}_n^Y(\tau)\,\mathrm{d}W_n(\tau) + \sum_{i,j=1}^{D} \int_0^t \Gamma_{i,j}^Y(\tau)\,\mathrm{d}M_{ij}(\tau) \quad \text{(A.8)}$$

Then squaring both sides and using the identity

$$\begin{aligned} &(Z_1 + Z_2 + Z_3 + Z_4)^2 \leq Z_1^2 + Z_2^2 + Z_3^2 + Z_4^2, \quad \forall\, z_1, z_2, z_3, z_4 \in \mathbb{R} \\ &\text{and} \qquad \left| \int_0^t \Upsilon(\tau)\,\mathrm{d}\tau \right| \leq \int_0^t |\Upsilon(\tau)|\,\mathrm{d}\tau \end{aligned} \quad \text{(A.9)}$$

We get

$$\begin{aligned} |Y(t)|^2 \leq 4|Y_0|^2 + 4\left(\int_0^t |\Upsilon(\tau)|\,\mathrm{d}\tau\right)^2 + 4\left|\sum_{i,j=1}^{n} \int_0^t \xi_n^Y(\tau)\,\mathrm{d}W_n(\tau)\right|^2 \\ + 4\left|\sum_{i,j=1}^{n} \int_0^t \Gamma_{i,j}^Y(\tau)\,\mathrm{d}M_{ij}(\tau)\right|^2 \end{aligned} \quad \text{(A.10)}$$

Taking the supremum over $t \in [0, T]$, and taking expectations, we get

$$\begin{aligned} |Y(t)|^2 \leq 4|Y_0|^2 + 4E\left(\int_0^t |\Upsilon(\tau)|\,\mathrm{d}\tau\right)^2 \\ + 4E\left(\sup_{t\in[0,T]} \left|\sum_{i,j=1}^{n} \int_0^t \xi_n^Y(\tau)\,\mathrm{d}W_n(\tau)\right|^2\right) \\ + 4E\left(\sup_{t\in[0,T]} \left|\sum_{i,j=1}^{n} \int_0^t \Gamma_{i,j}^Y(\tau)\,\mathrm{d}M_{ij}(\tau)\right|^2\right) \end{aligned} \quad \text{(A.11)}$$

Applying Doob's L_2-*inequality* to the last two terms on the right-hand side, both of which are martingales since $\boldsymbol{\xi}^Y \in L_2(\mathbf{W})$ and $\boldsymbol{\Gamma}^Y \in L_2(\mathbf{M})$, followed by the Itô isometry we get

$$\begin{aligned} |Y(t)|^2 \leq 4|Y_0|^2 + 4E\left(\int_0^t |\Upsilon(\tau)|\,\mathrm{d}\tau\right)^2 \\ + 16E\sum_{i,j=1}^{n} \int_0^T \left|\xi_n^Y(\tau)\right|^2\,\mathrm{d}\tau \\ + 16E\sum_{i,j=1}^{n} \int_0^T |\Gamma_{i,j}^Y(\tau)|^2\,\mathrm{d}[M_{ij}](\tau) \end{aligned} \quad \text{(A.12)}$$

Then as all the terms on the left-hand side above are finite, we obtain

$$E\left(\sup_{t\in[0,T]}|Y(t)|^2\right)<\infty \tag{A.13}$$

Proof of Proposition 2.2.6. First we show that $X^{\boldsymbol{\pi}}\in\mathbb{A}$ implies that $\boldsymbol{\pi}\in L_2(\mathbf{W})$. From the wealth equation (2.28) for the portfolio process $\boldsymbol{\pi}$, we have

$$X^{\boldsymbol{\pi}}\equiv(x_0, rX^{\boldsymbol{\pi}}+\boldsymbol{\pi}^\top\boldsymbol{\sigma}\boldsymbol{\theta}, \boldsymbol{\sigma}^\top\boldsymbol{\pi})\in\mathbb{R}\times L_{21}\times L_2(\mathbf{W}) \tag{A.14}$$

Therefore, $\boldsymbol{\sigma}^\top\boldsymbol{\pi}\in L_2(\mathbf{W})$. From the uniform boundedness of $\boldsymbol{\sigma}$, it follows that $\boldsymbol{\pi}\in L_2(\mathbf{W})$.

Next we show that $\boldsymbol{\pi}\in L_2(\mathbf{W})$ implies that $X^{\boldsymbol{\pi}}\in\mathbf{A}$

First note that by the uniform boundedness of r assumed in Condition 2.1.9,the bank account price process $S_0(t)=\exp\int_0^t r(\tau)\,\mathrm{d}\tau$, given by (2.19), is also uniformly bounded. Let κ_{S_0} be a uniform upper bound on $S_0(t)$. By the nonnegativity of the risk-free interest rate process $\{r(t)\}$, we have the $S_0^{-1}(t)$ is bounded above by one. Then from (2.29), we get

$$|X^{\boldsymbol{\pi}}(t)|=\kappa_{S_0}\left\{x_0+\int_0^t\boldsymbol{\pi}^\top(\tau)\boldsymbol{\sigma}(\tau)\boldsymbol{\theta}(\tau)\,\mathrm{d}\tau+\int_0^t\boldsymbol{\pi}^\top(\tau)\boldsymbol{\sigma}(\tau)\mathrm{d}\mathbf{W}(\tau)\right\} \tag{A.15}$$

Then using the identity

$$|z_1+z_2+z_3|^2\le 3z_1^2+3z_2^2+3z_3^2,\quad\forall z_1,z_2,z_3\in\mathbb{R} \tag{A.16}$$

and $|\int f\,\mathrm{d}\mu|^2\le\int|f|^2\,\mathrm{d}\mu$, we get from (A.15) that

$$\begin{aligned}E|\sup_{t\in[0,T]}X^{\boldsymbol{\pi}}(t)|^2\le 3\kappa_{S_0}^2&\left\{x_0^2+E\left(\sup_{t\in[0,T]}\int_0^t|\boldsymbol{\pi}^\top(\tau)\boldsymbol{\sigma}(\tau)\boldsymbol{\theta}(\tau)|^2\,\mathrm{d}\tau\right)\right.\\&\left.+E\left(\sup_{t\in[0,T]}|\int_0^t\boldsymbol{\pi}^\top(\tau)\boldsymbol{\sigma}(\tau)\mathrm{d}\mathbf{W}(\tau)|^2\right)\right\}\end{aligned} \tag{A.17}$$

Applying Doob's L_2-inequality and the Itô isometry, and noting that the supremum of the second term on the right-hand side of (A.17) occurs at $t=T$, we get

$$\begin{aligned}&E|\sup_{t\in[0,T]}X^{\boldsymbol{\pi}}(t)|^2\\&\overset{Doob}{\le}3\kappa_{S_0}^2\left\{x_0^2+E\left(\int_0^T|\boldsymbol{\pi}^\top(\tau)\boldsymbol{\sigma}(\tau)\boldsymbol{\theta}(\tau)|^2\,\mathrm{d}\tau\right)+4E\left(\int_0^T\boldsymbol{\pi}^\top(\tau)\boldsymbol{\sigma}(\tau)\mathrm{d}\mathbf{W}(\tau)\right)^2\right\}\\&\overset{Itô}{=}3\kappa_{S_0}^2\left\{x_0^2+E\left(\int_0^T|\boldsymbol{\pi}^\top(\tau)\boldsymbol{\sigma}(\tau)\boldsymbol{\theta}(\tau)|^2\,\mathrm{d}\tau\right)+4E\int_0^T\|\,\boldsymbol{\sigma}^\top(\tau)\boldsymbol{\pi}(\tau)\,\|^2\,\mathrm{d}\tau\right\}\end{aligned} \tag{A.18}$$

Using the bounds in Remark 2.1.15 and Remark 2.1.18, and the assumption that $\boldsymbol{\pi} \in L_2(\mathbf{W})$, we obtain

$$E\int_0^T |\boldsymbol{\pi}^\top(t)\boldsymbol{\sigma}(t)\boldsymbol{\theta}(t)|^2\,\mathrm{d}t \le \kappa_{\boldsymbol{\sigma}}^2\kappa_{\boldsymbol{\theta}}^2 E\int_0^T \parallel \boldsymbol{\pi}(t) \parallel^2 \,\mathrm{d}t < \infty \tag{A.19}$$

and

$$E\int_0^T \parallel \boldsymbol{\sigma}^\top(t)\boldsymbol{\pi}(t) \parallel^2 \,\mathrm{d}t \le \kappa_{\boldsymbol{\sigma}}^2\, E\int_0^T \parallel \boldsymbol{\pi}(t) \parallel^2 \,\mathrm{d}t < \infty \tag{A.20}$$

Applying these bounds to (A.17), we get $E\left(\sup_{t\in[0,T]}|X^{\boldsymbol{\pi}}(t)|^2\right) < \infty$. It follows from this square-integrability, (A.19) and the uniform boundedness of r that $rX^{\boldsymbol{\pi}} + \boldsymbol{\pi}^\top\boldsymbol{\sigma}\boldsymbol{\theta} \in L_{21}$. From (A.20), we have $\boldsymbol{\sigma}^\top\boldsymbol{\pi} \in L_2(\mathbf{W})$. Thus $X^{\boldsymbol{\pi}} \equiv (x_0, rX^{\boldsymbol{\pi}} + \boldsymbol{\pi}^\top\boldsymbol{\sigma}\boldsymbol{\theta}, \boldsymbol{\pi}^\top\boldsymbol{\sigma}) \in \mathbb{A}$.

Proof of Lemma 2.2.15. First we show that $\eta < \infty$. Since $\mathbf{0} \in K$ then $\mathcal{A} \neq \emptyset$. Choose $\boldsymbol{\pi} \in \mathcal{A}$. Then from Proposition 2.2.6, the solution $X^{\boldsymbol{\pi}}$ of the wealth equation (2.28) for the portfolio process $\boldsymbol{\pi}$ is such that $X^{\boldsymbol{\pi}} \in \mathbb{A}$. Using the bounds on random variable a (see Condition 2.2.9), the square integrability of $X^{\boldsymbol{\pi}}$ (Lemma 2.2.4) and Hölder's inequality, we find,

$$\begin{aligned} E(J(X^{\boldsymbol{\pi}}(T))) &\overset{(2.42)}{=} \frac{1}{2}E\left(a(X^{\boldsymbol{\pi}}(T))^2\right) + E\left(cX^{\boldsymbol{\pi}}(T)\right) + q \\ &\le \frac{1}{2}\sup_{\omega\in\Omega}\{a(\omega)\}E\left((X^{\boldsymbol{\pi}}(T))^2\right) + \left(E(c^2)\right)^{\frac{1}{2}}\left(E((X^{\boldsymbol{\pi}}(T))^2)\right)^{\frac{1}{2}} + q \\ &< \infty. \end{aligned} \tag{A.21}$$

Taking the infimum over $\boldsymbol{\pi} \in \mathcal{A}$, we obtain $\eta < \infty$.

To show $\eta > -\infty$, we show that $E(J(X^{\boldsymbol{\pi}}(T)))$ is bounded from below. Using the strict positivity of random variable a (see Condition 2.2.9), we get

$$\begin{aligned} E(J(X^{\boldsymbol{\pi}}(T))) &= \frac{1}{2}E\left(a(X^{\boldsymbol{\pi}}(T))^2 + 2cX^{\boldsymbol{\pi}}(T)\right) + q \\ &= \frac{1}{2}E\left(a\left(X^{\boldsymbol{\pi}}(T) + \frac{c}{a}\right)^2 - \frac{c^2}{a}\right) + q \\ &\ge \frac{1}{2}\inf_{\omega\in\Omega}\{a(\omega)\}E\left(\left(X^{\boldsymbol{\pi}}(T) + \frac{c}{a}\right)^2\right) - E\left(\frac{c^2}{2a}\right) + q \\ &\ge -E\left(\frac{c^2}{2a}\right) + q. \end{aligned} \tag{A.22}$$

Taking the infimum over $\boldsymbol{\pi} \in \mathcal{A}$, we obtain $\eta \ge -E\left(\frac{c^2}{2a}\right) + q > -\infty$.

Proof of Lemma 4.1.4:

Claim (1): First we show that map is linear by showing that it is additive and homogeneous.

Additivity and homogeneity of the map follows easily from the additivity homogeneity of the stochastic integral. Therefore the map Ξ is linear.

Claim (2): Fix $(y, \boldsymbol{\lambda}, \boldsymbol{\gamma}) \in \mathbb{R} \times L_2(\mathbf{W}) \times L_2(\mathbf{M})$. Setting $Y \equiv \Xi(y, \boldsymbol{\lambda}, \boldsymbol{\gamma})$, we have a.s. for all $t \in [0, T]$,

$$\begin{aligned} Y(t) &= \Xi(y, \boldsymbol{\lambda}, \boldsymbol{\gamma})(t) \\ &\overset{(4.6)}{=} \beta(t) \left\{ y + \int_0^t \beta^{-1}(\tau) \boldsymbol{\lambda}^\top(\tau) \, \mathrm{d}\mathbf{W}(\tau) + \sum_{i,j=1}^{D} \int_0^t \beta^{-1}(\tau) \gamma_{i,j}(\tau) \, \mathrm{d}M_{ij}(\tau) \right\} \end{aligned} \tag{A.23}$$

We begin by showing that Y is square-integrable. Expanding (A.23), squaring and using the fact that $\beta(t) \leq 1$ a.s., we get

$$\begin{aligned} |Y(t)|^2 &\leq \left| y + \int_0^t \beta^{-1}(\tau) \boldsymbol{\lambda}^\top(\tau) \, \mathrm{d}\mathbf{W}(\tau) + \sum_{i,j=1}^{D} \int_0^t \beta^{-1}(\tau) \gamma_{i,j}(\tau) \, \mathrm{d}M_{ij}(\tau) \right|^2 \\ &\leq 3|y|^2 + 3 \left| \int_0^t \beta^{-1}(\tau) \boldsymbol{\lambda}^\top(\tau) \, \mathrm{d}\mathbf{W}(\tau) \right|^2 + 3 \left| \sum_{i,j=1}^{D} \int_0^t \beta^{-1}(\tau) \gamma_{i,j}(\tau) \, \mathrm{d}M_{ij}(\tau) \right|^2 \end{aligned} \tag{A.24}$$

Let $\kappa_{\boldsymbol{\beta}}$ be a uniform upper bound on $\{\beta^{-1}(t)\}$. Applying this bound, taking the supremum over $[0, T]$ and then expectations, we obtain

$$\begin{aligned} E\left(\sup_{t \in [0,T]} |Y(t)|^2 \right) \leq 3|y^2| + 3\kappa_{\boldsymbol{\beta}}^2 E \left(\sup_{t \in [0,T]} \left| \int_0^t \boldsymbol{\lambda}^\top(\tau) \, \mathrm{d}\mathbf{W}(\tau) \right|^2 \right) \\ + 3\kappa_{\boldsymbol{\beta}}^2 E \left(\sup_{t \in [0,T]} \left| \sum_{i,j=1}^{D} \int_0^t \gamma_{i,j}(\tau) \, \mathrm{d}M_{ij}(\tau) \right|^2 \right) \end{aligned} \tag{A.25}$$

Applying Doob's L_2-inequality to the second and third terms of the right-hand side of the above inequality, followed by the Itô isometry, this becomes

$$
\begin{aligned}
E\left(\sup_{t\in[0,T]}|Y(t)|^2\right) &\overset{Doob}{\leq} 3|y|^2 + 12\kappa_{\boldsymbol{\beta}}^2 E\left|\int_0^T \boldsymbol{\lambda}^\top(\tau)\,\mathrm{d}\mathbf{W}(\tau)\right|^2 \\
&\qquad + 12\kappa_{\boldsymbol{\beta}}^2 E\left|\sum_{i,j=1}^{D}\int_0^T \gamma_{i,j}(\tau)\,\mathrm{d}M_{ij}(\tau)\right|^2 \\
&\overset{\text{Itô}}{=} 3|y|^2 + 12\kappa_{\boldsymbol{\beta}}^2 E\int_0^T \parallel \boldsymbol{\lambda}(\tau) \parallel^2\,\mathrm{d}\tau \\
&\qquad + 12\kappa_{\boldsymbol{\beta}}^2 E\sum_{i,j=1}^{D}\int_0^T |\gamma_{i,j}(\tau)|^2\,\mathrm{d}M_{ij}(\tau) \\
&< \infty. \qquad \text{(A.26)}
\end{aligned}
$$

The finiteness comes from the facts that $\boldsymbol{\lambda} \in L_2(\mathbf{W})$ and $\boldsymbol{\gamma} \in L_2(\mathbf{M})$. Thus we have shown that Y is square-integrable.

Using the integration-by-parts formula (Theorem B.2.48) to expand (A.23), we get

$$
\begin{aligned}
Y(t) &= y + \int_0^t \beta(\tau)\left(\beta^{-1}(\tau)\boldsymbol{\lambda}^\top(\tau)\,\mathrm{d}\mathbf{W}(\tau) + \sum_{i,j=1}^{D}\beta^{-1}(\tau)\gamma_{ij}(\tau)\,\mathrm{d}M_{ij}(\tau)\right) \\
&\qquad - \int_0^t r(\tau)Y(\tau_-)\,\mathrm{d}\tau \\
&= y - \int_0^t r(\tau)Y(\tau_-)\mathrm{d}\tau + \int_0^t \boldsymbol{\lambda}^\top(\tau)\,\mathrm{d}\mathbf{W}(\tau) + \sum_{i,j=1}^{D}\int_0^t \gamma_{ij}(\tau)\,\mathrm{d}M_{ij}(\tau). \qquad \text{(A.27)}
\end{aligned}
$$

By the uniform boundedness of r and the square integrability of Y, shown above, we have $rY_- \in L_{21}$. Then, as $(y, \boldsymbol{\lambda}, \boldsymbol{\gamma}) \in \mathbb{R} \times L_2(\mathbf{W}) \times L_2(\mathbf{M})$, we get

$$
Y \equiv (y, -rY_-, \boldsymbol{\lambda}, \boldsymbol{\gamma}) \in \mathbb{B}, \qquad \text{(A.28)}
$$

so that (4.7) holds. Finally, as $Y(t) \neq Y(t_-)$ only on a set of Lebesgue measure zero, we find that

$$
-r(t)Y(t_-) = -r(t)Y(t), \qquad (\mathbb{P} \otimes Leb) - \text{a.e.}, \qquad \text{(A.29)}
$$

So that from the definition of $\mathbb{B}_1$ in (4.4), we get $Y \in \mathbb{B}_1$. Hence (4.8) follows.

Claim (3): To show that the map Ξ is bijective, we need to show that it is both injective and surjective.

The map Ξ is injective if and only if for all $(y^m, \boldsymbol{\lambda}^m, \boldsymbol{\gamma}^m) \in \mathbb{R} \times L_2(\mathbf{W}) \times L_2(\mathbf{M})$, $m = 1, 2$, we have that

$$\Xi(y^1, \boldsymbol{\lambda}^1, \boldsymbol{\gamma}^1) = \Xi(y^2, \boldsymbol{\lambda}^2, \boldsymbol{\gamma}^2) \tag{A.30}$$

implies $y^1 = y^2$, $\boldsymbol{\lambda}^1 = \boldsymbol{\lambda}^2\, (\mathbb{P} \otimes Leb) - a.e.$ and $\boldsymbol{\gamma}^1 = \boldsymbol{\gamma}^2\, \nu_{[M_{ij}]} - a.e.$ Where $\nu_{[M_{ij}]}$ is defined by (4.2.6).

From (4.6) and (A.30), we have for all $t \in [0, T]$

$$\begin{aligned} &\beta(t)\left\{y^1 + \int_0^t \beta^{-1}(\tau)(\boldsymbol{\lambda}^1)^\top(\tau)\,\mathrm{d}\mathbf{W}(\tau) + \sum_{i,j=1}^{D}\int_0^t \beta^{-1}(\tau)(\gamma^1)_{i,j}(\tau)\,\mathrm{d}M_{ij}(\tau)\right\} \\ &\quad= \beta(t)\left\{y^2 + \int_0^t \beta^{-1}(\tau)(\boldsymbol{\lambda}^2)^\top(\tau)\,\mathrm{d}\mathbf{W}(\tau) + \sum_{i,j=1}^{D}\int_0^t \beta^{-1}(\tau)(\gamma^2)_{i,j}(\tau)\,\mathrm{d}M_{ij}(\tau)\right\} \end{aligned} \tag{A.31}$$

Thus,

$$\begin{aligned} &y^1 + \int_0^t \beta^{-1}(\tau)(\boldsymbol{\lambda}^1)^\top(\tau)\,\mathrm{d}\mathbf{W}(\tau) + \sum_{i,j=1}^{D}\int_0^t \beta^{-1}(\tau)(\gamma^1)_{i,j}(\tau)\,\mathrm{d}M_{ij}(\tau) \\ &\quad= y^2 + \int_0^t \beta^{-1}(\tau)(\boldsymbol{\lambda}^2)^\top(\tau)\,\mathrm{d}\mathbf{W}(\tau) + \sum_{i,j=1}^{D}\int_0^t \beta^{-1}(\tau)(\gamma^2)_{i,j}(\tau)\,\mathrm{d}M_{ij}(\tau). \end{aligned}$$

From Proposition 2.2.3, we must then have

$$y^1 = y^2,\ \boldsymbol{\lambda}^1 = \boldsymbol{\lambda}^2\ (\mathbb{P} \otimes Leb) - a.e. \tag{A.32}$$

$$\boldsymbol{\gamma}^1 = \boldsymbol{\gamma}^2\ \nu_{[M_{ij}]} - a.e. \tag{A.33}$$

Hence the map Ξ is injective.

The map Ξ is surjective if and only if for each $Y \in \mathbb{B}_1$, there exists at least one triple $(y, \boldsymbol{\lambda}, \boldsymbol{\gamma}) \in \mathbb{R} \times L_2(\mathbf{W}) \times L_2(\mathbf{M})$ such that $\Xi(y, \boldsymbol{\lambda}, \boldsymbol{\gamma}) = Y$. Fix $Y \equiv (Y_0, \boldsymbol{\Upsilon}_Y, \boldsymbol{\xi}_Y, \boldsymbol{\Gamma}_Y) \in \mathbb{B}_1$. From the definition of $\mathbb{B}_1$ in (4.4), we see that Y has the particular integral form

$$Y(t) = Y_0 - \int_0^t r(\tau)Y(\tau)\,\mathrm{d}t + \int_0^t \boldsymbol{\xi}_Y^\top(\tau)\,\mathrm{d}\mathbf{W}(\tau) + \sum_{i,j=1}^{D}\int_0^t \Gamma_{i,j}^Y(\tau)\,\mathrm{d}M_{ij}(\tau) \tag{A.34}$$

Now consider an arbitrary triple $(y, \boldsymbol{\lambda}, \boldsymbol{\gamma}) \in \mathbb{R} \times L_2(\mathbf{W}), L_2(\mathbf{M})$. By (4.7), we can express $\Xi(y, \boldsymbol{\lambda}, \boldsymbol{\gamma})$ as

$$\Xi(y,\boldsymbol{\lambda},\boldsymbol{\gamma})(t)$$
$$\shortparallel$$
$$y - \int_0^t r(\tau)\Xi(y,\boldsymbol{\lambda},\boldsymbol{\gamma})(\tau_-)\,\mathrm{d}\tau + \int_0^t \boldsymbol{\lambda}^\top(\tau)\,\mathrm{d}\mathbf{W}(\tau) + \sum_{i,j=1}^{D}\int_0^t \gamma_{ij}(\tau)\,\mathrm{d}M_{ij}(\tau). \quad \text{(A.35)}$$

Setting $Y \equiv \Xi(y, \boldsymbol{\lambda}, \boldsymbol{\gamma})$ and applying Proposition 2.2.3, we must have $y = Y_0$, $\boldsymbol{\lambda} = \boldsymbol{\xi}_Y$ $(\mathbb{P} \otimes Leb) - a.e.$ and $\boldsymbol{\gamma} = \boldsymbol{\Gamma}_Y\ \nu_{[\mathbf{M}]} - a.e.$ Since $(y, \boldsymbol{\lambda}, \boldsymbol{\gamma}) = (Y_0, \boldsymbol{\xi}_Y, \boldsymbol{\Gamma}_Y) \in \mathbb{R} \times L_2(\mathbf{W}) \times L_2(\mathbf{M})$, we have shown that the map Ξ is surjective. It follows that the map Ξ is bijective.

Proof of Lemma 4.1.6. Substituting (2.42) into (4.14), we obtain

$$\begin{aligned}
J^*(y,\omega) &= \sup_{x\in\mathbb{R}}[xy - J(x,\omega)] \\
&= \sup_{x\in\mathbb{R}}\left\{xy - \frac{1}{2}a(\omega)x^2 - c(\omega)x\right\} - q \\
&= -\frac{1}{2}a(\omega)\inf_{x\in\mathbb{R}}\left\{-\frac{2xy}{a(\omega)} + x^2 + \frac{2c(\omega)x}{a(\omega)}\right\} - q \quad \text{since } a(\omega) > 0 \\
&= -\frac{1}{2}a(\omega)\inf_{x\in\mathbb{R}}\left\{\left(x - \frac{y-c(\omega)}{a(\omega)}\right)^2 - \frac{(y-c(\omega))^2}{a^2(\omega)}\right\} - q \\
&= -\frac{1}{2}a(\omega)\inf_{x\in\mathbb{R}}\left\{\left(x - \frac{y-c(\omega)}{a(\omega)}\right)^2\right\} + \frac{(y-c(\omega))^2}{a^2(\omega)} - q.
\end{aligned}$$

As $(x - \frac{y-c(\omega)}{a(\omega)})^2 \geq 0$ and, by the strict positivity of the random variable a posited in Condition 2.2.9, $-\frac{a(\omega)}{2} < 0$, it follows that the infimum equals zero. Thus we get

$$J^*(y,\omega) = \frac{(y-c(\omega))^2}{2a(\omega)} - q. \quad \text{(A.36)}$$

In order to establish Proposition 4.1.7 (explicit calculation of the Lagrangian defined by (4.11)) we need the following immediate consequence of a technical result due to Rockafellar (see Theorem F.2.3):

Proposition A.0.28 For the loss function J and its conjugate J^* we have

$$\sup_{u\in L_2}\left(E[u\eta] - \int_\Omega J(u(\omega),\omega)\,\mathrm{d}\mathbb{P}(\omega)\right) = \int_\Omega J^*(\eta(\omega),\omega)\,\mathrm{d}\mathbb{P}(\omega) \ \text{ for } \eta \in L_2. \quad \text{(A.37)}$$

Proof. Since L_2 is decomposable and the loss function J is a normal convex integrand, the result follows from Theorem F.2.3.

Proof of Proposition 4.1.7: Fix some $X \in \mathbb{A} - \mathbb{D}_1$. From the definition of $\mathbb{D}_1$ (i.e. (4.13)), we have either

$$X \notin \mathbb{A}_1 \text{ or } X \notin \mathbb{D}. \tag{A.38}$$

Suppose $X \notin \mathbb{A}_1$: then, from (4.12), we have

$$P[X(T) - B < u_2] > 0 \text{ for each } u_2 \in L_\infty, \tag{A.39}$$

so that (4.2) gives

$$F(X, u) = +\infty, \quad \forall\, u \in \mathbb{U}. \tag{A.40}$$

On the other hand, when $X \notin \mathbb{D}$, (4.2) again gives

$$F(X, u) = +\infty, \quad \forall\, u \in \mathbb{U}. \tag{A.41}$$

Since

$$\langle (u_1, u_2)(Y, Z)\rangle \in \mathbb{R}, \tag{A.42}$$

for all $(u_1, u_2) \in \mathbb{U}$ and $(Y, Z) \in \mathbb{Y}$, we conclude from (4.11) that

$$K(X, (Y, Z)) = +\infty, \tag{A.43}$$

for all $X \in \mathbb{A} - \mathbb{D}_1$ and $(Y, Z) \in \mathbb{Y}$.

Now fix

$$X \in \mathbb{D}_1 \text{ (i.e. } X \in \mathbb{D} \text{ and } X(T) > B + u_2). \tag{A.44}$$

From the definition of Lagrangian (4.11), bilinear form $\langle \cdot, \cdot \rangle$ (4.10) and perturbation function (4.2) we get for each $(Y, Z) \in \mathbb{Y}$,

$$\begin{aligned} K(X, (Y, Z)) &= \inf_{(u_1,u_2)\in\mathbb{U}} [\langle (u_1, u_2), (Y, Z)\rangle + F(X, u_1, u_2)] \\ &= \inf_{(u_1,u_2)\in\mathbb{U}} [E[u_1 Y(T)] + Z(u_2) + EJ(X(T) - u_1)] \\ &= \inf_{u_1 \in L_2} \{E[u_1 Y(T)] + E[J(X(T) - u_1)]\} + \inf_{\substack{u_2 \in L_\infty \\ u_2 \le X(T) - B}} Z(u_2). \end{aligned} \tag{A.45}$$

From Remark D.1.13,

$$\inf_{\substack{u_2 \in L_\infty \\ u_2 \le X(T)-B}} Z(u_2) = -\infty, \tag{A.46}$$

for all $Z \in L_\infty^*$ such that $Z \nleq 0$.

Moreover, for each $(X, Y) \in \mathbb{D} \times \mathbb{B}_1$ we also have

$$\begin{aligned} &\inf_{u_1 \in L_2} \{E[u_1 Y(T)] + EJ(X(T) - u_1)\} \\ &\qquad = E[X(T)Y(T)] - \sup_{v_1 \in L_2} \{E[v_1 Y(T)] - EJ(v_1)\}. \end{aligned} \tag{A.47}$$

This equality follows upon defining $v_1 := X(T) - u_1$, and using $X(T) \in L_2$.

From Proposition A.0.28, it follows that

$$\sup_{v_1 \in L_2} \{E[v_1 Y(T)] - EJ(v_1)\} = E[J^*(Y(T))]. \tag{A.48}$$

From (A.47) and (A.48),

$$\inf_{u_1 \in L_2} \{E[u_1 Y(T)] + EJ(X(T) - u_1)\} = E[X(T)Y(T)] - E[J^*(Y(T))]. \tag{A.49}$$

From (A.45) and (A.49), we get the required expression for Lagrangian for $X \in \mathbb{D}_1$ and $Z \le 0$ as

$$K(X, (Y, Z)) = E[X(T)Y(T)] - E[J^*(Y(T))] + \inf_{\substack{u_2 \in L_\infty \\ u_2 \le X(T)-B}} Z(u_2). \tag{A.50}$$

Proof of Lemma 4.2.1. Fix $\zeta \in L_2$, and put

$$\zeta_1 := S_0(T)\zeta. \tag{A.51}$$

By Condition 2.1.9, r is uniformly bounded. Recalling equation (2.19), it follows that

$$\zeta_1 \in L_2. \tag{A.52}$$

Hence, from (2.8), (2.9) and the martingale representation theorem (Theorem G.3.5), there is some $\bar{\boldsymbol{\xi}} \in L_2(\mathbf{W})$ and $\bar{\boldsymbol{\Gamma}} \in L_2(\mathbf{M})$ such that

$$\zeta_1 = \bar{y} + \sum_{n=1}^{N} \int_0^T \bar{\boldsymbol{\xi}}_n(\tau)\, \mathrm{d}W_n(\tau) + \sum_{i,j=1}^{D} \int_0^T \bar{\Gamma}_{i,j}(\tau)\, \mathrm{d}M_{ij}(\tau) \qquad \text{a.s.,} \tag{A.53}$$

for $\bar{y} := E\zeta_1$.
Let

$$\bar{\boldsymbol{\lambda}}(t) := [S_0(t)]^{-1}\bar{\boldsymbol{\xi}}(t), \tag{A.54}$$

$$\text{and} \qquad \bar{\boldsymbol{\gamma}}(t) := [S_0(t)]^{-1}\bar{\boldsymbol{\Gamma}}(t). \tag{A.55}$$

Hence, from Condition 2.1.9, the fact that $\bar{\boldsymbol{\xi}} \in L_2(\mathbf{W})$ and $\bar{\boldsymbol{\Gamma}} \in L_2(\mathbf{M})$ we get

$$\bar{\boldsymbol{\lambda}} \in L_2(\mathbf{W}), \tag{A.56}$$

$$\text{and} \qquad \bar{\boldsymbol{\gamma}} \in L_2(\mathbf{M}). \tag{A.57}$$

From (A.51)(A.53), (A.54) and (A.55) we have

$$\zeta = \Xi(\bar{y}, \bar{\boldsymbol{\lambda}}, \bar{\boldsymbol{\gamma}})(T), \tag{A.58}$$

where Ξ is given by (4.6), and we have used

$$\beta(t) = [S_0(t)]^{-1},$$

(see (4.5)). Put

$$Y := \Xi(\bar{y}, \bar{\boldsymbol{\lambda}}, \bar{\boldsymbol{\gamma}}). \tag{A.59}$$

By Lemma 4.1.4

$$Y \in \mathbb{B}_1. \tag{A.60}$$

Proof of Proposition 4.2.8. Fix some $(X, (Y, Z)) \in \mathbb{A} \times \mathbb{Y}$. Since $E[X(T)Y(T)]$ and $EJ^*(Y(T))$ are $\mathbb{R}$ valued (recall Remark 4.1.8) from (4.19) and (4.16) we get

$$g(Y, Z) = K(X, (Y, Z)) \in \mathbb{R} \tag{A.61}$$

iff

$$\left.\begin{array}{l} X \in \mathbb{D}_1, \\ Z \leq 0, \\ \varkappa(Y, Z) \in \mathbb{R}, \\ E[X(T)Y(T)] + \varkappa(Y, Z) + \inf\limits_{\substack{u_2 \in L_\infty \\ u_2 \leq X(T) - B}} Z(u_2) = 0. \end{array}\right\} \tag{A.62}$$

Moreover from Definition 2.2.26 and Equation (4.16), we have

$$f(X) = K(X, (Y, Z)) \in \mathbb{R} \tag{A.63}$$

iff

$$\left.\begin{array}{l} X(T) - B \geq 0, \\ X \in \mathbb{D}_1, \\ Z \leq 0, \\ E[J(X(T)) + J^*(Y(T)) - X(T)Y(T)] = \inf\limits_{\substack{u_2 \in L_\infty \\ u_2 \leq X(T) - B}} Z(u_2). \end{array}\right\} \tag{A.64}$$

From Equation (4.14) it follows that

$$J(X(T)) + J^*(Y(T)) - X(T)Y(T) \geq 0 \quad \text{a.s.}. \tag{A.65}$$

Moreover $Z \leq 0$ and $X(T) \geq B$ imply

$$\inf_{\substack{u_2 \in L_\infty \\ u_2 \leq X(T)-B}} Z(u_2) \leq 0. \tag{A.66}$$

From (A.64), (A.65) and (A.66) we get

$$E[J(X(T)) + J^*(Y(T)) - X(T)Y(T)] = 0, \tag{A.67}$$

$$\inf_{\substack{u_2 \in L_\infty \\ u_2 \leq X(T)-B}} Z(u_2) = 0. \tag{A.68}$$

Because of the non-negativity of integrand at (A.67), we see that (A.67) holds if and only if

$$J(X(T)) + J^*(Y(T)) - X(T)Y(T) = 0 \quad \text{a.s.}. \tag{A.69}$$

It is also immediate from Lemma B.4.15 (with $N = 1$) that (A.69) holds if and only if

$$X(T) = (\partial J^*)(Y(T)). \tag{A.70}$$

We then get

$$f(X) = K(X, (Y, Z)) \in \mathbb{R} \tag{A.71}$$

$$\text{iff}$$

$$\left.\begin{array}{l} X(T) - B \geq 0, \\ X \in \mathbb{D}_1, \\ Z \leq 0, \\ X(T) = (\partial J^*)(Y(T)), \\ \displaystyle\inf_{\substack{u_2 \in L_\infty \\ u_2 \leq X(T)-B}} Z(u_2) = 0. \end{array}\right\} \tag{A.72}$$

From Definition 2.2.26, we know that

$$f(X) \in (-\infty, \infty], \quad \text{for all}\, X \in \mathbb{D}. \tag{A.73}$$

Since $\mathbb{D}_1 \neq \emptyset$ (see Equation (4.13)), from Equation (4.18), we get

$$\varkappa(Y, Z) \in (-\infty, \infty], \quad \text{for all}\, (Y, Z) \in \mathbb{Y}, \tag{A.74}$$

and since $J^*(Y(T))$ is P-integrable for each $\mathbb{Y} \in \mathbb{B}_1$ (recall Remark 4.1.8), Equation (4.19) gives

$$g(Y,Z) \in [-\infty, \infty) \quad \text{for all}(Y,Z) \in \mathbb{Y}. \tag{A.75}$$

Thus for each $X \in \mathbb{A}$ and each $(Y,Z) \in \mathbb{Y}$, we obtain from Remark 4.1.11 the equivalence that

$$\begin{aligned} f(X) &= g(Y,Z) \\ &\text{iff} \\ f(X) &= K(X,(Y,Z)) \in \mathbb{R} \\ \text{and} \quad g(Y,Z) &= K(X,(Y,Z)) \in \mathbb{R}. \end{aligned}$$

This equivalence, together with (A.72) and (A.62), proves the result.

Proof of Proposition 4.3.6. From Proposition 2.2.6, for each $\boldsymbol{\pi} \in L_2(\mathbf{W})$ we obtain

$$X^{\boldsymbol{\pi}} \in \mathbb{A}, \tag{A.76}$$

$$X^{\boldsymbol{\pi}}(0) = x_0. \tag{A.77}$$

$$\tag{A.78}$$

Let

$$\Upsilon^{X^{\boldsymbol{\pi}}} = rX^{\boldsymbol{\pi}} + \boldsymbol{\pi}^\top \boldsymbol{\sigma}\boldsymbol{\theta}, \tag{A.79}$$

$$\text{and} \qquad (\boldsymbol{\xi}^{X^{\boldsymbol{\pi}}})^\top = \boldsymbol{\pi}^\top \boldsymbol{\sigma}. \tag{A.80}$$

Recall that for $Y \in \mathbb{B}_1$ we have

$$\Upsilon^Y = -rY. \tag{A.81}$$

Substituting the values of Equations (A.77)-(A.81) into Proposition 4.3.2 we get for every $Y \in \mathbb{B}_1$ and $\boldsymbol{\pi} \in L_2(\mathbf{W})$

$$\mathcal{M}(X^{\boldsymbol{\pi}}, Y)(T) = X^{\boldsymbol{\pi}}(T)Y(T) - x_0Y_0 - \int_0^T \boldsymbol{\pi}^\top(\tau)\boldsymbol{\sigma}(\tau)[\boldsymbol{\theta}(\tau)Y(\tau) + \boldsymbol{\xi}^Y(\tau)]\mathrm{d}\tau. \tag{A.82}$$

It also follows from Proposition 4.3.2 that

$$E[\mathcal{M}(X^{\boldsymbol{\pi}}, Y)(T)] = 0, \tag{A.83}$$

for all $Y \in \mathbb{B}_1$ and $\boldsymbol{\pi} \in L_2(\mathbf{W})$. It then follows from (A.82) and (A.83) that

$$E[X^{\boldsymbol{\pi}}(T)Y(T)] = x_0 Y_0 + E\int_0^T \boldsymbol{\pi}^\top(\tau)\boldsymbol{\sigma}(\tau)[\boldsymbol{\theta}(\tau)Y(\tau) + \boldsymbol{\xi}^Y(\tau)]\mathrm{d}\tau, \tag{A.84}$$

for all $Y \in \mathbb{B}_1$ and $\boldsymbol{\pi} \in L_2(\mathbf{W})$.

We now calculate the supremum on the left side of (4.45). Using the definition of $\mathbb{D}$ (2.47) we get

$$\begin{aligned}
\sup_{X\in\mathbb{D}} E[-X(T)Y(T)] &= \sup_{\substack{\boldsymbol{\pi}\in L_2(\mathbf{W})\\ \boldsymbol{\pi}\in K}} E[-X^{\boldsymbol{\pi}}(T)Y(T)] \\
&\overset{(4.43)}{=} \sup_{\boldsymbol{\pi}\in L_2(\mathbf{W})} \{E[-X^{\boldsymbol{\pi}}(T)Y(T)] - E\int_0^T \bar{\delta}_{\mathbb{R}^N}(\boldsymbol{\pi}(\tau)|K)\mathrm{d}\tau\} \\
&\overset{(A.84)}{=} -x_0Y_0 + \sup_{\boldsymbol{\pi}\in L_2(\mathbf{W})} E\int_0^T \{-\boldsymbol{\pi}^\top(\tau)\boldsymbol{\sigma}(\tau)[\boldsymbol{\theta}(\tau)Y(\tau) + \boldsymbol{\xi}^Y(\tau)] \\
&\qquad -\bar{\delta}_{\mathbb{R}^N}(\boldsymbol{\pi}(\tau)|K)\}\mathrm{d}\tau,
\end{aligned} \tag{A.85}$$

for each $Y \in \mathbb{B}_1$.

We next evaluate the supremum on the right of (A.85). From Lemma F.1.2 it follows that the characteristic function $\bar{\delta}_{\mathbb{R}^N}(\cdot|K)$ is a normal convex integrand.

Since $L_2(\mathbf{W})$ is decomposable and $\delta_{\mathbb{R}^N}(\cdot|K)$ is the convex conjugate of $\bar{\delta}_{\mathbb{R}^N}(\cdot|K)$, from Theorem F.2.3 we have

$$\sup_{\boldsymbol{\pi}\in L_2(\mathbf{W})} E\int_0^T \{\boldsymbol{\pi}^\top(\tau)\mathcal{V}(\tau) - \bar{\delta}_{\mathbb{R}^N}(\cdot|K)\}\mathrm{d}\tau = E\int_0^T \delta(\mathcal{V}(\tau)|K)\mathrm{d}\tau, \tag{A.86}$$

for $\mathcal{V} \in L_2(\mathbf{W})$.

Now (4.45) follows from (A.85) and (A.86).

Notation A.0.29 1. We use the symbol $\bullet$ to indicate stochastic integration, so that

$$(-\boldsymbol{\theta} \bullet \mathbf{W})(t) := \int_0^t -\boldsymbol{\theta}^\top(\tau)\,\mathrm{d}\mathbf{W}(\tau). \tag{A.87}$$

2. Denote by $\mathcal{E}(-\boldsymbol{\theta} \bullet \mathbf{W})$ the Doléans-Dade exponential of the continuous martingale $(-\boldsymbol{\theta} \bullet \mathbf{W})$ which by Remark B.2.51 and Equation (B.35) satisfies

$$\mathcal{E}(-\boldsymbol{\theta} \bullet \mathbf{W})(t) = \exp\left(-\int_0^t \boldsymbol{\theta}^\top(\tau)\,\mathrm{d}\mathbf{W}(\tau) - \frac{1}{2}\int_0^t \parallel \boldsymbol{\theta}(\tau) \parallel^2 \,\mathrm{d}\tau\right), \tag{A.88}$$

We can then rewrite (4.53) in the more compact form (recall (4.5))

$$H(t) = \beta(t)\mathcal{E}(-\boldsymbol{\theta} \bullet \mathbf{W})(t). \tag{A.89}$$

The boundedness properties of the exponential at (A.88) are given in the next proposition.

Proposition A.0.30 For any $p \in \mathbb{R}$ we have $\mathcal{E}(-p\boldsymbol{\theta} \bullet \mathbf{W}) \in \mathcal{M}_2(\{\mathcal{F}_t\}, \mathbb{P})$.

Proof of Proposition A.0.30. Fix an arbitrary $p \in \mathbb{R}$. We begin by showing that $\mathcal{E}(-p\boldsymbol{\theta} \bullet \mathbf{W})$ satisfies Novikov's Criterion (Theorem B.2.53). Consider

$$Z(t) := \int_0^t -p\boldsymbol{\theta}^\top(\tau)\mathrm{d}\mathbf{W}(\tau). \tag{A.90}$$

The square-bracket quadratic variation process $[Z]$ of Z is

$$[Z] = p^2 \int_0^t \parallel \boldsymbol{\theta}(\tau) \parallel^2 \mathrm{d}(\tau) \quad \text{a.s.,} \quad \forall t \in [0, T]. \tag{A.91}$$

Using the constant $\kappa_{\boldsymbol{\theta}} \in (0, \infty)$ which satisfies (2.25), we obtain the bound $[Z](t) \leq (p\kappa_{\boldsymbol{\theta}})^2 t$ a.s. for each $t \in [0, T]$. Hence

$$E\left(\exp\left\{\frac{1}{2}[Z](T)\right\}\right) \leq \exp\left\{\frac{1}{2}(p\kappa_{\boldsymbol{\theta}})^2 T\right\} < \infty. \tag{A.92}$$

Thus $\mathcal{E}(-p\boldsymbol{\theta} \bullet \mathbf{W})$ satisfies Novikov's Criterion (see Theorem B.2.53), and so is a uniformly integrable martingale for all $p \in \mathbb{R}$.

Using Corollary B.2.52, we have

$$(\mathcal{E}(-p\boldsymbol{\theta} \bullet \mathbf{W}))^2 = \exp\left\{p^2 \int_0^t \parallel \boldsymbol{\theta}(\tau) \parallel^2 \mathrm{d}\tau\right\} \mathcal{E}(-2p\boldsymbol{\theta} \bullet \mathbf{W})(t), \tag{A.93}$$

and as we have just shown that $\mathcal{E}(-2p\boldsymbol{\theta} \bullet \mathbf{W})$ is a martingale, we get for all $t \in [0, T]$,

$$E|\mathcal{E}(-p\boldsymbol{\theta} \bullet \mathbf{W})(t)|^2 \leq \exp\left\{(p\kappa_{\boldsymbol{\theta}})^2 T\right\} E|\mathcal{E}(-2p\boldsymbol{\theta} \bullet \mathbf{W})(t)| < \infty. \tag{A.94}$$

Thus $\mathcal{E}(-p\boldsymbol{\theta} \bullet \mathbf{W})$ is a square-integrable martingale, which holds for all $p \in \mathbb{R}$.

For later use we note that

$$|\mathcal{E}(-\boldsymbol{\theta} \bullet \mathbf{W})(t)|^p = \exp\left(-p\int_0^t \boldsymbol{\theta}^\top(\tau)\,\mathrm{d}\mathbf{W}(\tau) - \frac{1}{2}p\int_0^t \parallel \boldsymbol{\theta}(\tau) \parallel^2 \mathrm{d}\tau\right). \tag{A.95}$$

Proof of Proposition 4.3.9. Fix $t \in [0,T]$. From the nonnegativity of the interest rate process {r(t)}, we have $[S_0(t)]^{-1} \leq 1$ a.s. for all $t \in [0,T]$. Expanding $H(t)^p$ using (A.95), we get

$$\begin{aligned}
|H(t)|^p &\overset{(A.88)}{=} |[S_0(t)]^{-1}\mathcal{E}(-\boldsymbol{\theta} \bullet \mathbf{W})(t)|^p \\
&\overset{[S_0(t)]^{-1}\leq 1}{\leq} |\mathcal{E}(-\boldsymbol{\theta} \bullet \mathbf{W})(t)|^p \\
&\overset{(A.95)}{=} \exp\left(-p\int_0^t \boldsymbol{\theta}^\top(\tau)\,\mathrm{d}\mathbf{W}(\tau) - \frac{1}{2}p\int_0^t \parallel \boldsymbol{\theta}(\tau) \parallel^2 \,\mathrm{d}\tau\right) \\
&\overset{(B.36)}{=} \exp\left\{\frac{1}{4}p(p-2)\int_0^t \parallel \boldsymbol{\theta}(\tau) \parallel^2 \,\mathrm{d}\tau\right\}\left|\mathcal{E}\left(-\frac{p}{2}\boldsymbol{\theta} \bullet \mathbf{W}\right)(\tau)\right|^2 .
\end{aligned} \tag{A.96}$$

Recalling that the process $\{\boldsymbol{\theta}(t)\}$ is uniformly bounded, let $\kappa_{\boldsymbol{\theta}} \in (0,\infty)$ be the uniform bound for this process. Taking supremum over $t \in [0,T]$ in (A.96), we get

$$\sup_{t\in[0,T]} |H(t)|^p \leq \max\left[1, \exp\left\{\frac{1}{4}p(p-2)\kappa_{\boldsymbol{\theta}}^2 T\right\}\right] \sup_{t\in[0,T]}\left|\mathcal{E}\left(-\frac{p}{2}\boldsymbol{\theta} \bullet \mathbf{W}\right)(\tau)\right|^2, \tag{A.97}$$

where the maximum will equal one if $p \in (0,2)$ and will otherwise equal

$$\exp\left\{\frac{1}{4}p(p-2)\kappa_{\boldsymbol{\theta}}^2 T\right\}. \tag{A.98}$$

Upon taking expectations, we get

$$E\left(\sup_{t\in[0,T]} |H(t)|^p\right) \leq \max\left[1, \exp\left\{\frac{1}{4}p(p-2)\kappa_{\boldsymbol{\theta}}^2 T\right\}\right] E\left(\sup_{t\in[0,T]}\left|\mathcal{E}\left(-\frac{p}{2}\boldsymbol{\theta} \bullet \mathbf{W}\right)(\tau)\right|^2\right) \tag{A.99}$$

By Proposition A.0.30, $\mathcal{E}\left(-\frac{p}{2}\boldsymbol{\theta} \bullet \mathbf{W}\right)$ is a square integrable martingale, so upon applying Doob's L_2-Inequality, we get

$$E\left(\sup_{t\in[0,T]} |H(t)|^p\right) < \infty. \tag{A.100}$$

Proof of Proposition 4.3.10. Setting $p = 2$ in Proposition 4.3.9 shows that H is square-integrable. Expanding $H(t) = [S_0(t)]^{-1}\mathcal{E}(-\boldsymbol{\theta} \bullet \mathbf{W})(t)$ using the integration by parts formula (see Theorem B.2.48), we get

$$\begin{aligned}
H(t) &= [S_0(t)]^{-1}\mathcal{E}(-\boldsymbol{\theta} \bullet \mathbf{W})(t) \\
&= 1 + \int_0^t \mathcal{E}(-\boldsymbol{\theta} \bullet \mathbf{W})(\tau)\,\mathrm{d}[S_0(\tau)]^{-1} + \int_0^t [S_0(\tau)]^{-1}\,\mathrm{d}\mathcal{E}(-\boldsymbol{\theta} \bullet \mathbf{W})(\tau) \\
&\quad + [S_0^{-1}, \mathcal{E}(-\boldsymbol{\theta} \bullet \mathbf{W})](t).
\end{aligned} \tag{A.101}$$

Since $[S_0(t)]^{-1}$ is a continuous, finite variation process, the square bracket quadratic co-variation term is given by

$$[S_0^{-1}, \mathcal{E}(-\boldsymbol{\theta} \bullet \mathbf{W})](t) = 0 \quad \text{a.s.} \tag{A.102}$$

From (B.33), the Doléans-Dade exponential $\mathcal{E}(-\boldsymbol{\theta} \bullet \mathbf{W})$ satisfies

$$\begin{aligned} \mathcal{E}(-\boldsymbol{\theta} \bullet \mathbf{W})(t) &= 1 + \int_0^t \mathcal{E}(-\boldsymbol{\theta} \bullet \mathbf{W})(\tau)\, \mathrm{d}(-\boldsymbol{\theta} \bullet \mathbf{W})(\tau) \\ &= 1 - \int_0^t \mathcal{E}(-\boldsymbol{\theta} \bullet \mathbf{W})(\tau)\, \boldsymbol{\theta}^\top(\tau)\, \mathrm{d}\mathbf{W} \end{aligned} \tag{A.103}$$

From the definition of $S_0(t)$ in (2.19),

$$\mathrm{d}[S_0(\tau)]^{-1} = r(\tau) S_0(\tau)\, \mathrm{d}\tau \tag{A.104}$$

Substituting (A.103) and (A.104) in (A.101), we find

$$\begin{aligned} H(t) &= 1 - \int_0^t r(\tau)[S_0(\tau)]^{-1} \mathcal{E}(-\boldsymbol{\theta} \bullet \mathbf{W})(\tau)\, \mathrm{d}\tau \\ &\qquad - \int_0^t [S_0(\tau)]^{-1} \mathcal{E}(-\boldsymbol{\theta} \bullet \mathbf{W})(\tau) \boldsymbol{\theta}^\top(\tau)\, \mathbf{W}(\tau) \\ &= 1 - \int_0^t r(\tau) H(\tau)\, \mathrm{d}\tau - \int_0^t H(\tau) \boldsymbol{\theta}^\top(\tau)\, \mathrm{d}\mathbf{W}(\tau). \end{aligned} \tag{A.105}$$

From the square-integrability of $\{H(t)\}$ and the fact that the risk free interest rate process is uniformly bounded, we get $-rH \in L_{21}$. Similarly from the square-integrability of $\{H(t)\}$ and the uniform boundedness of $\{\boldsymbol{\theta}(t)\}$, we obtain $-H\boldsymbol{\theta} \in L_2(\mathbf{W})$. Hence $H \equiv (1, -rH, -H\boldsymbol{\theta}, 0) \in \mathbb{A}$.

Proof of Proposition 4.3.11.

Claim (a): The following proof carries over directly from the argument of (Labbé and Heunis [25], proof of Lemma 5.1) in which the filtration is generated only by the Wiener process $\mathbf{W}$. We include all the details, as we wish to ensure that it continues to hold for our larger filtration generated jointly by the Wiener process $\mathbf{W}$ and the Markov chain α (see (2.8) and (2.9)).

From (4.56)

$$X(t) = E[H(t)^{-1} \zeta H(T) | \mathcal{F}_t],\ t \in [0, T]. \tag{A.106}$$

From Proposition 4.3.9 and Hölder's inequality, for any $p \in \mathbb{R}$ we have

$$E|H(T)H(t)^{-1}|^p \le \left(E|H(T)|^{2p}\right)^{\frac{1}{2}} \left(E|H(t)|^{-2p}\right)^{\frac{1}{2}} < \infty. \tag{A.107}$$

Setting $p = 2$ in the above equation, we see that $\frac{H(T)}{H(t)}$ is square-integrable. Using Hölder's inequality,

$$E|\zeta H(T)H(t)^{-1}| \leq \left(E|\zeta|^2\right)^{\frac{1}{2}} \left(E|H(T)H(t)^{-1}|^2\right)^{\frac{1}{2}} < \infty. \quad \text{(A.108)}$$

Then $\zeta H(T)H(t)^{-1}$ is integrable and it follows that $X(t)$ is also integrable.

Now fix real numbers $p \in (2, \infty)$ and $q \in (1, 2)$ which together satisfy $\frac{1}{p} + \frac{1}{q} = 1$. Applying Hölder's inequality for conditional expectation at (A.106) gives

$$|X(t)| \leq \left(E(|H(T)H(t)^{-1}|^p|\mathcal{F}_t)\right)^{\frac{1}{p}} \left(E(|\zeta|^q|\mathcal{F}_t)\right)^{\frac{1}{q}} \quad \text{a.s.} \quad \text{(A.109)}$$

Raising all the terms to the power of q, we obtain

$$|X(t)|^q \leq \left(E(|H(T)H(t)^{-1}|^p|\mathcal{F}_t)\right)^{\frac{q}{p}} \left(E(|\zeta|^q|\mathcal{F}_t)\right) \quad \text{a.s.} \quad \text{(A.110)}$$

We next upper-bound separately the two factors in the product on the right-hand side.

Expand $|H(t)|^p$ using (A.95) and (A.88), we get

$$|H(t)|^p = S_0(t)^{-p} \exp\left\{\frac{1}{2}p(p-1)\int_0^t \parallel \boldsymbol{\theta}(\tau) \parallel^2 \, \mathrm{d}\tau\right\} \mathcal{E}(-p\boldsymbol{\theta} \bullet \mathbf{W})(t). \quad \text{(A.111)}$$

By the nonnegativity of the risk-free interest rate process $\{r(t)\}$, we have $[S_0(t)]^{-1} \leq 1$ a.s., Consequently for all $t \in [0, T]$,

$$|H(t)|^p \leq \exp\left\{\frac{1}{2}p(p-1)\int_0^t \parallel \boldsymbol{\theta}(\tau) \parallel^2 \, \mathrm{d}\tau\right\} \mathcal{E}(-p\boldsymbol{\theta} \bullet \mathbf{W})(t). \quad \text{(A.112)}$$

Then, recalling the constant $\kappa_{\boldsymbol{\theta}}$ which satisfies (2.25), we have

$$\begin{aligned} |H(T)|^p|H(t)|^{-p} &\leq \exp\left\{\frac{1}{2}p(p-1)\int_0^T \parallel \boldsymbol{\theta} \parallel^2 \, \mathrm{d}\tau + \frac{1}{2}p(p+1)\int_0^t \parallel \boldsymbol{\theta} \parallel^2 \, \mathrm{d}\tau\right\} \\ &\quad \mathcal{E}(-p\boldsymbol{\theta} \bullet \mathbf{W})(T)\mathcal{E}(p\boldsymbol{\theta} \bullet \mathbf{W})(t) \\ &\leq \exp\{p^2\kappa_{\boldsymbol{\theta}}^2 T\}\mathcal{E}(-p\boldsymbol{\theta} \bullet \mathbf{W})(T)\mathcal{E}(p\boldsymbol{\theta} \bullet \mathbf{W})(t). \end{aligned} \quad \text{(A.113)}$$

From Proposition A.0.30 $\mathcal{E}(-p\boldsymbol{\theta} \bullet \mathbf{W})$ is an $\{\mathcal{F}_t\}$-martingale. Using this fact and Corollary B.2.52, we obtain

$$\begin{aligned} E\left(|H(T)|^p|H(t)|^{-p}|\mathcal{F}_t\right) &\leq \exp\{p^2\kappa_{\boldsymbol{\theta}}^2 T\}\mathcal{E}(-p\boldsymbol{\theta} \bullet \mathbf{W})(t)\mathcal{E}(p\boldsymbol{\theta} \bullet \mathbf{W})(t) \\ &= \exp\{p^2\kappa_{\boldsymbol{\theta}}^2 T\}\exp\left\{-p^2\int_0^t \parallel \boldsymbol{\theta} \parallel^2 \, \mathrm{d}\tau\right\} \quad \text{(A.114)} \\ &\leq \exp\{p^2\kappa_{\boldsymbol{\theta}}^2 T\} \quad \text{(A.115)} \end{aligned}$$

Substituting this bound into (A.110), we get

$$|X(t)|^q \leq \exp\{pq\kappa_{\boldsymbol{\theta}}^2 T\}\left(E(|\zeta|^q|\mathcal{F}_t)\right) \quad \text{a.s.} \tag{A.116}$$

We now show that $(E(|\zeta|^q|\mathcal{F}_t))$ is finite. Note first that $q \in (1,2)$, so that we have

$$E(|\zeta|^q) < E(|\zeta|^2) < \infty. \tag{A.117}$$

Now define

$$N(t) := E(|\zeta|^q|\mathcal{F}_t). \tag{A.118}$$

Then of course $N \in \mathcal{M}(\{\mathcal{F}_t\}, \mathbb{P})$. Set

$$p_1 := \frac{2}{q} \tag{A.119}$$

and note that $p_1 > 1$ as $q \in (1,2)$. Now apply Jensen's inequality for conditional expectations to obtain to obtain for all $t \in [0,T]$,

$$E|N(t)|^{p_1} \overset{(A.118)}{=} E|E(|\zeta|^q|\mathcal{F}_t)|^{p_1} \tag{A.120}$$
$$\leq E|E(|\zeta|^{qp_1}|\mathcal{F}_t)| \tag{A.121}$$
$$\overset{(A.119)}{=} E|E(|\zeta|^2|\mathcal{F}_t)| \tag{A.122}$$
$$= E|\zeta|^2 < \infty. \tag{A.123}$$

As N is a martingale which is bounded in $L_{p_1}(\Omega, \mathcal{F}, \mathbb{P})$ and $p_1 > 1$, we can apply Doob's L_{p_1} inequality to get

$$E\left(\sup_{t\in[0,T]} |N(t)|^{p_1}\right) \leq \left(\frac{p_1}{p_1 - 1}\right)^{p_1} E|N(T)|^{p_1} \overset{(A.123)}{<} \infty \tag{A.124}$$

Substituting $N(t) = E(|\zeta|^q|\mathcal{F}_t)$ into (A.116) and raising both sides of (A.116) to the power $p_1 = \frac{2}{q}$, we get

$$|X(t)|^2 \leq \exp\{2p\kappa_{\boldsymbol{\theta}}^2 T\} N(t)^{p_1}. \quad \text{a.s.} \tag{A.125}$$

Taking the supremum over $t \in [0,T]$, followed by expectations, we find

$$E\left(\sup_{t\in[0,T]} |X(t)|^2\right) \leq \exp\{2p\kappa_{\boldsymbol{\theta}}^2 T\} E\left(\sup_{t\in[0,T]} |N(t)|^{p_1}\right) \overset{(A.124)}{<} \infty. \tag{A.126}$$

Claim (b): From Hölder's inequality and the square integrability of ζ and $H(T)$,

$$E|\zeta H(T)| \leq \left[E|\zeta|^2\right]^{\frac{1}{2}} \left[E|H(T)|^2\right]^{\frac{1}{2}} < \infty. \tag{A.127}$$

From (4.56) and (A.127) it is immediate that

$$XH \in \mathcal{M}\left(\{\mathcal{F}_t, \mathbb{P}\}\right). \tag{A.128}$$

For all $m \in \mathbb{N}$, let

$$T^m := \inf\{t > 0 : |H(t)| > m\} \wedge T. \tag{A.129}$$

Then T^m is an $\{\mathcal{F}_t\}$-stopping time (by Proposition B.2.23) and $T^m \Uparrow T$ a.s. (see Definition B.2.11), since $\sup_{t\in[0,T]}\{H(t)\}$ is finite a.s. by the pathwise continuity of H on the compact interval $[0, T]$.

Fix $m \in \mathbb{N}$. Then for all $t \in [0, T]$,

$$E|X(t \wedge T^m)H(t \wedge T^m)|^2 \overset{\text{(A.129)}}{\leq} m^2 E\left|\sup_{t\in[0,T]} X(t)^2\right| \overset{(4.57)}{<} \infty, \tag{A.130}$$

showing that XH is locally square integrable. Thus from (A.128) and (A.130) XH is locally square integrable martingale. As any martingale is also a local martingale, we have

$$XH \in \mathcal{M}_2^{loc}\left(\{\mathcal{F}_t, \mathbb{P}\}\right). \tag{A.131}$$

Claim (c): Follows from (A.131) and the Martingale representation theorem for locally square-integrable martingales, Theorem G.3.5.

Claim (d): The predictability of the $\tilde{\boldsymbol{\pi}}$ is immediate from (4.62). To prove (4.63), recall the constants $\kappa_{\boldsymbol{\sigma}} \in (0, \infty)$ and $\kappa_{\boldsymbol{\theta}} \in (0, \infty)$ from (2.23) and (2.25), respectively. Then

$$\begin{aligned}
\int_0^T \parallel \tilde{\boldsymbol{\pi}}(t) \parallel^2 \, \mathrm{d}t &\overset{(4.62)}{=} \int_0^T \parallel [\boldsymbol{\sigma}^\top]^{-1}\{H^{-1}(t)\tilde{\boldsymbol{\xi}}(t) + X(t_-)\boldsymbol{\theta}(t)\} \parallel^2 \, \mathrm{d}t \\
&\leq 2\kappa_{\boldsymbol{\sigma}}^2 \int_0^T \parallel H^{-1}(t)\tilde{\boldsymbol{\xi}} \parallel^2 \, \mathrm{d}t + 2\kappa_{\boldsymbol{\sigma}}^2 \int_0^T \parallel X(t_-)\boldsymbol{\theta}(t) \parallel^2 \, \mathrm{d}t \\
&\leq 2\kappa_{\boldsymbol{\sigma}}^2 \sup_{t\in[0,T]}\{H^{-2}(t)\} \int_0^T \parallel \tilde{\boldsymbol{\xi}} \parallel^2 \, \mathrm{d}t + 2\kappa_{\boldsymbol{\sigma}}^2\kappa_{\boldsymbol{\theta}}^2 T \sup_{t\in[0,T]} |X(t)|^2.
\end{aligned} \tag{A.132}$$

We show that the last line of the above inequality is finite. Since H^{-2} is pathwise continuous function on compact interval $[0,T]$, then not only is the set $\{H^{-2}(t) : t \in [0,T]\}$ bounded, but it also attains its bounds. Therefore,

$$\sup_{t\in[0,T]} \{H^{-2}(t)\} < \infty \quad \text{a.s.} \tag{A.133}$$

We also have from (4.59) that $\tilde{\boldsymbol{\xi}} \in L_2^{loc}(\mathbf{W})$, so there exists a sequence $(S^m)_{m\in\mathbb{N}}$ of $\{\mathcal{F}_t\}$-stopping times such that $S^m \Uparrow T$ a.s. and $\tilde{\boldsymbol{\xi}}[0,S^m] \in L_2(\mathbf{W})$ for all $m \in \mathbb{N}$. Then for each $m \in \mathbb{N}$,

$$E\int_0^{T\wedge S^m} \| \tilde{\boldsymbol{\xi}}(t) \|^2 \;\mathrm{d}t < \infty \Rightarrow \int_0^{T\wedge S^m} \| \tilde{\boldsymbol{\xi}}(t) \|^2 \;\mathrm{d}t < \infty \text{ a.s.} \tag{A.134}$$

Since $S^m \Uparrow T$ a.s., there exists $M(\omega) \in \mathbb{N}$ such that $S^m(\omega) = T$ for all $m > M(\omega)$ for all $\omega \in \Omega$. Then, letting $m \to \infty$ in (A.134), we obtain

$$\int_0^T \| \tilde{\boldsymbol{\xi}}(t) \|^2 \;\mathrm{d}t < \infty \quad \text{a.s.} \tag{A.135}$$

Moreover, from (4.57), we get

$$\sup_{t\in[0,T]} |X(t)|^2 < \infty \quad \text{a.s.} \tag{A.136}$$

From (A.133), (A.135) and (A.136) it follows that last line of Equation (A.132) is finite.

Claim (e): Begin by defining for $t \in [0,T]$,

$$\Lambda(t) := X(t)H(t). \tag{A.137}$$

Applying the integration-by-parts formula (Theorem B.2.48) to the product $X(t) = H(t)^{-1}\Lambda(t)$ and using the continuity of H gives

$$X(t) = X_0 + \int_0^t H(\tau)^{-1}\,\mathrm{d}\Lambda(\tau) + \int_0^t \Lambda(\tau_-)\,\mathrm{d}H^{-1}(\tau) + [H^{-1},\Lambda](t). \tag{A.138}$$

Expanding H^{-1} using Itô's formula (Theorem B.2.49), we find

$$H(t)^{-1} = 1 - \int_0^t H(\tau)^{-2}\,\mathrm{d}H(\tau) + \int_0^t H(\tau)^{-3}\,\mathrm{d}[H,H](\tau). \tag{A.139}$$

From Proposition 4.3.10,

$$\mathrm{d}H(\tau) = -r(\tau)H(\tau)\,\mathrm{d}\tau - H(\tau)\boldsymbol{\theta}^\top(\tau)\mathrm{d}\mathbf{W}(\tau) \tag{A.140}$$

and

$$\mathrm{d}[H,H](\tau) = H^2(\tau) \parallel \boldsymbol{\theta}(\tau) \parallel^2 \mathrm{d}\tau. \tag{A.141}$$

Substituting (A.140) and (A.141) into (A.139) gives,

$$H(t)^{-1} = 1 + \int_0^t H(\tau)^{-1}(r(\tau) + \parallel \boldsymbol{\theta}(\tau) \parallel^2)\,\mathrm{d}\tau + \int_0^t H(\tau)^{-1}\boldsymbol{\theta}^\top(\tau)\,\mathrm{d}\mathbf{W}(\tau) \tag{A.142}$$

From (A.137) and part (c), we get

$$\mathrm{d}\Lambda(t) = \sum_{n=1}^{N} \tilde{\xi}_n(\tau)\,\mathrm{d}W_n(\tau) + \sum_{i,j=1}^{D} \tilde{\Gamma}_{i,j}(\tau)\,\mathrm{d}M_{ij}(\tau) \tag{A.143}$$

Substituting (A.142) and (A.143) into equation (A.138) and using the continuity of H, we find that

$$\begin{aligned}
X(t) &= X_0 + \int_0^t H(\tau)^{-1}\left(\sum_{n=1}^{N} \tilde{\xi}_n(\tau)\,\mathrm{d}W_n(\tau) + \sum_{i,j=1}^{D} \tilde{\Gamma}_{i,j}(\tau)\,\mathrm{d}M_{ij}(\tau)\right) \\
&\quad + \int_0^t X(\tau_-)H(\tau)\left\{H(\tau)^{-1}(r(t) + \parallel \boldsymbol{\theta}(\tau) \parallel^2)\,\mathrm{d}\tau + H(\tau)^{-1}\boldsymbol{\theta}^\top(\tau)\,\mathrm{d}\mathbf{W}(\tau)\right\} \\
&\quad + \Bigg[1 + \int_0 H(\tau)^{-1}(r(\tau) + \parallel \boldsymbol{\theta}(\tau) \parallel^2)\,\mathrm{d}\tau + \int_0 H(\tau)^{-1}\boldsymbol{\theta}^\top(\tau)\,\mathrm{d}\mathbf{W}(\tau)\,, \\
&\qquad X_0 + \sum_{n=1}^{N} \int_0 \tilde{\xi}_n(\tau)\,\mathrm{d}W_n(\tau) + \sum_{i,j=1}^{D} \int_0 \tilde{\Gamma}_{i,j}(\tau)\,\mathrm{d}M_{ij}(\tau)\Bigg](t) \\
&= X_0 + \int_0^t H(\tau)^{-1}\tilde{\boldsymbol{\xi}}^\top(\tau)\,\mathrm{d}\mathbf{W}(\tau) + \sum_{i,j=1}^{D} \int_0^t H(\tau)^{-1}\tilde{\Gamma}_{ij}(\tau)\,\mathrm{d}M_{ij}(\tau) \\
&\quad + \int_0^t X(\tau_-)r(\tau)\mathrm{d}\tau + \int_0^t X(\tau_-) \parallel \boldsymbol{\theta}(\tau) \parallel^2\, \mathrm{d}\tau + \int_0^t X(\tau_-)\boldsymbol{\theta}^\top(\tau)\,\mathrm{d}\mathbf{W}(\tau) \\
&\quad + \int_0^t H(\tau)^{-1}\tilde{\boldsymbol{\xi}}^\top(\tau)\boldsymbol{\theta}(\tau)\,\mathrm{d}\tau \\
&\quad + \sum_{i,j=1}^{D} \sum_{n=1}^{N} \int_0^t H(\tau)^{-1}\theta_n(\tau)\Gamma_{ij}(\tau)\,\mathrm{d}[W_n, M_{ij}](\tau).
\end{aligned} \tag{A.144}$$

From Lemma G.2.18, we have $[M_{ij}, W_n](t) = 0 \quad \text{for all } t \in [0,T]$. Equation (A.144) therefore becomes

$$\begin{aligned} X(t) &= X_0 \\ &+ \int_0^t \Big\{ r(\tau)X(\tau_-) + \Big(H(\tau)^{-1}\tilde{\boldsymbol{\xi}}^\top(\tau) + X(\tau_-)\boldsymbol{\theta}^\top(\tau) \Big) \boldsymbol{\theta}(\tau) \Big\} \, \mathrm{d}\tau \\ &+ \int_0^t \Big(H(\tau)^{-1}\tilde{\boldsymbol{\xi}}^\top(\tau) + X(\tau_-)\boldsymbol{\theta}^\top \Big) \, \mathrm{d}\mathbf{W}(\tau) \\ &+ \sum_{i,j=1}^{D} \int_0^t H(\tau)^{-1}\tilde{\Gamma}_{ij}(\tau) \, \mathrm{d}M_{ij}(\tau). \end{aligned} \tag{A.145}$$

Substituting the value of $\tilde{\boldsymbol{\pi}}$ from Equation (4.62) into Equation (A.145) we get

$$\begin{aligned} X(t) &= X_0 + \int_0^t \{ r(\tau)X(t) + \tilde{\boldsymbol{\pi}}^\top(\tau)\boldsymbol{\sigma}(\tau)\boldsymbol{\theta}(\tau)\}\mathrm{d}\tau + \int_0^t \tilde{\boldsymbol{\pi}}^\top \boldsymbol{\sigma}(\tau) \, \mathrm{d}\mathbf{W}(\tau) \\ &+ \sum_{i,j=1}^{D} \int_0^t H(\tau)^{-1}\tilde{\Gamma}_{ij}(\tau) \, \mathrm{d}M_{ij}(\tau). \end{aligned} \tag{A.146}$$

Claim (f): For each $m \in \mathbb{N}$, let

$$R^m := \inf\left\{ t > 0 : \int_0^t \parallel \tilde{\boldsymbol{\pi}}(\tau) \parallel^2 \, \mathrm{d}\tau > m \right\} \wedge T. \tag{A.147}$$

Then R^m is an $\mathcal{F}_t$-stopping time (by Proposition B.2.23) and $R^m \Uparrow T$ a.s., since $\int_0^T \parallel \tilde{\boldsymbol{\pi}}(\tau) \parallel^2 \, \mathrm{d}\tau < \infty$ a.s. by (4.63).

For each $m \in \mathbb{N}$, let

$$S^m := \inf\{ t > 0 : |X(t_-)|^2 > m \} \wedge T. \tag{A.148}$$

Then S^m is an $\mathcal{F}_t$-stopping time since $X(t_-)$ is locally bounded and $S^m \Uparrow T$ a.s., since $\sup_{t\in[0,T]} |X(t)|^2 < \infty$ a.s. by (4.57).

For $m \in \mathbb{N}$, let

$$T^m := \inf\{ t > 0 : |H(t)^{-1}| > m \} \wedge T. \tag{A.149}$$

Then T^m is an $\{\mathcal{F}_t\}-$ stopping time (by Proposition B.2.23) and $T^m \Uparrow T$ a.s., since $\sup_{t\in[0,T]}\{H(t)^{-2}\}$ is finite a.s. by the pathwise continuity of H^{-2} on the compact interval $[0, T]$.

Since $\tilde{\Gamma} \in L_2^{loc}(\mathbf{M})$, there exists a sequence $\{U^m\}_{m\in\mathbb{N}}$, of $\{\mathcal{F}_t\}$-stopping times such that $U^m \Uparrow T$ a.s. and $\tilde{\mathbf{\Gamma}}[0, U^m] \in L_2(\mathbf{M})$ for all $m \in \mathbb{N}$. Then for all $m \in \mathbb{N}$,

$$E \sum_{i,j=1}^{D} \int_0^{T\wedge U^m} |\tilde{\Gamma}_{ij}(t)|^2 \, \mathrm{d}[M_{ij}](t) < \infty. \tag{A.150}$$

Finally, define

$$V^m := R^m \wedge S^m \wedge T^m \wedge U^m. \tag{A.151}$$

Then V^m is an $\{\mathcal{F}_t\}$-stopping time and $V^m \Uparrow T$ a.s.

Applying the integration by parts formula (Theorem B.2.48) to (4.64), we can expand the mapping $t \to X(t)^2$. Evaluating the expansion at time $t \wedge V^m$, gives for all $t \in [0, T]$,

$$\begin{aligned} X^2(t\wedge V^m) &= X^2(0) + 2\int_0^{t\wedge V^m} X(\tau_-)\,\mathrm{d}X(\tau) + [X,X](t\wedge V^m) \\ &= X^2(0) + 2\int_0^{t\wedge V^m} X(\tau_-)\left(r(\tau)X(\tau_-) + \tilde{\boldsymbol{\pi}}^\top(\tau)\boldsymbol{\sigma}(\tau)\boldsymbol{\theta}(\tau)\right)\,\mathrm{d}\tau \\ &\quad + 2\int_0^{t\wedge V^m} X(\tau_-)\tilde{\boldsymbol{\pi}}^\top(\tau)\boldsymbol{\sigma}(\tau)\,\mathrm{d}\mathbf{W}(\tau) \\ &\quad + 2\sum_{i,j=1}^{D}\int_0^{t\wedge V^m} H(\tau)^{-1}X(\tau_-)\tilde{\Gamma}_{ij}(\tau)\,\mathrm{d}M_{ij}(\tau) \\ &\quad + \int_0^{t\wedge V^m} \parallel \boldsymbol{\sigma}^\top(\tau)\tilde{\boldsymbol{\pi}}(\tau) \parallel^2 \,\mathrm{d}\tau \\ &\quad + \sum_{i,j=1}^{D}\int_0^{t\wedge V^m} H(\tau)^{-2}|\tilde{\Gamma}_{ij}(\tau)|^2\,\mathrm{d}[M_{ij}](\tau). \end{aligned} \tag{A.152}$$

We show that the $\mathrm{d}\mathbf{W}(\tau)$- and $\mathrm{d}M_{ij}(\tau)$-stochastic integrals on the right side of (A.152) are square-integrable martingales. Recall the constant $\kappa_{\boldsymbol{\sigma}} \in (0, \infty)$ satisfying (2.23). For the fourth to last term, we have

$$\begin{aligned} E\int_0^{t\wedge V^m} \parallel X(\tau_-)\boldsymbol{\sigma}^\top(\tau)\tilde{\boldsymbol{\pi}}(\tau) \parallel^2 \,\mathrm{d}\tau \quad &\leq \quad \kappa_{\boldsymbol{\sigma}}^2 E\left(\sup_{t\in[0,T]} |X(t)|^2 \int_0^{t\wedge V^m} \parallel \tilde{\boldsymbol{\pi}}(\tau) \parallel^2 \,\mathrm{d}\tau\right) \\ &\overset{(A.147)}{\leq} \quad \kappa_{\boldsymbol{\sigma}}^2 m E\left(\sup_{t\in[0,T]} |X(t)|^2\right) \\ &\overset{(4.57)}{<} \quad \infty. \end{aligned} \tag{A.153}$$

Then for each $m \in \mathbb{N}$, the fourth to last term in (A.152) is a square integrable martingale, which is clearly null at the origin, so that for all $t \in [0, T]$,

$$E \int_0^{t \wedge V^m} X(\tau_-) \boldsymbol{\sigma}^\top(\tau) \tilde{\boldsymbol{\pi}}(\tau) \, \mathrm{d}\mathbf{W}(\tau) = 0 \tag{A.154}$$

For the third to last term in (A.152), we have

$$\begin{aligned} & E \sum_{i,j=1}^{D} \int_0^{t \wedge V^m} |H(\tau)^{-1} X(\tau_-) \tilde{\Gamma}_{ij}(\tau)|^2 \, \mathrm{d}[M_{ij}](\tau) \\ & \qquad = \quad E \sum_{i,j=1}^{D} \int_0^{t \wedge V^m} |H(\tau)|^{-2} |X(\tau_-)| |\tilde{\Gamma}_{ij}(\tau)|^2 \, \mathrm{d}[M_{ij}](\tau) \\ & \quad \overset{\text{(A.148),(A.149)}}{\leq} \quad m^2 E \sum_{i,j=1}^{D} \int_0^{t \wedge V^m} |\tilde{\Gamma}_{ij}(\tau)|^2 \, \mathrm{d}[M_{ij}](\tau) \\ & \qquad \overset{\text{(A.150)}}{<} \quad \infty. \end{aligned} \tag{A.155}$$

Then for each $m \in \mathbb{N}$, the third to last term in (A.152) is a square integrable martingale, which is clearly null at the origin, so that for all $t \in [0, T]$,

$$\sum_{i,j=1}^{D} E \int_0^{t \wedge V^m} H(\tau)^{-1} X(\tau_-) \tilde{\Gamma}_{ij}(\tau) \, \mathrm{d}M_{ij}(\tau) = 0. \tag{A.156}$$

Hence the sum of the third and fourth to last terms of (A.152)is a maringale, null at the origin, and thus has zero expectation for all $t \in [0, T]$. Evaluating (A.152) at time $t = T$, noting that $T \wedge V^m = V^m$, and taking expectations, we get

$$\begin{aligned} EX^2(V^m) \quad = \quad & EX^2(0) + E \int_0^{V^m} 2X(\tau_-) \left(r(\tau) X(\tau_-) + \tilde{\boldsymbol{\pi}}^\top(\tau) \boldsymbol{\sigma}(\tau) \boldsymbol{\theta}(\tau) \right) \, \mathrm{d}\tau \\ & + E \int_0^{V^m} \| \, \boldsymbol{\sigma}(\tau)^\top \tilde{\boldsymbol{\pi}}(\tau) \, \|^2 \, \mathrm{d}\tau + \sum_{i,j=1}^{D} E \int_0^{V^m} H(\tau)^{-2} |\tilde{\Gamma}(\tau)|^2 \, \mathrm{d}[M_{ij}](\tau). \end{aligned} \tag{A.157}$$

From the nonnegativity of the risk-free interest rate process $\{r(t)\}$ and X_-^2, upon rearranging (A.157) we obtain the inequality

$$\begin{aligned} & EX^2(V^m) - E \int_0^{V^m} 2X(\tau_-) \tilde{\boldsymbol{\pi}}^\top(\tau) \boldsymbol{\sigma}(\tau) \boldsymbol{\theta}(\tau) \, \mathrm{d}\tau \\ & \quad \geq E \int_0^{V^m} \| \, \boldsymbol{\sigma}^\top(\tau) \tilde{\boldsymbol{\pi}}(\tau) \, \|^2 \, \mathrm{d}\tau + \sum_{i,j=1}^{D} E \int_0^{V^m} H(\tau)^{-2} |\tilde{\Gamma}_{ij}(\tau)|^2 \, \mathrm{d}[M_{ij}](\tau). \end{aligned} \tag{A.158}$$

Now, for arbitrary $\boldsymbol{\nu}_1, \boldsymbol{\nu}_2 \in \mathbb{R}^N$, we have the inequality

$$\boldsymbol{\nu}_1^\top \boldsymbol{\nu}_2 \leq \frac{1}{2} \parallel \boldsymbol{\nu}_1 \parallel^2 + \frac{1}{2} \parallel \boldsymbol{\nu}_2 \parallel^2 . \tag{A.159}$$

Setting $\boldsymbol{\nu}_1 = \boldsymbol{\sigma}^\top(\tau)\tilde{\boldsymbol{\pi}}(\tau)$ and $\boldsymbol{\nu}_2 = -2X(\tau_-)\boldsymbol{\theta}(\tau)$, we then get

$$-2X(\tau_-)\boldsymbol{\sigma}^\top(\tau)\tilde{\boldsymbol{\pi}}(\tau)\boldsymbol{\theta}(\tau) \leq \frac{1}{2} \parallel \boldsymbol{\sigma}^\top(\tau)\tilde{\boldsymbol{\pi}}(\tau) \parallel^2 + 2X^2(\tau_-) \parallel \boldsymbol{\theta}(\tau) \parallel^2 . \tag{A.160}$$

Integrating, taking expectations and adding $EX^2(V^m)$ to each side of the above inequality, we get

$$\begin{aligned} &EX^2(V^m) - E\int_0^{V^m} 2X(\tau_-)\boldsymbol{\sigma}^\top(\tau)\tilde{\boldsymbol{\pi}}(\tau)\boldsymbol{\theta}(\tau)\,\mathrm{d}\tau \\ &\quad \leq EX^2(V^m) + \frac{1}{2}E\int_0^{V^m} \parallel \boldsymbol{\sigma}^\top(\tau)\tilde{\boldsymbol{\pi}}(\tau) \parallel^2 \,\mathrm{d}\tau + 2E\int_0^{V^m} X^2(\tau_-) \parallel \boldsymbol{\theta}(\tau) \parallel^2 \,\mathrm{d}\tau. \end{aligned} \tag{A.161}$$

Combining (A.158) and (A.161), we get

$$\begin{aligned} &E\int_0^{V^m} \parallel \boldsymbol{\sigma}^\top(\tau)\tilde{\boldsymbol{\pi}}(\tau) \parallel^2 \,\mathrm{d}\tau + \sum_{i,j=1}^{D} E\int_0^{V^m} H(\tau)^{-2}|\tilde{\Gamma}_{ij}(\tau)|^2\,\mathrm{d}[M_{ij}](\tau) \\ &\quad \leq EX^2(V^m) + \frac{1}{2}E\int_0^{V^m} \parallel \boldsymbol{\sigma}^\top(\tau)\tilde{\boldsymbol{\pi}}(\tau) \parallel^2 \,\mathrm{d}\tau + 2E\int_0^{V^m} X^2(\tau_-) \parallel \boldsymbol{\theta}(\tau) \parallel^2 \,\mathrm{d}\tau. \end{aligned} \tag{A.162}$$

Recall the constant $\kappa_{\boldsymbol{\theta}} \in (0,\infty)$ satisfying (2.25). Rearranging (A.162),

$$\begin{aligned} E\int_0^{V^m} \parallel \boldsymbol{\sigma}^\top(\tau)\tilde{\boldsymbol{\pi}}(\tau) \parallel^2 \,\mathrm{d}\tau &+ \sum_{i,j=1}^{D} E\int_0^{V^m} H(\tau)^{-2}|\tilde{\Gamma}_{ij}(\tau)|^2\,\mathrm{d}[M_{ij}](\tau) \\ &\leq (1 + 2T\kappa_{\boldsymbol{\theta}}^2)E\left(\sup_{t\in[0,T]} |X(t)|^2\right) \\ &\overset{(4.57)}{<} \infty. \end{aligned} \tag{A.163}$$

Since $V^m \Uparrow T$ a.s., and the upper-bound at the first inequality is uniform with respect to the integer m, upon letting $m \to \infty$ in (A.163), we obtain

$$\boldsymbol{\sigma}^\top \tilde{\boldsymbol{\pi}} \in L_2(\mathbf{W}) \quad \text{and} \quad \frac{1}{H}\tilde{\boldsymbol{\Gamma}} \in L_2(\mathbf{M}). \tag{A.164}$$

Proof of Lemma 4.3.15. The proof which follows is taken from Donnelly and Heunis [7], and is included here only for completeness. Existence and uniqueness of solutions are obtained from the Banach contraction mapping theorem.

As in (2.31) define the norm $\| \cdot \|_{L_2(\mathbf{W})}$ on $L_2(\mathbf{W})$ by

$$\| \boldsymbol{\lambda} \|^2_{L_2(\mathbf{W})} := E \int_0^T \| \boldsymbol{\lambda}(t) \|^2 \, \mathrm{d}t. \tag{A.165}$$

Here $\| \cdot \|$ after the integral on the right side denotes the usual Euclidean norm on $\mathbb{R}^N$. With the $\| \cdot \|_{L_2(\mathbf{W})}$-norm it is immediate that $L_2(\mathbf{W})$ is a Banach space.

Put

$$\eta(t) := \sum_{i,j=1}^{D} \int_0^t \gamma_{ij}(s) \, \mathrm{d}M_{ij}(s) \tag{A.166}$$

$$\text{and} \qquad \boldsymbol{\Lambda}(t) := \boldsymbol{\rho}(t) - \boldsymbol{\theta}(t)\eta(t_-), \quad t \in [0,T]. \tag{A.167}$$

For each $\boldsymbol{\lambda} \in L_2(\mathbf{W})$ put

$$\mathbb{G}\boldsymbol{\lambda}(t) := \boldsymbol{\Lambda}(t) - \boldsymbol{\theta}(t) \int_0^t \boldsymbol{\lambda}^\top(s) \, \mathrm{d}\mathbf{W}(s), \qquad t \in [0,T]. \tag{A.168}$$

Let $\kappa_{\boldsymbol{\theta}} \in (0, \infty)$ satisfy (2.25), so that $\kappa_{\boldsymbol{\theta}}$ is uniform bound on $\| \boldsymbol{\theta} \|$. Then

$$\begin{aligned}
& E \int_0^T \| \mathbb{G}\boldsymbol{\lambda}(t) \|^2 \, \mathrm{d}t \\
& \overset{(A.167),(A.168)}{=} E \int_0^T \| \boldsymbol{\rho}(t) - \boldsymbol{\theta}(t)\eta(t_-) - \boldsymbol{\theta}(t) \int_0^t \boldsymbol{\lambda}^\top(s) \, \mathrm{d}\mathbf{W}(s) \|^2 \, \mathrm{d}t \qquad (A.169) \\
& \overset{(A.166)}{\leq} 3E \int_0^T |\boldsymbol{\rho}(t)|^2 \, \mathrm{d}t + 3\kappa^2_{\boldsymbol{\theta}} E \int_0^T \left| \sum_{i,j=1}^{D} \int_0^{t_-} \gamma_{ij}(s) \, \mathrm{d}M_{ij}(s) \right|^2 \mathrm{d}t \\
& \qquad + 3\kappa^2_{\boldsymbol{\theta}} E \int_0^T \left| \int_0^t \boldsymbol{\lambda}^\top(s) \, \mathrm{d}\mathbf{W}(s) \right|^2 \mathrm{d}t. \qquad (A.170)
\end{aligned}$$

Using Doob's L_2-inequality (see Theorem B.2.9) for the càdlàg martingale $M(t) =$

$\eta(t_-)$ and the Itô isometry for the second last integral of (A.170), we get

$$\begin{aligned} E\int_0^T |M(\tau)|^2\,\mathrm{d}\tau &\leq TE\left(\sup_{s\in[0,T]} |M(t)|^2\right) \\ &\overset{\text{Doob}}{\leq} 4TE|M(T)|^2 \\ &= 4TE\left|\sum_{i,j=1}^{D}\int_0^{T_-}\gamma_{ij}(\tau)\,\mathrm{d}M_{ij}(\tau)\right|^2 \\ &\overset{\text{Itô}}{=} 4T\sum_{i,j=1}^{D} E\int_0^{T_-}|\gamma_{ij}(\tau)|^2\,\mathrm{d}[M_{ij}](\tau) \\ &= 4T \parallel \boldsymbol{\gamma} \parallel^2_{L_2(\mathbf{M})} < \infty. \end{aligned} \tag{A.171}$$

Again, using Doob's L_2-inequality and the Itô isometry to evaluate the last integral of (A.170), we get

$$\begin{aligned} E\int_0^T \left|\sum_{i,j=1}^{D}\int_0^t \boldsymbol{\lambda}^\top(\tau)\,\mathrm{d}\mathbf{W}(\tau)\right|^2 \mathrm{d}t &\leq TE\left(\sup_{s\in[0,T]}\left|\boldsymbol{\lambda}^\top(\tau)\,\mathrm{d}\mathbf{W}(\tau)\right|^2\right) \\ &\overset{\text{Doob}}{\leq} 4TE\left|\int_0^T \boldsymbol{\lambda}^\top(\tau)\,\mathrm{d}\mathbf{W}(\tau)\right|^2 \\ &\overset{\text{Itô}}{=} 4TE\int_0^T \parallel \boldsymbol{\lambda}^\top(\tau) \parallel^2\ \mathrm{d}\tau \\ &= 4T \parallel \boldsymbol{\lambda} \parallel^2_{L_2(\mathbf{W})}. \end{aligned} \tag{A.172}$$

From (A.171), (A.172) and (A.170), we get

$$\mathbb{G}\boldsymbol{\lambda} \in L_2(\mathbf{W}) \tag{A.173}$$

Now fix $\boldsymbol{\lambda}_1$ and $\boldsymbol{\lambda}_2 \in L_2(\mathbf{W})$. Recalling (2.25), it follows from (A.168) and the Itô isometry that

$$E \parallel \mathbb{G}\boldsymbol{\lambda}_1(t) - \mathbb{G}\boldsymbol{\lambda}_2(t) \parallel^2 \leq \kappa^2_{\boldsymbol{\theta}} \parallel \boldsymbol{\lambda}_1 - \boldsymbol{\lambda}_2 \parallel^2_{L_2(\mathbf{W})}. \tag{A.174}$$

and also for all $t \in [0,T]$ and $m = 1, 2, \ldots$ we get

$$E \parallel \mathbb{G}^{m+1}\boldsymbol{\lambda}_1(t) - \mathbb{G}^{m+1}\boldsymbol{\lambda}_2(t) \parallel^2 \leq \kappa^2_{\boldsymbol{\theta}} \int_0^t E \parallel \mathbb{G}^m\boldsymbol{\lambda}_1(t) - \mathbb{G}^m\boldsymbol{\lambda}_2(t) \parallel^2\ \mathrm{d}t. \tag{A.175}$$

By induction, for each $t \in [0,T]$ and $m = 1, 2, \ldots$ we then have the bound

$$E \parallel \mathbb{G}^m \boldsymbol{\lambda}_1(t) - \mathbb{G}^m \boldsymbol{\lambda}_2(t) \parallel^2 \leq \kappa_{\boldsymbol{\theta}}^{2m} \frac{t^{m-1}}{(m-1)!} \parallel \boldsymbol{\lambda}_1 - \boldsymbol{\lambda}_2 \parallel^2_{L_2(\mathbf{W})} \tag{A.176}$$

and this in turn gives

$$\parallel \mathbb{G}^m \boldsymbol{\lambda}_1 - \mathbb{G}^m \boldsymbol{\lambda}_2 \parallel^2_{L_2(\mathbf{W})} \leq \kappa_{\boldsymbol{\theta}}^{2m} \frac{T^m}{m!}, \tag{A.177}$$

for all $m = 1, 2, \ldots$ and $\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2 \in L_2(\mathbf{W})$.

Now one can fix some positive integer m such that

$$\kappa_{\boldsymbol{\theta}}^{2m} \frac{T^m}{m!} < 1. \tag{A.178}$$

Then $\mathbb{G}^m$ is a contraction on the Banach space $L_2(\mathbf{W})$ and the generalized Banach contraction principle (Theorem B.1.3) establishes that $\boldsymbol{\lambda}(t) = \mathbb{G}\boldsymbol{\lambda}(t)$, $(\mathbb{P} \otimes Leb)$-a.e. for some unique $\boldsymbol{\lambda} \in L_2(\mathbf{W}.)$ The result follows since $\eta(t_-) = \eta(t)$, $(\mathbb{P} \otimes Leb)$-a.e.

Proof of Lemma 4.3.19. From Remark 4.3.16, (4.71) and Corollary 4.3.13 we obtain

$$\bar{X} \equiv (\bar{X}_0, r\bar{X}_- + \bar{\boldsymbol{\pi}}\boldsymbol{\sigma}\boldsymbol{\theta}, \boldsymbol{\sigma}^\top \bar{\boldsymbol{\pi}}, H^{-1}\bar{\boldsymbol{\Gamma}}) \in \mathbb{B}. \tag{A.179}$$

From (4.4), each $Y \in \mathbb{B}_1$ can be written as

$$Y \equiv (Y_0, \Upsilon^Y, \boldsymbol{\xi}^Y, \boldsymbol{\Gamma}^Y) = (Y_0, -rY, \boldsymbol{\xi}^Y, \boldsymbol{\Gamma}^Y) \in \mathbb{B}. \tag{A.180}$$

From Proposition 4.3.2, (A.179) and (A.180), we find that for

$$\begin{aligned} \mathbb{M}(\bar{X}, Y)(t) \;&:=\; \bar{X}(T)Y(T) - \bar{X}_0 Y_0 - \int_0^t \bar{\boldsymbol{\pi}}^\top(\tau)\boldsymbol{\sigma}(\tau)[\boldsymbol{\theta}(\tau)Y(\tau) + \boldsymbol{\xi}^Y(\tau)]\,\mathrm{d}\tau \\ &\quad - \sum_{i,j=1}^{D} \int_0^t H(\tau)^{-1}\bar{\Gamma}_{ij}(\tau)\Gamma^Y_{ij}(\tau)\,\mathrm{d}[M_{ij}](\tau), \end{aligned} \tag{A.181}$$

we have $\mathbb{M}(\bar{X}, Y) \in \mathcal{M}_0(\{\mathcal{F}_t\}, \mathbb{P})$. Taking expectations at $t = T$ in (A.181) and using the fact that $E(\mathbb{M}(\bar{X}, Y)(T)) = 0$, we find

$$\begin{aligned} E(\bar{X}(T)Y(T)) \;&=\; \bar{X}_0 Y_0 + E\int_0^t \bar{\boldsymbol{\pi}}^\top(\tau)\boldsymbol{\sigma}(\tau)[\boldsymbol{\theta}(\tau)Y(\tau) + \boldsymbol{\xi}^Y(\tau)]\,\mathrm{d}\tau \\ &\quad + E\sum_{i,j=1}^{D} \int_0^t H(\tau)^{-1}\bar{\Gamma}_{ij}(\tau)\Gamma^Y_{ij}(\tau)\,\mathrm{d}[M_{ij}](\tau) \end{aligned} \tag{A.182}$$

From (4.71) we have $\bar{X}(T) = \partial J^*(\bar{Y}(T))$. Hence, from (4.46) of Proposition 4.3.7 and (A.182), we get

$$(\bar{X}_0 - x_0)Y_0 + E\int_0^t \bar{\boldsymbol{\pi}}^\top(\tau)\boldsymbol{\sigma}(\tau)[\boldsymbol{\theta}(\tau)Y(\tau) + \boldsymbol{\xi}^Y(\tau)]\,\mathrm{d}\tau$$
$$+E\sum_{i,j=1}^D \int_0^t H(\tau)^{-1}\bar{\Gamma}_{ij}(\tau)\Gamma^Y_{ij}(\tau)\,\mathrm{d}[M_{ij}](\tau)$$
$$+E\int_0^T \delta(\boldsymbol{\sigma}(\tau)[\boldsymbol{\theta}(\tau)Y(\tau) + \boldsymbol{\xi}^Y(\tau)]|K)\mathrm{d}\tau \geq 0. \tag{A.183}$$

Proof of Proposition 4.3.20. **Claim (a)**: Fix an arbitrary $y \in \mathbb{R}$. From the uniform boundedness of $\boldsymbol{\theta} \in L_2(\mathbf{W})$, we have that $-y\boldsymbol{\theta} \in L_2(\mathbf{W})$. Applying Lemma 4.3.15 to $(\boldsymbol{\rho}, \boldsymbol{\gamma}) := (-Y_0\boldsymbol{\theta}, \mathbf{0}) \in L_2(\mathbf{W}) \times L_2(\mathbf{M})$, there exists $\boldsymbol{\lambda} \in L_2(\mathbf{W})$ such that

$$\boldsymbol{\lambda}(t) + \boldsymbol{\theta}(t)\int_0^t \boldsymbol{\lambda}^\top(\tau)\,\mathrm{d}\mathbf{W}(t) = -Y_0\boldsymbol{\theta}(t) \qquad \text{a.s.} \tag{A.184}$$

For all $t \in [0, T]$, set

$$\bar{\boldsymbol{\lambda}}(t) := \beta(t)\boldsymbol{\lambda}(t), \tag{A.185}$$

where $\beta(t)$ is given by (4.5). As β is uniformly bounded and $\boldsymbol{\lambda} \in L_2(\mathbf{W})$, we have $\bar{\boldsymbol{\lambda}} \in L_2(\mathbf{W})$. Substituting $\boldsymbol{\lambda} = \beta^{-1}\bar{\boldsymbol{\lambda}}$ into (A.184) and multiplying across by $\beta(t)$, we get

$$\bar{\boldsymbol{\lambda}}(t) + \beta(t)\boldsymbol{\theta}(t)\int_0^t \beta^{-1}(\tau)\bar{\boldsymbol{\lambda}}^\top(\tau)\,\mathrm{d}\mathbf{W}(\tau) = -Y_0\beta(t)\boldsymbol{\theta}(t) \qquad \text{a.s.} \tag{A.186}$$

Rearranging, this becomes

$$\bar{\boldsymbol{\lambda}}(t) = -\boldsymbol{\theta}(t)\left(y\beta(t) + \beta(t)\int_0^t \beta^{-1}(\tau)\bar{\boldsymbol{\lambda}}^\top(\tau)\,\mathrm{d}\mathbf{W}(\tau)\right) \qquad \text{a.s.} \tag{A.187}$$

Recalling Equation (4.6), set

$$\begin{aligned} Y(t) &:= \Xi(Y_0, \bar{\boldsymbol{\lambda}}, \mathbf{0})(t) \\ &= \beta(t)\left\{Y_0 + \int_0^t \beta^{-1}(\tau)\boldsymbol{\lambda}^\top(\tau)\,\mathrm{d}\mathbf{W}(\tau)\right\}. \end{aligned} \tag{A.188}$$

From Lemma 4.1.4, we have

$$Y \equiv (Y_0, -rY, \bar{\boldsymbol{\lambda}}, \mathbf{0}) \in \mathbb{B}_1. \tag{A.189}$$

Using (A.188) to replace the term in brackets in (A.187) gives

$$\bar{\boldsymbol{\lambda}}(t) = -\boldsymbol{\theta}(t)Y(t). \tag{A.190}$$

Substituting $(Y_0, \boldsymbol{\xi}^Y, \boldsymbol{\Gamma}^Y) := (Y_0, \bar{\boldsymbol{\lambda}}, \mathbf{0})$ and $\bar{\boldsymbol{\lambda}}(t) = -\boldsymbol{\theta}(t)Y(t)$ into equation (4.82), we obtain

$$(\bar{X}_0 - x_0)Y_0 \geq 0. \tag{A.191}$$

Since (A.191) holds for all $Y_0 \in \mathbb{R}$ we must have that $\bar{X}_0 = x_0$.

Claim (b) Using the fact that $\bar{X}_0 = x_0$, which has just been established in Claim (a), we can simplify (4.82) to get

$$\begin{aligned} &E\int_0^t \bar{\boldsymbol{\pi}}^\top(\tau)\boldsymbol{\sigma}(\tau)[\boldsymbol{\theta}(\tau)Y(\tau) + \boldsymbol{\xi}^Y(\tau)]\,\mathrm{d}\tau \\ &\quad + E\sum_{i,j=1}^{D}\int_0^t H(\tau)^{-1}\bar{\Gamma}_{ij}(\tau)\Gamma_{ij}^Y(\tau)\,\mathrm{d}[M_{ij}](\tau) \\ &\quad + E\int_0^T \delta(\boldsymbol{\sigma}(\tau)[\boldsymbol{\theta}(\tau)Y(\tau) + \boldsymbol{\xi}^Y(\tau)]|K)\mathrm{d}\tau \geq 0. \end{aligned} \tag{A.192}$$

which holds for all $(Y_0, \boldsymbol{\xi}^Y, \boldsymbol{\Gamma}^Y) \in \mathbb{R} \times L_2(\mathbf{W}) \times L_2(\mathbf{M})$, for $Y = \Xi(Y_0, \boldsymbol{\xi}^Y, \boldsymbol{\Gamma}^Y)$.

Define

$$B := \{(\omega, t) \in \Omega \times [0, T] : \bar{\boldsymbol{\pi}}(\omega, t) \in K\,(\mathbb{P} \otimes Leb) - \text{a.e.}\}. \tag{A.193}$$

From Lemma B.1.1 with $\mathbf{p} := \bar{\boldsymbol{\pi}}$, there exists a predictable mapping $\boldsymbol{\nu}^{\bar{\boldsymbol{\pi}}} : \Omega \times [0, T] \to \mathbb{R}^N$ such that $\| \boldsymbol{\nu}^{\bar{\boldsymbol{\pi}}}(t) \| \leq 1$ $(\mathbb{P} \otimes Leb)$-a.e., $|\delta(\boldsymbol{\nu}^{\bar{\boldsymbol{\pi}}}(t)|K)| \leq 1$ $(\mathbb{P} \otimes Leb)$-a.e. and

$$\delta(\boldsymbol{\nu}^{\bar{\boldsymbol{\pi}}}(t)|K) + \bar{\boldsymbol{\pi}}^\top(t)\boldsymbol{\nu}^{\bar{\boldsymbol{\pi}}}(t) = 0 \quad (\mathbb{P} \otimes Leb) - \text{a.e. on } B, \tag{A.194}$$

$$\delta(\boldsymbol{\nu}^{\bar{\boldsymbol{\pi}}}(t)|K) + \bar{\boldsymbol{\pi}}^\top(t)\boldsymbol{\nu}^{\bar{\boldsymbol{\pi}}}(t) < 0 \quad (\mathbb{P} \otimes Leb) - \text{a.e. on } \Omega \times [0, T] \setminus B. \tag{A.195}$$

In order to obtain $\bar{\boldsymbol{\pi}} \in K\,(\mathbb{P} \otimes Leb) -$ a.e., it is sufficient to show $(\mathbb{P} \otimes Leb)(B^c) = 0$.

Suppose $(\mathbb{P} \otimes Leb)(B^c) > 0$. It then follows from (A.195) that

$$E\int_0^T \{\bar{\boldsymbol{\pi}}^\top(\tau)\boldsymbol{\nu}^{\bar{\boldsymbol{\pi}}}(\tau) + \delta(\boldsymbol{\nu}^{\bar{\boldsymbol{\pi}}}(t)|K)\}\mathrm{d}\tau < 0. \tag{A.196}$$

We are now going to use Lemma 4.3.15 to construct some $Y \in \mathbb{B}_1$ such that

$$Y_0 = 0, \tag{A.197}$$

$$\boldsymbol{\nu}^{\bar{\boldsymbol{\pi}}}(t) = \boldsymbol{\sigma}(t)[Y(t)\boldsymbol{\theta}(t) + \boldsymbol{\xi}^Y(t)], \tag{A.198}$$

$$\boldsymbol{\Gamma}^Y = \mathbf{0}, \tag{A.199}$$

since it then follows from (A.192) that

$$E\int_0^T \{\bar{\boldsymbol{\pi}}^\top(\tau)\boldsymbol{\nu}^{\bar{\boldsymbol{\pi}}}(\tau)+\delta(\boldsymbol{\nu}^{\bar{\boldsymbol{\pi}}}(t)|K)\}\mathrm{d}\tau \geq 0, \tag{A.200}$$

which contradicts (A.196). We therefore get $(\mathbb{P}\otimes Leb)(B^c)=0$, and therefore $\bar{\boldsymbol{\pi}}\in K\,(\mathbb{P}\otimes Leb)-$ a.e., as required to establish the first assertion at (4.84). The second assertion, that $\bar{\boldsymbol{\pi}}\in\mathcal{A}$, is then immediate from this together with (4.79) and Definition 2.2.12.

It therefore remains to construct some $Y\in\mathbb{B}_1$ such that (A.197) - (A.199) holds. For this we shall again use Lemma 4.3.15 on integral equations. Put $\boldsymbol{\rho}=\beta^{-1}(t)\boldsymbol{\sigma}^{-1}(t)\boldsymbol{\nu}^{\bar{\boldsymbol{\pi}}}(t)$. By the boundedness of $\beta,\boldsymbol{\sigma},$ and $\boldsymbol{\nu}^{\bar{\boldsymbol{\pi}}}$, we have $\boldsymbol{\rho}\in L_2(\mathbf{W})$.

Applying Lemma 4.3.15 to $(\boldsymbol{\rho},\boldsymbol{\gamma}):=(\beta^{-1}(t)\boldsymbol{\sigma}^{-1}\boldsymbol{\nu}^{\bar{\boldsymbol{\pi}}},\mathbf{0})\in L_2(\mathbf{W})\times L_2(\mathbf{M})$, there exists $\boldsymbol{\lambda}\in L_2(\mathbf{W})$ such that for all $t\in[0,T]$,

$$\boldsymbol{\lambda}(t)+\boldsymbol{\theta}(t)\int_0^t \boldsymbol{\lambda}^\top(\tau)\mathrm{d}\mathbf{W}(\tau)=\beta^{-1}(t)\boldsymbol{\sigma}^{-1}(t)\boldsymbol{\nu}^{\bar{\boldsymbol{\pi}}}(t) \qquad \text{a.s.} \tag{A.201}$$

For all $t\in[0,T]$, define

$$\bar{\boldsymbol{\lambda}}(t):=\beta(t)\boldsymbol{\lambda}(t), \tag{A.202}$$

Then as β is uniformly bounded and $\boldsymbol{\lambda}\in L_2(\mathbf{W})$, we have $\bar{\boldsymbol{\lambda}}\in L_2(\mathbf{W})$. Substituting $\boldsymbol{\lambda}=\beta^{-1}(t)\bar{\boldsymbol{\lambda}}(t)$ into (A.201), multiplying across by $\boldsymbol{\sigma}(t)\beta(t)$ and rearranging we get

$$\begin{aligned}\boldsymbol{\nu}^{\bar{\boldsymbol{\pi}}}(t) &= \boldsymbol{\sigma}(t)\left(\bar{\boldsymbol{\lambda}}(t)+\beta(t)\boldsymbol{\theta}(t)\int_0^t \beta^{-1}(\tau)\bar{\boldsymbol{\lambda}}^\top(\tau)\mathrm{d}\mathbf{W}(\tau)\right)\\ &= \boldsymbol{\sigma}(t)\left(\bar{\boldsymbol{\lambda}}(t)+\boldsymbol{\theta}(t)\Xi(0,\bar{\boldsymbol{\lambda}},\mathbf{0})(t)\right). \end{aligned} \tag{A.203}$$

Now put

$$Y:=\Xi(0,\bar{\boldsymbol{\lambda}},\mathbf{0}). \tag{A.204}$$

Applying Lemma 4.1.4 to $(0,\bar{\boldsymbol{\lambda}},\mathbf{0})\in\mathbf{S}$, we have

$$Y\equiv(0,-rY_-,\bar{\boldsymbol{\lambda}},\mathbf{0})\in\mathbb{B}_1, \tag{A.205}$$

with $Y_0=0$ and $\boldsymbol{\xi}^Y(t)=\bar{\boldsymbol{\lambda}}(t)$. Thus

$$\boldsymbol{\nu}^{\bar{\boldsymbol{\pi}}}(t) = \boldsymbol{\sigma}(t)[Y(t)\boldsymbol{\theta}(t)+\boldsymbol{\xi}^Y(t)] \quad \text{a.e.} \tag{A.206}$$

This proves Claim (b).

Claim (c) Since we have established $(\mathbb{P} \otimes Leb)(B^c) = 0$ for Claim (b), it follows from (A.194) that

$$\delta(\boldsymbol{\nu}^{\bar{\boldsymbol{\pi}}}(t)|K) + \bar{\boldsymbol{\pi}}^\top(t)\boldsymbol{\nu}^{\bar{\boldsymbol{\pi}}}(t) = 0 \quad (\mathbb{P} \otimes Leb) - \text{a.e.} \tag{A.207}$$

Applying Lemma 4.3.15 to $(\boldsymbol{\rho}, \boldsymbol{\gamma}) := (\beta^{-1}\boldsymbol{\sigma}^{-1}\boldsymbol{\nu}^{\bar{\boldsymbol{\pi}}}, \beta^{-1}H^{-1}\bar{\boldsymbol{\Gamma}}) \in L_2(\mathbf{W}) \times L_2(\mathbf{M})$, where β is defined by (4.5), there exists $\boldsymbol{\eta} \in L_2(\mathbf{W})$ such that

$$\begin{aligned}\boldsymbol{\eta}(t) &+ \boldsymbol{\theta}(t)\int_0^t \boldsymbol{\eta}^\top(\tau)\,\mathrm{d}\mathbf{W}(\tau) \\ &= \beta^{-1}(t)\boldsymbol{\sigma}^{-1}(t)\boldsymbol{\nu}^{\bar{\boldsymbol{\pi}}}(t) - \boldsymbol{\theta}(t)\sum_{i,j=1}^{D}\int_0^t \beta^{-1}(\tau)H^{-1}(\tau)\bar{\Gamma}_{ij}(\tau)\,\mathrm{d}M_{ij}(\tau). \end{aligned} \tag{A.208}$$

For all $t \in [0, T]$, define

$$\bar{\boldsymbol{\eta}}(t) := \beta(t)\boldsymbol{\eta}(t). \tag{A.209}$$

Since β is uniformly bounded and $\boldsymbol{\eta} \in L_2(\mathbf{W})$, we have $\bar{\boldsymbol{\eta}} \in L_2(\mathbf{W})$. Substituting $\boldsymbol{\eta} = \beta^{-1}\bar{\boldsymbol{\eta}}$ into (A.208), multiplying across by $\boldsymbol{\sigma}(t)\beta(t)$ and rearranging, we get

$$\begin{aligned}\boldsymbol{\nu}^{\bar{\boldsymbol{\pi}}}(t) = \boldsymbol{\sigma}(t)\Bigg(&\bar{\boldsymbol{\eta}}(t) + \beta(t)\boldsymbol{\theta}(t)\int_0^t \beta^{-1}(\tau)\bar{\boldsymbol{\eta}}^\top(\tau)\,\mathrm{d}\mathbf{W}(\tau) \\ &+\beta(t)\boldsymbol{\theta}(t)\sum_{i,j=1}^{D}\int_0^t \beta^{-1}(\tau)H(\tau)^{-1}\bar{\Gamma}_{ij}(\tau)\,\mathrm{d}M_{ij}(\tau)\Bigg) \quad \text{a.s.}\end{aligned} \tag{A.210}$$

Recalling Equation (4.6), set

$$\begin{aligned} Y(t) &:= \Xi(0, \bar{\boldsymbol{\eta}}, H^{-1}\bar{\boldsymbol{\Gamma}})(t) \\ &= \beta(t)\left\{\int_0^t \beta^{-1}(\tau)\bar{\boldsymbol{\eta}}^\top(\tau)\,\mathrm{d}\mathbf{W}(\tau) + \sum_{i,j=1}^{D}\int_0^t \beta^{-1}(\tau)H(\tau)^{-1}\bar{\Gamma}_{ij}(\tau)\,\mathrm{d}M_{ij}(\tau)\right\}. \end{aligned} \tag{A.211}$$

From Lemma 4.1.4, we have

$$Y \equiv (0, -rY_-, \bar{\boldsymbol{\eta}}, H^{-1}\bar{\boldsymbol{\Gamma}}) \in \mathbb{B}_1. \tag{A.212}$$

Substituting (A.211) into (A.210), we get

$$\boldsymbol{\nu}^{\bar{\boldsymbol{\pi}}}(t) = \boldsymbol{\sigma}(t)(\bar{\boldsymbol{\eta}}(t) + \boldsymbol{\theta}(t)Y(t)) \qquad \text{a.s.} \tag{A.213}$$

Then substituting $\boldsymbol{\nu}^{\bar{\boldsymbol{\pi}}}(t)$ from (A.213) into (A.194), we obtain $(\mathbb{P} \otimes Leb)$-a.e.,

$$\delta(\boldsymbol{\sigma}(t)(\bar{\boldsymbol{\eta}}(t) + \boldsymbol{\theta}(t)Y(t))|K) + \bar{\boldsymbol{\pi}}^{\top}(t)\boldsymbol{\sigma}(t)(\bar{\boldsymbol{\eta}}(t) + \boldsymbol{\theta}(t)Y(t)) = 0. \tag{A.214}$$

Substituting $(Y_0, \boldsymbol{\xi}^Y, \boldsymbol{\Gamma}^Y) := (0, \bar{\boldsymbol{\eta}}, -H^{-1}\bar{\boldsymbol{\Gamma}})$ in equation (A.192) and using (A.214), we find

$$E \sum_{i,j=1}^{D} \int_0^T H(t)^{-2} |\bar{\boldsymbol{\Gamma}}_{ij}(t)|^2 \, \mathrm{d}[M_{ij}](t) \leq 0. \tag{A.215}$$

However, the left-hand side of (A.215) is the sum of nonnegative terms, so we must have equality in (A.215), that is

$$E \sum_{i,j=1}^{D} \int_0^T H(t)^{-2} |\bar{\boldsymbol{\Gamma}}_{ij}(t)|^2 \, \mathrm{d}[M_{ij}](t) = 0. \tag{A.216}$$

Hence,

$$H^{-1}\bar{\boldsymbol{\Gamma}} = 0 \qquad \nu_{[\mathbf{M}]} - \text{a.e.}, \tag{A.217}$$

which in turn implies (since H^{-1} is strictly positive),

$$\bar{\boldsymbol{\Gamma}} = 0 \qquad \nu_{[\mathbf{M}]} - \text{a.e.} \tag{A.218}$$

Appendix B

Standard definitions and Results

In this Appendix we summarize for convenience of reference a miscellany of technical results, terminology and definitions that are needed in the main part of the thesis.

B.1 Miscellaneous Results

B.1.1 Measurable Selection for Support Functions

The following lemma on measurable selections for support functions is a minor variant of a result of Karatzas and Shreve [22], Lemma 5.4.2. This lemma is an essential tool for establishing Proposition 4.3.20(b).

Lemma B.1.1 Let K be a nonempty, closed, convex set of $\mathbb{R}^N$ and let δ be the support function of the convex set K defined by

$$\delta(\mathbf{z}|K) := \sup_{\boldsymbol{\pi}\in K}\{-\boldsymbol{\pi}^\top \mathbf{z}\} \qquad \forall \mathbf{z} \in \mathbb{R}^N. \tag{B.1}$$

Fix some predictable process $\mathbf{p} : \Omega \times [0,T] \to \mathbb{R}^N$ and put

$$B := \{(\omega, t) \in \Omega \times [0,T] : \mathbf{p}(\omega,t) \in K\}. \tag{B.2}$$

Then there exists an $\mathbb{R}^N$-valued, predictable process $\boldsymbol{\nu}(\cdot)$ such that a.s.

$$\parallel \boldsymbol{\nu}(t) \parallel \leq 1, \qquad |\delta(\boldsymbol{\nu}(t)|K)| \leq 1, \qquad \forall t \in [0,T], \tag{B.3}$$

and

$$\delta(\boldsymbol{\nu}(t)|K) + \mathbf{p}^\top(t)\boldsymbol{\nu}(t) = 0 \qquad (\mathbb{P}\otimes Leb) - \text{a.e. on } B, \tag{B.4}$$
$$\delta(\boldsymbol{\nu}(t)|K) + \mathbf{p}^\top(t)\boldsymbol{\nu}(t) < 0 \qquad (\mathbb{P}\otimes Leb) - \text{a.e. on } \Omega\times[0,T]\setminus B. \tag{B.5}$$

Proof. The proof follows that of Karatzas and Shreve [22], Lemma 5.4.2. The main difference is that Karatzas and Shreve [22], Lemma 5.4.2 is for an $\{\mathcal{F}_t^{\mathbf{W}}\}$-progressively measurable process $\mathbf{p}$, in place of our $\{\mathcal{F}_t\}$-predictable $\mathbf{p}$. However, upon examining their proof, the measurability of $\mathbf{p}$ is used only to determine the measurability of $\boldsymbol{\nu}$. Hence, $\boldsymbol{\nu}$ inherits the measurability of $\mathbf{p}$, so we can safely state that $\boldsymbol{\nu}(\cdot)$ is predictable.

B.1.2 Generalized Contraction Mapping Principle

Definition B.1.2 Let (Y,d) be a complete metric space and F be a mapping from Y into itself. Then y is called a *fixed point* of F if $F(y)=y$, i.e. if F carries y into itself. Suppose there exists a number $\alpha<1$ such that

$$d(F(x),F(y)) \le \alpha d(x,y)$$

for every pair of points $x,y\in Y$. Then F is said to be a contraction mapping.

We next state the *generalized Banach contraction principle*, the proof of which can be found in Kolmogorov and Fomin [23], Theorem 1', page 70. This result is used to establish Lemma 4.3.15.

Theorem B.1.3 (Banach contraction principle). Let (Y,d) be a complete metric space and suppose that the mapping $F:Y\to Y$ is such that $F^k:Y\to Y$ is a *contraction mapping* for some fixed integer $k\ge 1$ (here F^k denotes the k-fold composition of F). Then F has a unique fixed point $u\in Y$, and $F^n(y)\to u$ for each $y\in Y$.

B.1.3 Conditional expectation results

For completeness we first recall Hölder's inequality (see Theorem 3.2.1 of Friedman [13]):

Theorem B.1.4 (Hölder's inequality)
Let $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}})$ be a measure space. Let p and q be extended real numbers, $1 \leq p \leq \infty$, $1 \leq q \leq \infty$, $\frac{1}{p} + \frac{1}{q} = 1$. If $f \in L_p(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}})$, $g \in L_q(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}})$, then $fg \in L_1(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}})$ and

$$\int_{\tilde{\Omega}} |fg| \, \mathrm{d}\mu \leq \left(\int_{\tilde{\Omega}} |f|^p \, \mathrm{d}\mu \right)^{\frac{1}{p}} \left(\int_{\tilde{\Omega}} |g|^q \, \mathrm{d}\mu \right)^{\frac{1}{q}}. \tag{B.6}$$

We now state a conditional version of Hölder's inequality for integration over a probability space (see Chow and Teicher [4], Theorem 7.2.4):

Theorem B.1.5 (Hölder's inequality for conditional expectations)
Let $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}})$ be a probability space and let $\mathcal{G}$ be any sub-σ-field of $\tilde{\mathcal{F}}$. Let X, Y be random variables on $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}})$ and let $p > 1$ and q be real numbers such that $\frac{1}{p} + \frac{1}{q} = 1$. Then

$$E(|XY| \,|\, \mathcal{G}) \leq (E(|X|^p \,|\, \mathcal{G}))^{\frac{1}{p}} \, (E(|Y|^q \,|\, \mathcal{G}))^{\frac{1}{q}}. \tag{B.7}$$

From Elliot [10], Lemma 1.9, we have *Jensen's inequality* for conditional expectations:

Lemma B.1.6 Jensen's inequality for conditional expectations:
Let $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}})$ be a probability space and let $\mathcal{G}$ be any sub-σ-field of $\tilde{\mathcal{F}}$. Suppose $f : \mathbb{R} \to \mathbb{R}$ is a convex map and suppose X is an integrable random variable defined on $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}})$ such that $f(X)$ is integrable. Then

$$f\left(E(X|\mathcal{G})\right) \leq E\left(f(X)|\mathcal{G}\right). \tag{B.8}$$

B.2 Various Classes of Stochastic Processes

B.2.1 Càdlàg stochastic processes

Definition B.2.1 A process $X = \{X(t) : \mathrm{t} \in [0, \mathrm{T}]\}$ is càdlàg 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.

Remark B.2.2 If X is càdlàg then we define

$$\begin{aligned} & X(0_-) := X(0) \\ \text{and} \qquad & X(t_-) := \lim_{\substack{s \to t \\ s < t}} X(s), \quad \forall\, t \in [0,T]. \end{aligned} \tag{B.9}$$

and we define the *jump process* $\{\triangle X(t) : t \in [0,T]\}$ as

$$\triangle X(t) := X(t) - X(t_-), \qquad \forall\, t \in [0,T]. \tag{B.10}$$

Definition B.2.3 A *filtered probability space* is a pair $((\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}}), \{\tilde{\mathcal{F}}_t\})$ consisting of a probability space $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}})$ and a filtration $\{\tilde{\mathcal{F}}_t\}_{t\in[0,T]}$ on $\tilde{\mathcal{F}}$. A *standard filtered probability space* is a filtered probability space $((\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}}), \{\tilde{\mathcal{F}}_t\})$ with the following additional properties:

- $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}})$ is a complete probability space;
- $\tilde{\mathcal{F}}_0$ includes all $\tilde{P}$-null events in $\tilde{\mathcal{F}}$;
- the filtration $\{\tilde{\mathcal{F}}_t\}$ is right-continuous, that is

$$\tilde{\mathcal{F}}_t = \bigcap_{u>t} \tilde{\mathcal{F}}_u \qquad \text{for all } t \in [0,T).$$

Notation B.2.4 We write E to denote expectation with respect to the measure $\mathbb{P}$. If there is any ambiguity about the measure $\mathbb{P}$, we will write $E_{\mathbb{P}}$. If the expectation is with respect to another measure $\tilde{\mathbb{P}}$, we will write $E_{\tilde{\mathbb{P}}}$.

Definition B.2.5 A process $X = \{X(t) : t \in [0,T]\}$ defined on a filtered probability space $((\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}}), \{\tilde{\mathcal{F}}_t\})$ is *non-decreasing* if the mappings $t \to X(t,\omega)$ are non-decreasing on $[0,T]$ for all $\omega \in \tilde{\Omega}$.

B.2.2 Spaces of martingales

In this section we formulate spaces of martingales and local martingales restricted to the finite closed interval $[0,T]$ rather than the semi-infinite line $[0,\infty)$ since all processes in this thesis are limited to the interval $[0,T]$.

Definition B.2.6 A real-valued, $\{\tilde{\mathcal{F}}_t\}$-adapted processes $M = \{M(t) : t \in [0,T]\}$ on $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}})$ is called a martingale if

- $E_{\tilde{\mathbb{P}}}|M(t)| < \infty, \ \forall\, t \in [0,T]$; and
- $E_{\tilde{\mathbb{P}}}\left(M(t)|\tilde{\mathcal{F}}_s\right) = M(s)$ $\tilde{\mathbb{P}}$-a.s., for all $0 \leq s \leq t \leq T$

We shall use $\mathcal{M}((\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}}), \{\tilde{\mathcal{F}}_t\})$ to denote the set of $\{\tilde{\mathcal{F}}_t\}$-adapted martingales on $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}})$.

Remark B.2.7 Usually the measure space on which the space of martingales $\mathcal{M}((\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}}), \{\tilde{\mathcal{F}}_t\})$ is defined is clear. In this case we specify only the filtration and the probability measure and use the notation $\mathcal{M}(\{\tilde{\mathcal{F}}_t\}, \tilde{\mathbb{P}})$ instead of $\mathcal{M}((\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}}), \{\tilde{\mathcal{F}}_t\})$.

Definition B.2.8 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.*

A very useful property of L_p-bounded martingales is:

Theorem B.2.9 Doob's L_p-inequality :
Let $p \in (1, \infty)$. Let M be a càdlàg martingale relative to the filtered probability space $((\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}}), \{\tilde{\mathcal{F}}_t\})$ which is L_p-bounded. Then

$$E_{\tilde{\mathbb{P}}}\left(\sup_{t\in[0,T]} |M(t)|^p\right) \leq \left(\frac{p}{p-1}\right)^p E_{\tilde{\mathbb{P}}}|M(T)|^p. \tag{B.11}$$

We next introduce useful notation for various classes of martingales:

Notation B.2.10 (a) $\mathcal{M}_0(\{\tilde{\mathcal{F}}_t\}, \tilde{\mathbb{P}})$ denotes the set of $M \in \mathcal{M}(\{\tilde{\mathcal{F}}_t\}, \tilde{\mathbb{P}})$ which are $\tilde{\mathbb{P}}$-a.s. null at the origin i.e. $M(0) = 0$ $\tilde{\mathbb{P}}$-a.s.

(b) $\mathcal{M}^c(\{\tilde{\mathcal{F}}_t\}, \tilde{\mathbb{P}})$ denotes the set of $M \in \mathcal{M}(\{\tilde{\mathcal{F}}_t\}, \tilde{\mathbb{P}})$ which are sample-path continuous.

(c) $\mathcal{M}_0^c(\{\tilde{\mathcal{F}}_t\}, \tilde{\mathbb{P}}) := \mathcal{M}_0(\{\tilde{\mathcal{F}}_t\}, \tilde{\mathbb{P}}) \cap \mathcal{M}^c(\{\tilde{\mathcal{F}}_t\}, \tilde{\mathbb{P}})$.

(d) $\mathcal{M}_2(\{\tilde{\mathcal{F}}_t\}, \tilde{\mathbb{P}})$ denotes the set of $M \in \mathcal{M}(\{\tilde{\mathcal{F}}_t\}, \tilde{\mathbb{P}})$ which are square-integrable.

(e) $\mathcal{M}_{0,2}(\{\tilde{\mathcal{F}}_t\}, \tilde{\mathbb{P}}) := \mathcal{M}_0(\{\tilde{\mathcal{F}}_t\}, \tilde{\mathbb{P}}) \cap \mathcal{M}_2(\{\tilde{\mathcal{F}}_t\}, \tilde{\mathbb{P}})$.

B.2.3 Spaces of local martingales

Definition B.2.11 For a sequence $\{T^m\}_{m\in\mathbb{N}}$ of $\{\tilde{\mathcal{F}}_t\}$-stopping times, we write

$$T^m \Uparrow T \tag{B.12}$$

to mean that

- $0 \le T^m(\omega) \le T^{m+1}(\omega) \le T$ for all $\omega \in \tilde{\Omega}$ and for all $m \in \mathbb{N}$;
- there exists $M(\omega) \in \mathbb{N}$ such that $T^m(\omega) = T$, for all $m \ge M(\omega)$ and for all $\omega \in \tilde{\Omega}$.

Remark B.2.12 The preceding notion of increasing stopping times ensures that the right end-point T of the interval $[0, T]$ is included in the localization, and rules out the possibility that the T^m are all *strictly* less than T (i.e. $T^m < T$) for all $m = 1, 2, \ldots$ Increasing sequences of stopping times in the sense of Definition B.2.11 occur quite naturally in several proofs.

Definition B.2.13 A real-valued process $M = \{M(t) : t \in [0, T]\}$ on $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}})$ is called an $\{\tilde{\mathcal{F}}_t\}$-local martingale if there is a sequence $\{T^m\}_{m\in\mathbb{N}}$ of $\{\tilde{\mathcal{F}}_t\}$-stopping times such that

1. $T^m \Uparrow T \qquad \tilde{\mathbb{P}}$-a.s.; and
2. $\{M(t \wedge T^m) : t \in [0, T]\} \in \mathcal{M}((\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}}), \{\tilde{\mathcal{F}}_t\})$ for each $m \in \mathbb{N}$.

We shall use $\mathcal{M}^{loc}((\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}}), \{\tilde{\mathcal{F}}_t\})$ to denote the set of $\{\tilde{\mathcal{F}}_t\}$-local martingales on $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}})$.

Definition B.2.14 A real-valued process $M = \{M(t) : t \in [0, T]\}$ on $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}})$ is called a $\{\tilde{\mathcal{F}}_t\}$-locally square integrable martingale if there is a sequence $\{T^m\}_{m\in\mathbb{N}}$ of $\{\tilde{\mathcal{F}}_t\}$-stopping times such that

1. $T^m \Uparrow T \qquad \tilde{\mathbb{P}}$-a.s.; and
2. $\{M(t \wedge T^m) : t \in [0, T]\} \in \mathcal{M}_2((\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}}), \{\tilde{\mathcal{F}}_t\})$ for each $m \in \mathbb{N}$.

We shall use $\mathcal{M}_2^{loc}((\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}}), \{\tilde{\mathcal{F}}_t\})$ to denote the set of $\{\tilde{\mathcal{F}}_t\}$-locally square integrable martingales on the probability space $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}})$.

Remark B.2.15 With reference to Definitions B.2.13 and B.2.14 we say that the sequence $\{T^m\}_{m\in\mathbb{N}}$ of $\{\tilde{\mathcal{F}}_t\}$-stopping times is a localizing sequence for M.

Remark B.2.16 Usually the measure space on which the space of local Martingales $\mathcal{M}^{loc}((\tilde{\Omega},\tilde{\mathcal{F}},\tilde{\mathbb{P}}),\{\tilde{\mathcal{F}}_t\})$ is defined is clear. In this case we specify only the filtration and the probability measure and use the notation $\mathcal{M}^{loc}(\{\tilde{\mathcal{F}}_t\},\tilde{\mathbb{P}})$ instead of $\mathcal{M}^{loc}((\tilde{\Omega},\tilde{\mathcal{F}},\tilde{\mathbb{P}}),\{\tilde{\mathcal{F}}_t\})$.

Notation B.2.17 (a) $\mathcal{M}_0^{loc}(\{\tilde{\mathcal{F}}_t\},\tilde{\mathbb{P}})$ denotes the set of $M\in\mathcal{M}^{loc}(\{\tilde{\mathcal{F}}_t\},\tilde{\mathbb{P}})$ which are $\tilde{\mathbb{P}}$-a.s. null at the origin.

(b) $\mathcal{M}^{c,loc}(\{\tilde{\mathcal{F}}_t\},\tilde{\mathbb{P}})$ denotes the set of $M\in\mathcal{M}^{loc}(\{\tilde{\mathcal{F}}_t\},\tilde{\mathbb{P}})$ which are sample-path continuous.

(c) $\mathcal{M}_0^{c,loc}(\{\tilde{\mathcal{F}}_t\},\tilde{\mathbb{P}}) := \mathcal{M}_0^{loc}(\{\tilde{\mathcal{F}}_t\},\tilde{\mathbb{P}})\cap\mathcal{M}^{c,loc}(\{\tilde{\mathcal{F}}_t\},\tilde{\mathbb{P}})$.

(d) $\mathcal{M}_{0,2}^{loc}(\{\tilde{\mathcal{F}}_t\},\tilde{\mathbb{P}}) := \mathcal{M}_0^{loc}(\{\tilde{\mathcal{F}}_t\},\tilde{\mathbb{P}})\cap\mathcal{M}_2^{loc}(\{\tilde{\mathcal{F}}_t\},\tilde{\mathbb{P}})$.

From Jacod and Shiryaev [18], Definition I.4.11 we have the following definitions.

Definition B.2.18 Two local martingales N and M are called *orthogonal* if their product $L = MN$ is a local martingale.

Definition B.2.19 A local martingale M is called a *purely discontinuous local martingale* if $M(0) = 0$ and if it is orthogonal to all *continuous* local martingales. That is, $M\in\mathcal{M}^{loc}(\{\tilde{\mathcal{F}}_t\},\tilde{\mathbb{P}})$ is a purely discontinuous local martingale when $M(0) = 0$ and $MN\in\mathcal{M}^{loc}(\{\tilde{\mathcal{F}}_t\},\tilde{\mathbb{P}})$ for each $N\in\mathcal{M}^{c,loc}(\{\tilde{\mathcal{F}}_t\},\tilde{\mathbb{P}})$.

B.2.4 Finite variation processes

Definition B.2.20 A process $A = \{A(t) : t\in[0,T]\}$ is a process of *finite variation* if it is an $\{\tilde{\mathcal{F}}_t\}$-adapted, càdlàg process such that each path $t\to A(\omega,t)$ is of finite variation on $[0,T]$, in other words for all $\omega\in\Omega$ the variation $V_A(\omega,T)$ of $t\to A(\omega,t)$ over $(0,T]$ is finite, in which the *variation process* is defined by

$$V_A(\omega,t) := \int_{(0,t]} |\mathrm{d}A(\omega,s)| = \sup\sum_{i=1}^{n}|A(\omega,s_i)-A(\omega,s_{i-1})| < \infty. \tag{B.13}$$

The supremum is taken over all *finite* partitions $0 = s_0 < s_1 < \ldots < s_n = t$ of $[0,t]$.

Definition B.2.21 A process $A = \{A(t) : t \in [0,T]\}$ is a process of *integrable variation* if it is a process of finite variation such that

$$E(V_A(\omega, T)) < \infty, \tag{B.14}$$

for V_A given by (B.13).

Notation B.2.22 The definition of processes of finite variation is very general, and in particular does not postulate any sample-path regularity of the process. In practice we are usually interested in processes of finite variation which also have càdlàg sample paths. We build this into the following notation:

(a) We denote by $\mathcal{FV}(\{\tilde{\mathcal{F}}_t\})$ the set of all real-valued, $\{\tilde{\mathcal{F}}_t\}$-adapted, càdlàg processes on the filtered probability space $((\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}}), \{\tilde{\mathcal{F}}_t\})$ which are of finite variation.

(b) We denote by $\mathcal{FV}_0(\{\tilde{\mathcal{F}}_t\})$ the set of $A \in \mathcal{FV}(\{\tilde{\mathcal{F}}_t\})$ which are $\tilde{\mathbb{P}}$-a.s. null at the origin.

(c) We denote by $\mathcal{FV}^+(\{\tilde{\mathcal{F}}_t\})$ the set of all real-valued, $\{\tilde{\mathcal{F}}_t\}$-adapted, càdlàg processes on the filtered probability space $((\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}}), \{\tilde{\mathcal{F}}_t\})$ which are non-decreasing (and are therefore processes of finite variation i.e. $\mathcal{FV}^+(\{\tilde{\mathcal{F}}_t\}) \subset \mathcal{FV}(\{\tilde{\mathcal{F}}_t\})$).

(d) We denote by $\mathcal{FV}_0^+(\{\tilde{\mathcal{F}}_t\})$ the set of $A \in \mathcal{FV}^+(\{\tilde{\mathcal{F}}_t\})$ which are $\tilde{\mathbb{P}}$-a.s. null at the origin.

(e) We denote by $\mathcal{IV}(\{\tilde{\mathcal{F}}_t\})$ the set of $A \in \mathcal{FV}(\{\tilde{\mathcal{F}}_t\})$ which are of integrable variation.

(f) We denote by $\mathcal{IV}_0(\{\tilde{\mathcal{F}}_t\})$ the set of $A \in \mathcal{IV}(\{\tilde{\mathcal{F}}_t\})$ which are $\tilde{\mathbb{P}}$-a.s. null at the origin.

(g) We denote by $\mathcal{IV}^+(\{\tilde{\mathcal{F}}_t\})$ the set of $A \in \mathcal{FV}^+(\{\tilde{\mathcal{F}}_t\})$ which are integrable, that is $E(A(T)) < \infty$.

(h) We denote by $\mathcal{IV}_0^+(\{\tilde{\mathcal{F}}_t\})$ the set of $A \in \mathcal{IV}^+(\{\tilde{\mathcal{F}}_t\})$ which are $\tilde{\mathbb{P}}$-a.s. null at the origin.

From Jacod and Shiryaev [18], Proposition I.1.28(a), Definition I.1.11(a) and Definition I.1.20a, we have the following proposition which is used repeatedly in the thesis to construct stopping times:

Proposition B.2.23 If X is an $\mathbb{R}^n$-valued adapted càdlàg process on the *standard* filtered probability space $((\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}}), \{\tilde{\mathcal{F}}_t\})$, and if B is an open subset of $\mathbb{R}^n$, then $S := \inf\{t : X(t) \in B\}$ is an $\{\tilde{\mathcal{F}}_t\}$-stopping time.

B.2.5 Angle bracket processes for locally square integrable martingales

The definition of the angle bracket quadratic variation and co-variation processes is motivated by the following theorem from Jacod and Shiryaev [18], Theorem I.4.2:

Theorem B.2.24 For each pair $N, M \in \mathcal{M}_2^{loc}(\{\tilde{\mathcal{F}}_t\}, \tilde{\mathbb{P}})$, there exists a real-valued, càdlàg , $\{\tilde{\mathcal{F}}_t\}$-adapted, finite variation process $\langle N, M\rangle$, which is unique up to indistinguishability, such that

1. $\langle N, M\rangle(0) = 0$ a.s.;
2. $\langle N, M\rangle$ is predictable; and
3. $NM - \langle N, M\rangle \in \mathcal{M}_2^{loc}(\{\tilde{\mathcal{F}}_t\}, \tilde{\mathbb{P}})$.

If $N, M \in \mathcal{M}_2(\{\tilde{\mathcal{F}}_t\}, \tilde{\mathbb{P}})$ then we also have $\langle N, M\rangle \in \mathcal{IV}_0(\{\tilde{\mathcal{F}}_t\}, \tilde{\mathbb{P}})$ and $NM - \langle N, M\rangle$ is a uniformly integrable martingale. Furthermore, $\langle N, M\rangle$ is non-decreasing when $N = M$ (in which case we denote $\langle N, M\rangle$ by $\langle M, M\rangle$ or (more briefly) by $\langle M\rangle$; see Remark B.2.26).

Remark B.2.25 We call $\langle N, M\rangle$, which is uniquely determined by Theorem B.2.24, the angle-bracket quadratic co-variation process of N and M.

Remark B.2.26 For any $M \in \mathcal{M}_2^{loc}(\{\tilde{\mathcal{F}}_t\}, \tilde{\mathbb{P}})$, the process $\langle M, M\rangle$ (i.e. the co-quadratic variation of M with itself) is called the angle-bracket quadratic variation process of M. We often write $\langle M\rangle$ for $\langle M, M\rangle$. On the basis of Theorem B.2.24 one can relate the co-quadratic variation $\langle N, M\rangle$ (for $N, M \in \mathcal{M}_2^{loc}(\{\tilde{\mathcal{F}}_t\}, \tilde{\mathbb{P}})$) to the quadratic variation processes of $N + M$ and $N - M$ as follows:

$$\langle N, M\rangle = \frac{1}{4}(\langle N + M, N + M\rangle) - \langle N - M, N - M\rangle. \tag{B.15}$$

Remark B.2.27 A continuous local martingale is locally bounded and therefore of course locally square-integrable, that is we have

$$\mathcal{M}^{c,loc}(\{\tilde{\mathcal{F}}_t\}, \tilde{\mathbb{P}}) \subset \mathcal{M}_2^{loc}(\{\tilde{\mathcal{F}}_t\}, \tilde{\mathbb{P}}). \tag{B.16}$$

It follows that the angle-bracket quadratic covariation $\langle N, M\rangle$ of the continuous local martingales $N, M \in \mathcal{M}^{c,loc}(\{\tilde{\mathcal{F}}_t\}, \tilde{\mathbb{P}})$ is also given by Theorem B.2.24.

B.2.6 Square bracket processes for local martingales

Remark B.2.28 In Theorem B.2.24 it is essential that M and N be *locally square integrable* martingales, that is $N, M \in \mathcal{M}_2^{loc}(\{\tilde{\mathcal{F}}_t\}, \tilde{\mathbb{P}})$. If we simply hypothesized that $N, M \in \mathcal{M}^{loc}(\{\tilde{\mathcal{F}}_t\}, \tilde{\mathbb{P}})$ i.e. N and M are just local martingales without any square integrability, then there generally will not exist an angle-bracket quadratic co-variation process $\langle M, N\rangle$ with the properties stated in Theorem B.2.24. In this section we shall formulate the so-called *square-bracket* processes which do exist even when the local martingales are not necessarily square integrable.

From Jacod and Shiryaev [18], equation I.4.46 and Proposition I.4.50 and Rogers and Williams [32], Theorem VI.36.6 and Theorem VI.37.8, we have the following theorem.

Theorem B.2.29 For each pair $N, M \in \mathcal{M}^{loc}(\{\tilde{\mathcal{F}}_t\}, \tilde{\mathbb{P}})$, there exists a càdlàg , $\{\tilde{\mathcal{F}}_t\}$-adapted process $[N, M]$ of finite variation, which is unique up to indistinguishability, such that

1. $[N, M](0) = 0$ a.s;
2. $\triangle[N, M](t) = \triangle N(t)\triangle M(t)$ for all $t > 0$; and
3. $NM - [N, M] \in \mathcal{M}^{loc}(\{\tilde{\mathcal{F}}_t\}, \tilde{\mathbb{P}})$.

If $N, M \in \mathcal{M}_2(\{\tilde{\mathcal{F}}_t\}, \tilde{\mathbb{P}})$ then $[N, M] \in \mathcal{IV}_0(\{\tilde{\mathcal{F}}_t\}, \tilde{\mathbb{P}})$ and $NM - [N, M]$ is a uniformly integrable martingale. Furthermore, $[N, M]$ is non-decreasing when $N = M$, in which case we denote $[N, M]$ by $[M, M]$ or (more briefly) by $[M]$ (see Remark B.2.31).

Remark B.2.30 We call [N,M] the *square bracket quadratic co-variation process* of N and M.

Remark B.2.31 For any $M \in \mathcal{M}^{loc}(\{\tilde{\mathcal{F}}_t\}, \tilde{\mathbb{P}})$, the process $[M, M]$ is called the *square bracket quadratic variation process* of the local martingale M. We will often write $[M]$ for $[M, M]$. On the basis of Theorem B.2.29 one can relate the square-bracket $[N, M]$ (for $N, M \in \mathcal{M}^{loc}(\{\tilde{\mathcal{F}}_t\}, \tilde{\mathbb{P}})$) to the square-bracket processes of $N + M$ and $N - M$ as follows:

$$[N, M] = \frac{1}{4}([N + M, N + M] - [N - M, N - M]). \tag{B.17}$$

Remark B.2.32 If $N, M \in \mathcal{M}_2^{loc}(\{\tilde{\mathcal{F}}_t\}, \tilde{\mathbb{P}})$ then of course $N, M \in \mathcal{M}^{loc}(\{\tilde{\mathcal{F}}_t\}, \tilde{\mathbb{P}})$, so that both the angle-bracket process $\langle M, N\rangle$ and the square-bracket process $[M, N]$ exist, and it is natural to enquire how these angle-bracket processes might be related. Clearly $\langle N, M\rangle$ has stronger properties than $[M, N]$ in that it is a *predictable* process (a very strong measurability property) whereas $[M, N]$ is only adapted (a much weaker measurability property). Furthermore, the angle-bracket process partially "smooths out" jumps and therefore does not have any property comparable with Theorem B.2.29(2). In fact, there are examples of *discontinuous* $M \in \mathcal{M}_2^{loc}(\{\tilde{\mathcal{F}}_t\}, \tilde{\mathbb{P}})$ (i.e. $\Delta M \neq 0$) such that $\langle M, M\rangle$ is continuous (see e.g. Problem 4 on page 60 of Liptser and Shiryayev [26]). It can in fact be shown that $\langle M, N\rangle$ is the so-called *dual predictable projection* of $[M, N]$. This result plays no role in the thesis so we shall say no more about it here. Moreover, if M and N are *continuous* local martingales (and consequently $N, M \in \mathcal{M}_2^{loc}(\{\tilde{\mathcal{F}}_t\}, \tilde{\mathbb{P}})$), then it turns out that $\langle M, N\rangle$ and $[M, N]$ are actually identical: If $M, N \in \mathcal{M}^{c,loc}(\{\tilde{\mathcal{F}}_t\}, \tilde{\mathbb{P}})$, then

$$[M, N] = \langle M, N\rangle. \tag{B.18}$$

Again, we shall not require this result in the thesis. Despite the fact that the square-bracket process lacks some of the properties of the angle-bracket process it is in fact the preferred entity for general stochastic calculus, where general local martingales are very common whereas square-integrable local martingales are quite rare. In particular, it is fortunately possible to formulate the most important single result in stochastic calculus (the Itô product formula, see Theorem B.2.48 to follow) just in terms of the square-bracket process.

From Protter [30], Chapter II, Section 6, corollary 3, page 73, we have the following corollary which gives conditions in terms of the square-bracket process for a local martingale to be a square integrable martingale:

Corollary B.2.33 Let $M \in \mathcal{M}^{loc}(\{\tilde{\mathcal{F}}_t\}, \tilde{\mathbb{P}})$. We then have the equivalence $M \in \mathcal{M}_2(\{\tilde{\mathcal{F}}_t\}, \tilde{\mathbb{P}})$ iff $E_{\tilde{\mathbb{P}}}[M, M](T) < \infty$, and in this case $E_{\tilde{\mathbb{P}}}|M(t)|^2 = E_{\tilde{\mathbb{P}}}[M, M](t)$, $t \in [0, T]$.

From Protter [30], Chapter II, Section 6, Corollary 4, page 74, we also have the next corollary. In our terminology, an L_2-bounded martingale corresponds to Protter's definition of a square integrable martingale. We use our terminology to state the corollary.

Corollary B.2.34 If $M = \{M(t),\ t \in [0,T]\} \in \mathcal{M}^{loc}(\{\tilde{\mathcal{F}}_t\}, \tilde{\mathbb{P}})$ and $E[M,M](T) < \infty$, then M is a L_2-bounded martingale (that is $\sup_{t\in[0,T]} E|M(t)|^2 = E|M(\infty)|^2 < \infty$). Moreover

$$E_{\tilde{\mathbb{P}}}|M(t)^2| = E_{\tilde{\mathbb{P}}}[M,M](t) \tag{B.19}$$

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

B.2.7 Semimartingales and their decomposition

Definition B.2.35 A real-valued $\{\tilde{\mathcal{F}}_t\}$-adapted process $\{X(t) : t \in [0,T]\}$ on $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}})$ is called a semimartingale if it can be written in the form

$$X = X(0) + M + A, \tag{B.20}$$

for some $M \in \mathcal{M}_0^{loc}((\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}}), \{\tilde{\mathcal{F}}_t\})$ and $A \in \mathcal{FV}_0((\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}}), \{\tilde{\mathcal{F}}_t\})$.

Notation B.2.36 The notation $\mathcal{SM}((\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}}), \{\tilde{\mathcal{F}}_t\})$ denotes the set of all $\{\tilde{\mathcal{F}}_t\}$-adapted semimartingales on $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}})$.

As for the spaces of martingales, if there is no ambiguity about the measurable space on which the space of semimartingales $\mathcal{SM}((\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}}), \{\tilde{\mathcal{F}}_t\})$ is defined, we will write $\mathcal{SM}(\{\tilde{\mathcal{F}}_t\}, \tilde{\mathbb{P}})$ instead of $\mathcal{SM}((\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}}), \{\tilde{\mathcal{F}}_t\})$.

The decomposition on the right side of (B.20) is generally not unique so that we can have

$$X = X(0) + \tilde{M} + \tilde{A} \tag{B.21}$$

for some $\tilde{M} \in \mathcal{M}_0^{loc}((\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}}), \{\tilde{\mathcal{F}}_t\})$, which is distinct from the M at (B.20), and for some $\tilde{A} \in \mathcal{FV}_0((\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}}), \{\tilde{\mathcal{F}}_t\})$ which is distinct from the A at (B.20). In a decomposition of the form (B.20) we call M a *local martingale part* and call A a *finite variation part* of the semimartingale X, and we see that the local martingale part and finite variation part of a semimartingale are generally non-unique.

Although the decomposition at (B.20) is generally not unique there is nevertheless a type of uniqueness which can still be associated with every such decomposition. We discuss this next. We need the following result from Jacod and Shiryaev ([18], Theorem I.4.18) which gives a unique decomposition of a local martingale into a continuous and a purely discontinuous part:

Theorem B.2.37 Any local martingale $M \in \mathcal{M}^{loc}((\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}}), \{\tilde{\mathcal{F}}_t\})$ admits a unique (up to indistinguishability) decomposition

$$M = M(0) + M^c + M^d, \tag{B.22}$$

where $M^c(0) = M^d(0)$, $M^c \in \mathcal{M}^{loc}((\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}}), \{\tilde{\mathcal{F}}_t\})$ is a *continuous* local martingale and $M^d \in \mathcal{M}^{loc}((\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}}), \{\tilde{\mathcal{F}}_t\})$ is a *purely discontinuous* local martingale.

Remark B.2.38 We call M^c the continuous part of the local martingale M and we call M^d the purely discontinuous part of the local martingale M. From Theorem B.2.37 it follows that these entities are unique to within indistinguishability.

With the decomposition of Theorem B.2.37 in mind we have the following result (see Jacod and Shiryaev [18], Proposition I.4.27):

Proposition B.2.39 Let $X \in \mathcal{SM}(\{\tilde{\mathcal{F}}_t\}, \tilde{\mathbb{P}})$ and fix any two arbitrary decompositions

$$X = X(0) + M + A \tag{B.23}$$

$$\text{and} \qquad X = X(0) + \tilde{M} + \tilde{A}, \tag{B.24}$$

for $M, \tilde{M} \in \mathcal{M}_0^{loc}(\{\tilde{\mathcal{F}}_t\}, \tilde{\mathbb{P}})$ and $A, \tilde{A} \in \mathcal{FV}_0(\{\tilde{\mathcal{F}}_t\}, \tilde{\mathbb{P}})$. Decompose the local martingales $M, \tilde{M} \in \mathcal{M}_0^{loc}(\{\tilde{\mathcal{F}}_t\}, \tilde{\mathbb{P}})$ (according to Theorem B.2.37) to get the unique decompositions

$$M = M(0) + M^c + M^d, \tag{B.25}$$

$$\text{and} \qquad \tilde{M} = \tilde{M}(0) + \tilde{M}^c + \tilde{M}^d, \tag{B.26}$$

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.

This proposition can also be stated as follows:

Proposition B.2.40 Let $X \in \mathcal{SM}(\{\tilde{\mathcal{F}}_t\}, \tilde{\mathbb{P}})$. Then there exists some continuous $M^* \in \mathcal{M}_0^{c,loc}(\{\tilde{\mathcal{F}}_t\}, \tilde{\mathbb{P}})$ with the following property: For every decomposition

$$X = X(0) + M + A \tag{B.27}$$

with $M \in \mathcal{M}_0^{loc}(\{\tilde{\mathcal{F}}_t\}, \tilde{\mathbb{P}})$ and $A \in \mathcal{FV}_0(\{\tilde{\mathcal{F}}_t\}, \tilde{\mathbb{P}})$, the continuous part of the local martingale M is indistinguishable from M^*.

Remark B.2.41 According to Proposition B.2.40 to each $X \in \mathcal{SM}(\{\tilde{\mathcal{F}}_t\}, \tilde{\mathbb{P}})$, there corresponds a uniquely defined $M^* \in \mathcal{M}_0^{c,loc}(\{\tilde{\mathcal{F}}_t\}, \tilde{\mathbb{P}})$; this continuous local martingale is is denoted by X^c, that is X^c denotes the unique (up to indistinguishability) member of $\mathcal{M}_0^{c,loc}(\{\tilde{\mathcal{F}}_t\}, \tilde{\mathbb{P}})$ given by Proposition B.2.40. The continuous local martingale X^c is called the *continuous local martingale part* of the semimartingale X. As we shall see, this process is essential for writing down Itô 's formula for general semimartingales.

Another essential ingredient for writing down Itô 's formula for general semimartingales is the so-called *square bracket process* of a pair of semimartingales which we now proceed to define. We need the following technical result (see Section 2.1 of Liptser and Shiryayev [26]):

Proposition B.2.42 *If $X, Y \in \mathcal{SM}(\{\tilde{\mathcal{F}}_t\}, \tilde{\mathbb{P}})$ then for each $t \in [0, \infty)$ one has*

$$\sum_{0 \leq s \leq t} |\Delta X(s) \Delta Y(s)| < \infty \qquad a.s. \tag{B.28}$$

Definition B.2.43 Given any $X, Y \in \mathcal{SM}(\{\tilde{\mathcal{F}}_t\}, \tilde{\mathbb{P}})$ put

$$[X, Y](t) := \langle X^c, Y^c \rangle(t) + \sum_{0 \leq s \leq t} \Delta X(s) \Delta Y(s), \tag{B.29}$$

for all $t \in [0, \infty)$. Here X^c and Y^c denote the continuous local martingale parts X and Y respectively, which are of course uniquely defined (see Remark B.2.41).

Remark B.2.44 Since $X^c, Y^c \in \mathcal{M}_0^{c,loc}(\{\tilde{\mathcal{F}}_t\}, \tilde{\mathbb{P}})$ the first term on the right of (B.29) is given by Theorem B.2.24 (see Remark B.2.27). As for the series in the second term on the right of (B.29), this is absolutely convergent in view of Proposition B.2.42.

From Protter [30], Chapter II, Seciton 6, page 71 we have the following definition and theorem.

Definition B.2.45 Let X be a semimartingale and let X^c denote its continuous local martingale part. Then X is called a purely discontinuous semimartingale if $\langle X^c, X^c \rangle = 0$.

Theorem B.2.46 If a semimartingale X is adapted, càdlàg , with paths of finite variation then X is a purely discontinuous semimartingale.

Protter uses the term quadratic pure jump for purely discontinuous.

As a special case of Jacod and Shiryaev [18], Theorem I.4.52, we have the following theorem.

Theorem B.2.47 let X be a purely discontinuous semimartingale. Then for any semimartingale Y, we have

$$[X, Y](t) = \sum_{0 \leq s \leq t} \triangle X(s) \triangle Y(s). \tag{B.30}$$

B.2.8 Itô formula for general semimartingales

From Rogers and Williams [32], Theorem VI.38.3, we have the following *Itô integration by parts formula* for semimartingales:

Theorem B.2.48 Let X and Y be semimartingales, that is members of $\mathcal{SM}(\{\tilde{\mathcal{F}}_t\}, \tilde{\mathbb{P}})$. Then

$$X(t)Y(t) = X(0)Y(0) + \int_0^t X(\tau_-)\,\mathrm{d}Y(\tau) + \int_0^t Y(\tau_-)\,\mathrm{d}X(\tau) + [X, Y](t). \tag{B.31}$$

From Rogers and Williams [32], Theorem VI.39.1, we have the following *Itô's formula for semimartingales*, which includes the integration by parts formula of Theorem B.2.48 as a special case:

Theorem B.2.49 Let $f : \mathbb{R}^N \to \mathbb{R}$ be a function which has continuous derivatives up to order two. Suppose $\mathbf{X} = (X_1, \ldots, X_N)$ is a semimartingale in $\mathbb{R}^N$, that is $X_i \in \mathcal{SM}(\{\tilde{\mathcal{F}}_t\}, \tilde{\mathbb{P}})$ for all $i = 1, 2, \ldots, N$. Then

$$\begin{aligned} f(\mathbf{X}(t)) - f(\mathbf{X}(0)) &= \sum_{i=0}^{N} \int_0^t \frac{\partial f}{\partial X_i} \mathbf{X}(\tau_-)\,\mathrm{d}\mathbf{X}(\tau) \\ &+ \frac{1}{2} \sum_{i=0}^{N} \sum_{j=0}^{N} \int_0^t \frac{\partial^2 f}{\partial X_i X_j} \mathbf{X}(\tau_-)\,\mathrm{d}[(X_i)^c, (X_j)^c](\tau) \\ &+ \sum_{0 \leq \tau \leq t} \left(f(\mathbf{X}(\tau)) - f(\mathbf{X}(\tau_-) - \sum_{i=1}^{N} \frac{\partial f}{\partial X_i} \mathbf{X}(\tau_-) \triangle X_i(\tau) \right), \end{aligned} \tag{B.32}$$

$(X_i)^c$ denoting the continuous local martingale part of the semimartingale X_i.

B.2.9 Doléans-Dade exponential results

From Elliott [10], Chapter 13, Theorem 13.5 and Remark 13.6, we have

Theorem B.2.50 Suppose $X = \{X(t) : t \geq 0\}$ is a semimartingale which is null at the origin. Let X^c denote its continuous local martingale part. Then there is a unique semimartingale $Z = \{Z(t) : t \geq 0\}$ such that

$$Z(t) = 1 + \int_0^t Z(\tau_-)\,\mathrm{d}X(\tau). \tag{B.33}$$

Furthermore, $Z(t)$ is given by the expression

$$Z(t) = \exp\left\{X(t) - \frac{1}{2}\langle X^c, X^c\rangle\right\} \prod_{\tau\in[0,t]} (1 + \triangle X(\tau)) \exp\{-\triangle X(\tau)\}\,, \tag{B.34}$$

for $t \geq 0$, where the infinite product is absolutely convergent almost surely.

Remark B.2.51 We will use the notation $\mathcal{E}(X)(t)$ to represent $Z(t)$, that is $Z(t) = \mathcal{E}(X)(t)$, and we call $\mathcal{E}(X)$ the *Doléans-Dade exponential* of the semimartingale X.

Clearly, from (B.33), if X is a local martingale then $\mathcal{E}(X)$ is also a local martingale. Furthermore, from (B.34), $\mathcal{E}(X)$ is strictly positive if and only if $\triangle X(t) > -1$ a.s. for all $t \geq 0$. In particular, if X is continuous then, by Remark B.2.32, $[X, X] = \langle X, X\rangle$, and hence

$$Z(t) = \exp\left\{X(t) - \frac{1}{2}[X, X](t)\right\} \quad a.s. \tag{B.35}$$

From Elliott [10], Chapter 13, Corollary 13.58, we also have the following result.

Corollary B.2.52 If X and Y are semimartingales, then

$$\mathcal{E}(X)\mathcal{E}(Y) = \mathcal{E}(X + Y + [X, Y]) \tag{B.36}$$

From Protter [30], Chapter III, Section 8, Theorem 45, page 141, we have *Novikov's Criterion,* which gives conditions for the Doléans-Dade exponential of a continuous local martingale to be a martingale.

Theorem B.2.53 Novikov's Criterion. Let M be a continuous local martingale and suppose that

$$E\left(\exp\left\{\frac{1}{2}[M, M](\infty)\right\}\right) < \infty \tag{B.37}$$

Then $\mathcal{E}(M)$ is a uniformly integrable martingale.

B.3 Compensator Results

In the following definitions we localize the notion of integrable variation process (see Definition B.2.21 and Notation B.2.22).

Definition B.3.1 Recalling Notation B.2.22 we denote by $\mathcal{IV}_{loc}(\{\tilde{\mathcal{F}}_t\})$ the set of all $A \in \mathcal{FV}(\{\tilde{\mathcal{F}}_t\})$ for which there exists a sequence of $\{\tilde{\mathcal{F}}_t\}$-stopping times $\{T^m\}_{m\in N}$ (depending on A) such that $T^m \Uparrow T$ and each stopped process $A[0,T^m] \in \mathcal{IV}(\{\tilde{\mathcal{F}}_t\})$. Members of $\mathcal{IV}_{loc}(\{\tilde{\mathcal{F}}_t\})$ are processes of *locally integrable variation.*

It is immediately clear from this definition that

$$\mathcal{IV}(\{\tilde{\mathcal{F}}_t\}) \subset \mathcal{IV}_{loc}(\{\tilde{\mathcal{F}}_t\}). \tag{B.38}$$

Notation B.3.2 Put

$$\mathcal{IV}^+_{0,loc}(\{\tilde{\mathcal{F}}_t\}) = \mathcal{FV}^+_0(\{\tilde{\mathcal{F}}_t\}) \cap \mathcal{IV}_{loc}(\{\tilde{\mathcal{F}}_t\}). \tag{B.39}$$

That is, members of $\mathcal{IV}^+_{0,loc}(\{\tilde{\mathcal{F}}_t\})$ are processes of locally integrable variation with sample paths which are null at $t = 0$ and non-decreasing.

From Jacod and Shiryaev [18], Theorem I.3.17, we have the following basic result:

Theorem B.3.3 Let $A \in \mathcal{IV}^+_{0,\mathrm{loc}}(\{\tilde{\mathcal{F}}_t\}, \tilde{\mathbb{P}})$. Then there exists a *predictable* process $A^p \in \mathcal{IV}^+_{0,\mathrm{loc}}(\{\tilde{\mathcal{F}}_t\}, \tilde{\mathbb{P}})$ such that $A - A^p \in \mathcal{M}^{loc}_0(\{\tilde{\mathcal{F}}_t\}, \tilde{\mathbb{P}})$. Moreover, A^p is unique in the sense that, for any predictable process $\hat{A} \in \mathcal{IV}^+_{0,\mathrm{loc}}(\{\tilde{\mathcal{F}}_t\}, \tilde{\mathbb{P}})$ such that $A - \hat{A} \in \mathcal{M}^{loc}_0(\{\tilde{\mathcal{F}}_t\}, \tilde{\mathbb{P}})$, the process A^p and $\hat{A}$ are indistinguishable.

Remark B.3.4 The process A^p is called the *compensator* (or *dual predictable projection*) of the given process A, and is unique to within indistinguishability.

Theorem B.3.5 Let $A \in \mathcal{IV}^+_{0,\mathrm{loc}}(\{\tilde{\mathcal{F}}_t\}, \tilde{\mathbb{P}})$. For each predictable process $\bar{A} \in \mathcal{IV}^+_{0,\mathrm{loc}}(\{\tilde{\mathcal{F}}_t\}, \tilde{\mathbb{P}})$ the following are equivalent:

1. $\bar{A}$ is the compensator of A, that is $\bar{A} = A^p$;

2. for all nonnegative predictable processes H one has

$$E\int_0^\infty H(\tau)\,\mathrm{d}A(\tau) = E\int_0^\infty H(\tau)\,\mathrm{d}\bar{A}(\tau). \tag{B.40}$$

Given a local martingale $M \in \mathcal{M}^{loc}(\{\tilde{\mathcal{F}}_t\}, \tilde{\mathbb{P}})$ we know that $[M] \in \mathcal{FV}_0^+(\{\tilde{\mathcal{F}}_t\}, \tilde{\mathbb{P}})$ (see Theorem B.2.29). The following theorem (from Rogers and Williams [32], Theorem VI.34.2) establishes conditions on M which ensure that $[M] \in \mathcal{IV}_{0,\mathrm{loc}}^+(\{\tilde{\mathcal{F}}_t\}, \tilde{\mathbb{P}})$ and shows that $[M]^p$ (the compensator of $[M]$) and $\langle M \rangle$ are identical in this case:

Theorem B.3.6 Let $M \in \mathcal{M}_0^{loc}(\{\tilde{\mathcal{F}}_t\}, \tilde{\mathbb{P}})$. Then the following statements are equivalent:

1. $M \in \mathcal{M}_{0,2}^{loc}(\{\tilde{\mathcal{F}}_t\}, \tilde{\mathbb{P}})$;

2. the increasing process $[M]$ is locally integrable, that is $[M] \in \mathcal{IV}_{0,\mathrm{loc}}^+(\{\tilde{\mathcal{F}}_t\}, \tilde{\mathbb{P}})$.

Under these equivalent conditions we have

$$\langle M \rangle = [M]^p,$$

that is $\langle M \rangle$ is the compensator of $[M]$ (recall Theorem B.2.24).

B.4 Convex analysis

In this section we summarize some basic definitions and results on convex analysis in the simplest setting of general vector spaces (without any norm or topology on the vector spaces involved).

Definition B.4.1 Let U be a real vector space and suppose that $\psi : U \to [-\infty, +\infty]$ is a given function. The *epigraph* of ψ is the subset of the vector space $\mathbb{R} \times U$ defined by

$$\mathrm{epi}\psi := \{(\alpha, u) \in \mathbb{R} \times U \quad : \quad \alpha \geq \psi(u)\} \tag{B.41}$$

and the *domain* of ψ is the subset of U defined by

$$\mathrm{dom}\psi := \{u \in U \quad : \quad \psi(u) < +\infty\}. \tag{B.42}$$

The function ψ is said to be *c*onvex when epiψ is a convex subset of $\mathbb{R} \times U$, that is

$$\lambda(\alpha_1, u_1) + (1-\lambda)(\alpha_2, u_2) \in \mathrm{epi}\psi \text{ for all } \lambda \in [0,1] \text{ and } (\alpha_i, u_i) \in \mathrm{epi}\psi. \tag{B.43}$$

Remark B.4.2 It is easily seen that a function $\psi : U \to [-\infty, +\infty]$ is convex if and only if

$$\psi(\epsilon u + (1-\epsilon)\bar{u}) \le \epsilon\psi(u) + (1-\epsilon)\psi(\bar{u}), \tag{B.44}$$

for all $\epsilon \in [0,1]$ and all u and $\bar{u}$ in U such that the right hand side is defined (i.e. u and $\bar{u}$ are such that we do not get $\infty - \infty$).

Remark B.4.3 Elementary definitions of convexity are typically only for $\mathbb{R}$-valued functions. However, this is not quite general enough for us since we will want to deal with convex functions which can take the value $+\infty$ to account for constraints in our optimization problem (see e.g. the primal function $f(\cdot)$ at Definition 2.2.26). On the other hand, there is usually little interest in convex functions which take the value $-\infty$ anywhere on U since these functions are degenerate or pathological in the following sense: if $\psi : U \to [-\infty, +\infty]$ is convex and $\psi(\bar{u}) = -\infty$ for some $\bar{u} \in U$ then, for every half-line in U starting at $\bar{u}$ with direction $v \in U$, that is

$$\mathrm{HL}(\bar{u}, v) = \{\bar{u} + \alpha v \;\; : \;\; \alpha \in [0, \infty)\},$$

one has either $\psi(\bar{u} + \alpha v) = -\infty$ for all $\alpha \in [0, \infty)$, or there exists some $\alpha_0 \in (0, \infty)$ such that $\psi(\bar{u} + \alpha v) = -\infty$ for all $0 < \alpha < \alpha_0$ and $\psi(\bar{u} + \alpha v) = +\infty$ for all $\alpha > \alpha_0$ (see page 8 of Ekeland and Témam [9] for further discussion of this). Such pathological convex functions are generally of no interest. Likewise, the function ψ defined by $\psi(u) = +\infty$ for *all* $u \in U$, that is $\mathrm{dom}\psi = \emptyset$ (recall (B.42)), is undoubtedly convex but also rather pathological and of very little interest. We next single out a particularly important class of convex functions on U which avoid these two pathologies (see Definition 7.1 of Aliprantis and Border [1]):

Definition B.4.4 A convex function $\psi : U \to [-\infty, +\infty]$ defined on the vector space U is a *proper* convex function when $\mathrm{dom}\psi \neq \emptyset$ and $\psi(u) > -\infty$ for all $u \in U$.

Definition B.4.5 Given two vector spaces U and Y, a mapping $\psi : U \times Y \to \mathbb{R}$ is a bilinear form on $U \times Y$ when

(a) The mapping $u \to \psi(u, y) : U \to \mathbb{R}$ is linear for each fixed $y \in Y$,

(b) The mapping $y \to \psi(u, y) : Y \to \mathbb{R}$ is linear for each fixed $u \in U$.

Remark B.4.6 If U and Y are vector spaces and $\psi : U \times Y \to \mathbb{R}$ is linear, that is

$$\text{i.e.} \qquad \psi(\alpha_1(u_1, y_1) + \alpha_2(u_2, y_2)) = \alpha_1\psi(u_1, y_1) + \alpha_2\psi(u_2, y_2)$$
$$\text{for all } \alpha_1 \,\&\, \alpha_2 \in \mathbb{R}, \text{ all } (u_i, y_i) \in U \times Y,$$

then ψ is bilinear form on $U \times Y$. However, there are bilinear forms on $U \times Y$ which are not necessarily linear functions on $U \times Y$.

Remark B.4.7 It is customary to use the notation $\langle \cdot, \cdot \rangle$ to denote a given bilinear form on $U \times Y$ for vector spaces U and Y, that is $\langle u, y \rangle$ denotes $\psi(u, y)$, $(u, y) \in U \times Y$, in the notation of Definition B.4.5. The triple $(U, Y, \langle \cdot, \cdot \rangle)$ is called a *dual system*.

Remark B.4.8 A particularly simple but important special case of a dual system $(U, Y, \langle \cdot, \cdot \rangle)$ corresponds to $U = Y = \mathbb{R}^N$ with $\langle u, y \rangle$ being the usual inner product (or "dot" product) of the vectors $u, y \in \mathbb{R}^N$.

Definition B.4.9 Suppose that $(U, Y, \langle \cdot, \cdot \rangle)$ is a given dual system. The $\langle U, Y \rangle$ - convex conjugate of a given function $\psi : U \to [-\infty, \infty]$ is the *convex* function $\psi^* : Y \to (-\infty, +\infty]$ defined by

$$\psi^*(y) := \sup_{u \in U}[\langle u, y \rangle - \psi(u)], \qquad y \in Y. \tag{B.45}$$

Remark B.4.10 Note that the function ψ in Definition B.4.9 need not be convex. However, it is immediate from (B.45) that ψ^* is always a convex function on Y, regardless of the properties of ψ, although ψ^* is not necessarily a *proper* convex function. In fact, if ψ is such that $\psi(\bar{u}) = -\infty$ for some $\bar{u} \in U$ then it is immediate that $\psi^*(y) = +\infty$ for all $y \in Y$, that is $\text{dom}\psi^* = \emptyset$, and therefore ψ^* fails to be a proper convex function. On the other hand, if $\psi(u) = +\infty$ for all $u \in U$, then it is immediate that $\psi^*(y) = -\infty$ for all $y \in Y$, that is ψ^* is again a (highly) non-proper convex function.

Definition B.4.11 Suppose that $(U, Y, \langle \cdot, \cdot \rangle)$ is a given dual system. Then the $\langle U, Y \rangle$-bi-conjugate of the given function $\psi : U \to [-\infty, \infty]$ is the convex function $\psi^{**} : U \to [-\infty, +\infty]$ defined by

$$\psi^{**}(u) := \sup_{y \in Y}[\langle u, y \rangle - \psi^*(y)], \qquad u \in U. \tag{B.46}$$

Remark B.4.12 Observe that for a given function $\psi : U \to [-\infty, \infty]$ (not necessarily convex), we always have

$$\psi^{**}(u) \leq \psi(u), \qquad u \in U. \tag{B.47}$$

Definition B.4.13 Suppose that $(U, Y, \langle\cdot,\cdot\rangle)$ is a given dual system and $\psi : U \to [-\infty, +\infty]$ is a given mapping (not necessarily convex). At each $u \in U$ such that $-\infty < \psi(u) < +\infty$ we define the *subgradient* of ψ at u to be the subset of Y given by

$$\partial\psi(u) := \{y \in Y \mid \psi(u') \geq \psi(u) + \langle u' - u, y\rangle \textit{ all } u' \in U\}, \tag{B.48}$$

and ψ is said to be *sub-differentiable at* u when $\partial\psi(u) \neq \emptyset$.

Remark B.4.14 Suppose that $\psi : \mathbb{R}^N \to \mathbb{R}$ is a *smooth* convex function with usual the vector of partial derivatives $D\psi(u) \in \mathbb{R}^N$ at each $u \in \mathbb{R}^N$, and $(U, Y, \langle\cdot,\cdot\rangle)$ is the simple duality system at Remark B.4.8. Then it is immediate that the subgradient of ψ at each u is just the single-point set $\{D\psi(u)\}$, that is

$$\partial\psi(u) = \{D\psi(u)\}, \qquad \text{for all } u \in \mathbb{R}^N. \tag{B.49}$$

In this situation we will use $\partial\psi(u)$ to denote the actual vector $D\psi(u)$ itself, rather than the single-point set which contains the vector $D\psi(u)$.

The next result, which is a very special case of Proposition 5.1 and Corollary 5.2, page 21-22 of Ekeland and Témam[9], is used in this thesis to construct *transversality conditions* in Kuhn-Tucker optimality relations:

Lemma B.4.15 Suppose that $\psi : U := \mathbb{R}^N \to \mathbb{R}$ is a smooth convex function, and $(U, Y, \langle\cdot,\cdot\rangle)$ is the simple duality system at Remark B.4.8. Then $\psi^* : Y := \mathbb{R}^N \to \mathbb{R}$ is also a smooth convex function, with the vector of partial derivatives denoted by $\partial\psi^*(y) \in \mathbb{R}^N$ at each $y \in \mathbb{R}^N$ (in accordance with Remark B.4.14). Moreover, for each $u \in U := \mathbb{R}^N$ and $y \in Y := \mathbb{R}^N$ we have the equivalence

$$\begin{gathered}\psi(u) + \psi^*(y) = \langle u, y\rangle \\ \text{iff} \\ u = \partial\psi^*(y).\end{gathered} \tag{B.50}$$

Appendix C

Topological Vector Spaces

This appendix is for supplementary information only, and is not necessary for reading the thesis. We give some elementary definitions and remarks on topological vector spaces which are used throughout the thesis. For background purposes only we also state the Mackey-Arens theorem and Mackey theorem since these may help to promote a greater understanding of the choice of the vector space of dual variables given by (3.11) and the choice of the bilinear form given by (3.12). With this choice of bilinear form, the Mackey-Arens theorem together with the Mackey theorem states that the norm-topology on $\mathbb{U}$ is the Mackey topology, as asserted at Remark 3.1.6.

Definition C.1.1 A *topology* on a set X is a collection $\mathcal{T}$ of subsets of X having the following properties.

1. $\emptyset$ and X are in $\mathcal{T}$,
2. The union of the elements of any subcollection of $\mathcal{T}$ is in $\mathcal{T}$,
3. The intersection of the elements of any finite subcollection of $\mathcal{T}$ is in $\mathcal{T}$.

A *Topological space* is an ordered pair $(X, \mathcal{T})$ consisting of a set X and a topology $\mathcal{T}$ on X. Any element of $\mathcal{T}$ is called an $\mathcal{T}$-open (or open) subset of X. A subset of X which is the complement of an element of $\mathcal{T}$ is called $\mathcal{T}$-closed.

Definition C.1.2 Let $(X, \mathcal{T})$ be a topological space and let $A \subset X$. The $\mathcal{T}$-interior of A, denoted by A° is defined to be the union of all subsets of A which are $\mathcal{T}$-open. The $\mathcal{T}$-closure of A, denoted by $\bar{A}$, is defined to be the intersection of all $\mathcal{T}$-closed supersets of A.

Remark C.1.3 Given a topological space $(X, \mathcal{T})$ and some $A \subset X$, we see that A° is the largest open subset of A and $\bar{A}$ is the smallest closed superset of A.

Remark C.1.4 Given a topological space $(X, \mathcal{T})$ and some $x \in X$, a set $U \subset X$ is a $\mathcal{T}$-neighborhood of x when $U^\circ \neq \phi$ and $x \in U^\circ$.

Remark C.1.5 $(X, \mathcal{T})$ is a *Hausdorff space*, and $\mathcal{T}$ is a *Hausdorff topology* if distinct points of X have disjoint neighborhoods. A collection $\mathcal{T}' \subset \mathcal{T}$ is a base for $\mathcal{T}$ if every member of $\mathcal{T}$ is a union of members of $\mathcal{T}'$. A collection $\mathcal{B}$ of neighborhoods at point $x \in X$ is a *local base* at x if every neighborhood of x contains a member of $\mathcal{B}$.

Definition C.1.6 Suppose $\mathcal{T}$ is a topology on a vector space X such that the vector space operations of vector addition and scalar multiplication are continuous with respect to $\mathcal{T}$. Then $\mathcal{T}$ is said to be a *vector topology* on X, and $(X, \mathcal{T})$ is a topological vector space.

To say that vector addition is continuous means, by definition, that the mapping

$$(x, y) \to x + y \tag{C.1}$$

of the cartesion product $X \times X$ into X is continuous. This means that, if $x_i \in X$ for $i = 1, 2$, and V is a neighborhood of $x_1 + x_2$, then there should exist neighborhoods V_i of x_i, such that

$$V_1 + V_2 \subset V. \tag{C.2}$$

Similarly, the assumption that scalar multiplication is continuous means that the mapping

$$(\alpha, x) \to \alpha x \tag{C.3}$$

of $\phi \times X$ into X is continuous: That is, if $x \in X$, α is a scalar, and V is a neighborhood of αx, then for some $r > 0$ and some neighborhood W of x we have $\beta W \subset V$ whenever $|\beta - \alpha| < r$.

Definition C.1.7 If $(X, \mathcal{T})$ is a topological vector space then the vector topology $\mathcal{T}$ is called locally convex when for each $\mathcal{T}$-neighborhood G of $0 \in X$ there exists a convex $\mathcal{T}$-neighborhood N of $0 \in X$ such that $N \subset G$. That is, $\mathcal{T}$ has a local base of convex sets at $0 \in X$.

Remark C.1.8 X is *locally convex* if there is local base $\mathcal{B}$ whose members are convex.

Remark C.1.9 Suppose that $(\mathbb{U}, \mathbb{Y}, \langle\cdot,\cdot\rangle)$ is a dual system in the sense of Remark B.4.7. We do not as yet have any natural topology on either of the vector spaces $\mathbb{U}$ or $\mathbb{Y}$. We denote by $\mathfrak{S}(\mathbb{U}, \mathbb{Y})$ the topology on $\mathbb{U}$ which is generated by the mappings

$$u \to \langle u, y\rangle : \mathbb{U} \to \mathbb{R} \qquad \text{for all } y \in \mathbb{Y},$$

and call $\mathfrak{S}(\mathbb{U}, \mathbb{Y})$ the *weak topology* generated on $\mathbb{U}$ through the bilinear form $\langle\cdot,\cdot\rangle$. In an exactly symmetric way we can define $\mathfrak{S}(\mathbb{Y}, \mathbb{U})$, the *weak topology* generated on $\mathbb{Y}$ through the bilinear form $\langle\cdot,\cdot\rangle$. It is an elementary exercise to check that $\mathfrak{S}(\mathbb{U}, \mathbb{Y})$ is a locally convex vector topology on $\mathbb{U}$, and similarly for $\mathfrak{S}(\mathbb{Y}, \mathbb{U})$ on $\mathbb{Y}$.

Definition C.1.10 Suppose that $(\mathbb{U}, \mathbb{Y}, \langle\cdot,\cdot\rangle)$ is a given dual system. A locally convex Hausdorff topology $\mathcal{U}$ on $\mathbb{U}$ is called compatible with the bilinear form $\langle\cdot,\cdot\rangle$ on $\mathbb{U} \times \mathbb{Y}$ (or $\langle\mathbb{U}, \mathbb{Y}\rangle$-compatible for short) when

(a) The mapping $u \to \langle u, y\rangle : \mathbb{U} \to \mathbb{R}$ is $\mathcal{U}$-continuous for each $y \in \mathbb{Y}$,

(b) If $\Sigma : \mathbb{U} \to \mathbb{R}$ is any $\mathcal{U}$-continuous linear functional then there exists some $y \in Y$ such that $\Sigma(u) = \langle u, y\rangle$ for all $u \in \mathbb{U}$.

Remark C.1.11 Given the dual system $(\mathbb{U}, \mathbb{Y}, \langle\cdot,\cdot\rangle)$ it is natural to ask how one characterizes the $\langle\mathbb{U}, \mathbb{Y}\rangle$-compatible topologies on $\mathbb{U}$. It is immediate from Definition C.1.10 and Remark C.1.9 that $\mathfrak{S}(\mathbb{U}, \mathbb{Y})$ is a $\langle\mathbb{U}, \mathbb{Y}\rangle$-compatible topology on $\mathbb{U}$, and if $\mathcal{U}$ is another $\langle\mathbb{U}, \mathbb{Y}\rangle$-compatible topology on $\mathbb{U}$ then $\mathfrak{S}(\mathbb{U}, \mathbb{Y}) \subset \mathcal{U}$, that is $\mathfrak{S}(\mathbb{U}, \mathbb{Y})$ is the weakest among all $\langle\mathbb{U}, \mathbb{Y}\rangle$-compatible topologies on $\mathbb{U}$. It turns out that there is also a *strongest* locally convex Hausdorff topology on $\mathbb{U}$ which is $\langle\mathbb{U}, \mathbb{Y}\rangle$-compatible, called the *Mackey topology* which is defined as follows: Put

$$\mathcal{S} = \{A \subset \mathbb{Y} \mid A \text{ is balanced, convex and } \mathfrak{S}(\mathbb{Y}, \mathbb{U})\text{-compact}\}.$$

For each $A \in \mathcal{S}$ define the seminorm $q_A : \mathbb{U} \to [0, \infty)$ as follows:

$$q_A(u) = \sup\{|\langle u, y\rangle| \quad | \quad y \in A\}, \qquad \text{for all } u \in \mathbb{U},$$

and let $\tau(\mathbb{U}, \mathbb{Y})$ be the locally convex Hausdorff topology on $\mathbb{U}$ with the subbase

$$\{u \in \mathbb{U} \quad | \ q_A(u) < \epsilon\} \qquad \text{for all } \epsilon \in (0, \infty) \text{ and } A \in \mathcal{S}.$$

Then $\tau(\mathbb{U}, \mathbb{Y})$ is called the *Mackey topology* on $\mathbb{U}$ derived from the dual system $(\mathbb{U}, \mathbb{Y}, \langle\cdot,\cdot\rangle)$.

The significance of this topology is given by the following result of Mackey and Arens (see e.g. Theorem 9-2-3 of Wilansky [43] or Theorem 5.113 of Aliprantis and Border [1]):

Theorem C.1.12 (Mackey-Arens) *Suppose that* $(\mathbb{U}, \mathbb{Y}, \langle \cdot, \cdot \rangle)$ *is a given dual system. Then the weak topology* $\mathfrak{S}(\mathbb{U}, \mathbb{Y})$ *and the Mackey topology* $\tau(\mathbb{U}, \mathbb{Y})$ *are both* $\langle \mathbb{U}, \mathbb{Y} \rangle$*-compatible topologies on* $\mathbb{U}$*. Moreover, if* $\mathcal{U}$ *is any* $\langle \mathbb{U}, \mathbb{Y} \rangle$*-compatible topology on* $\mathbb{U}$ *then*

$$\mathfrak{S}(\mathbb{U}, \mathbb{Y}) \subset \mathcal{U} \subset \tau(\mathbb{U}, \mathbb{Y}).$$

Effectively, Theorem C.1.12 says that the weak topology $\mathfrak{S}(\mathbb{U}, \mathbb{Y})$ and the Mackey topology $\tau(\mathbb{U}, \mathbb{Y})$ on $\mathbb{U}$ are, respectively, the weakest and strongest among all $\langle \mathbb{U}, \mathbb{Y} \rangle$-compatible topologies on $\mathbb{U}$, and if $\mathcal{U}$ is any $\langle \mathbb{U}, \mathbb{Y} \rangle$-compatible topology on $\mathbb{U}$ then $\mathcal{U}$ is stronger than $\mathfrak{S}(\mathbb{U}, \mathbb{Y})$ and weaker than $\tau(\mathbb{U}, \mathbb{Y})$.

In Remark 3.1.6 it is noted that the most appropriate choice of topology $\mathcal{U}$ on $\mathbb{U}$ when using Theorem 3.1.4 is the Mackey topology since it has the largest collection of open sets. Theorem C.1.13 which follows shows that the norm topology on $\mathcal{U}$ is in fact the Mackey topology (see e.g. Corollary 6.23 of Aliprantis and Border [1]):

Theorem C.1.13 (Mackey) Suppose that $(\mathbb{U}, \|\cdot\|)$ is a normed vector space, let $\mathbb{Y} = \mathbb{U}^\star$ be the norm dual of $\mathbb{U}$, and define the dual system $(\mathbb{U}, \mathbb{Y}, \langle \cdot, \cdot \rangle)$ as

$$\langle u, y \rangle = y(u), \qquad u \in \mathbb{U}, \;\; y \in \mathbb{Y}.$$

The Mackey topology $\tau(\mathbb{U}, \mathbb{Y})$ and the norm-topology on $\mathbb{U}$ are identical.

It is Theorem C.1.13 which makes the norm topology the most appropriate choice of topology on $\mathcal{U}$.

Appendix D

The Yosida - Hewitt Decomposition of $(L_{\infty}^{*}(S, \Sigma, \nu))$

In this appendix we look at the dual space of the normed vector space $L_{\infty}(S, \Sigma, \nu)$ of essentially bounded random variables on the probability space (S, Σ, ν) with the usual essential-supremum norm. This dual space is rather subtle and is characterized in full by the *Yosida-Hewitt* decomposition theorem. First we recall a simple definition:

Definition D.1.1 For each constant $p \in [1, \infty]$ the set of all norm-continuous linear functionals on $L_p(S, \Sigma, \nu)$ is called the dual or adjoint space, and is denoted by $L_p^*(S, \Sigma, \nu)$.

For $1 \leq p < \infty$, let $g \in L_q(S, \Sigma, \nu)$, where $q = \frac{p}{(p-1)}$ if $1 < p < \infty$ and $q = \infty$ if $p = 1$. Let

$$Z_g(f) = \int fg \, \mathrm{d}\nu, \quad f \in L^p(S, \Sigma, \nu) \tag{D.1}$$

It is immediate from Holder's inequality that Z_g is a norm-continuous linear functional on $L_p(S, \Sigma, \nu)$. The next result asserts that *every* norm-continuous linear functional on $L_p(S, \Sigma, \nu)$ is of the form (D.1) for some $g \in L_q(S, \Sigma, \nu)$:

Theorem D.1.2 *(Riesz Representation Theorem)* Let $1 \leq p < \infty$. Let $Z : L_p(S, \Sigma, \nu) \to \mathbb{R}$ be linear and norm-continuous. Then there exists a unique g in $L_q((S, \Sigma, \nu)$ such that $Z = Z_g$, i.e.

$$Z(f) = Z_g(f) := \int fg \, \mathrm{d}\nu \quad \text{for all } f \in L_p(S, \Sigma, \nu), \tag{D.2}$$

where, $q = \frac{p}{p-1}$ for $1 < p < \infty$ and $q = \infty$ if $p = 1$.

See Theorem 8 Section 6.4 of Royden [38], Chapter 11, section 7, page 284.

Remark D.1.3 It is immediate from Holder's inequality that, for each $g \in L_1(S, \Sigma, \nu)$, the linear functional Z defined by

$$Z(f) = \int fg \, \mathrm{d}\nu, \qquad \text{for all } f \in L_\infty(S, \Sigma, \nu), \tag{D.3}$$

is a norm-continuous linear functional on $L_\infty(S, \Sigma, \nu)$ so that

$$L_1(S, \Sigma, \nu) \text{ is a vector subspace of } L^*_\infty(S, \Sigma, \nu). \tag{D.4}$$

Example D.1.4 which follows demonstrates that $L_1(S, \Sigma, \nu)$ is a *strict* vector subspace of $L^*_\infty(S, \Sigma, \nu)$. That is, there are linear functionals $Z \in L^*_\infty(S, \Sigma, \nu)$ which are not in the form of (D.3) for some $g \in L_1(S, \Sigma, \nu)$.

Example D.1.4 The fact that there exist linear functionals $Z \in L^*_\infty(S, \Sigma, \nu)$ which cannot be represented in the form (D.3) for some $g \in L_1(S, \Sigma, \nu)$ can be seen from a simple example discussed in Rudin [40], Chapter 6, Exercise 13, page 134. To see this consider a linear functional

$$Z : C[0,1] \to \mathbb{R} \tag{D.5}$$

$$\text{such that} \qquad Z(f) = f(0). \tag{D.6}$$

$$\text{Then} \qquad \| Z \| = 1. \tag{D.7}$$

Z is bounded linear functional on $C[0,1]$ and $C[0,1]$ is a norm-closed linear subspace of $L_\infty[0,1]$. By the Hahn-Banach extension theorem, Z can be extended to a bounded linear functional (also denoted by Z) on $L_\infty[0,1]$ with identical norm.

Suppose there exists $g \in L_1[0,1]$ such that

$$Z(f) = \int_0^1 f(t)g(t)\, dt \qquad \forall f \in C[0,1] \tag{D.8}$$

Now, let $\{f_n\}$ be a sequence of continuous functions on $[0,1]$ that are norm-bounded by 1, $f_n(0) = 1$, and are such that $f_n(t) \to 0 \ \forall t \neq 0$.

Then $Z(f_n) = \int_0^1 f_n(t)g(t)\, dt \to 0$ (by the dominated convergence theorem), while $Z(f_n) = f_n(0) = 1$ for all $n = 1, 2, \ldots$, giving a contradiction.

Remark D.1.5 From Example (D.1.4) we see that in general

$$L_1(S,\Sigma,\nu) \subsetneqq (L_\infty^*(S,\Sigma,\nu)). \tag{D.9}$$

With this remark in mind the question of a complete representation of $(L_\infty^*(S,\Sigma,\nu))$ arises. This representation is given by the *Yosida-Hewitt decomposition* [46].

For stating the Yosida-Hewitt decomposition theorem we must first define *singular* continuous linear functionals on $L_\infty(S,\Sigma,\nu)$:

Definition D.1.6 We say that $Z \in L_\infty^*(S,\Sigma,\nu)$ is a *singular* norm-continuous linear functional on $L_\infty(S,\Sigma,\nu)$ if there is a sequence $A_n \in \Sigma$, $n = 1,2,\ldots$, s.t.

$$\begin{aligned} A_{n+1} &\subset A_n, \\ \lim_{n\to\infty} \nu(A_n) &= 0, \\ \text{and} \qquad Z(u) &= Z(uI_{A_n}), \end{aligned}$$

for each $u \in L_\infty(S,\Sigma,\nu)$ and $n = 1,2,\ldots$.

Remark D.1.7 It is clear from Definition D.1.6 that, if $\{A_n\}$ is a decreasing sequence of sets associated with a *non-trivial* $Z \in L_\infty^*(S,\Sigma,\nu)$ (i.e. $Z(u) \neq 0$ for some $u \in L_\infty(S,\Sigma,\nu)$), then $\nu(A_n) > 0$ for each $n = 1,2,\ldots$.

Notation D.1.8 We denote by $\mathcal{Q}(S,\Sigma,\nu)$ the set of all singular norm-continuous linear functionals on $L_\infty(S,\Sigma,\nu)$.

It is easily verified that $\mathcal{Q}(S,\Sigma,\nu)$ is a *linear* subspace of $L_\infty^*(S,\Sigma,\nu)$.

Remark D.1.9 The members of $\mathcal{Q}(S,\Sigma,\nu)$ seem almost counter-intuitive and are rather hard to grasp intuitively. Indeed, suppose that we take some non-trivial $Z \in \mathcal{Q}(S,\Sigma,\nu)$ with corresponding sequence $\{A_n\} \subset \Sigma$ "shrinking to zero" in accordance with Definition D.1.6 (i.e. $\nu(A_n) \to 0$), that is the sequence $\{I_{A_n}\} \subset L_\infty(S,\Sigma,\nu)$ monotonically decreases to the zero element of $L_\infty(S,\Sigma,\nu)$. Intuitively one would then expect that $Z(I_{A_n}) \to 0$. However, taking $u \in L_\infty(S,\Sigma,\nu)$ to be the function with constant unit value (i.e. $u \equiv 1$ on S) in Definition D.1.6 we see that in fact

$$Z(I_{A_n}) = Z(1) \neq 0 \qquad \text{for all } n = 1,2,\ldots,$$

contrary to what seems to be intuitively reasonable. Example D.1.10 due to Yosida and Hewitt [46] nevertheless shows that non-trivial singular linear functionals definitely do exist, that is the vector subspace $\mathcal{Q}(S,\Sigma,\nu)$ is non-trivial.

Example D.1.10 Suppose that (S, Σ, ν) is a probability space which is not finitely atomic. Then we can fix a sequence $A_n \in \Sigma$, such that $A_{n+1} \subset A_n$, $\nu(A_n) > 0$ and $\nu(A_n) \to 0$ as $n \to \infty$.

For $u \in L_\infty(S, \Sigma, \nu)$ Define

$$Z_n(u) := E[uI_{A_n}]/\nu(A_n), \tag{D.10}$$

where $n = 1, 2, \ldots$ Then $\{Z_n,\ n = 1, 2, \ldots\}$ is a subset of the unit ball of $\{L^*_\infty(S, \Sigma, \nu)\}$, hence (by Alaoglu theorem) has a $\sigma(L^*_\infty, L_\infty)$-accumulation point $Z^\circ \in L^*_\infty(S, \Sigma, \nu)$. Moreover $Z_n(1) = 1$, and, for each fixed m, we have

$$Z_n(uI_{A_m}) = Z_n(u), \tag{D.11}$$

for all $n \geq m$ and $u \in L^*_\infty(S, \Sigma, \nu)$.

Thus for $u \in L_\infty(S, \Sigma, \nu)$ we get

$$\begin{aligned} & Z^\circ(1) = 1 \\ \text{and} \qquad & Z^\circ(uI_{A_m}) = Z^\circ(u), \end{aligned} \tag{D.12}$$

that is Z° is a non-trivial member of $\mathcal{Q}(S, \Sigma, \nu)$.

The next theorem gives the celebrated *Yosida-Hewitt decomposition* of the norm-dual space $L^*_\infty(S, \Sigma, \nu)$.

Theorem D.1.11 (Yosida-Hewitt) The adjoint space $L^*_\infty(S, \Sigma, \nu)$ is given by the direct sum

$$L^*_\infty(S, \Sigma, \nu) = L_1(S, \Sigma, \nu) \oplus \mathcal{Q}(S, \Sigma, \nu), \tag{D.13}$$

in the sense that if $Z \in L^*_\infty(S, \Sigma, \nu)$, then there exists unique

$$\begin{aligned} & Z_r \in L_1(S, \Sigma, \nu) \\ \text{and} \qquad & Z^\circ \in \mathcal{Q}(S, \Sigma, \nu), \\ \text{s.t.} \qquad & Z(u) = E[Z_r u] + Z^\circ(u), \end{aligned} \tag{D.14}$$

for all $u \in L_\infty(S, \Sigma, \nu)$.

The decomposition given by Theorem D.1.11 shows that the linear space $\mathcal{Q}(S, \Sigma, \nu)$ complements the space $L_1(S, \Sigma, \nu)$ in giving a full characterization of $L^*_\infty(S, \Sigma, \nu)$.

Definition D.1.12 *The indicator function,* $I_A : S \to \{0,1\}$ for a subset A of S is defined as

$$I_A(x) := \begin{cases} 1, & \text{if } x \in A, \\ 0, & \text{if } x \notin A. \end{cases} \tag{D.15}$$

Remark D.1.13 When $Z \in L_\infty^*(S, \Sigma, \nu)$, the notation $Z \leq 0$ indicates that $Z(u) \leq 0$ for each ν-a.s. non-negative $u \in L_\infty(S, \Sigma, \nu)$.

Remark D.1.14 We shall write $Z \in L_\infty^*(S, \Sigma, \nu)$ in the form $Z = (Z_r, Z^\circ)$ to indicate that (Z_r, Z°) is the unique pair in $L_1(S, \Sigma, \nu) \times \mathcal{Q}(S, \Sigma, \nu)$ for which the representation at (D.14) holds. Following the terminology of [37], Z_r is the *regular part* and Z° is the *singular part* of continuous linear functional Z.

The following simple and useful result on non-positive members of $L_\infty^*(S, \Sigma, \nu)$ (see Remark D.1.13) is typically stated without proof. Since the proof is not completely trivial we include it here for completeness:

Lemma D.1.15 For each $Z = (Z_r, Z^\circ) \in L_\infty^*(S, \Sigma, \nu)$, one has

$$Z \leq 0 \tag{D.16}$$

$$\text{iff}$$

$$Z_r \leq 0 \ \nu - \text{a.s. and } Z^\circ \leq 0. \tag{D.17}$$

Proof. Suppose (D.17) holds.

Then for $u \in L_\infty$, $u \geq 0$, we get

$$\begin{aligned} & Z_r \leq 0 \ \nu - \text{a.s. and } Z^\circ \leq 0. \\ & \Rightarrow E[Z_r] + Z^\circ(u) \leq 0. \\ & \Rightarrow Z(u) \leq 0. \\ & \Rightarrow Z \leq 0. \end{aligned}$$

Hence

$$\text{(D.17)} \Rightarrow \text{(D.16)}. \tag{D.18}$$

Now suppose (D.16) holds($Z \leq 0$).

When $Z^\circ = 0$ we have for all $u \in L_\infty$, $u \geq 0$,

$$Z(u) = E[Z_r u] \leq 0. \tag{D.19}$$

Thus $Z_r \leq 0$-a.s.

Thus (D.16)⇒(D.17) when $Z^\circ = 0$.

Next suppose $Z \leq 0$ and $Z^\circ \neq 0$.

Then there exists $\{A_n\} \subset \Sigma$, s.t.

$$\begin{aligned} A_{n+1} \subset A_n,\ \nu(A_n) > 0,\ \nu(A_n) \to 0, \\ Z^\circ(u) = Z^\circ(uI_{A_n}), \end{aligned} \tag{D.20}$$

for all $u \in L_\infty$, $n = 1, 2, \ldots$

Then for $u \in L_\infty$, $n = 1, 2, \ldots$ we have

$$Z(u) \overset{(D.14)}{=} E[Z_r] + Z^\circ(uI_{A_n}) \tag{D.21}$$

Therefore

$$\begin{aligned} Z(uI_{A_n^c}) &\overset{(D.21)}{=} E[Z_r uI_{A_n^c}] + Z^\circ(uI_{A_n^c} I_{A_n}) \\ &= E[Z_r I_{A_n^c}]. \end{aligned} \tag{D.22}$$

for all $u \in L_\infty$, $n = 1, 2, \ldots$

Since $Z \leq 0$, for $u \in L_\infty \,\&\, u \geq 0$ we get

$$\begin{aligned} & Z(uI_{A_n^c}) \leq 0 \\ \overset{(D.22)}{\Rightarrow}\ & E[Z_r uI_{A_n^c}] \leq 0. \end{aligned} \tag{D.23}$$

Since $Z_r \in L_1$, using the fact that $P(A_n^c) \to 1$ together with Lebesgue dominated convergence theorem we get

$$\lim_{n\to\infty} E[Z_r uI_{A_n^c}] = E[Z_r u], \tag{D.24}$$

for $u \in L_\infty$.

From (D.23) and (D.24) we get

$$\begin{aligned} E[Z_r u] &\leq 0 \quad \text{for all } u \in L_\infty,\ u \geq 0 \\ \text{i.e.} \qquad Z_r &\leq 0 \quad \text{a.s.} \end{aligned} \tag{D.25}$$

Moreover for $u \in L_\infty$, $n = 1, 2, \ldots$

$$\begin{aligned} Z(uI_{A_n}) &= E[Z_r uI_{A_n}] + Z^\circ(uI_{A_n}) \\ &= E[Z_r uI_{A_n}] + Z^\circ(u) \end{aligned} \tag{D.26}$$

Since $Z_r \in L_1$ and $\nu(A_n) \to 0$, from Lebesgue dominated convergence theorem we get

$$\lim_{n\to\infty} E[Z_r u I_{A_n}] = 0, \quad u \in L_\infty \tag{D.27}$$

From (D.26) and (D.27)

$$\lim_{n\to\infty} Z(uI_{A_n}) = Z^\circ(u) \qquad u \in L_\infty. \tag{D.28}$$

Now

$$\begin{aligned}
&u \geq 0 \,\&\, u \in L_\infty \\
&\overset{Z\leq 0}{\Rightarrow} Z(uI_{A_n}) \leq 0, \quad n = 1, 2, \ldots \\
&\overset{\text{(D.28)}}{\Rightarrow} Z^\circ(u) \leq 0. \\
&\text{i.e.} \qquad Z^\circ \leq 0.
\end{aligned} \tag{D.29}$$

From (D.25) and (D.29) we get

$$\text{(D.16)} \Rightarrow \text{(D.17)}. \tag{D.30}$$

Appendix E

Supplement to Chapter 3

In the first section of this appendix we are going to work through an example which illustrates the application of the Rockafellar-Moreau approach in a very simple setting. In the second section we discuss proof of Rockafellar-Moreau Theorem.

E.1 Example: Application of R-M Approach

The example that we discuss in this section is motivated by a work of Rockafellar and Wets [37] on static problems of stochastic convex optimization, and serves as a guide "in miniature" for addressing the stochastic control QLM Problem 2.2.16. In fact the application of the Rockafellar-Moreau approach to the simple example of this section exactly mimics the application of this approach to the much more challenging QLM Problem 2.2.16 (we point out the similarities in the course of the following discussion). The example also demonstrates another central point, namely that singular Lagrange multipliers unavoidably arise in even the simplest optimization problems with almost-sure inequality constraints, and that these singular multipliers are in fact the singular parts of elements of $L_\infty^*(S, \Sigma, \nu)$ (recall Remark D.1.14). Finally, this simple example also illustrates the effectiveness of Theorem 3.1.4 in securing existence of Lagrange multipliers.

Given the function

$$J_1(x) := \frac{1}{2}x^2, \quad x \in \mathbb{R}, \tag{E.1}$$

together with a random variable ς on a probability space (S, Σ, ν), which is uniformly distributed over the unit interval $[0, 1]$.

Define

$$\eta := \inf_{x \geq \varsigma \text{ a.s.}} J_1(x). \tag{E.2}$$

The problem is

$$\text{determine } \bar{x} \in \mathbb{R} \text{ such that } \bar{x} \geq \varsigma \text{ a.s. and } \eta = J_1(\bar{x}). \tag{E.3}$$

Of course this is a completely trivial problem since one sees by inspection that the minimizer is $\bar{x} = 1$ and the corresponding value of the problem is $\eta = 1/2$. Despite the simplicity of this problem the structure of the Lagrange multiplier which corresponds to the a.s. constraint $x \geq \varsigma$ is far from evident. We are going to see that the Rockafellar-Moreau approach establishes that this Lagrange multiplier is in fact a *singular* linear functional on the space $L_\infty(S, \Sigma, \nu)$, and that there are clear intuitive reasons why one gets this "strange" Lagrange multiplier.

In problem (E.3) the space of primal variables $\mathbb{X}_1$ is clearly

$$\mathbb{X}_1 = \mathbb{R}. \tag{E.4}$$

Define the "primal function" $f_1 : \mathbb{X}_1 \to (-\infty, +\infty]$ for problem (E.3) as follows:

$$f_1(x) := \begin{cases} J_1(x), & \text{when } x \geq \varsigma \text{ a.s.,} \\ +\infty, & \text{otherwise,} \end{cases} \tag{E.5}$$

for all $x \in \mathbb{X}_1$. The function $f_1(\cdot)$ is clearly convex on $\mathbb{X}_1$. The primal problem (E.3) amounts to minimization of $f_1(x)$ over all $x \in \mathbb{X}_1$.

Remark E.1.1 The primal function f_1 at (E.5) clearly mimics (but is obviously much simpler than) the primal function at (2.48) for the QLM Problem 2.2.16.

We next implement Step I of the Rockafellar-Moreau approach (recall Section 3.1), that is we define a linear space $\mathbb{U}_1$ of perturbations and a perturbation function $F_1 : \mathbb{X}_1 \times \mathbb{U}_1 \to [-\infty, \infty]$. According to ([36], Example 1 on page 7 and Example 4 on page 8), this is a matter of perturbing the "infinitely many" constraints at (E.5). Define

$$\mathbb{U}_1 := L_\infty(S, \Sigma, \nu) \tag{E.6}$$

$$\text{and} \qquad F_1(x, u) := \begin{cases} J_1(x), & \text{when } x \geq \varsigma + u \text{ a.s.,} \\ +\infty, & \text{otherwise,} \end{cases} \tag{E.7}$$

for all $(x, u) \in \mathbb{X}_1 \times \mathbb{U}_1$. From (E.5) and (E.7) we have the consistency relation of the form (3.3), namely

$$f_1(x) = F_1(x, 0), \; x \in \mathbb{X}_1. \tag{E.8}$$

Remark E.1.2 The choice of perturbation space $\mathbb{U}_1$ and perturbation function F_1 at (E.6) - (E.7) for the primal function (E.5) clearly mimics the perturbation space $\mathbb{U}$ at (4.1) and perturbation function F at (4.2) for the primal function at (2.48). Notice that, at (E.6) - (E.7), we have nothing comparable to the factor space L_2 at (4.1) and there is nothing comparable to the perturbation by members $u_1 \in L_2$ that one has at (4.2). The reason for this is that the very simple primal function f_1 at (E.5) does not feature any constraint analogous to the admissible wealth constraint $X \in \mathbb{D}$ built into the primal function f at (2.48); the role of the variable $u_1 \in L_2$ at (4.2) is precisely to perturb the constraint $X \in \mathbb{D}$, and since there is no analogue of this constraint in the simple primal function (E.5) there is no corresponding perturbation by $u_1 \in L_2$ either. On the other hand, the almost-sure constraint $x \geq \varsigma$ a.s. is obviously very analogous to the constraint $X(T) \geq B$ a.s. at (2.48), and is likewise perturbed by functions which are *essentially bounded*. As we shall see shortly, it is the use of essentially bounded perturbations of almost-sure constraints which enables us to use Theorem 3.1.4 to establish existence of Lagrange multipliers. We emphasize that the need for *essentially bounded perturbations* of almost-sure constraints, in order to be able to use Theorem 3.1.4 to secure existence of Lagrange multipliers, is one of the more profound discoveries of Rockafellar and Wets. In [37] this principle is established for static problems of stochastic convex optimization, and one of the main goals of this thesis has been to generalize this to the dynamic QLM Problem 2.2.16.

We now implement Step II the Rockafellar-Moreau approach (see Section 3.1), that is we define a vector space of *dual variables* $\mathbb{Y}_1$, a duality pairing of the space of perturbations $\mathbb{U}_1$ with the space of dual variables $\mathbb{Y}_1$, a Lagrangian on $\mathbb{X}_1 \times \mathbb{Y}_1$, and a dual function on $\mathbb{Y}_1$. To this end put

$$\begin{aligned} \mathbb{Y}_1 &:= L_\infty^*(S, \Sigma, \nu), && \text{(E.9)} \\ \langle u, z \rangle &:= z(u), \qquad (u, z) \in \mathbb{U}_1 \times \mathbb{Y}_1, && \text{(E.10)} \\ K_1(x, z) &:= \inf_{u \in \mathbb{U}_1} [\langle u, z \rangle + F_1(x, u)] \\ &= J_1(x) + \inf_{\substack{u \in L_\infty \\ u \leq x - \varsigma}} z(u), \qquad (x, z) \in \mathbb{X}_1 \times \mathbb{Y}_1. && \text{(E.11)} \end{aligned}$$

The definition of $K_1(x, z)$ is motivated by (3.4), and the (second) equality at (E.11) follows from (E.7). From Remark D.1.13, the notation $z \leq 0$ (for $z \in \mathbb{Y}_1$) indicates that $z(u) \leq 0$ for all $u \in \mathbb{U}_1$ such that $u \geq 0$. We then have

$$\inf_{\substack{u \in L_\infty \\ u \leq x - \varsigma}} z(u) = \begin{cases} z(x - \varsigma), & \text{when } z \leq 0, \\ -\infty, & \text{when } z \not\leq 0. \end{cases} \qquad \text{(E.12)}$$

Using (E.12) in (E.11) then gives the Lagrangian

$$K_1(x,z) = \begin{cases} z(x-\varsigma) + J_1(x), & \text{when } z \le 0, \\ -\infty, & \text{otherwise,} \end{cases} \tag{E.13}$$

for each $(x,z) \in \mathbb{X}_1 \times \mathbb{Y}_1$.

Next define the dual function on g_1 on $\mathbb{Y}_1$ in accordance with (3.5) namely

$$g_1(z) := \inf_{x\in\mathbb{R}} K_1(x,z) = \begin{cases} -z(\varsigma) - J_1^*(-z(1)), & \text{when } z \le 0, \\ -\infty, & \text{otherwise,} \end{cases} \tag{E.14}$$

for each $z \in \mathbb{Y}_1$. Here

$$\begin{aligned} J_1^*(\alpha) &:= \sup_{x\in\mathbb{R}}[\alpha x - J_1(x)] \\ &= \frac{1}{2}\alpha^2, \ \alpha \in \mathbb{R}, \end{aligned} \tag{E.15}$$

is the usual convex conjugate of $J_1(\cdot)$, and the equality at (E.14) is an immediate consequence of (E.13) together with the elementary identity

$$z(x) = xz(1) \quad (x,z) \in \mathbb{X}_1 \times \mathbb{Y}_1. \tag{E.16}$$

In view of (E.14) the dual function is defined by

$$g_1(z) := \begin{cases} -z(\varsigma) - J_1^*(-z(1)), & \text{when } z \le 0, \\ -\infty, & \text{otherwise,} \end{cases} \tag{E.17}$$

for all $z \in \mathbb{Y}_1$.

In accordance with the discussion at Remark 3.1.3 we must establish that the duality gap is zero and that there exists some $\bar{z} \in \mathbb{Y}_1$ which maximizes g_1 over the space of dual variables $\mathbb{Y}_1$. Actually, it is not at all clear from the expression for g_1 at (E.17) that there exists some $\bar{z} \in \mathbb{Y}_1$ which maximizes g_1. We are now going to use Theorem 3.1.4 to establish that such a maximizer indeed exists (later it will be seen that this maximizer is necessarily a *singular* member of $\mathbb{Y}_1$, that is a member of $\mathcal{Q}(S, \Sigma, \nu)$, see Notation D.1.8). We first establish that the conditions of Theorem 3.1.4 hold:

Fixing some $\tilde{x} \in (1, \infty)$ and $\epsilon \in (0, \tilde{x} - 1)$, it follows from the U[0,1]-distribution of ς that

$$\tilde{x} \ge \epsilon + \varsigma. \tag{E.18}$$

From (E.18) and (E.7) it is immediate that

$$\sup_{\substack{u\in\mathbb{U}_1\\ ||u||_{L_\infty}<\epsilon}} F_1(\tilde{x},u) = J_1(\tilde{x}) < \infty. \tag{E.19}$$

Now the norm-topology $\mathcal{U}$ on $\mathbb{U}_1$ is $\langle \mathbb{U}_1, \mathbb{Y}_1\rangle$-compatible (see (E.6) and (E.9)), thus Theorem 3.1.4 indeed applies, and gives

$$\inf_{x\in\mathbb{R}} f_1(x) = \sup_{z\in\mathbb{Y}_1} g_1(z) = g_1(\bar{z}) \in \mathbb{R} \qquad \text{for some } \bar{z}\in\mathbb{Y}_1. \tag{E.20}$$

Remark E.1.3 The argument that we have just used to establish (E.20) mimics, on the "miniature scale" of problem (E.3), exactly the central ideas that were used to establish Proposition 4.2.3. Of course, it is significantly less technical than the proof of Proposition 4.2.3, but this just mirrors the fact that problem (E.3) (with associated dual function given by (E.17)) is much simpler than the QLM problem 2.2.16 (with associated dual function given by (4.19)). In particular, the choice of the perturbation space $\mathbb{U}_1$ of *essentially bounded* functions at (E.6), together with the norm-topology $\mathcal{U}$ on $\mathbb{Y}_1$, is indispensable for verifying (E.19), which in turn is the key to using Theorem 3.1.4. This is exactly analogous to the central role played by the *essentially bounded* perturbations u_2 in verifying the bound (4.30), which is similarly the key to using Theorem 3.1.4 in the proof of Proposition 4.2.3 (recall Remark 4.2.4). It is worthwhile to note that the Slater Condition 2.2.22 was essential to establishing the bound (4.30) (again see Remark 4.2.4), for it was on the basis of this condition that we obtained a primal variable $\tilde{X}$ such that (4.30) holds. One can reasonably ask why we have not needed to postulate a Slater condition comparable with Condition 2.2.22 when verifying the bound (E.19). The answer of course is that problem (E.3) is so simple that *any* $\tilde{x} > 1$ will suffice for verifying (E.19), as is immediately apparent from (E.18).

We are now going to establish a set of *Kuhn-Tucker* optimality relations between general primal variables $x \in \mathbb{X}_1$ and dual variables $z \in \mathbb{Y}_1$ such that $f_1(x) = g_1(z)$ (see Remark 3.1.5 no.3). That is, we show that the condition $f_1(x) = g_1(z)$ (for arbitrary $(x,z) \in \mathbb{X}_1 \times \mathbb{Y}_1$) is equivalent to some Kuhn-Tucker optimality relations.

To this end observe from (E.5) and (E.17) that, for each $(x,z) \in \mathbb{X}_1 \times \mathbb{Y}_1$, we have

$$f_1(x) \in (-\infty,\infty] \tag{E.21}$$

$$\text{and} \qquad g_1(z) \in [-\infty,\infty). \tag{E.22}$$

Exactly as in Remark 3.1.3 it follows from (E.8) and (E.17) that

$$f_1(x) \geq K_1(x,z) \geq g_1(z). \tag{E.23}$$

In light of (E.21), (E.22) and (E.23) we then get

$$f_1(x) \;=\; g_1(z) \tag{E.24}$$

$$\Updownarrow$$

$$f_1(x) = K_1(x,z) \in \mathbb{R} \quad \& \quad g_1(z) = K_1(x,z) \in \mathbb{R}, \tag{E.25}$$

for each $(x,z) \in \mathbb{X}_1 \times \mathbb{Y}_1$.

In view of (E.13) and (E.17) we see that

$$g_1(z) = K_1(x,z) \in \mathbb{R} \tag{E.26}$$

$$\Updownarrow$$

$$\begin{cases} (1) & z \leq 0, \\ (2) & J_1(x) + J_1^*(-z(1)) = -xz(1) \end{cases} \tag{E.27}$$

$$\Updownarrow$$

$$\begin{cases} (1') & z \leq 0, \\ (2') & x = (\partial J_1^*)(-z(1)), \end{cases} \tag{E.28}$$

where $\partial J_1^*(\cdot)$ is the derivative of $J_1^*(\cdot)$ and the equivalence of (2) of (E.27) and (2′) of (E.28) follows from Lemma B.4.15 (with $N = 1$).

Moreover, from (E.5) and (E.13) we also have

$$f_1(x) = K_1(x,z) \in \mathbb{R} \tag{E.29}$$

$$\Updownarrow$$

$$\begin{cases} (1) & x \geq \varsigma \\ (2) & z \leq 0 \\ (3) & z(x - \varsigma) = 0. \end{cases} \tag{E.30}$$

Combining (E.24), (E.28) and (E.30), we then obtain

$$f_1(x) = g_1(z) \tag{E.31}$$

$$\Updownarrow$$

$$\begin{cases} (1) & x \geq \varsigma \text{ a.s.}, \\ (2) & z \leq 0, \\ (3) & z(x - \varsigma) = 0, \\ (4) & x = (\partial J_1^*)(-z(1)), \end{cases} \tag{E.32}$$

for each $(x,z) \in \mathbb{X}_1 \times \mathbb{Y}_1$.

Remark E.1.4 Items (1)-(4) of (E.32) are *Kuhn-Tucker optimality relations.* In particular (E.32)(1)-(2) are *feasibility conditions* on the primal and dual variables x and z respectively . On the other hand, (E.32)(3) relates the primal variable x, the dual variable z and the constraint $x \geq \varsigma$ a.s. of the primal problem (E.3), and is therefore a *complementary slackness condition.* Finally, (E.32)(4) relates the primal variable x, the dual variable z and the cost functional J_1 of the primal problem (E.3), and is therefore a *transversality condition.*

Remark E.1.5 There are clear similarities between the relations (E.32)(1)-(4) established for the primal problem (E.3) and the Kuhn-Tucker relations (4.36)-(4.39) that were established at Proposition 4.2.8 for the QLM Problem 2.2.16. On comparing (1)-(4) of (E.32) with Equations (4.36)-(4.39) we see that condition similar to $X \in \mathbb{D}_1$ (in (4.36)) does not appear in (E.32). We also do not have condition given by Equation (4.38) in (E.32). These conditions do not appear in (E.32) because we do not have any constraint similar to portfolio constraint of the form $X \in \mathbb{D}_1$ (see (4.13)) built into primal function given by (E.5) and (4.38) is a complementary slackness relation for the portfolio constraint $X \in \mathbb{D}$ (see (2.47)).

It remains to construct an $\bar{x} \in \mathbb{R}$ in terms of the dual maximizer $\bar{z}$ given by (E.20) such that $(\bar{x}, \bar{z})$ satisfies (1)-(4) of (E.32), for then $f_1(\bar{x}) = g_1(\bar{z})$ and thus $\bar{x}$ solves the primal problem. Although it is immediate from problem (E.3) that $\bar{x} = 1$, we shall nevertheless give the construction of $\bar{x}$ in terms of $\bar{z}$ in some detail. We do this because the construction closely mimics the essential features of the construction in Section 4.3 of the optimal wealth process $\bar{X}$ in terms of the dual maximizer $(\bar{Y}, \bar{Z}) \in \mathbb{Y}$ (see Proposition 4.2.3) such that $(\bar{X}, (\bar{Y}, \bar{Z}))$ satisfies the optimality relations (4.36) - (4.39) of Proposition 4.2.8. However, in the present instance, this construction is in the very simple setting of the problem (E.3), and therefore involves much less technical effort than is required in Section 4.3, in which we must construct a stochastic process $\bar{X}$ instead of just a scalar $\bar{x}$.

To construct $\bar{x}$ we first obtain *necessary conditions* which result from the fact that $\bar{z} \in \mathbb{Y}_1$ maximizes the dual function g_1 (much as we obtained necessary conditions resulting from the optimality of $(\bar{Y}, \bar{Z})$ in the construction of $\bar{X}$, see Proposition 4.3.7, Proposition 4.3.22, and Proposition 4.3.25). Motivated by (4) of (E.32) define

$$\bar{x} := (\partial J_1^*)(-\bar{z}(1)). \tag{E.33}$$

From (E.5) we have

$$\inf_{x \in \mathbb{R}} f_1(x) > -\infty. \tag{E.34}$$

It then follows from (E.17) and (E.20) that

$$\bar{z} \leq 0. \tag{E.35}$$

Also from (E.20), we have the optimality relation

$$\frac{1}{\epsilon}[g_1(\bar{z}) - g_1(\bar{z} + \epsilon z)] \geq 0, \tag{E.36}$$

for all $\epsilon > 0$ and $z \in \mathbb{Y}_1$.

Fix some $z \in \mathbb{G}_1 := \{z \in \mathbb{Y}_1 : z \leq 0\}$; then, from (E.35) we find that $\bar{z} + \epsilon z \leq 0$ for $\epsilon > 0$, and therefore from (E.17) and (E.36) it follows that

$$\frac{1}{\epsilon}[\partial J_1^*(-\bar{z}(1) - \epsilon z(1)) - \partial J_1^*(-\bar{z}(1))] + z(\varsigma) \geq 0, \tag{E.37}$$

for all $\epsilon > 0$.

Taking $\epsilon \to 0$ at (E.37) and recalling (E.33), we obtain

$$z(\varsigma - \bar{x}) \geq 0 \text{ for all } z \in \mathbb{G}_1, \tag{E.38}$$

and thus

$$\bar{x} \geq \varsigma \quad \text{a.s.} \tag{E.39}$$

In view of (E.35) and (E.39) we have verified (1) and (2) of (E.32), and it remains to verify the relation (3) of (E.32). From (E.35) we have that

$$(1 - \epsilon)\bar{z} \leq 0, \tag{E.40}$$

for all $\epsilon \in (0, 1)$.

Upon taking $z = -\bar{z}$ and using (E.40), we get the following inequality from (E.36)

$$\frac{1}{\epsilon}[\partial J_1^*(-(1 - \epsilon)\bar{z}(1)) - \partial J_1^*(-\bar{z}(1))] - z(\varsigma) \geq 0 \tag{E.41}$$

Taking $\epsilon \to 0$ and using (E.33), we obtain

$$\bar{z}(\bar{x} - \varsigma) \geq 0. \tag{E.42}$$

and thus from (E.35), (E.39) and (E.42) we have

$$\bar{z}(\bar{x} - \varsigma) = 0, \tag{E.43}$$

which verifies (3) of (E.32).

From (E.33), (E.35), (E.39), (E.43) we observe that Kuhn-Tucker optimality relations given by (1)-(4) of (E.32) are satisfied for $(\bar{x}, \bar{z}) \in \mathbb{R} \times \mathbb{Y}_1$. Hence from (E.31) and (E.32) we get

$$f_1(\bar{x}) = g_1(\bar{z}). \tag{E.44}$$

Thus $\bar{x}$ defined at (E.33) in terms of $\bar{z}$ minimizes $f_1(\cdot)$, and therefore solves problem (E.3).

Observe that Lagrange multiplier $\bar{z}$ is a *singular* element of $\mathbb{Y}_1$. In fact, with $\bar{z} = (\bar{z}_r, \bar{z}^\circ)$, from (E.35) and Remark D.1.15, we obtain

$$\bar{z}_r \leq 0 \,\text{a.s. and}\, \bar{z}^\circ \leq 0, \tag{E.45}$$

while (E.43) gives

$$E[\bar{z}_r(\bar{x} - \varsigma)] + \bar{z}^\circ(\bar{x} - \varsigma) = 0. \tag{E.46}$$

From (E.45) and (E.39) we get

$$\bar{z}_r(\bar{x} - \varsigma) \leq 0 \text{ a.s. and } \bar{z}^\circ(\bar{x} - \varsigma) \leq 0, \tag{E.47}$$

and combining this with (E.46) gives

$$\bar{z}_r(\bar{x} - \varsigma) = 0 \text{ a.s. and } \bar{z}^\circ(\bar{x} - \varsigma) = 0. \tag{E.48}$$

Now by inspection we actually know that $\bar{x} = 1$, thus in fact $\bar{x} > \varsigma$ a.s. (since ς is $U[0,1]$-distributed), whence (E.48) gives

$$\bar{z}_r = 0 \quad \text{a.s.} \tag{E.49}$$

Moreover, using $\bar{x} = 1$ together with (E.49), (E.15) and (E.33) we observe that

$$\begin{aligned} 1 = \bar{x} &= \partial J_1^*(-\bar{z}(1)) \\ &= -\bar{z}(1) \\ &= -\bar{z}^\circ(1) \end{aligned} \tag{E.50}$$

The multiplier $\bar{z}$ is therefore non-trivial and a *singular* member of $\mathbb{Y}_1$, that is

$$\bar{z} = \bar{z}^\circ \in \mathcal{Q}(S, \Sigma, \nu). \tag{E.51}$$

Remark E.1.6 The constraint $x \geq \varsigma$ a.s. in the problem (E.3) is certainly active, in the sense that removal of the constraint from the problem leads to an unconstrained problem with strictly smaller optimal value

$$\eta := \inf_{x \in \mathbb{R}} J_1(x) = 0, \tag{E.52}$$

(c.f. (E.2)), but nevertheless is active *only on a set of zero probability* (since $\bar{x} = 1 > \varsigma$ a.s. from the $U[0,1]$-distribution of ς). An essential general insight of Rockafellar and Wets [37] in the study of static problems of stochastic convex optimization is that such "singularly binding" constraints necessarily lead to singular Lagrange multipliers, and we see this insight clearly evident in the case of problem (E.3).

E.2 Proof of the Rockafellar-Moreau Theorem 3.1.4

The Rockafellar-Moreau theorem is established in Rockafellar [36] in extremely compressed and bare-bones form. Since this result is so important in this thesis, for completeness we shall now give a more expanded proof along the lines of the proof in [36] but with all relevant details included.

Definition E.2.1 For the convex perturbation function $F : \mathbb{X} \times \mathbb{U} \to [-\infty, \infty]$, define the perturbed value function $\varphi : \mathbb{U} \to [-\infty, \infty]$ as follows:

$$\varphi(u) := \inf_{x \in \mathbb{X}} F(x, u), \qquad u \in \mathbb{U}. \tag{E.53}$$

From (3.3), we have

$$\varphi(0) = \inf_{x \in \mathbb{X}} f(x). \tag{E.54}$$

Lemma E.2.2 Recall Definition B.4.9. For the perturbed value function $\varphi : \mathbb{U} \to [-\infty, \infty]$, given by Definition E.2.1 and the dual function $g : \mathbb{Y} \to [-\infty, \infty]$ given by ((3.5)), one has

$$-\varphi^*(-y) \;\; = \;\; g(y). \tag{E.55}$$

Proof: From (E.53) and (B.45),

$$\begin{aligned}\varphi^*(y) &= \sup_{u\in\mathbb{U}}[\langle u,y\rangle - \inf_{x\in\mathbb{X}} F(x,u)] \\ &= \sup_{u\in\mathbb{U}}[\langle u,y\rangle + \sup_{x\in\mathbb{X}}(-F(x,u))] \\ &= \sup_{u\in\mathbb{U}}\sup_{x\in\mathbb{X}}[\langle u,y\rangle - F(x,u)] \\ &= \sup_{x\in\mathbb{X}}\sup_{u\in\mathbb{U}}[\langle u,y\rangle - F(x,u)].\end{aligned}$$

Now, $\mathbb{Y}$ is a vector space, therefore Replacing y with $-y$, we get

$$\varphi^*(-y) = \sup_{x\in\mathbb{X}}\sup_{u\in\mathbb{U}}[\langle u,-y\rangle - F(x,u)].$$

Using the definition of $K(x,y)$(3.4) and $g(y)$(3.5), we find that

$$\begin{aligned}-\varphi^*(-y) &= \inf_{x\in\mathbb{X}}\inf_{u\in U}[\langle u,y\rangle + F(x,u)] \\ &= \inf_{x\in\mathbb{X}} K(x,y) = g(y). \\ \text{i.e.} \quad -\varphi^*(-y) &= g(y), \qquad y\in\mathbb{Y}. \qquad (E.56)\end{aligned}$$

Recall the Definitions B.4.9, B.4.11 and E.2.1, then we have

$$\begin{aligned}\varphi^{**}(0) &= \sup_{y\in\mathbb{Y}}[-\varphi^*(y)] \\ &\overset{\mathbb{Y}\text{ is a vector space}}{=} \sup_{y\in\mathbb{Y}}[-\varphi*(-y)] \\ &\overset{\text{Lemma } E.2.2}{=} \sup_{y\in\mathbb{Y}} g(y). \qquad (E.57)\end{aligned}$$

We next state without proof the following results on convex analysis from Rockafellar [36].

Theorem E.2.3 Recall Definition B.4.9, B.4.11 and B.4.13. Suppose that $(U, Y, \langle\cdot,\cdot\rangle)$ is a given dual system and $\psi : U \to [-\infty,\infty]$ is a given function (not necessarily convex). We then have the following:

(a) If $\psi(u)\in\mathbb{R}$ and $\partial\psi(u)\neq\emptyset$ at some $u\in U$ then $\psi(u)=\psi^{**}(u)$,

(b) If $\psi(0)=\psi^{**}(0)\in R$ then
$\partial\psi(0) = \{y\in Y \mid \psi^*(y')\geq\psi^*(y) \ \textit{all} \ \ y'\in Y\}$.

Lemma E.2.4 For the perturbed value function $\varphi : \mathbb{U} \to [-\infty, \infty]$, given by Definition E.2.1 and the dual function $g : \mathbb{Y} \to [-\infty, \infty]$ given by ((3.5)),

$$\text{if } \partial\varphi(0) \neq \emptyset, \tag{E.58a}$$

then

$$\inf_{x \in \mathbb{X}} f(x) = \sup_{y \in \mathbb{Y}} g(y) = g(\hat{y}) \text{ for some } \hat{y} \in \mathbb{Y}. \tag{E.58b}$$

Proof:
Suppose

$$\partial\varphi(0) \neq \emptyset. \tag{E.59}$$

From (E.59) and Theorem E.2.3

$$\varphi(0) = \varphi^{**}(0). \tag{E.60}$$

From (E.59) and Theorem E.2.3,

$$\partial\varphi(0) = \{y \in \mathbb{Y} | \varphi^*(y') \geq \varphi^*(y) \, \forall y' \in \mathbb{Y}\}. \tag{E.61}$$

From (E.60),(E.54) and (E.57),

$$\inf_{x \in \mathbb{X}} f(x) = \sup_{y \in \mathbb{Y}} g(y). \tag{E.62}$$

Moreover,

$$\begin{aligned}
\hat{y} \in -\partial\varphi(0) \quad & \Rightarrow \quad -\hat{y} \in \partial\varphi(0). \\
\therefore \quad \varphi^*(y') \quad & \overset{(E.61)}{\geq} \quad \varphi^*(-\hat{y}) \quad \forall y' \in \mathbb{Y}. \\
\therefore \quad \varphi^*(-y') \quad & \geq \quad \varphi^*(-\hat{y}) \quad \forall y' \in \mathbb{Y}. \\
\therefore \quad -\varphi^*(-y') \quad & \leq \quad -\varphi^*(-\hat{y}) \quad \forall y' \in \mathbb{Y}. \\
\therefore \quad g(y') \quad & \overset{\text{from Lemma (E.2.2)}}{\leq} \quad g(\hat{y}) \quad \forall y' \in \mathbb{Y}. \\
\text{i.e.} \quad g(\hat{y}) \quad & = \quad \sup_{y \in \mathbb{Y}} g(y) \text{ for each } \hat{y} \in -\partial\varphi(0).
\end{aligned} \tag{E.63}$$

From (E.63), (E.62) and (E.59),

$$\text{if } \partial\varphi(0) \neq \emptyset \text{ then } \inf_{x \in \mathbb{X}} f(x) = \sup_{y \in \mathbb{Y}} g(y) = g(\hat{y}) \text{ for some } \hat{y} \in \mathbb{Y}. \tag{E.64}$$

It therefore remains to establish sufficient conditions which ensure that $\partial\varphi(0) \neq \emptyset$ holds in order to use the previous result. This calls upon two results from convex analysis: From Rockafellar ([36], Theorem 8) we have

Theorem E.2.5 Suppose that $(U, Y, \langle\cdot,\cdot\rangle)$ is a given dual system and $\mathcal{U}$ is a $\langle\cdot,\cdot\rangle$-compatible topology on U. Suppose that $\psi : U \to \bar{\mathbb{R}}$ is convex, and $\psi(0) \in \mathbb{R}$. If there exists some $\mathcal{U}$-neighborhood G of $0 \in U$ such that $\sup_{u\in G} \psi(u) < \infty$ then ψ is $\mathcal{U}$-continuous at $u = 0$.

From Rockafellar ([36], Theorem 11) we also have

Theorem E.2.6 Suppose that $(U, Y, \langle\cdot,\cdot\rangle)$ is a given dual system and $\mathcal{U}$ is a $\langle U, Y\rangle$-compatible topology on U. Suppose that $\psi : U \to \bar{\mathbb{R}}$ is convex and $\psi(0) \in \mathbb{R}$. If ψ is $\mathcal{U}$-continuous at $u = 0$ then $\partial\psi(0) \neq \emptyset$.

On combining Theorem E.2.5 and Theorem E.2.6, we get the next result which gives sufficient conditions to ensure that $\partial\varphi(0) \neq \emptyset$:

Corollary E.2.7 Suppose that $(U, Y, \langle\cdot,\cdot\rangle)$ is a given dual system and $\mathcal{U}$ is a $\langle U, Y\rangle$-compatible topology on U. Suppose that $\psi : U \to \bar{\mathbb{R}}$ is convex and $\psi(0) \in \mathbb{R}$. If there exists some $\mathcal{U}$-neighborhood G of $0 \in U$ such that $\sup_{u\in G} \psi(u) < \infty$ then $\partial\psi(0) \neq \emptyset$.

Using Corollary E.2.7 we are finally able to establish the Rockafellar-Moreau theorem:

Proof of Theorem 3.1.4: From (E.53),

$$\varphi(u) \le F(x_1, u) \qquad \text{all}\, u \in \mathbb{U}. \tag{E.65}$$

Then

$$\sup_{u\in G} \varphi(u) \le \sup_{u\in G} F(x_1, u) < \infty. \tag{E.66}$$

From Corollary E.2.7 with $\phi := \varphi$, we get

$$\partial\varphi(0) \neq 0. \tag{E.67}$$

Result follows from (E.64).

Appendix F

Integrals which are Convex Functionals

In this section we examine integrals having certain convexity properties which can be analyzed in the light of the theory of conjugate convex functions. The results in this section are applicable in the study of problems in control theory, as well as for dealing with integrals of convex or concave loss functions. The arguments in this section are taken from R.T. Rockafellar [33].

F.1 Normal Integrands

Let T denote a measure space with σ-finite measure μ. Suppose that L is a designated real vector space of measurable functions u from T to $\mathbb{R}^n$, and L^* is another designated real vector space of measurable functions from T to $\mathbb{R}^n$ with the property that the inner product $(u(t))'u^*(t)$ of the $\mathbb{R}^N$-vectors $u(t)$ and $u^*(t)$ gives a summable function of t for every $u \in L$ and $u^* \in L^*$. It is then immediate that

$$\langle u, u^* \rangle = \int_T (u(t))'u^*(t)\,\mathrm{dt}, \tag{F.1}$$

for $u \in L$, $u^* \in L^*$, defines a duality pairing between L and L^*, that is $(L, L^*, \langle \cdot, \cdot \rangle)$ is a dual system.

Quite often we need to study functionals of the form

$$I_f(u) = \int_T f(t, u(t))\,\mathrm{dt}, \quad u \in L, \tag{F.2}$$

where $f : T \times \mathbb{R}^n \to (-\infty, \infty]$ is a *normal convex integrand* in the following sense:

Definition F.1.1 The function $f : T \times \mathbb{R}^n \to (-\infty, \infty]$ is a *normal convex integrand* when $f(t, x)$ is proper, convex and lower semi-continuous in $x \in \mathbb{R}^n$ for each $t \in T$, and if further there exists a *countable* collection U of measurable functions u from T to $\mathbb{R}^n$ having the following properties:

1. for each $u \in U$, $f(t, u(t))$ is measurable in t;
2. for each t, $U_t \cap \mathrm{dom} f_t$ is dense in $\mathrm{dom} f_t$, where

$$U_t = \{u(t) | u \in U\}. \tag{F.3}$$

The proof of Lemma F.1.2 and F.1.3 below can be found in Rockafellar,[33], page 528.

Lemma F.1.2 Suppose $f(t, x) = F(x)$ for all t, where F is a lower semi-continuous proper convex function from $\mathbb{R}^n$ to $(-\infty, \infty]$. Then f is a normal convex integrand.

Lemma F.1.3 Suppose f is a convex integrand such that $f(t, x)$ is measurable in t for each fixed x, and such that, for each t, $f(t, x)$ is lower semi-continuous in x and has interior points in its effective domain $\{x | f(t, x) < +\infty\}$. Then f is a normal convex integrand.

F.2 Conjugate Convex Integrals.

Considering the spaces L and L^* defined in the beginning of this appendix, the main question treated in this section is whether the conjugate of a convex functional $I_f : L \to (-\infty, \infty]$ (see (F.2)) is in the form of a functional $I_g : L^* \to (-\infty, \infty]$ for some normal convex integrand $g : T \times \mathbb{R}^n \to (-\infty, \infty]$. The question is significant, because this theory of convex functions is concerned with conjugates.

Suppose that $f(t, x)$ is a convex integrand which is proper and lower semi-continuous in x for each t. Define $f^*(t, x^*)$ by taking conjugates in $x \in \mathbb{R}^n$ for each t, that is

$$f^*(t, x^*) = \sup_{x \in \mathbb{R}^n} [\langle x, x^* \rangle - f(t, x)], \quad x^* \in \mathbb{R}^n.$$

Then f^* is another convex integrand, proper and lower semi-continuous in its convex argument. We call it the *integrand conjugate to* f. The conjugate of the conjugate is the original integrand f.

The principal fact brought out in Theorem F.2.3 is that conjugate integrands f and f^* usually furnish conjugate functionals of L and L^*.

Definition F.2.1 We shall say that the vector space L is *decomposable* when it satisfies the following conditions:

1. L contains every bounded measurable function from T to R^n which vanishes outside a set of finite measure;
2. if $u \in L$ and E is a set of finite measure in T, then L contains $I_E \cdot u$, where I_E is the indicator function of E.

An identical definition of decomposability holds for the vector space L^*.

These conditions guarantee that one can alter functions in L arbitrarily in a bounded manner on sets of finite measure. (subtract $I_E \cdot u$ from u, and then add any bounded measurable function vanishing outside E).

Remark F.2.2 It can be shown that the vector spaces $L_p(S, \Sigma, \nu)$, $p \in [1, \infty]$ are decomposable, and in fact these are the most important cases of decomposable vector spaces. On the other hand, if T is a compact Hausdorff space carrying a finite Borel measure μ on its Borel σ-algebra then the space $C(T)$ of all real-valued continuous functions on T is definitely *not* decomposable.

From Rockafellar,[33], Theorem 2, page 532, we have the following.

Theorem F.2.3 Let spaces L and L^* defined in the beginning of Section F.1 be decomposable. Let $f : T \times \mathbb{R}^n \to (-\infty, \infty]$ be a normal convex integrand such that $f(t, u(t))$ is summable in t for at least one $u \in L$, and $f^*(t, u^*(t))$ is summable in t for at least one $u^* \in L^*$. Then

$$\sup_{u \in L} \left[\int_T (u(t))' u^*(t)\, \mathrm{dt} - \int_T f(t, u(t))\, \mathrm{dt} \right] = \int_T f^*(t, u^*(t)) \mathrm{dt}, \tag{F.4}$$

for each $u^* \in L^*$.

Remark F.2.4 Theorem F.2.3 really says formally that the outside supremum over $u \in L$ on the left can be shifted to *after* the integral over T and becomes a supremum over $x \in \mathbb{R}^n$, that is

$$\begin{aligned} \sup_{u \in L} \left[\int_T [(u(t))' u^*(t)\, \mathrm{dt} - f(t, u(t))]\, \mathrm{dt} \right] &= \int_T \sup_{x \in \mathbb{R}^n} [\langle x, u^*(t) \rangle - f(t, x)] \mathrm{dt} \\ &= \int_T f^*(t, u^*(t)) \mathrm{dt}, \end{aligned}$$

for each $u^* \in L^*$. In this result it is absolutely essential that spaces L together with L^* be decomposable.

Appendix G

The canonical martingales of the Markov chain

The basic market model is formulated in terms of the joint filtration of an independent Brownian motion $\mathbf{W}$ and continuous time finite state Markov chain α defined on the common probability space $(\Omega, \mathcal{F}, \mathbb{P})$ over the interval $t \in [0, T]$ (see Section 2.1). The goal of this appendix is to summarize basic properties of finite state Markov chains needed for the thesis, and to indicate that the basic results on Markov chains, in particular Dynkin's formula, trivially extend from the self-filtration of the Markov chain to the joint filtration of the Markov chain and the Brownian motion. We also state martingale representation theorem for square integrable martingales and locally square integrable martingales since it is used in some of the results.

G.1 The finite state space Markov chain

Throughout this thesis the basic filtration is defined by

$$\mathcal{F}_t := \mathcal{F}_t^{\circ} \vee \mathcal{N}(\mathbb{P}), \quad \forall t \in [0, T], \tag{G.1}$$

in which

$$\mathcal{F}_t^{\circ} := \sigma\{\mathbf{W}(s), \alpha(s) : s \in [0, t]\} \quad \forall t \in [0, T]. \tag{G.2}$$

(see (2.9) and (2.8)), and we define the augmented self-filtration of the Markov chain α in the usual way namely

$$\mathcal{F}_t^{\alpha} := \sigma\{\alpha(s) : s \in [0, t]\} \vee \mathcal{N}(\mathbb{P}), \quad \forall t \in [0, T]. \tag{G.3}$$

Since α and $\mathbf{W}$ are independent, for any $\mathcal{F}_T^\alpha$-measurable and integrable random variable ξ on $(\Omega, \mathcal{F}, \mathbb{P})$ we have, by elementary properties of conditioning and independence, that

$$E(\xi|\mathcal{F}_t) = E(\xi|\mathcal{F}_t^\alpha), \qquad \text{for all } t \in [0,T]. \tag{G.4}$$

We recall from Section 2.1 that the continuous time finite state Markov chain $\alpha = \{\alpha(t) : t \in [0,T]\}$ defined on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ takes values in the finite state space

$$I = \{1, 2, \ldots, D\} \tag{G.5}$$

and that the Markov chain α is always assumed to start in a non-random initial state $i_0 \in I$, so that

$$\alpha(0) = i_0 \quad a.s. \tag{G.6}$$

The Markov property of α is expressed by the usual relation

$$E[f(\alpha(t))|\mathcal{F}_s^\alpha] = E[f(\alpha(t))|\alpha(s)], \qquad \text{for all } 0 \leq s \leq t \leq T, \tag{G.7}$$

for every function $f : I \to \mathbb{R}$. In view of (G.4) we can write this as

$$E[f(\alpha(t))|\mathcal{F}_s] = E[f(\alpha(t))|\alpha(s)], \qquad \text{for all } 0 \leq s \leq t \leq T. \tag{G.8}$$

Associated with the Markov chain α is a generator Q which is a $D \times D$ matrix $Q = (q_{ij})_{i,j=1}^D$ with the properties

$$q_{i,j} \geq 0, \quad \forall i \neq j \quad and \quad -q_{ii} = \sum_{j \neq i} q_{ij}. \tag{G.9}$$

Definition G.1.1 The Markov transition function $\{P_t\}$ on I is defined in terms of generator Q as

$$P_t := \exp\{tQ\} \quad \forall t \in [0,T]. \tag{G.10}$$

In particular, P_0 is the $D \times D$ identity matrix.

Remark G.1.2 From Rogers and Williams [32], equation IV.21.11, the Markov chain α makes finitely many jumps in the finite time interval $[0,T]$. Thus the Lebesgue measure of the set of times where $\alpha(t) \neq \alpha(t_-)$ is zero and this observation will allow us to write, for example,

$$\int_0^T f(\alpha(s_-))\,\mathrm{d}s = \int_0^T f(\alpha(s))\,\mathrm{d}s \tag{G.11}$$

for any Borel measurable integrable function $f : I \to \mathbb{R}$.

Proposition G.1.3 The Markov property of the Markov chain α can be formulated as follows: for all functions $f : I \to \mathbb{R}$, for all $0 \le s \le t \le T$, one has

$$E(f(\alpha(t))|\mathcal{F}_s) = (P_{t-s}f)(\alpha(s)), \tag{G.12}$$

where $(P_t f)(j) = \sum_{k\in I} P_t(j,k)f(k)$.

Proof. First we prove the case when f is an indicator function.
Suppose $f = I_j$, then

$$\begin{aligned} E(f(\alpha(t))|\mathcal{F}_s) &= E(I_j(\alpha(t))|\mathcal{F}_s) \\ &= P\left(I_j(\alpha(t)) = 1|\mathcal{F}_s\right) \\ &= P\left(\alpha(t) = j|\mathcal{F}_s\right) \\ &= P\left(\alpha(t) = j|\alpha(s)\right) \quad \text{(see (G.8))} \\ &= P_{t-s}(\alpha(s), j). \end{aligned} \tag{G.13}$$

$$\begin{aligned} (P_{t-s}f)(\alpha(s)) &= (P_{t-s}I_j)(\alpha(s)) \\ &= \sum_{k\in I} P_{t-s}(\alpha(s),k)I_j(k) \\ &= P_{t-s}(\alpha(s), j). \end{aligned} \tag{G.14}$$

From (G.13) and (G.14) the relation (G.12) holds when f is an indicator function. When f is not an indicator function, f can be written as

$$f(\alpha(s)) = \beta_j I_j(\alpha(s)) = \begin{cases} \beta_j, & \text{when}\, \alpha(s) = j, \\ 0, & \text{otherwise.} \end{cases} \tag{G.15}$$

$$\therefore \quad f(\alpha(s)) = \sum_{j\in I} \beta_j I_j(\alpha(s)). \tag{G.16}$$

$$\begin{aligned} (P_{t-s}f)(\alpha(s)) &= \sum_{k\in I} P_{t-s}(\alpha(s),k)f(k) \\ &= \sum_{k\in I} P_{t-s}(\alpha(s),k) \sum_{k\in I} \beta_j I_j(k) \\ &= \sum_{k\in I} P_{t-s}(\alpha(s),k)\beta_k. \end{aligned} \tag{G.17}$$

$$
\begin{aligned}
E[f(\alpha(t))|\mathcal{F}_s] &= E\left[\sum_{j\in I}\beta_j I_j(\alpha(t))|\mathcal{F}_s\right] \\
&= \sum_{j\in I}\beta_j E\left[I_j(\alpha(t))|\mathcal{F}_s\right] \\
&= \sum_{j\in I}\beta_j P(I_j(\alpha(t))=1|\mathcal{F}_s) \\
&= \sum_{j\in I}\beta_j P(\alpha(t)=j|\mathcal{F}_s) \\
&= \sum_{j\in I}\beta_j P_{t-s}(\alpha(s),j). \qquad (G.18)
\end{aligned}
$$

Hence from (G.17) and (G.18), the result holds for any $f : I \to \mathbb{R}$.

Proposition G.1.4

$$\frac{\mathrm{d}}{\mathrm{d}t}(P_t) = QP_t = P_t Q. \qquad (G.19)$$

Proof. See Norris [29], Theorem 2.1.1.

The following Proposition is a simple consequence of Proposition G.1.4.

Proposition G.1.5 For every $f : I \to \mathbb{R}$ and $t \in [0,T]$ we have

$$\frac{\mathrm{d}}{\mathrm{d}t}(P_t f) = QP_t f, \qquad (G.20)$$

where $(P_t f)(j) = \sum_{j\in I} P_t(j,k)f(k)$ and $(QP_t f)(i) = \sum_{j\in I} q_{ij}(P_t f)(j)$.

Proof.

$$
\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}((P_t f)(j)) &= \sum_{k\in I}\frac{\mathrm{d}}{\mathrm{d}t}(P_t(j,k)f(k)) \\
&\overset{(G.19)}{=} \sum_{k\in I}(QP_t)(j,k)f(k) = (QP_t f)(j).
\end{aligned} \qquad (G.21)
$$

Theorem G.1.6 Dynkin's formula. If $f : I \to \mathbb{R}$ then

$$f(\alpha(t)) - f(\alpha(0)) - \int_0^t (Qf)(\alpha(s))\,\mathrm{d}s \in \mathcal{M}_0((\Omega, \mathcal{F}, \mathbb{P}), \mathcal{F}_t). \tag{G.22}$$

Proof. Fix a function $f : I \to \mathbb{R}$ and define

$$M(t) := f(\alpha(t)) - f(\alpha(0)) - \int_0^t (Qf)(\alpha(s))\,\mathrm{d}s. \tag{G.23}$$

By the boundedness of the function f, $M(t)$ is integrable for all $t \in [0, T]$.

We first show that the martingale property holds for M. First note that upon integrating (G.20), we obtain for all $t \in [0, T]$ and each $i \in I$,

$$(P_t f)(i) - f(i) = \int_0^t (QP_u f)(i)\,\mathrm{d}u. \tag{G.24}$$

From (G.23) it follows that for $0 \leq s \leq t \leq T$, we have

$$\begin{aligned}
E(M(t) - M(s)|\mathcal{F}_s) &= E(f(\alpha(t)) - f(\alpha(s))|\mathcal{F}_s) - E\left(\int_s^t (Qf)(\alpha(u))\,\mathrm{d}u|\mathcal{F}_s\right) \\
&\overset{\text{(G.12)}}{=} (P_{t-s} f)(\alpha(s)) - f(\alpha(s)) - E\left(\int_s^t (Qf)(\alpha(u))\,\mathrm{d}u|\mathcal{F}_s\right) \\
&\overset{\text{(G.24)}}{=} \int_0^{t-s} (QP_u f)(\alpha(s))\,\mathrm{d}u - E\left(\int_s^t (Qf)(\alpha(u))\,\mathrm{d}u|\mathcal{F}_s\right).
\end{aligned}$$

Applying Fubini's theorem for conditional expectations (see Ethier and Kurtz[12], Chapter 2 Proposition 4.6 and Remark 4.7), we obtain

$$\begin{aligned}
E(M(t) - M(s)|\mathcal{F}_s) &= \int_0^{t-s} (QP_u f)(\alpha(s))\,\mathrm{d}u - \int_s^t E((Qf)(\alpha(u))|\mathcal{F}_s)\,\mathrm{d}u \\
&\overset{\text{(G.12)}}{=} \int_0^{t-s} (QP_u f)(\alpha(s))\,\mathrm{d}u - \int_s^t (P_{u-s} Qf)(\alpha(s))\,\mathrm{d}u \\
&\overset{\text{(G.19)}}{=} \int_0^{t-s} (QP_u f)(\alpha(s))\,\mathrm{d}u - \int_s^t (QP_{u-s} f)(\alpha(s))\,\mathrm{d}u \\
&= \int_0^{t-s} (QP_u f)(\alpha(s))\,\mathrm{d}u - \int_0^{t-s} (QP_u f)(\alpha(s))\,\mathrm{d}u.
\end{aligned}$$

Thus $M \in \mathcal{M}_0((\Omega, \mathcal{F}, \mathbb{P}), \mathcal{F}_t)$.

G.2 Martingale properties of finite state Markov chains

In this section we summarize a variety of useful martingale properties possessed by a finite state Markov chain α.

Definition G.2.1 For each $i, j = 1, \ldots, D$, define a mapping $R_{ij} : \Omega \times [0, \mathrm{T}] \to \mathbb{N}_0$ by

$$R_{ij}(\omega, t) := \begin{cases} \sum_{0<s\leq t} \chi[\alpha(s_-) = i](\omega)\chi[\alpha(s) = j](\omega), & \text{for } i \neq j, \\ 0, & \text{for } i = j. \end{cases} \tag{G.25}$$

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

Remark G.2.2 For $i \neq j$, $R_{ij}(t)$ counts the number of jumps between distinct states i and j up to time t.

Remark G.2.3 $R_{ij} = \{R_{ij}(t) : \mathrm{t} \in [0, T]\}$ is an $\{\mathcal{F}_t\}$-adapted, non-decreasing, càdlàg process which is null at the origin.

Definition G.2.4 For each $i, j = 1, \ldots, D$, define a mapping $\tilde{R}_{ij} : \Omega \times [0, \mathrm{T}] \to [0, \infty)$ by

$$\tilde{R}_{ij}(\omega, t) := \begin{cases} q_{ij} \int_0^t \chi[\alpha(s) = i](\omega)\,\mathrm{d}s, & \text{for } i \neq j, \\ 0, & \text{for } i = j. \end{cases} \tag{G.26}$$

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

Remark G.2.5 For $i \neq j$, $\tilde{R}_{ij}/q_{ij}$ measures the time that the Markov chain α spends in state i up to time t.

Remark G.2.6 $\tilde{R}_{ij} = \{\tilde{R}_{ij}(t) : t \in [0, T]\}$ is an $\{\mathcal{F}_t\}$-adapted, non-decreasing, continuous process which is null at the origin. Since $\tilde{R}_{ij}$ is continuous, it is predictable.

Finally we define the set of processes $\{M_{ij}\}$ which turn out to be the canonical martingales of the Markov chain α.

Definition G.2.7 For each $i, j \in I$, define a process $M_{ij} : \Omega \times [0, \mathrm{T}] \to [0, \infty)$ by

$$M_{ij}(\omega, t) = R_{ij}(\omega, t) - \tilde{R}_{ij}(\omega, t). \tag{G.27}$$

Remark G.2.8 As R_{ij} and $\tilde{R}_{ij}$ are $\{\mathcal{F}_t\}$-adapted, càdlàg processes which are null at the origin, then M_{ij} is an $\{\mathcal{F}_t\}$-adapted, càdlàg processes which is null at the origin.

Remark G.2.9 Upon expanding (G.27) by using the definitions of R_{ij} and $\tilde{R}_{ij}$ given in (G.2.1) and (G.2.4), we obtain for $i = 1, 2, \ldots, D$

$$M_{ij}(t) = \begin{cases} \sum_{0<s\leq t} \chi[\alpha(s_-)]\chi[\alpha(s) = j] - q_{ij} \int_0^t \chi[\alpha(s) = i]\, \mathrm{d}s & \text{for } i \neq j \\ 0 & \text{for } i = j \end{cases} \tag{G.28}$$

We next state a standard result which asserts that the M_{ij} are local martingales (see Lemma IV(21.12) of Rogers and Williams [32]).

Lemma G.2.10 For all $i, j = 1, \ldots, D$,

$$M_{ij} \in \mathcal{M}_0^{loc}((\Omega, \mathcal{F}, \mathbb{P}), \{\mathcal{F}_t\}). \tag{G.29}$$

Now we examine the integrability properties of the processes R_{ij} and $\tilde{R_{ij}}$. We will see that these are finite-variation processes which have strong integrability properties, and in particular are square integrable. This will enable us to strengthen the previous result and show that the M_{ij} are actually square-integrable martingales.

Lemma G.2.11 For all $i, j \in I$, $E\left(\tilde{R}_{ij}(t)\right)^n < \infty$ for all $t \in [0, T]$ and $n \in \mathbb{N}$.

Proof. For $i = j$, $R_{ij} = 0$ and the result is trivial. So assume that $i \neq j$. Then

$$\tilde{R}_{ij}(t) = q_{ij} \int_0^t \chi[\alpha(s) = i](\omega)\, \mathrm{d}s \leq tq_{ij} \leq Tq_{ij} \quad a.s. \tag{G.30}$$

Thus we obtain $E\left(\tilde{R}_{ij}(t)\right)^n \leq T^n q_{ij}^n < \infty$.

Remark G.2.12 It is immediate from Lemma G.2.11 and the fact that $\tilde{R}_{ij}$ is a non-decreasing process that $\tilde{R}_{ij}$ has paths of finite variation over compact intervals, that is

$$V_{\tilde{R}_{ij}}(t) = \tilde{R}_{ij}(t) < \infty, \quad t \in [0, T]. \tag{G.31}$$

The proof of the next result is standard, and can be found for example in Rogers and Williams [32], Section IV.21.

Lemma G.2.13 For all $i, j \in I$ we have $E\left(R_{ij}(t)\right)^n < \infty$ for all $t \in [0, T]$ and each $n \in \mathbb{N}$.

Remark G.2.14 It is immediate from Lemma G.2.13 and the fact that R_{ij} is a non-decreasing process that R_{ij} has paths of finite variation over compact intervals, namely

$$V_{R_{ij}}(t) = R_{ij}(t) < \infty, \quad \text{a.s.} \tag{G.32}$$

Remark G.2.15 As R_{ij} and $\tilde{R}_{ij}$ have paths of finite variation over compact intervals so also does $M_{ij} = R_{ij} - \tilde{R}_{ij}$.

We can now improve Lemma G.2.10 and establish that the M_{ij} are actually square integrable martingales:

Lemma G.2.16 For all $i, j \in I, M_{ij} \in \mathcal{M}_{0,2}((\Omega, \mathcal{F}, \mathbb{P}), \mathcal{F}_t)$.

Proof. From Lemma G.2.10, M_{ij} is a local martingale which is null at the origin. We show that M_{ij} is L_2-bounded. For all $t \in [0, T]$,

$$\begin{aligned} E|M_{ij}(t)|^2 &\overset{\text{(G.29)}}{=} E|R_{ij}(t) - \tilde{R}_{ij}(t)|^2 \\ &\leq 2E(R_{ij}(t))^2 + 2E(\tilde{R}_{ij}(t))^2 \qquad \text{(G.33)} \\ &\leq 2E(R_{ij}(t))^2 + 2E(\tilde{R}_{ij}(t))^2. \qquad \text{(G.34)} \end{aligned}$$

Thus

$$\sup_{t\in[0,T]} E|M_{ij}(t)|^2 \leq 2E(R_{ij}(T))^2 + 2E(\tilde{R}_{ij}(T))^2. \tag{G.35}$$

The finiteness of $E(R_{ij}(T))^2$ and $E(\tilde{R}_{ij}(T))^2$ coming from Lemmas (G.2.13) and (G.2.11). Applying Corollary B.2.34, we have that M_{ij} is an L_2-bounded martingale.

Remark G.2.17 Applying Theorem B.2.46 to M_{ij}, which is $\{\mathcal{F}_t\}$-adapted, càdlàg and has paths of finite variation on compact intervals, we have that M_{ij} is a purely discontinuous square integrable martingale. Then, from Theorem B.2.47,

$$[M_{ij}, M_{ij}](t) = \sum_{0 \leq s \leq t} (\triangle M_{ij}(s))^2 \quad a.s. \tag{G.36}$$

Lemma G.2.18 For all $t \in [0, T]$,

$$[M_{ij}, W_n](t) = \langle M_{ij}, W_n \rangle(t) = 0 \quad a.s. \tag{G.37}$$

for $i, j = 1, 2, \ldots, D$ and $n = 1, \ldots, N$.

Proof. For $i = j$, $M_{ij} = 0$ and trivially $[M_{ii}, W_n](t) = \langle M_{ii}, W_n \rangle(t) = 0$ for $i = 1, \ldots, D$ and $n = 1, \ldots, N$, for all $t \in [0, T]$. So assume $i \neq j$. Applying Theorem B.2.47 to the martingale W_n and the purely discontinuous martingale M_{ij}, we have

$$[W_n, M_{ij}](t) = \sum_{0<s\leq t} \triangle W_n(s) \triangle M_{ij}(s) \tag{G.38}$$

As W_n is continuous, $\triangle W_n = 0$ and substituting this into (G.38), we get the result.

Lemma G.2.19

1. $[M_{ij}, M_{ij}](t) = R_{ij}(t) \, a.s.$
2. For all $t \in [0, T]$ and $i, j, a, b \in I$, $[M_{ij}, M_{ab}](t) = 0$ a.s. if $\{(i, j)\} \neq \{(a, b)\}$.

Proof. Suppose first that $i \neq j$. Recalling from Remark G.2.6 that $\tilde{R}_{ij}$ is a continuous process, we have

$$\begin{aligned} \triangle M_{ij}(t) &\overset{(G.27)}{=} \triangle R_{ij}(t) - \triangle \tilde{R}_{ij}(t) \\ &= \triangle R_{ij}(t) \overset{(G.25)}{=} \chi[\alpha(t_-) = i] \chi[\alpha(t) = j]. \end{aligned} \tag{G.39}$$

Note from (G.39) that $(\triangle M_{ij}(t))^2 = \triangle M_{ij}(t)$. Substituting (G.39) into (G.36), we obtain

$$\begin{aligned} [M_{ij}, M_{ij}](t) &= \sum_{0\leq s\leq t} (\triangle M_{ij}(s))^2 \\ &\overset{(G.39)}{=} \sum_{0\leq s\leq t} \chi[\alpha(s_-) = i] \chi[\alpha(s) = j] \overset{(G.25)}{=} R_{ij}(t) \quad \text{a.s.} \end{aligned} \tag{G.40}$$

Since $M_{ii}(t) = 0$ a.s., trivially $[M_{ii}, M_{ii}](t) = 0$ a.s. for all $t \in [0, T]$. Now consider a square-bracket quadratic co-variation process of M_{ij} and M_{ab}, for $\{(i, j)\} \neq \{(a, b)\}$.

$$\begin{aligned} [M_{ij}, M_{ab}](t) &\overset{(G.38)}{=} \sum_{0\leq s\leq t} \triangle M_{ij}(s) \triangle M_{ab}(s) \\ &\overset{(G.39)}{=} \sum_{0\leq s\leq t} \chi[\alpha(s_-) = i] \chi[\alpha(s) = j] \chi[\alpha(s_-) = a] \chi[\alpha(s) = b] \\ &= 0 \quad \text{a.s.} \end{aligned} \tag{G.41}$$

Remark G.2.20 From Lemma G.2.19 and Theorem B.2.29, $M_{ij}^2 - R_{ij}$ is a uniformly integrable martingale.

Having shown that R_{ij} is the square-bracket quadratic variation process of M_{ij}, we show next that $\tilde{R}_{ij}$ is the angle-bracket quadratic variation process of M_{ij}.

Lemma G.2.21 For all $t \in [0, T]$ and $i, j, a, b \in I$, the following hold

1. $\langle M_{ij}, M_{ij}\rangle(t) = \tilde{R}_{ij}(t)$ a.s.
2. $\langle M_{ij}, M_{ij}\rangle(t) = 0$ a.s. for $\{(i, j)\} \neq \{(a, b)\}$.

Proof. Suppose $i \neq j$ and note that the angle-bracket quadratic variation process $\langle M_{ij}, M_{ij}\rangle$ of M_{ij} exists by the result in Lemma G.2.16 applied to Theorem B.3.6. We show that $\tilde{R}_{ij}$ satisfies all the conditions of being the angle-bracket quadratic variation process of M_{ij}, as given by Theorem B.2.24. We have the $\tilde{R}_{ij}$ is predictable, continuous, $\{\mathcal{F}_t\}$-adapted, non-decreasing and null at the origin. It remains to show that $M_{ij}^2 - \tilde{R}_{ij}$ is a martingale.

From Remark G.2.20, $M_{ij}^2 - R_{ij}$ is a martingale. From Lemma G.2.16, $M_{ij} = R_{ij} - \tilde{R}_{ij} \in \mathcal{M}_{0,2}(\{\mathcal{F}_t\}, \mathbb{P})$, so $R_{ij} - \tilde{R}_{ij}$ is certainly a martingale. Then

$$\left(M_{ij}^2 - \tilde{R}_{ij}\right) + \left(R_{ij} - \tilde{R}_{ij}\right) = M_{ij}^2 - \tilde{R}_{ij} \tag{G.42}$$

is also a martingale.

Since $M_{ii}(t) = 0$ a.s., trivially $\langle M_{ii}, M_{ii}\rangle(t) = 0$ a.s.

Similarly, as $[M_{ij}, M_{ab}](t) = 0$ a.s. then trivially $\langle M_{ij}, M_{ab}\rangle(t) = 0$ a.s. for $\{(i, j)\} \neq \{(a, b)\}$.

Remark G.2.22 The processes R_{ij} and $\tilde{R}_{ij}$ were useful in constructing the canonical martingales $\{M_{ij}\}$ of the Markov chain α. However, as we see from Lemma G.2.19 and Lemma G.2.21, R_{ij} is the square bracket quadratic variation process of M_{ij} and $\tilde{R}_{ij}$ is the angle bracket quadratic variation process of M_{ij}. From now on, we will cease to use the notation R_{ij} and $\tilde{R}_{ij}$, and, instead we will use the standard notation to represent these processes. In other words, we will use $[M_{ij}]$ to represent the square-bracket quadratic variation process of M_{ij} and $\langle M_{ij}\rangle$ to represent the angle-bracket quadratic variation process of M_{ij}.

Remark G.2.23 From Lemma G.2.16, $M_{ij} \in \mathcal{M}_{0,2}((\Omega, \mathcal{F}, \mathbb{P}), \{\mathcal{F}_t\})$. Then from Theorem B.3.6, $\langle M_{ij} \rangle$ is the compensator of $[M_{ij}]$. It then follows from Theorem B.3.5 that for any nonnegative predictable process H we have

$$E \int_0^T H(t)\, \mathrm{d}[M_{ij}](t) = E \int_0^T H(t)\, \mathrm{d}\langle M_{ij} \rangle(t). \tag{G.43}$$

Remark G.2.24 The measures that we define next are used to specify the degree of uniqueness of the stochastic integrands of the stochastic integrals which have M_{ij} as integrator.

Definition G.2.25 On the measurable space $(\Omega \times [0,T], \mathcal{P}^*)$ and for each $i, j \in I$, $i \neq j$, define a measure $\nu_{\langle M_{ij} \rangle}$ by

$$\nu_{\langle M_{ij} \rangle}[A] := E \int_0^T \chi_A(\omega, t)\, \mathrm{d}\langle M_{ij} \rangle(\omega, t) \quad \forall\, A \in \mathcal{P}^* \tag{G.44}$$

and a measure $\nu_{[M_{ij}]}$ by

$$\nu_{[M_{ij}]}[A] := E \int_0^T \chi_A(\omega, t)\, \mathrm{d}[M_{ij}](\omega, t) \quad \forall\, A \in \mathcal{P}^*. \tag{G.45}$$

Lemma G.2.26 Recalling that Leb represents Lebesgue measure, we have for each $i, j \in I$, $i \neq j$,

$$\nu_{[M_{ij}]} = \nu_{\langle M_{ij} \rangle} \text{ on } \mathcal{P}^*, \tag{G.46}$$

$$\nu_{[M_{ij}]} << \mathbb{P} \otimes \text{Leb} \quad \text{on } \mathcal{P}^*. \tag{G.47}$$

Proof. Fix $i \neq j$. Applying Fubini's theorem to obtain the joint measure, we get from (G.44), for all $A \in \mathcal{P}^*$

$$\begin{aligned}
\nu_{\langle M_{ij} \rangle}[A] &= E \int_0^T \chi_A(\omega, t)\, \mathrm{d}\langle M_{ij} \rangle(\omega, t) \\
&\overset{\text{Lemma} G.2.21}{=} E \int_0^T \chi_A(\omega, t)\, \mathrm{d}\tilde{R}_{ij}(\omega, t) \\
&\overset{(\text{G.2.4})}{=} E \int_0^T \chi_A(\omega, t) q_{ij} \chi[\alpha(t) = i](\omega)\, \mathrm{d}t \\
&= \int_{\Omega \times [0,T]} \chi_A(\omega, t) q_{ij} \chi[\alpha(t) = i](\omega)\, \mathrm{d}(\mathbb{P} \otimes Leb) \\
&= \int_A q_{ij} \chi[\alpha(t) = i](\omega)\, \mathrm{d}(\mathbb{P} \otimes Leb).
\end{aligned} \tag{G.48}$$

This shows that the measure $\nu_{\langle M_{ij}\rangle}$ is absolutely continuous with respect to the measure $(\mathbb{P}\otimes Leb)$ on $\mathcal{P}^*$, in other words

$$\nu_{\langle M_{ij}\rangle} << (\mathbb{P}\otimes Leb) \qquad \text{on}\, \mathcal{P}^*. \tag{G.49}$$

For all $A\in\mathcal{P}^*$, we have

$$\begin{aligned} \nu_{\langle M_{ij}\rangle}[A] &\overset{(G.44)}{=} E\int_0^T \chi_A(\omega,t)\,\mathrm{d}\langle M_{ij}\rangle(\omega,t) \\ &\overset{(G.43)}{=} E\int_0^T \chi_A(\omega,t)\,\mathrm{d}[M_{ij}](\omega,t) \\ &= \nu_{[M_{ij}]}[A]. \end{aligned} \tag{G.50}$$

Hence $\nu_{\langle M_{ij}\rangle} = \nu_{[M_{ij}]}$ on $\mathcal{P}^*$ and it immediately follows from this and (G.49) that $\nu_{[M_{ij}]} << (\mathbb{P}\otimes Leb)$ on $\mathcal{P}^*$.

G.3 Martingale Representation Theorem

Notation G.3.1 For a process H and $\{\mathcal{F}_t\}$-stopping time S, put $H[0,S](\omega,t) := H(\omega,t)$ when $t\in[0,S(\omega)]$, and $H[0,S](\omega,t) := 0$ when $t > S(\omega)$.

Notation G.3.2 $S^{(m)} \Uparrow T$ indicates that $(S^{(m)})_{m\in\mathbb{N}}$ is a sequence $0 \le S^{(m)} \le S^{(m+1)} \le T$ of $\{\mathcal{F}_t\}$-stopping times and for each ω there is an integer $M(\omega)$ such that $S^{(m)}(\omega) = T$ for all $m \ge M(\omega)$. Such increasing sequences of stopping times arise naturally in later arguments.

Definition G.3.3 Define the spaces of integrands

$$L^2_{loc}(\mathbf{W}) := \{\boldsymbol{\lambda} : \Omega\times[0,T]\to\mathbb{R}^N |\, \text{there exists a sequence of}\, \{\mathcal{F}_t\} - \text{stopping times}\\ \left(S^{(m)}\right)_{m\in\mathbb{N}}\, \text{such that}\, S^{(m)} \Uparrow T\, \text{and}\, \boldsymbol{\lambda}[0,S^{(m)}] \in L_2(\mathbf{W})\, \text{for all}\, m\in\mathbb{N}\Big\}$$

$$L^2_{loc}(\mathbf{M}) := \{\boldsymbol{\gamma} = \{\gamma_{ij}\}_{i,j=1}^D : \Omega\times[0,T]\to\mathbb{R}^{D\times D} |\, \text{there exists a sequence of}\, \{\mathcal{F}_t\}-\\ \text{stopping times}\, \left(S^{(m)}\right)_{m\in\mathbb{N}}\, \text{such that}\, S^{(m)} \Uparrow T\, \text{and}\, \boldsymbol{\gamma}[0,S^{(m)}] \in\\ L_2(\mathbf{M})\, \text{for all}\, m\in\mathbb{N}\}$$

Definition G.3.4 The $\mathbb{R}$-valued process $\{Z(t); t \in [0,T]\}$ is a locally-square integrable $\{\mathcal{F}_t\}$-martingale when there exists a sequence of $\{\mathcal{F}_t\}$-stopping times $\left(S^{(m)}\right)_{m\in\mathbb{N}}$ such that $S^{(m)} \Uparrow T$ and $\{Z(t \wedge S^{(m)}), t \in [0,T]\}$ is a square integrable $\{\mathcal{F}_t\}$-martingale for each $m \in \mathbb{N}$.

We shall need the following martingale representation theorem which is an immediate consequence of Definition G.3.4 and Elliot ([11], 1976, Theorem 5.1)

Theorem G.3.5 **(a)**(MRT for locally-square integrable martingale.)

Suppose $\{Z(t); t \in [0,T]\}$ is a locally-square integrable $\{\mathcal{F}_t\}$-martingale and null at the origin. Then there exists a process $\boldsymbol{\xi} = (\xi_1, \ldots, \xi_N)^\top \in L^2_{loc}(\mathbf{W})$ and $\boldsymbol{\Gamma} = (\Gamma_{ij})_{i,j=1}^D \in L^2_{loc}(\mathbf{M})$ such that Z has the stochastic integral representation

$$Z(t) = \sum_{n=1}^{N} \int_0^t \xi_n(s)\,\mathrm{d}W_n(s) + \sum_{i,j=1}^{D} \int_0^t \Gamma_{i,j}(s)\,\mathrm{d}M_{ij}(s), \quad \forall\, t \in [0,T],\ a.s. \tag{G.51}$$

In view of the orthogonality relations in Lemma G.2.18 the integrands $\boldsymbol{\xi}$ and $\boldsymbol{\Gamma}$ at (G.51) are $(\mathbb{P} \otimes Leb)$-a.e. unique and $\nu_{[\mathbf{M}]}$-a.e. unique respectively.

(b)(MRT for square integrable martingale.)

Suppose $Y \in \mathcal{M}_{0,2}\left((\Omega, \mathcal{F}, \mathbb{P}), \{\mathcal{F}_t\}\right)$. Then there exists

$$\boldsymbol{\xi} = (\xi_1, \ldots, \xi_N)^\top \in L_2(\mathbf{W}) \text{ and } \boldsymbol{\Gamma} = (M_{ij})_{i,j=1}^D \in L_2(\mathbf{M}) \tag{G.52}$$

such that Y has the representation

$$Y(t) = \sum_{n=1}^{N} \int_0^t \xi_n(\tau)\,\mathrm{d}W_n(\tau) + \sum_{i,j=1}^{D} \int_0^t \Gamma_{ij}(\tau)\,\mathrm{d}M_{ij}(\tau) \quad \text{a.s.}, \quad \forall t \in [0,T], \tag{G.53}$$

with the square-bracket quadratic variation process of Y given by

$$[Y](t) = \sum_{n=1}^{N} \int_0^t \xi_n^2(\tau)\,\mathrm{d}\tau + \sum_{i,j=1}^{D} \int_0^t \Gamma_{ij}^2(\tau)\,\mathrm{d}[M_{ij}](\tau) \quad \text{a.s.}, \quad \forall t \in [0,T]. \tag{G.54}$$

Moreover, $\boldsymbol{\xi}$ and $\boldsymbol{\Gamma}$ are unique in the sense that if $\bar{\boldsymbol{\xi}} = (\bar{\xi}_1, \ldots, \bar{\xi}_N)^\top \in L_2(\mathbf{W})$ and $\bar{\boldsymbol{\Gamma}} = \left(\bar{\Gamma}_{ij}\right)_{i,j=1}^D \in L_2(\mathbf{M})$ are such that (G.53) holds with ξ_n replaced by $\bar{\xi}_n$ and Γ_{ij} replaced by $\bar{\Gamma}_{ij}$, then $\bar{\boldsymbol{\xi}} = \boldsymbol{\xi}\,(\mathbb{P} \otimes Leb) -$ a.e. and $\bar{\boldsymbol{\Gamma}} = \boldsymbol{\Gamma}\,\nu_{[\mathbf{M}]} -$ a.e.

Bibliography

[1] C.D. Aliprantis and K.C. Border. *Infinite Dimensional Analysis - A Hitchhker's Guide, 3rd Ed.* Springer, Berlin, 2005.

[2] T.R. Bielecki, H. Jin, S.R. Pliska and X.Y. Zhou. Continuous-time mean-variance portfolio selection with bankruptcy prohibition. *Math. Finance*, pp. 213–244, v.15, 2005.

[3] J-M. Bismut. Conjugate convex functions in optimal stochastic control. *Journal of Mathematical Analysis and Applications*, 44(2):384-404, 1973

[4] Y.S. Chow and H. Teicher. *Probability Theory: Independence, Inter-Changeability, Martingales, 3rd Ed.* Springer, New York, 1997.

[5] J. Cvitanic and I. Karatzas. Convex Duality in Constrained Portfolio Optimization. *The Annals of Applied Probability*, pp. 767-818, v.2(4), 1992.

[6] C. Donnelly. *Convex duality in constrained mean-variance portfolio optimization under a regime-switching model.* PhD thesis, University of Waterloo, Canada, 2008.

[7] C. Donnelly and A. Heunis. Quadratic risk minimization in a regime-switching model with portfolio constraints. *SIAM J. Control Optim.*, pp. 2431 - 2461, v.50(4), 2012.

[8] A. Ya. Dubovitskii and A. A. Mil'yutin. Necessary conditions for a weak extremum in problems of optimal control with mixed inequality constraints. *Zhur. Vychislitel. Mat. i Mat. Fys.*, pp. 725–779, v.8 1968 (*USSR Comp. Math. and Math. Phys.*, v.8, pp. 24–98).

[9] I. Ekeland and R. Témam. *Convex Analysis and Variational Problems.* North-Holland, Amsterdam, 1976.

[10] R. J. Elliott. *Stochastic Calculus and Applications.* Springer, New York, 1982.

[11] R.J. Elliott. *Double Martingales.* Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete, 34:17-18, 1976

[12] S. N. Ethier and T. G. Kurtz. *Markove Processes.* John Wiley and Sons, New York, 1986.

[13] A. Friedman. *Foundations of Modern Analysis.* Dover, New York, 1982.

[14] A.J. Heunis. Quadratic risk minimisation with portfolio and terminal wealth constraints, *Annals of Finance* (published online 30 July 2014).

[15] A.J. Heunis. *Notes on Stochastic Calculus.* Lecture Notes, University of Waterloo, 2001.

[16] A.D. Ioffe and V.L. Levin, Subdifferentials of convex functions, *Trans. Moscow Math. Soc.*, (translation by *Amer. and London Math. Socs.*), pp. 1–72, v.26, (1972).

[17] K. Ito and K. Kunisch, *Lagrange Multiplier Approach to Variational Problems and Applications*, SIAM Books, Philadelphia, (2008).

[18] J. Jacod and A. N. Shiryaev. *Limit Theorems for Stochastic Processes.* Springer, Berlin, 1987.

[19] O. Kallenberg. *Foundations of Modern Probability, 2nd Ed.* Springer, New York, 2002.

[20] I. Karatzas, J.P. Lehoczky and S.E. Shreve. Optimal portfolio and consumption decisions for a small investor on a finite horizon. *SIAM Journal on Control and Optimization,* 25(6), 1987.

[21] I. Karatzas and S.E. Shreve. *Brownian Motion and Stochastic Calculus.* Springer, New York, 1991.

[22] I. Karatzas and S.E. Shreve. *Methods of Mathematical Finance.* Springer, New York, 1998.

[23] A.N. Kolmogorov and S.V.Fomin. *Introductory Real Analysis.* Dover Publications, New York, 1970.

[24] C. Labbé . *Contributions to the Theory of Constrained Portfolio Optimization.* PhD thesis, University of Waterloo, Canada, 2004.

[25] C. Labbé and A.J. Heunis. Convex duality in constrained mean variance portfolio optimization. *Advances in Applied Probability,* 39(1):77-104, March 2007.

[26] R. Sh. Liptser and A.N. Shiryayev. *Theory of Martingales.* Kluwer Academic, Dordrecht, 1989.

[27] K. Makowski and L.W. Neustadt. Optimal control with mixed control-phase variable equality and inequality constraints. *SIAM J. Control Optim.,* pp.184–228, v.12, 1974.

[28] H. Markowitz. Portfolio selection. *Jour. Finance,* pp. 77–91, v.7, 1952.

[29] J. R. Norris. *Markov Chains.* Cambridge University Press, Cambridge (UK), 1997.

[30] P. E. Protter. *Stochastic Integration and Differential Equations, 2nd Ed.* Springer, New York, 2005.

[31] D. Revuz and M. Yor. *Continuous Martingales and Brownian Motion.* Springer, Berlin, 1994.

[32] L.C.G. Rogers and D. Williams. *Diffusions, Markov Processes and Martingales, Volume II Ito Calculus.* Cambridge University Press, 2000.

[33] R.T. Rockafellar. Integrals which are convex functionals. *Pacific Journal of Mathematics,* Vol. 24, No.3, 1968.

[34] R.T. Rockafellar. *Convex Analysis.* Princeton University Press, Princeton, 1970.

[35] R.T. Rockafellar, State constraints in convex control problems of Bolza, *SIAM J. Control Optim.*, pp. 691–715, v.10, (1972).

[36] R.T. Rockafellar. *Conjugate Duality and Optimization.* SIAM (CBMS-NSF Series No. 16), Philadelphia, 1974.

[37] R.T. Rockafellar and R. J.B. Wets. Stochastic convex programming: singular multipliers and extended duality. *Pacific J. Math,* pp. 507–522, v.62, 1976.

[38] H.L. Royden. *Real Analysis, 3rd Ed.* Macmillan Publishing Co., New York, 1988.

[39] W. Rudin. *Principles of Mathematical Analsysis, 3rd Ed.* International Series in Pure and Applied Mathematics, McGraw-Hill Book Co., New York, 1976.

[40] W. Rudin. *Real and Complex Analysis, 3rd Ed.* McGraw-Hill Book Co., New York, 1987.

[41] M. Slater. Lagrange mulipliers revisited: a contribution to non-linear programming. Cowles Commission Discussion Paper, Math. 403, 1950.

[42] L.R. Sotomayor and A. Cadenillas. Explicit solutions of consumption-investment problems in financial markets with regime switching. *Math. Finance,* pp. 251–279, v.19, 2009.

[43] A. Wilansky, *Modern Methods in Topological Vector Spaces*, McGraw-Hill, New York, 1978.

[44] E. Wong and B. Hajek. *Stochastic Processes in Engineering Systems.* Springer, Berlin, 1985.

[45] G.L. Xu and S.E. Shreve. A duality method for optimal consumption and investment under short-selling prohibition. I. General market coefficients. *Ann. Appl. Probab.,* pp. 87–112, v.2, 1992.

[46] K. Yosida and E. Hewitt. Finitely additive measures. *Trans. Amer. Math. Soc.,* pp 46-66, v.72, 1952.

[47] X.Y Zhou and G. Yin. Markowitz's mean-variance portfolio selection with regime switching:a continuous-time model. *SIAM J. Control Optim.,* pp. 1466–1482, v.42, 2003.

Glossary

$[M]$, $[M, M]$	square-bracket quadratic variation process of M, page 110
$[N, M]$	square-bracket quadratic co-variation process of N and M, page 110
$\langle M \rangle$, $\langle M, M \rangle$	angle-bracket quadratic variation process of M, page 109
$\langle N, M \rangle$	angle-bracket quadratic co-variation process of N and M, page 109
a	random variable, page 24
$\mathcal{A}$	set of admissible portfolios, page 25
$\mathbb{A}$	subspace of $\mathbb{B}$, page 22
$\mathbb{A}_1$	set of Itô processes (subset of $\mathbb{A}$), page 40
α	markov chain, page 12
$\boldsymbol{b}$	mean rate of return process, page 18
β	discounting process, page 39
$\mathbb{B}$	product space of integrands, page 22
$\mathbb{B}_1$	subspace of the space $\mathbb{B}$, page 38
$\langle \cdot, \cdot \rangle$	bilinear form, page 40
c	random variable, page 24

$\mathbb{D}$ set of admissible wealth processes, page 30

$\mathbb{D}_1$ set of Itô processes ($\mathbb{D} \cap \mathbb{A}_1$), page 40

δ support function, page 101

$\bar{\delta}$ characteristic function, page 47

$\mathcal{E}$ Doléans-Dade exponential, page 80

η primal value, page 25

F perturbation function, page 38, 34

f primal function, page 31

$\mathcal{F}_s^{\circ,\mathbf{W}}$ raw filtration generated by $\mathbf{W}$over $[0, s]$, page 13

$\mathcal{F}_T^{\circ,\alpha}$ raw σ-algebra generated by α over $[0, T]$, page 13

$\mathcal{F}_T^{\circ,\mathbf{W}}$ raw σ-algebra generated by $\mathbf{W}$over $[0, T]$, page 13

$\mathcal{F}_t^{\circ}$ raw filtration, page 14

$\mathcal{F}_t$ standard filtration, page 14

$\mathcal{F}_t^{\mathbf{W}}$ augmented standard filtration generated by $\mathbf{W}$, page 14

$\mathcal{F}_t^{\alpha}$ augmented standard filtration generated by α, page 64

$\mathcal{FV}$ set of processes of finite variation, page 108

$\mathcal{FV}_0$ set of processes of finite variation and null at the origin, page 108

$\mathcal{FV}^+$ set of processes which are non-decreasing, page 108

$\mathcal{FV}_0^+$ set of processes which are non-decreasing and null at the origin, page 108

g	dual function, page 41, 34
$\mathbb{G}$	Subset of L_∞^*, page 38
H	State price density, page 49
$\mathcal{IV}$	set of processes in $\mathcal{FV}$ of integrable variation, page 108
$\mathcal{IV}_0$	set of processes in $\mathcal{FV}_0$ of integrable variation, page 108
$\mathcal{IV}^+$	set of processes in $\mathcal{FV}^+$ which are integrable, page 108
$\mathcal{IV}_0^+$	set of processes in $\mathcal{IV}^+$ which are null at the origin, page 108
$\mathcal{IV}_{loc}$	processes of locally integrable variation, 117
$\mathcal{IV}_{0,loc}^+$	set of processes in $\mathcal{FV}_0^+$ which are of locally integrable variation, 117
J	quadratic loss function, page 24
J^*	convex conjugate, page 40
K	Lagrangian, page 40
$\kappa_{\boldsymbol{\sigma}}$	uniform-bound for $\boldsymbol{\sigma}$, page 19
$\kappa_{\boldsymbol{\theta}}$	uniform-bound for $\boldsymbol{\theta}$, page 19
L_∞^*	dual space of L_∞, page 37, 129
L_p	Vector space, page 37
L_{21}	space of integrands, page 22
$L_2(\mathbf{W})$	space of L_2 integrands, page 22
$L_2(\mathbf{M})$	space of L_2 integrands, page 22
$\mathcal{M}$	set of martingales, page 104

$\mathcal{M}_0$	set of martingales null at the origin, page 105
$\mathcal{M}^c$	set of continuous martingales, page 105
$\mathcal{M}_0^c$	set of continuous martingales null at the origin, page 105
$\mathcal{M}_2$	set of square-integrable martingales, page 105
$\mathcal{M}_{0,2}$	set of square-integrable martingales null at the origin, page 105
$T^m \Uparrow T$	increasing stopping times, page 106
$\mathcal{M}^{loc}$	set of local martingales, page 106
$\mathcal{M}_0^{loc}$	set of local martingales null at the origin, page 107
$\mathcal{M}^{c,loc}$	set of continuous local martingales , page 107
$\mathcal{M}_0^{c,loc}$	set of continuous local martingales null at the origin, page 107
$\mathcal{M}_2^{loc}$	set of locally square-integrable martingales, 106
M_{ij}	canonical martingale associated with Markov chain α, page 15
$\|\| \cdot \|\|_{\mathbb{U}}$	maximum norm, page 43
$\nu_{[M_{ij}]}$	Doléans measure, page 16
$\mathcal{P}^*$	predictable σ-algebra, page 15
$\boldsymbol{\pi}$	portfolio process, page 21
$\bar{\boldsymbol{\pi}}$	candidate optimal portfolio process, page 53
q_{ij}	$(i,j)^{th}$ entry of the generator of the Markov chain, page 13
$\boldsymbol{\theta}$	market price of risk, page 19
$\boldsymbol{\sigma}$	volatility process, page 18
$\mathbb{U}$	linear space of perturbations, page 38

q a constant, page 24

Q generator of the Markov chain, page 13

$\mathcal{Q}$ set of singular norm-continuous linear functionals on L_∞, page 128

u_1 element of L_2, page 38

u_2 element of L_∞, page 38

$\varkappa$ function of (Y, Z), page 42

$\mathbf{W}$ Brownian Motion, page 13

$\bar{X}$ candidate optimal solution to the primal problem, page 53

$X^{\boldsymbol{\pi}}$ solution to the wealth equation for $\boldsymbol{\pi}$, page 21

Ξ linear, bijective map, page 39

$\mathbb{Y}$ linear space, page 40

Z an element of L_∞^*, page 40

$\bullet$ denotes stochastic integration, page 80

Index