Convex duality in constrained mean-variance portfolio optimization under a regime-switching model

by

Catherine Donnelly

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Doctor of Philosophy
in
Actuarial Science

Waterloo, Ontario, Canada, 2008

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 solve a mean-variance portfolio optimization problem with portfolio constraints under a regime-switching model. Specifically, we seek a portfolio process which minimizes the variance of the terminal wealth, subject to a terminal wealth constraint and convex portfolio constraints. The regime-switching is modeled using a finite state space, continuous-time Markov chain and the market parameters are allowed to be random processes. The solution to this problem is of interest to investors in financial markets, such as pension funds, insurance companies and individuals.

We establish the existence and characterization of the solution to the given problem using a convex duality method. We encode the constraints on the given problem as static penalty functions in order to derive the primal problem. Next, we synthesize the dual problem from the primal problem using convex conjugate functions. We show that the solution to the dual problem exists. From the construction of the dual problem, we find a set of necessary and sufficient conditions for the primal and dual problems to each have a solution. Using these conditions, we can show the existence of the solution to the given problem and characterize it in terms of the market parameters and the solution to the dual problem.

The results of the thesis lay the foundation to find an actual solution to the given problem, by looking at specific examples. If we can find the solution to the dual problem for a specific example, then, using the characterization of the solution to the given problem, we may be able to find the actual solution to the specific example.

In order to use the convex duality method, we have to prove a martingale representation theorem for processes which are locally square-integrable martingales with respect to the filtration generated by a Brownian motion and a finite state space, continuous-time Markov chain. This result may be of interest in problems involving regime-switching models which require a martingale representation theorem.

Acknowledgements

I wish to express my gratitude to my supervisor, Andrew Heunis. Without his time, support and patience, this thesis would not have been possible. It was a pleasure to have as a supervisor someone who believes in education and who is incredibly knowledgeable.

I would like to thank my thesis committee for their time and effort: Adam Kolkiewicz, Paul Marriott, Carole Bernard, Ken Vetzal and George Yin. I appreciated especially the helpful comments from Professor Kolkiewicz and the encouraging smiles of Professor Marriott during the oral defences.

I would also like to thank the staff in the department, who have been so helpful to me over the years.

Finally, a special thanks to Fred.

Dedication

This is dedicated to my parents, Liam and Gemma.

Contents

Chapter 1

Introduction

Portfolio theory concerns the allocation of investments between different asset classes. How to allocate one's investments optimally is a challenge with which all investors are faced, from banks to insurance companies to private individuals. There is no one-size-fits-all solution, since what constitutes an optimal allocation varies between investors, reflecting their different requirements. For example, speculators may be interested solely in profit-maximization and will seek an exceptional return on their stock market investments, regardless of the risk involved. A speculator may consider investing all of her wealth in one stock an optimal allocation. In contrast, a pension scheme may be only interested in meeting its pension liabilities and will seek to minimize the risk of not meeting those liabilities. The pension scheme will consider an allocation optimal if it matches the scheme's liabilities perfectly.

The birth of modern portfolio theory is generally attributed to the seminal work by Markowitz [36] published in 1952. He considered a problem of selecting a *portfolio of investments* which minimizes the variance of return for a given level of expected return on the investments. He realized that investors should consider the risk and return characteristics of portfolios of investments, and not just of single assets. This was a major insight as, before his paper, the risk and return characteristics of assets were considered in isolation and not as a group.

The problem of selecting a portfolio which satisfies certain risk and return requirements is one which continues to be of interest to investors. For example, an insurance company may specify a level of return on its investments in order to meet its liabilities and to satisfy a profit criterion. However, it may not want to follow an unduly risky investment strategy, where risk is measured by the variance of return, since it may face regulatory requirements that its assets must always be above some level. Thus it may seek an investment strategy, or portfolio of investments, which minimizes the risk of failing to meet its specified level of return.

Since Markowitz's paper [36], the *portfolio selection problems* which can be solved have grown in complexity. For example, the insurance company may also face regulatory requirements that it never has negative holdings in any stock (called short-selling a stock). As this requirement acts as a constraint on the optimal portfolio, we call it a *portfolio constraint*. Generally, the presence of portfolio constraints makes portfolio selection problems much more challenging to solve.

In this thesis, we solve a portfolio selection problem, called a *mean-variance portfolio optimization ("MVO") problem*. We allow for general portfolio constraints, so that we can accommodate investors who are restricted to which assets they can invest in and by how much they can invest in each asset. This would include the insurance company above and also investors such as pension funds, who, in addition to regulatory restrictions, may also have self-imposed restrictions on investments. For example, they may avoid investment in assets which are perceived to be risky or which they do not understand fully.

In order to solve any portfolio selection problem, it is necessary to construct a model of the financial market. The market model seeks to mimic the behavior of the stock market. The model should also be amenable to mathematical analysis so that investors can use it to find solutions to investment problems. Market models have become ever more sophisticated since Markowitz's original single-period model. They have developed from single-period models to multi-period models, and from these discrete-time models to continuous-time models.

The latest attempt to capture actual stock market behavior are *regime-switching models*. Regime-switching models allow for the market to undergo "shocks" at random times. At any time, the market is assumed to be in some regime. An example of such a regime would be a bull market, in which stock prices are generally rising. After a shock, the market's behavior fundamentally changes. The shock is represented as a switch, or change, of regime. An example of a shock would be a stock market crash, such as the Wall Street crash of 1929. After the shock, the market is in a new regime, for example a bear market, in which prices are generally falling.

We use a regime-switching model of the market. The model we use is quite general in nature and appears to offer significant advantages over existing regime-switching models. For example, the model allows the use of stochastic volatility models which existing regime-switching models do not. From a practical perspective, this allows for a more realistic model of the stock market. Added to the fact that investors are seeking to model the stock market as closely as possible, the solution to the MVO problem in the regime-switching model is a valuable piece of information.

The MVO problem with portfolio constraints has not yet been solved within the regime-switching model that we use. Indeed, it does not appear to have been solved even within a less general regime-switching model. Thus we believe that there is a need for the research contained in this thesis. However, we do note that Zhou and Yin [53] have solved the MVO problem *without portfolio constraints* in a regime-switching model.

We use a convex duality method to solve the MVO problem. Convex duality methods establish a connection between the original problem, called the "primal problem", and another problem, called the "dual problem". The hope is that the dual problem is easier to solve than the primal problem. The convexity properties of the primal problem are critical in establishing the connection between it and the dual problem. Using the solution to the dual problem, this connection may allow us to construct the solution to the primal problem.

In summary, we solve an MVO problem with portfolio constraints in a regime-switching market model by the application of a convex duality method.

1.1 Outline of the thesis

In Chapter 2, we give the background to the research. We also give the reasons for the specific choice of the regime-switching model, the MVO problem and the method of solution.

In Chapter 3, we define the market model and MVO problem in precise mathematical terms. In addition, we outline the results from Zhou and Yin's paper [53] since it is this work which motivates much of our own effort. They also solve a MVO problem in a regime-switching model. However, their MVO problem does not involve any portfolio constraints and their regime-switching model is less general than our model. Moreover, they use a different method of solution. Their paper also allows us to demonstrate the increased generality of our regime-switching model.

In Chapter 4, we set out the main stages in the solution to the MVO problem. Most of this chapter is concerned with the solution to the MVO problem without a *terminal wealth constraint*. The reason is that the solution to this uses convex duality, which is at the heart of the method. Solving the MVO problem with a terminal wealth constraint, which is done in the last section of the chapter, is a straightforward application of a Lagrange multiplier technique.

In Chapter 5, we give the conclusion and outline briefly some future areas of investigation related to the work in the thesis.

Some results which supplement Chapter 3 and Chapter 4 are given in Appendix A.

In Appendix B, we prove a martingale representation theorem. This theorem is required in order to solve the MVO problem and is tailored specifically to the regime-switching model.

Finally, in Appendix C, we give many of the standard results and definitions that we use throughout the thesis.

Chapter 2

Background

In this thesis, we find the solution to an MVO problem with portfolio constraints. The problem is to select a portfolio of assets which satisfies certain requirements. The market model that we use is called a regime-switching model. Once the regime-switching model is constructed, the problem can be precisely defined within its framework. The possible methods of solution to the problem can then be considered. Thus the work involved in the thesis can be divided quite naturally into three stages:

- describe the market model;
- define the problem; and
- implement the method of solution.

In this chapter, each of the three stages is considered in turn. As we shall see, each stage involves a choice of some kind. The purpose of the chapter is to justify these choices.

The first stage is to develop the market model. We use a regime-switching model, where the market switches between regimes, such as bear markets or bull markets. Starting with the historical development of the stock market model, we show that the market model is an improvement on prior market models.

After we have described the market model, we outline the problem. The full statement of the problem, given in Section 3.2.2, is highly mathematical. In this chapter, we have tried to avoid using mathematical terms as much as possible. However, the essence of the problem is easily garnered from the simple example that we outline in this chapter. We also demonstrate the validity of the problem posed (which is an MVO problem), since we could easily have posed another type of problem (a utility maximization problem).

Finally, we examine possible methods of solution and, based on this examination, we conclude that a convex duality method is most appropriate. However, as there is another approach, based on *stochastic control theory*, that we could have used, we give some of the background to both the convex duality method and the method based on stochastic control theory.

2.1 The market model

Many stock market investors seek to model the stock markets in which they invest. They do this to find solutions to questions that they have about the stock market. Perhaps they wish to know how secure their investments are, for example how likely they are to lose all their money tomorrow in a stock market crash, or how likely they are to lose a third of their money, or how likely they are to double their money. Perhaps they are not sure how much and where they should invest their money, for example, if they should invest all their money in one stock, or in many stocks, and, if so, in how many stocks? These are simple questions but they don't have simple answers. Many investors use a mathematical model of the stock market to answer these types of questions as well as others.

How do we capture the behavior of a stock market in a mathematical model? The prevailing model is based on the idea that stock prices move randomly. The classic example of a random movement is *Brownian motion*, a phenomenon observed in nature. The first recorded observation of Brownian motion was in 1827 by a botanist called Robert Brown. He used a microscope to observe pollen grains suspended in water and he noticed that the pollen grains were jittering about in a random fashion. The random movement was named Brownian motion in his honour.

Stock market prices also appear to jitter about in a random fashion. A French mathematician called Louis Bachelier noticed this and decided that stock prices moved like Brownian motion. Unfortunately for Bachelier, his paper on the subject [2], published in 1900 with an English translation given in [8], did not receive the attention it deserved at the time. It was over fifty years before Bachelier's paper [2] began to be widely-recognized as a ground-breaking piece of work. It began to gain recognition around the time of two significant papers by an economist called Robert Merton.

Merton [37] modeled the stock prices as *geometric Brownian motion*. This implies that stock prices are lognormally distributed. It was Merton who developed the continuous-time market model, which enabled him to model the stock prices as geometric Brownian motion, which is a continuous-time stochastic process. Expressed mathematically, for a stock price which is a geometric Brownian motion, its price $S(t)$ at time t satisfies the stochastic differential equation

$$\frac{\mathrm{d}S(t)}{S(t)} = \mu \ \mathrm{d}t + \sigma \ \mathrm{d}W(t), \tag{2.1.1}$$

where $\{W(t)\}$ is a Brownian motion and μ and σ are constants, called the *market coefficients*. μ is interpreted as the *expected rate of return* on the stock. σ is interpreted as the *volatility* of the stock. The stock price process $\{S(t) : t \geq 0\}$ that satisfies (2.1.1) is a stochastic process; it is a set of random variables indexed by time. As is standard practice, we suppress the notation showing the explicit dependence of $\{S(t)\}$ on the underlying probability set.

In a complementary paper [38], Merton modeled stock prices as stochastic processes with the market coefficients depending on the stock price, as in the following equation.

$$\frac{\mathrm{d}S(t)}{S(t)} = \mu(t, S(t))\ \mathrm{d}t + \sigma(t, S(t))\ \mathrm{d}W(t). \tag{2.1.2}$$

This model allows the market coefficients μ and σ to vary with time and the stock price, unlike (2.1.1), in which the market coefficients are fixed. This model allows for greater latitude when fitting actual market data to the model. For example, there is empirical evidence which suggests that a low stock price increases the stock price volatility more than a high stock price (see Black [5]). In that case, $\sigma(t, S(t))$ would increase more when $S(t)$ is small than when $S(t)$ is large. Merton's previous model, as in (2.1.1), cannot model a stock price volatility increasing with stock price.

Harrison and Kreps [18] and Harrison and Pliska [19], [20] developed the mathematics of continuous-time finance and allowed the stock price processes to be general stochastic processes. This permits the market coefficients μ and σ to be general random processes, such as in the following equation.

$$\frac{\mathrm{d}S(t)}{S(t)} = \mu(t)\ \mathrm{d}t + \sigma(t)\ \mathrm{d}W(t). \tag{2.1.3}$$

By random processes, we mean that μ and σ are functions of both time and the underlying probability set. Modeling the market coefficients as random processes increases the generality of the stock price model significantly. For example, it allows the use of stochastic volatility models, which provide a better fit of actual market data to the model (see Hull [23] and Hobson and Rogers [21] for references).

However, despite the flexibility of the model in (2.1.3), empirical evidence suggests that a modification of this model, called a regime-switching model, results in a more realistic model of the stock market (see empirical evidence by Gray [17] and Kalimipalli and Susmel [27] and the references therein).

Regime-switching market models are the most recent attempt to capture actual stock market behavior. The idea behind regime-switching market models is as follows. At any time, the market is in a regime. An example of such a regime would be a bull market, in which stock prices are generally rising. At some random time, the market suddenly switches to a new regime. For example, the market could switch from a bull market to a bear market, in which stock prices are generally falling. The cause of the regime switch may or may not be determinable. It could be due to changes in government regulations, an unexpected declaration of war, a meteorite crashing into the Earth, irrational exuberance from investors; there is a myriad of possibilities.

The action of regime-switching is modeled using a continuous-time Markov chain $\{\alpha(t)\}$, since it adequately describes the desired features of market regime-switching. Most literature to date (for example Jobert and Rogers [26], Stockbridge

[50], Zhou and Yin [53]) models the stock prices in a regime-switching model as either

$$\frac{\mathrm{d}S(t)}{S(t)} = \mu(\alpha(t))\ \mathrm{d}t + \sigma(\alpha(t))\ \mathrm{d}W(t) \tag{2.1.4}$$

or

$$\frac{\mathrm{d}S(t)}{S(t)} = \mu(t, \alpha(t))\ \mathrm{d}t + \sigma(t, \alpha(t))\ \mathrm{d}W(t), \tag{2.1.5}$$

where, in both cases, the market coefficients are Markov-modulated. This means that the market coefficients depend on the regime-switching Markov chain $\{\alpha(t)\}$. These models could be used to model a stock price whose mean rate of return, either $\mu(\alpha(t))$ or $\mu(t, \alpha(t))$, is higher in a bull market than in a bear market. However, the models in (2.1.4) and (2.1.5) have an obvious limitation. Stochastic volatility models can no longer be used with them, since the stock volatility is either constant (as in (2.1.4)) or deterministic (as in (2.1.5)) within each market regime. The ability to model the market coefficients as stochastic processes in their own right (between successive regime-switches) is lost.

In order to overcome this limitation, we use a regime-switching model where the market coefficients are random processes. For our model, the stock prices satisfy the following stochastic differential equation

$$\frac{\mathrm{d}S(t)}{S(t)} = \mu(t)\ \mathrm{d}t + \sigma(t)\ \mathrm{d}W(t), \tag{2.1.6}$$

where, as in (2.1.3), the market coefficients are random processes. The model in (2.1.6) looks identical to the model in (2.1.3). However, there is an important difference which is rather technical in nature: the filtrations to which the market coefficients are adapted are different in the two models. In (2.1.3), the filtration is the usual one that is generated by the Brownian motion $\{W(t)\}$. In (2.1.6), the filtration is generated jointly by the Brownian motion $\{W(t)\}$ and the regime-switching Markov chain $\{\alpha(t)\}$. This difference will have a very significant effect on the implementation of the solution method. We will explore this point later, in Chapter 3.

The model we use may otherwise be considered a standard one, similar to the one described in Karatzas and Shreve [31], Chapter 1. It consists of a finite number of stocks and a single bank account. The stocks will be considered as risky assets and their price processes will obey the dynamics in (2.1.6). The bank account will be considered the risk-free asset. The bank account's price process will increase in line with the risk-free interest rate process, which is a random process.

In Section 3.2.1, we define the market model in more precise terms.

2.2 Portfolio selection problems

The motivation for modeling a financial market is to enable us to answer questions concerning the market. We do this by posing the questions within the model of the financial market. The questions are usually either pricing questions ("How much should this derivative cost?") or portfolio selection questions ("Which stocks should I invest in?"). A portfolio represents how much wealth an investor has invested in each asset in the market. We will consider only the latter question, namely portfolio selection questions.

A portfolio selection problem involves the selection of a portfolio that satisfies the requirements of an investor. For example, consider an investor who has $25 000 in cash. She wishes to invest this sum of money in the stock market so that, in twenty years time, the average value of her investments has accumulated to $100 000. She wonders in which stocks she should invest and how much she should invest in each stock.

Portfolio selection problems can be broadly divided into two main categories: utility maximization and MVO problems. We study an MVO problem, which is not dissimilar to the example in the preceding paragraph. However, utility maximization problems are more popular in the financial literature. So why have we gone for the less popular option? Before we answer this, we give some of the background to these two kinds of portfolio selection problems.

Utility maximization is grounded in the economic theory of utility. For an investor given a range of investment choices, utility theory tells us which choice satisfies the investor's needs the most. The utility of $1 or an apple is how much it satisfies the investor's needs. In isolation, knowing the utility of $1 is not very informative. However, if we also know the investor's utility of an apple, then we can compare the investor's preference for $1 over an apple. Then, for an investor forced to choose between $1 and an apple, we can say which one the investor prefers. A utility function is used to measure the investor's utility of goods or services, including items such as $1 or an apple, by assigning it a numerical value. The higher the numerical value, the more the investor prefers the good or service.

A utility maximization problem involves the maximization of the investor's expected utility derived from future investment in the stock market. Generally, this involves the maximization of the expected utility of terminal wealth and consumption, where the idea of consumption is that the investor may spend, or consume, some of her wealth inter-temporally.

The appeal of utility maximization problems is that they can incorporate the risk attributes of an individual investor through the investor's utility function. Since modern portfolio theory seeks to model the real-world, or, at least, an economist's view of the world, this is an attractive notion.

Utility maximization problems were considered early on in the history of modern portfolio theory, for example by Mossin [39] and Samuelson [47] in a discrete-time model and by Merton [37] in a continuous-time model, and they have been studied ever since. As can be seen in the next section, most of the methods of solution to portfolio selection questions have been applied in the first instance to a utility maximization problem.

In contrast to utility maximization problems, classical MVO problems consider the trade-off between the variance of return and the expected return of a portfolio. The variance of return is referred to as the risk. The problem is usually to find the portfolio which minimizes the risk for a given level of expected return. The investor's preferences are only taken into account in so far as to determine what level of expected return is desired.

Unable to compete with the economic elegance of utility maximization problems, the appearance of MVO problems in the literature has been rather more lacklustre. Indeed, as we mentioned above, the development of modern portfolio theory has been done mostly through utility maximization problems. Aside from this, examples of MVO problems in the literature include Markowitz's paper [36], the papers by Duffie and Richardson [12], Hu and Zhou [22], Li, Zhou and Lim [34], Lim and Zhou [35], Schweizer [48] and, of most relevance to this thesis, Zhou and Yin [53].

In spite of MVO problems being less popular than utility maximization problems, they have a two-fold advantage. First, utility maximization problems require a utility function. Eliciting a utility function from an investor is a difficult task. MVO problems do not have this issue since they do not consider the preferences of the investor. Indeed, in this respect, it may be considered a more objective measure of risk. Second, the solution of a MVO problem explicitly shows the trade-off between risk and return. This allows any investor to utilize the solution of an MVO problem. In comparison, the risk-return trade-off is implicit in the solution of a utility maximization problem for a particular investor. This means that investors who do not share the same utility function as the particular investor cannot utilize the solution of the utility maximization problem.

We solve an MVO problem. Aside from the reasons given above, we have chosen an MVO problem over a utility maximization problem since it is a more tractable problem. In the MVO problem, the wealth processes are square-integrable Itô processes. In a utility maximization problem, the wealth processes are general Itô processes, making the implementation of the solution method more difficult. Moreover, with an eye to the future, since their mathematical structures are not fundamentally different (both have convexity properties), we hope that solving the MVO problem will aid in the solution of the analogous utility maximization problem in a market model with regime-switching.

The MVO problem that we consider asks an *existence and characterization question.* The full problem that we ultimately hope to solve is illustrated by the example in the following paragraph. As we stated in the introduction to this chapter, the precise definition of the problem, given in Section 3.2.2, is rather mathematical although the essence of the problem is contained in the following example.

An investor starts with a fixed initial wealth. She wishes to invest the initial wealth in the stock market in order to obtain some desired wealth at a fixed future time T. The desired wealth is called the *terminal wealth constraint.* At every time, the investor can choose how to invest her wealth in the assets of the stock market. This series of investment choices is called a *portfolio process*. The portfolio process represents how much wealth the investor has invested in each asset at each time. In making investment choices, the investor can operate in two possible trading environments, namely: (i) be free to distribute the current wealth at every instant among the assets without any constraints on this distribution; or (ii) be compelled by various externally-imposed trading regulations to respect some *portfolio constraints* at each instant in the course of allocating her wealth among the assets (for example, trading restrictions could insist that she never go *short* on any of the stocks, in which case she must trade so as to have a *nonnegative* amount of wealth invested in each stock at every instant of time). Typically, investment problems with portfolio constraints are significantly more challenging than those without such constraints. To make this example more interesting, we shall suppose that the investor is not permitted to short-sell any stock, that is she must always have nonnegative wealth invested in every stock. Later in this work, we shall deal with portfolio constraints which are considerably more general than this.

The investor realizes that she may not attain the desired wealth exactly, though she can expect to attain it on average. With this in mind, she would like to minimize the variance of her actual wealth at the future time T from the desired expected wealth. The variance at the future time T of her actual wealth from the desired wealth is called the risk.

The problem is to determine if there exists a portfolio process which satisfies both the terminal wealth and portfolio constraints and which minimizes the risk. We seek to characterize such a portfolio process in terms of the market coefficients, among other things. We do not find an exact expression for the portfolio process because the model is simply too general. Obtaining an exact expression for specific cases (such as the case with no portfolio constraints or with non-random or Markov-modulated market coefficients) continues to present challenges and is an area for future investigation.

When presenting the MVO problem in Section 3.2.2, we will drop the constraint on the expected terminal wealth. This is for the sake of clarity. The heart of the research is the solution of an MVO problem in a regime-switching model by the application of a convex duality method. A convex duality method is applied to the MVO problem without a terminal wealth constraint. Solving the problem with the terminal wealth constraint involves a second stage, where a Lagrange multiplier technique is applied. The second stage is a fairly routine application of standard Lagrange multiplier ideas, and is demonstrated in Section 4.9.

Even aside from issues of clarity, the problem without the terminal wealth constraint is a valid one and has been studied before by Schweizer [48]. An example of an MVO problem without a terminal wealth constraint would be a company-sponsored pension fund seeking to minimize the variance of its assets from its liabilities. The pension fund may not be interested in attaining a particular level of wealth. However, it may face regulatory penalties if it under-funds its liabilities. At the same time, the company sponsoring the pension fund will not want to over-fund the liabilities since it could have put the excess funds to better use elsewhere.

2.3 Methods of solution

The formulation of the MVO problem places it in a wider class of stochastic problems called stochastic control problems. Stochastic control problems appear in a variety of fields such as in manufacturing, insurance and finance. These and many other examples are given in the book of Yong and Zhou [52].

Stochastic control problems are dynamic problems, in that they change over time. A stochastic control problem allows for decisions to be made at each time. These decisions will constitute the solution to the problem. The aim is to maximize or minimize some quantity over the time horizon of the problem. For example, subject to meeting the demand for the factory's output, the aim of a factory owner may be to minimize the annual cost of running the factory. At any time, the factory owner can decide the production rate of the factory. By varying the production rate, the factory owner seeks to minimize the annual cost. There are usually constraints on the decisions. For example, there is a limit to how much can be produced by the factory. There are also generally exogenous factors affecting a stochastic control problem. For example, the demand for the factory's output, which will affect the production rate.

The MVO problem that we will formulate can be seen to fit easily into the framework of stochastic control problems. The aim is to minimize an investor's risk over a finite time horizon. The decisions that can be made over time can be represented as the portfolio process, since this specifies the investment decisions made at each time. Constraints on the portfolio process are the amount of initial wealth, how much can be invested in each stock over the time horizon and expectations on how much the portfolio should be worth at the end of the time horizon. Exogenous factors are represented by the fluctuations in the asset prices and the regime-switches of the market.

This natural fit leads to the supposition that MVO problems can be solved using stochastic control methods. Essentially, stochastic control methods are based on three possible approaches, namely

1. linear quadratic control;
2. dynamic programming; and
3. the Pontryagin Maximum Principle.

A major drawback of stochastic control methods is that it is far from clear how to apply them to problems involving a mixture of random coefficients and general portfolio constraints.

Linear quadratic control applies very successfully to problems with random coefficients but *no portfolio constraints*, and has been used by Lim and Zhou [35]. However, when portfolio constraints are added, it becomes very difficult to carry out the "completion of squares" calculations on which this approach depends.

On the other hand, dynamic programming may be applied when portfolio constraints are present and the market parameters are *non-random*, and has been so used by Li, Zhou and Lim [34], who solve an MVO problem with a no-shorting portfolio constraint and deterministic market parameters. However, when the market parameters are random then one ends up with a *random* Hamilton-Jacobi-Bellman equation, and such equations are still not very well understood.

Recently, Hu and Zhou [22] have extended stochastic linear quadratic control to include MVO problems with a no-shorting constraint, in which the mean rate of return and volatility are random and the short-rate is deterministic. However, this relies on the rather sophisticated mathematical technology of extended stochastic Riccati equations, which are (highly) nonlinear backward stochastic differential equations. It is not clear how this approach might be extended to deal with regime-switching market models (this problem could make a very interesting research topic in its own right).

Finally, the Pontryagin Maximum Principle is a result of great theoretical significance, but applies only with great difficulty to concrete problems, even in the non-random setting. These difficulties multiply considerably when one tries to apply the Pontryagin Maximum Principle to stochastic problems.

With the preceding in mind, we base our approach on the application of convex duality. Convex duality methods establish a connection between the original problem, called the *primal problem*, and another problem, called the *dual problem*. The hope is that the dual problem is easier to solve than the primal problem. The convexity properties of the primal problem are critical in establishing the connection between this problem and the corresponding dual problem. Using the solution to the dual problem, this connection allows us to construct the solution to the primal problem.

As with stochastic control methods, convex duality methods have many applications, such as in manufacturing, mechanics and economics. Convex duality methods are generally more powerful than stochastic control methods when the problem is convex. This is because they exploit the convexity aspect of the problem, which stochastic control methods do not.

Convex duality theory was first applied to portfolio selection problems by Bismut [4]. He solved a problem similar to the one solved by Merton [38] by applying the convex duality theory from Bismut [3]. Bismut [4] also used a stock price model similar to the one used by Merton, where the stock prices have market coefficients which depend on the stock price, as in (2.1.2).

It was the martingale methodology developed in the papers of Harrison and Kreps [18] and Harrison and Pliska [19], [20] which allowed the application of convex duality theory to a more general market model than the one used by Merton and Bismut, given by (2.1.2). This was first done by Pliska [41], who maximized the expected utility of terminal wealth. Cox and Huang [9], [10] and Karatzas, Lehoczky and Shreve [28] then used convex duality theory to solve a problem of maximizing the expected utility of consumption and terminal wealth. Since then, stochastic duality theory has proved remarkably successful as a method of solving portfolio selection problems, again because of its ability to exploit the underlying convexity.

Our comments have so far dealt with the application of convex duality to problems in which there are no portfolio constraints, that is at every instant the investor can *freely distribute* the wealth among all of the assets. We now turn to problems in which portfolio constraints are present, which, as we noted earlier, tends to make the problems much more challenging.

Karatzas, Lehoczky, Shreve and Xu [29] successfully tackled a problem of utility maximization in an incomplete market. The incomplete market consists of a risk-free asset and m stocks, where the price processes of the m stocks are driven by a d-dimensional Brownian motion. Incompleteness arises since m is assumed to be strictly smaller than d. This means that it is not possible to find a portfolio process to remove, or hedge, the risk inherent in every random payment.

Their method of solution involves a *completion* of the incomplete market. This is called a fictitious completion, since the market is completed with fictitious stocks. The fictitious stocks are carefully chosen so that the optimal portfolio will not be invested in them. The optimal portfolio process in the fictitious market will then be a potential solution in the original, incomplete market. The authors construct many fictitious markets and find the optimal portfolio process in each one. The optimal solution in the original, incomplete market is then the optimal portfolio process which minimizes the expected utility of terminal wealth.

A portfolio selection problem involving a no-short-selling portfolio constraint was solved in 1992 by Shreve and Xu [49] using convex duality. They solved a terminal wealth and consumption utility maximization problem with a no-short-selling portfolio constraint. Cvitanić and Karatzas [11] solved the same problem with a general portfolio constraint, where the portfolio process was constrained to lie in a closed, convex set. Effectively, their result subsumed all the previous results involving portfolio constraints. For example, incomplete markets and no-short-selling portfolio constraints become special cases of their result. They used the idea of fictitious completion to solve the problem without portfolio constraints in a fictitious market. The solution to the unconstrained problem was given in the paper by Karatzas, Lehoczky and Shreve [28]. Cvitanić and Karatzas [11] construct many fictitious markets and show that there is one fictitious market where the optimal portfolio process satisfies the portfolio constraint. This optimal portfolio process, which is optimal for the unconstrained problem in some fictitious market, is also optimal for the constrained problem in the original market.

These consistent successes in the application of convex duality to portfolio selection problems seem to render redundant the application of stochastic control theory. However, there is a major hurdle to be overcome when applying convex duality theory, which is why stochastic control theory remains a valuable alternative. As Rogers [44] points out, there has been a severe lack of transparency when applying convex duality methods. A candidate dual problem is produced and then subsequently shown to work as desired. How is the dual problem arrived at? The papers cited above evince very little clue. This is particularly the case for problems with portfolio constraints, for which a dual problem is established on the basis of the aforementioned fictitious markets. However, the origin of the fictitious markets is itself rather obscure and the introductory comments in Cvitanić and Karatzas [11] suggest that the construction of the fictitious markets is the outcome of a good deal of patient experimentation (the situation is not unlike that of solving a complicated differential equation, in which one might patiently experiment with different candidate solutions to eventually come up with the actual solution, the correctness of which is verified by substitution).

There is a way of applying convex duality theory in a more straightforward manner. Rogers [44] engineered a method which synthesizes the dual problem from the primal problem in a systematic way. He did this for a class of utility maximization problems by considering the wealth dynamics of the investor as a constraint. His methods give sufficient conditions for zero "duality gap" between the values of the primal problem and the dual problem. However, he did not address the issue of how to actually construct the optimal solution.

It was Labbé and Heunis [33] who, inspired by the work of Bismut [4], constructed the solution to the primal problem directly from the solution to the dual problem. They used *conjugate duality* to directly formulate the dual problem in terms of the primal problem. This is the key to generating mathematical relations between the solution to the primal problem and the solution to the dual problem. Having shown that the solution to the dual problem exists, they used these relations to construct the solution to the primal problem from that of the dual problem. It is this approach that we use when applying convex duality theory to solve the MVO problem with regime-switching.

By constructing the dual problem directly, we obviate the need to guess the solution to the primal problem. Indeed, we show that the solution to the primal problem arises quite naturally from the solution to the dual problem.

2.4 Summary

In this chapter, we examined the possible market models, portfolio selection problems and methods of solution that fall within the realm of the area of research.

The regime-switching market model with random market coefficients that we use offers significant advantages over regime-switching models with Markov-modulated market coefficients. The most obvious advantage is that random market coefficients allow stochastic volatility models to be used. The model is set out in full detail in Section 3.2.1.

The MVO problem with portfolio constraints can be considered a more objective problem that a utility maximization problem. The risk-return trade-off is explicit in the solution to the MVO problem. It is also an easier problem to solve and the path to its solution should aid in the path to the solution to the utility maximization problem for markets which incorporate regime-switching. The problem is set out in full detail in Section 3.2.2.

The method of solution, which uses convex duality theory, has been previously applied to similar portfolio selection problems without regime-switching to great effect. In Chapter 4, we apply it to the MVO problem with portfolio constraints in a market where regime-switching is present. An overview of the convex duality method is set out in Subsection 3.2.3. A more specific description of the steps required to solve our MVO problem using the convex duality method is set out in Subsection 3.2.4.

Chapter 3

Context and theoretical framework

Having given a rather non-technical summary of the thesis in Chapter 2, in this chapter we describe the main elements of the thesis in much more complete mathematical and technical detail. As in Chapter 2, we consider separately the market model, portfolio selection problem and method of solution. However, before we do this, we outline a paper by Zhou and Yin [53]. The rationale for this is that it illustrates the advantages and disadvantages of our market model, portfolio selection problem and method of solution. The work of Zhou and Yin [53] is concerned with a market model and portfolio optimization problem which has several similarities to the market model and portfolio optimization problem studied in this thesis. Indeed, the work [53] was the main motivation for our own efforts. There are, however, also significant differences. In particular, our market model allows for completely random market coefficients, and the optimization problem includes general convex portfolio constraints (as opposed to the Markov-modulated market coefficients and absence of portfolio constraints in Zhou and Yin [53]). The most obvious difference is the method of solution: Zhou and Yin [53] use a stochastic control approach, whereas we use convex duality.

3.1 A stochastic linear quadratic control theory approach

In the present section, we headline the main elements of the MVO problem with regime-switching as studied by Zhou and Yin [53]. We do so because this is the most impressive and comprehensive work so far on this problem, and directly motivates our own interest in this problem. This summary will also be useful in providing some points of comparison with our own approach to this problem.

3.1.1 The market model

Modeling the regime-switching

In order to model the regime-switching, Zhou and Yin [53] introduce a Markov chain $\{\alpha(t)\}$ taking values in a finite state space $M = \{1, \ldots, l\}$. The Markov chain has a generator $Q = (q_{ij})_{l\times l}$ and time-homogeneous, transition probabilities

$$p_{ij}(t) := \mathbb{P}\left[\alpha(t) = j |\, \alpha(0) = i\right], \quad \forall t \in [0,T], \quad \forall i,j = 1,\ldots,l. \tag{3.1.1}$$

The states of the Markov chain $\{\alpha(t)\}$ represent the market regimes. For example, suppose a market with two regimes is modeled, where the two regimes are a bull market and a bear market. Assign the bull market regime the number 1 and the bear market regime the number 2. If at time s the market is in a bull market then $\alpha(s) = 1$. Similarly, if the market is in a bear market at time $t > s$ then $\alpha(t) = 2$.

Defining the filtration

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a fixed complete probability space on which is defined a standard N-dimensional Brownian motion $\mathbf{W}(t) \equiv (W_1(t), \ldots, W_N(t))^\top$ and the continuous-time Markov chain $\{\alpha(t)\}$, whose structure is outlined above, such that the processes $\{\mathbf{W}(t)\}$ and $\{\alpha(t)\}$ are independent.

Define the filtration

$$\mathcal{F}_t = \sigma\{\mathbf{W}(s), \alpha(s) : s \in [0,t]\}. \tag{3.1.2}$$

The filtration $\{\mathcal{F}_t\}$ represents the information available to the investor at each time t.

Stocks and the bank account

Consider a market in which $N+1$ assets are traded continuously. One of the assets is a bank account whose price $P_0(t)$ is given by

$$\mathrm{d}P_0(t) = r(t, \alpha(t))\, P_0(t)\ \mathrm{d}t, \quad \forall t \in [0,T], \tag{3.1.3}$$

subject to $P_0(0) = p_0 > 0$, where $r(\cdot, i) : [0,T] \to [0,\infty)$ is the deterministic interest rate processes corresponding to regime i, for $i = 1, \ldots, l$.

The other N assets are stocks whose price processes $P_m(t)$, for $m = 1, \ldots, N$, satisfy the system of stochastic differential equations

$$\mathrm{d}P_m(t) = P_m(t) \left(b_m(t, \alpha(t))\ \mathrm{d}t + \sum_{n=1}^{N} \sigma_{mn}(t, \alpha(t))\ \mathrm{d}W_n(t) \right), \quad \forall t \in [0,T], \tag{3.1.4}$$

subject to $P_m(0) = p_m > 0$, where for each $i = 1, \ldots, l$, $b_m(\cdot, i) : [0, T] \to \mathbb{R}$ is the deterministic mean rate of return process and

$$\sigma_m(\cdot, i) := (\sigma_{m1}(\cdot, i), \sigma_{m2}(\cdot, i), \ldots, \sigma_{mN}(\cdot, i)) : [0, T] \to \mathbb{R}^N \tag{3.1.5}$$

is the deterministic volatility process of the mth stock, corresponding to the regime i.

We define the volatility matrix

$$\boldsymbol{\sigma}(t, i) := (\sigma_{mn}(t, i))_{N \times N} \quad \text{for each } i = 1, \ldots, l. \tag{3.1.6}$$

Throughout [53], there is the usual non-degeneracy assumption:

Condition 3.1.1. There exists a constant $\kappa \in (0, \infty)$ such that

$$\mathbf{z}^\top \boldsymbol{\sigma}(t, i) \boldsymbol{\sigma}^\top(t, i) \mathbf{z} \geq \kappa \|\mathbf{z}\|^2 \qquad \text{for all } (\mathbf{z}, t, i) \in \mathbb{R}^N \times [0, T] \times \{1, \ldots, l\}, \tag{3.1.7}$$

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

The following usual condition is also assumed in [53].

Condition 3.1.2. All the functions $r(t, i)$, $b_m(t, i)$ and $\sigma_{mn}(t, i)$ are Borel-measurable and uniformly bounded in t.

Suppose that the initial market mode is $\alpha(0) = i_0$. The trading of shares takes place continuously.

3.1.2 The mean-variance portfolio optimization problem

The investor

Consider an investor with an initial wealth $x_0 > 0$. We assume that a self-financing strategy is followed. The total wealth of an investor in the market at time t is denoted by $x(t)$. Let $u_m(t)$ represent the total market value of the investor's wealth in the mth asset at time t, for $m = 1, \ldots, N$.

Definition 3.1.3. $\boldsymbol{u}(\cdot) = (u_1(\cdot), \ldots, u_N(\cdot))^\top$ is called a portfolio of the investor.

Once $\boldsymbol{u}(\cdot)$ is determined, $u_0(\cdot)$, the asset in the bank account, is completely specified since $u_0(t) = x(t) - \sum_{i=1}^N u_i(t)$. Thus, in the analysis to follow, only $\boldsymbol{u}(\cdot)$ is considered.

Ignoring transaction costs and consumption, under the self-financing assumption the wealth of the investor satisfies the following well-known relation

$$\begin{aligned} \mathrm{d}x(t) = &\left(r(t, \alpha(t))\, x(t) + \sum_{m=1}^N (b_m(t, \alpha(t)) - r(t, \alpha(t)))\, u_m(t) \right) \mathrm{d}t \\ &+ \sum_{n=1}^N \sum_{m=1}^N \sigma_{mn}(t, \alpha(t))\, u_m(t)\, \mathrm{d}W_n(t), \end{aligned} \tag{3.1.8}$$

with initial conditions $x(0) = x_0 > 0$ and $\alpha(0) = i_0$.

Setting

$$\boldsymbol{B}(t,i) := (b_1(t,i) - r(t,i), \dots, b_N(t,i) - r(t,i)) \quad \text{for } i = 1, \dots, l, \tag{3.1.9}$$

the wealth equation (3.1.8) can be rewritten more compactly as

$$\mathrm{d}x(t) = \left(r(t,\alpha(t))\, x(t) + \boldsymbol{B}(t,\alpha(t))\, \boldsymbol{u}(t)\right)\, \mathrm{d}t + \boldsymbol{u}^\top(t)\, \boldsymbol{\sigma}(t,\alpha(t))\, \mathrm{d}\mathbf{W}(t), \tag{3.1.10}$$

with initial conditions $x(0) = x_0 > 0$ and $\alpha(0) = i_0$.

Definition 3.1.4. A portfolio $\boldsymbol{u}(\cdot)$ is said to be admissible if $\boldsymbol{u}(\cdot)$ is an $\mathbb{R}^N$-valued, $\{\mathcal{F}_t\}$–adapted stochastic process which is square-integrable, that is $\mathrm{E} \int_0^T \|\boldsymbol{u}(t)\|^2 \, \mathrm{d}t < \infty$. In this case, $(x(\cdot), \boldsymbol{u}(\cdot))$ is referred to as an admissible (wealth, portfolio) pair.

The implication of the portfolio process $\boldsymbol{u}(\cdot)$ being $\{\mathcal{F}_t\}$–adapted is that at each time t, the portfolio $\boldsymbol{u}(t)$ is a non-random function of the paths $\{\mathbf{W}(s) : s \in [0,t]\}$ and $\{\alpha(s) : s \in [0,t]\}$ (by Doob's theorem).

The investor's problem

The investor's objective is to find an admissible portfolio $\boldsymbol{u}(\cdot)$ among all the admissible portfolios whose expected terminal wealth is $\mathrm{E}(x(T)) = z$ for some specified $z \in \mathbb{R}$, so that the risk measured by the variance of the terminal wealth,

$$\mathrm{var}(x(T)) \equiv \mathrm{E}\left(x(T) - z\right)^2, \tag{3.1.11}$$

is minimized. Finding such a portfolio $\boldsymbol{u}(\cdot)$ is referred to as the mean-variance portfolio selection problem. The problem is formulated below.

Problem 3.1.5. The mean-variance portfolio selection is a constrained stochastic optimization problem, parametrized by $z \in \mathbb{R}$:

$$\begin{cases} \text{minimize} & J_{MV}(x_0, i_0, \boldsymbol{u}(\cdot)) := \mathrm{E}\left(x(T) - z\right)^2, \\ \text{subject to} & \mathrm{E}(x(T)) = z \text{ and } (x(\cdot), \boldsymbol{u}(\cdot)) \text{ admissible.} \end{cases} \tag{3.1.12}$$

Moreover, the problem is called feasible if there is at least one portfolio satisfying all the constraints. The problem is called finite if it is feasible and the infimum of $J_{MV}(x_0, i_0, \boldsymbol{u}(\cdot))$ over the set of admissible pairs $(x(\cdot), \boldsymbol{u}(\cdot))$ is finite.

3.1.3 A stochastic linear quadratic control method

The approach of Zhou and Yin [53] is summarized below.

Stage 1 Show that the problem with the terminal wealth constraint $\mathrm{E}(x(T)) = z$ is feasible.

Stage 2 Construct an unconstrained problem by using a Lagrange multiplier to "remove" the terminal wealth constraint $\mathrm{E}(x(T)) = z$.

Stage 3 Solve the unconstrained problem using stochastic linear quadratic control theory.

Stage 4 Use the solution to the unconstrained problem to solve the constrained problem.

Now consider each of the above stages in more detail.

Stage 1 Show that the problem with the terminal wealth constraint is feasible

Recall from Problem 3.1.5 that there is a terminal wealth constraint of the form $\mathrm{E}(x(T)) = z$. First, Zhou and Yin show that Problem 3.1.5 is feasible for every $z \in \mathbb{R}$. In other words, no matter what value of terminal wealth z we specify, we can find at least one portfolio which satisfies all the constraints in Problem 3.1.12. This requires the assumption that

$$\mathrm{E} \int_0^T |\boldsymbol{B}(t, \alpha(t))|^2 \, \mathrm{d}t > 0, \tag{3.1.13}$$

which holds if there is one stock m and one market mode i such that the appreciation-rate process $b_m(\cdot, i)$ is not equal to the interest rate process $r(\cdot, i)$ over a non-null Borel-measurable set.

Stage 2 Construct an unconstrained problem by using a Lagrange multiplier to "remove" the terminal wealth constraint

Next, note that Problem 3.1.5 is a dynamic optimization problem with a constraint $\mathrm{E}(x(T)) = z$. To handle the constraint, apply the Lagrange multiplier technique and define the usual Lagrangian

$$\begin{aligned} J(x_0, i_0, \boldsymbol{u}(\cdot), \lambda) :&= \mathrm{E}\left(|x(T) - z|^2 + 2\lambda\left(x(t) - z\right)\right) \\ &= \mathrm{E}\left(x(T) + \lambda - z\right)^2 - \lambda^2, \end{aligned} \tag{3.1.14}$$

in which $\lambda \in \mathbb{R}$ is the Lagrange multiplier which enforces the constraint $\mathrm{E}(x(T)) = z$.

Problem 3.1.6. The unconstrained problem can be stated as follows.

$$\begin{cases} \text{minimize} & J(x_0, i_0, \boldsymbol{u}(\cdot), \lambda) := \mathrm{E}\left(x(T) + \lambda - z\right)^2 - \lambda^2, \\ \text{subject to} & (x(\cdot), \boldsymbol{u}(\cdot)) \text{ admissible.} \end{cases} \tag{3.1.15}$$

Stage 3 Solve the unconstrained problem using stochastic linear quadratic control theory

Problem 3.1.6, which is the unconstrained problem, is a Markov-modulated stochastic LQ optimal control problem. Thus, we can use stochastic LQ control methods to solve it.

First we define the functions $H(t,i)$ and $P(t,i)$ to be the (unique) solutions to the two systems of ordinary differential equations given below.

$$\begin{cases} \dot{P}(t,i) = (\rho(t,i) - 2r(t,i))\ P(t,i) - \sum_{j=1}^{l} q_{ij} P(t,j) \\ P(T,i) = 1 \quad i = 1,2,\ldots,l, \end{cases} \tag{3.1.16}$$

and

$$\begin{cases} \dot{H}(t,i) = r(t,i)H(T,i) - \frac{1}{P(t,i)} \sum_{j=1}^{l} q_{ij} P(t,j)\,(H(t,i) - H(t,j)) \\ H(T,i) = 1 \quad i = 1,2,\ldots,l, \end{cases} \tag{3.1.17}$$

where

$$\rho(t,i) := \boldsymbol{B}(t,i) \left(\boldsymbol{\sigma}(t,i)\boldsymbol{\sigma}^\top(t,i)\right)^{-1} \boldsymbol{B}^\top(t,i), \quad i = 1,2,\ldots,l. \tag{3.1.18}$$

From (3.1.16), (3.1.17) and (3.1.18), we see that we can find values for $H(t,i)$ and $P(t,i)$ by knowing only the market model, which will give us the values of $\{\boldsymbol{B}(t,i)\}$, $\{\boldsymbol{\sigma}(t,i)\}$ and $\{q_{ij}\}$. We do not need to know anything about the investor.

Using the functions $H(t,i)$ and $P(t,i)$, the optimal solution to Problem 3.1.6 is given by the following theorem.

Theorem 3.1.7. *Problem 3.1.6 has an optimal feedback control*

$$\boldsymbol{u}^\star(t,x,i) = \left(\boldsymbol{\sigma}(t,i)\boldsymbol{\sigma}^\top(t,i)\right)^{-1} \boldsymbol{B}^\top(t,i)\ (x + (\lambda - z)H(t,i)). \tag{3.1.19}$$

Moreover, the corresponding optimal value is

$$\begin{aligned} \inf_{\boldsymbol{u}(\cdot)\,admissible} J(x_0, i_0, \boldsymbol{u}(\cdot), \lambda) &= \left(P(0,i_0)\, H^2(0,i_0) + \theta - 1\right)\ (\lambda - z)^2 \\ &\quad + 2\,(P(0,i_0)H(0,i_0)x_0 - z)\,(\lambda - z) + P(0,i_0)x_0^2 - z^2, \end{aligned} \tag{3.1.20}$$

where

$$\begin{aligned} \theta &:= E\int_0^T \sum_{j=1}^{l} q_{\alpha(t)j} P(t,j)\,(H(t,j) - H(t,\alpha(t)))^2\ \mathrm{d}t \\ &= \sum_{i=1}^{l}\sum_{j=1}^{l} \int_0^T P(t,j)p_{i_0 i}(t)q_{ij}\,(H(t,j) - H(t,i))^2\ \mathrm{d}t \\ &\geq 0, \end{aligned} \tag{3.1.21}$$

with the transition probabilities $p_{i_0 i}(t)$ given by (3.1.1).

Stage 4 Use the solution to the unconstrained problem to solve the constrained problem

The solution to Problem 3.1.5, which includes the terminal wealth constraint $\mathrm{E}(x(T)) = z$, is given by the following theorem.

Theorem 3.1.8. *Assume that (3.1.13) holds. Define*

$$\lambda^\star := z + \frac{z - P(0,i_0)H(0,i_0)x_0}{P(0,i_0)H(0,i_0)^2 + \theta - 1}. \tag{3.1.22}$$

The optimal portfolio for Problem 3.1.5 is

$$\boldsymbol{u}^\star(t,x,i) = -\left(\boldsymbol{\sigma}(t,i)\boldsymbol{\sigma}^\top(t,i)\right)^{-1} \boldsymbol{B}^\top(t,i)\ (x + (\lambda - z)H(t,i)), \tag{3.1.23}$$

where $P(0,i_0)H(0,i_0)^2 + \theta - 1 < 0$.

3.2 A convex duality theory approach

Now we turn to the MVO problem that we solve in Chapter 4. Recall that this is an MVO problem with portfolio constraints in a regime-switching model. We begin by describing the market model in which the MVO problem is set. The market consists of a bank account and a number of stocks. We assume that the market is subject to regime-switches from time to time. For example, suppose the market is in a bull market, in which stock prices are mostly rising. The bull market can be considered as a regime. Suddenly, there is a stock market crash. The market enters a bear market, in which stock prices are mostly falling. The bear market is another regime, so this is an example of a regime-switch; the market switches from the bull market regime to the bear market regime.

We make the assumption that at 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. A filtration is a structure which contains all the events which could have occurred up to each time t. In this context, the filtration is often called the information filtration.

We construct the filtration from a Brownian motion, which drives the stock prices, and a Markov chain, which models the regime-switching. In other words, we are assuming that the information available to the investor consists of these two items.

The financial market we describe is almost the same as the one in Zhou and Yin [53], except for one vital difference. In their market model, the market coefficients are exclusively Markov-modulated processes and otherwise have no randomness. We remove this restriction and allow the market coefficients to be genuine random processes. To highlight the advantages of using random market coefficients, we will later compare our price processes to those in Zhou and Yin [53].

After defining the market model, we want to specify the portfolio selection problem that we consider, which is an MVO problem with portfolio constraints. First we introduce the investor around whom the problem is set. We assume that the investor wishes to invest only in the market. Thus the investor will allocate all of her wealth among the bank account and the stocks. How much and where she allocates her wealth over time is called a portfolio process. We record how much wealth she invests in each of the assets; this is called the investor's portfolio. By adding up how much the investor has in each of the assets, we obtain the total wealth of the investor at any point in time. After introducing the investor, we then state the MVO problem.

Finally, we outline the convex duality method that we use to solve the MVO problem.

3.2.1 The market model

Before defining the asset price processes, it is necessary to define the regime-switching Markov chain and the Brownian motion which drives the stock price processes. We also need to specify the probability space on which the stock price processes are defined. Once we have the probability space, we define a filtration. We use the filtration to define a measurability property required of the stochastic integrands.

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

The probability space

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

Modeling the regime-switching

We model the regime-switching using a continuous-time Markov chain. We assume that there are only finitely many possible regimes, so the Markov chain will take values in a finite state space.

Mathematically, denote the regime-switching Markov chain as $\alpha = \{\alpha(t) : t \in [0, T]\}$. The Markov chain α is defined on the probability space $(\Omega, \mathcal{G}, \mathbb{P})$ and takes values in a finite state space

$$I = \{1, \dots, D\}. \tag{3.2.1}$$

The Markov chain α has a generator Q which is a $D \times D$ matrix $Q = (q_{ij})_{i,j=1}^{D}$ with the properties

$$q_{ij} \geq 0 \quad \forall i \neq j \quad \text{and} \quad -q_{ii} = \sum_{j \neq i} q_{ij}. \tag{3.2.2}$$

In Section B.1, we give some properties of the Markov chain α.

Defining the Brownian motion

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

Definition 3.2.1. A standard, N-dimensional Brownian motion is 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 $\mathcal{F}_s^W := \sigma\{\mathbf{W}(u) : u \in [0,s]\}$, where I_N is the $N \times N$ identity matrix.

Independence assumption

The following independence condition is necessary in order to define the joint filtration. However, we also believe that it 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 macro-economic effects which occur over much longer time periods.

Condition 3.2.2. The Brownian motion $\mathbf{W}$ is independent of the Markov chain α, in other words

$$\mathbb{P}[A \cap B] = \mathbb{P}[A]\mathbb{P}[B], \quad \forall A \in \mathcal{F}_T^\alpha \quad \forall B \in \mathcal{F}_T^{\mathbf{W}}, \tag{3.2.3}$$

for $\mathcal{F}_T^\alpha := \sigma\{\alpha(t) : t \in [0,T]\}$ and $\mathcal{F}_T^W := \sigma\{\mathbf{W}(t) : t \in [0,T]\}$.

Generating the filtration and defining previsibility

Now that we have defined the Brownian motion $\mathbf{W}$ and the regime-switching Markov chain α, we can construct the filtration.

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

$$\mathcal{F}_t^\circ := \sigma\{\alpha(s), \mathbf{W}(s) : s \in [0,t]\}, \quad \forall t \in [0,T]. \tag{3.2.4}$$

We want the filtration to have the usual regularity properties, namely that it contains all the $\mathbb{P}$-null sets in the σ–algebra $\mathcal{G}$ and is right-continuous. This allows us to use the usual results of stochastic calculus. Thus we define the standard filtration $\{\mathcal{F}_t : t \in [0,T]\}$ as

$$\mathcal{F}_t := \bigcap_{s\in(t,T]} \mathcal{F}_s^\circ \vee \mathcal{N}(\mathbb{P}), \quad \forall t \in [0,T], \tag{3.2.5}$$

where $\mathcal{N}(\mathbb{P}) := \{N \in \mathcal{G} : \mathbb{P}(N) = 0\}$ is the set of all $\mathbb{P}$-null subsets of Ω.

We define a σ–algebra $\mathcal{F}$ on Ω as

$$\mathcal{F} := \mathcal{F}_T. \tag{3.2.6}$$

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

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 previsibility. The previsible σ–algebra on $\Omega \times [0,T]$ associated with the filtration $\{\mathcal{F}_t : t \in [0,T]\}$ is defined next.

Definition 3.2.3. Let $\mathcal{P}^\star$ be the previsible σ–algebra on $\Omega \times [0,T]$ associated with the filtration $\{\mathcal{F}_t : t \in [0,T]\}$. Thus $\mathcal{P}^\star$ is the smallest σ–algebra on $\Omega \times [0,T]$ such that every $\{\mathcal{F}_t\}$–adapted, $\mathbb{R}$-valued process which is left-continuous with right-hand limits is $\mathcal{P}^\star$–measurable.

A process X is called *previsible* if it is $\mathcal{P}^\star$–measurable and we write $X \in \mathcal{P}^\star$.

The canonical martingales of the Markov chain

Associated with the Markov chain α are a set of canonical martingales $\{\mathcal{Q}_{ij} : i,j \in I, i \neq j\}$. Their construction and properties are detailed in Section B.3. We summarize the relevant findings here.

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

$$\mathcal{Q}_{ij}(t) := \begin{cases} \sum_{0<s\leq t} \chi[\alpha(s_-) = i]\,\chi[\alpha(s) = j] - q_{ij} \int_0^t \chi[\alpha(s) = i]\,\mathrm{d}s & \text{if } i \neq j \\ 0 & \text{if } i = j. \end{cases} \tag{3.2.7}$$

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

$$\chi[\alpha(s) = i] = \begin{cases} 1 & \text{if } \alpha(s) = i \\ 0 & \text{otherwise,} \end{cases} \tag{3.2.8}$$

(see Definition B.3.1, Definition B.3.4 and Definition B.3.7). We define $\mathcal{Q}_{ii}$ for notational convenience and we let $\mathcal{Q}_{ii} := 0$ to make it clear that the set $\{\mathcal{Q}_{ii} : i \in I\}$ is not part of the set of canonical martingales of the Markov chain.

We have $\mathcal{Q}_{ij} \in \mathcal{M}_0^2(\{\mathcal{F}_t\}, \mathbb{P})$ and $\mathcal{Q}_{ij}$ is a finite-variation process (see Remark B.3.15 and Lemma B.3.16). The set of martingales $\{\mathcal{Q}_{ij} : i,j \in I, i \neq j\}$ are the *canonical martingales of the Markov chain* α. We define the $D \times D$ matrix

$$\mathcal{Q} := (\mathcal{Q}_{ij})_{i,j=1}^D \tag{3.2.9}$$

We may loosely refer to $\mathcal{Q}$ as the set of canonical martingales of the Markov chain α. However, this should be understood as excluding the diagonal elements, which are all zero.

The square-bracket quadratic variation process $[\mathcal{Q}_{ij}]$ of $\mathcal{Q}_{ij}$ satisfies a.s.,

$$[\mathcal{Q}_{ij}](t) := \begin{cases} \sum_{0<s\leq t} \chi[\alpha(s_-) = i]\,\chi[\alpha(s) = j] & \text{if } i \neq j \\ 0 & \text{if } i = j, \end{cases} \quad \forall t \in [0,T], \tag{3.2.10}$$

and the square-bracket quadratic co-variation process $[\mathcal{Q}_{ab}, \mathcal{Q}_{ij}]$ satisfies a.s.,

$$[\mathcal{Q}_{ab}, \mathcal{Q}_{ij}](t) := \begin{cases} [\mathcal{Q}_{ij}](t) & \text{if } (a,b) = (i,j) \\ 0 & \text{if } (a,b) \neq (i,j), \end{cases} \quad \forall t \in [0,T], \tag{3.2.11}$$

(see Definition B.3.1 and Lemma B.3.19).

Almost everywhere

Sometimes, we wish to say that a statement holds almost everywhere ("a.e."). As we are dealing with stochastic integrals with either a Brownian motion or the canonical martingales of the Markov chain as integrator, we have to specify which measure the "a.e." is with respect to. The stochastic processes are defined on the measurable space $(\Omega \times [0,T], \mathcal{P}^\star)$ and when we say that a statement holds a.e., we state which measure on this measurable space we are referring to.

One measure on the measurable space $(\Omega \times [0,T], \mathcal{P}^\star)$ we will refer to is $\mathbb{P} \otimes Leb$, where Leb represents Lebesgue measure on the Borel σ–algebra on $[0,T]$.

Another measure on the measurable space $(\Omega \times [0,T], \mathcal{P}^\star)$ is $\nu_{[\mathcal{Q}_{ij}]}$, which is defined by the recipe

$$\nu_{[\mathcal{Q}_{ij}]}[A] := \mathrm{E} \int_0^T \chi_A(\omega, t)\ \mathrm{d}[\mathcal{Q}_{ij}](t), \quad \forall A \in \mathcal{P}^\star, \tag{3.2.12}$$

for each $i, j = 1, \ldots, D$, $i \neq j$.

Notation 3.2.4. By

$$\mathbf{G} = \mathbf{H} \quad \nu_{[\mathcal{Q}]}\text{-a.e.} \tag{3.2.13}$$

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 $\Omega \times [0,T]$, we mean that

$$G_{ij} = H_{ij} \quad \nu_{[\mathcal{Q}_{ij}]}\text{-a.e.}, \quad \forall i, j \in I,\ i \neq j, \tag{3.2.14}$$

and

$$G_{ii} = H_{ii} \quad (\mathbb{P} \otimes Leb)\text{-a.e.}, \quad \forall i \in I. \tag{3.2.15}$$

Stocks and the bank account

The assets in the market consist of a bank account and N stocks. In this section, we give the price processes of the assets and we compare them to the price processes in Zhou and Yin [53].

We start by considering the bank account. The price at time t of one unit holding in the bank account will be denoted $S_0(t)$, with the convention that $S_0(0) = 1$. The price process of the bank account is governed by the equation

$$\mathrm{d}S_0(t) = S_0(t)r(t)\,\mathrm{d}t, \quad \forall t \in [0, T], \tag{3.2.16}$$

where $r(t)$ is called the *risk-free interest rate process* at time t.

Condition 3.2.5. The risk-free interest rate process $\{r(t)\}$ is a uniformly bounded, nonnegative, previsible, $\mathbb{R}$-valued process on the set $\Omega \times [0, T]$.

We next compare the risk-free interest rate process found in (3.2.16), which is the one that we propose using, to that found in (3.1.3), from the paper of Zhou and Yin [53]. We will use this comparison to demonstrate the increased generality of our model.

The risk-free interest rate process in (3.1.3) is of the form $r(t, \alpha(t))$, so it is a function on the set $[0, T] \times I$. Thus its randomness is muted as it depends on the set Ω only through the Markov chain $\alpha(t)$. Within each regime, that is when the Markov chain $\alpha(t)$ takes a fixed value, the process $r(t, \alpha(t))$ becomes a deterministic function. By comparison, the risk-free interest rate process $r(t)$ given in (3.2.16) is a function on the set $\Omega \times [0, T]$, which means that it is a random process even within a regime.

Since the process $r(t, \alpha(t))$ is deterministic within each regime, then the number of these deterministic processes is limited by the number of regimes in the model. For our model, the risk-free interest rate process $r(t)$ is not deterministic within regimes and so is not restricted by the number of regimes. This allows much greater flexibility when modeling the evolution of future risk-free interest rates.

We give below a simple example which demonstrates the advantage of our model.

Example 3.2.6. Consider a market in which there are only two regimes; a high-interest rate environment, in which the Markov chain α takes the value 1, and a low-interest rate environment, in which the Markov chain α takes the value 2. Consider the kind of models that we can fit to the risk-free interest rate process.

Consider the risk-free interest rate process $\{r(t, \alpha(t))\}$ of Zhou and Yin [53]. As there are only two regime states in this example, this risk-free interest rate process can follow only one of two possible deterministic processes at any time. We can express this mathematically as

$$r(t, \alpha(\omega, t)) = \begin{cases} f(t) \text{ if } \alpha(\omega, t) = 1 \text{ (high-interest rate environment)} \\ g(t) \text{ if } \alpha(\omega, t) = 2 \text{ (low-interest rate environment),} \end{cases} \tag{3.2.17}$$

where $f(t) := r(t,1)$ and $g(t) := r(t,2)$ for all $t \in [0,T]$. The functions $f, g : [0,T] \to [0,\infty)$ are deterministic processes, so fitting a stochastic model of future risk-free interest rates within each regime is not possible. However, most interest rate models that are used in practice are stochastic (see Hunt and Kennedy [24] for examples). The inability of their Markov-modulated risk-free interest rate process to fit stochastic interest rate models is a limitation on its practical implementation.

Now consider our risk-free interest rate process $\{r(t)\}$, which is a stochastic process. We can certainly fit all the models that the model of Zhou and Yin [53] can fit, by defining

$$r(\omega,t) = \begin{cases} f(t) & \text{if } \alpha(\omega,t) = 1 \text{ (high-interest rate market)} \\ g(t) & \text{if } \alpha(\omega,t) = 2 \text{ (low-interest rate market)}, \end{cases} \tag{3.2.18}$$

where f and g are as the same functions as the ones in (3.2.17).

We can also fit interest rate models which are stochastic processes. For example, we could fit a one-factor model such as the Black-Karasinski model, as follows:

$$r(\omega,t) = \begin{cases} r_1(\omega,t) & \text{if } \alpha(\omega,t) = 1 \text{ (high-interest rate environment)} \\ r_2(\omega,t) & \text{if } \alpha(\omega,t) = 2 \text{ (low-interest rate environment)}, \end{cases} \tag{3.2.19}$$

where r_1 and r_2 satisfy the following stochastic differential equations

$$\mathrm{d}\left(\ln r_1(t)\right) = \left(h_1(t) - v_1 \ln r_1(t)\right)\, \mathrm{d}t + s_1(t)\,\mathrm{d}B_1(t) \tag{3.2.20}$$

$$\mathrm{d}\left(\ln r_2(t)\right) = \left(h_2(t) - v_2 \ln r_2(t)\right)\, \mathrm{d}t + s_2(t)\,\mathrm{d}B_2(t). \tag{3.2.21}$$

The functions $h_k, s_k, v_k : [0,T] \to \mathbb{R}$ are deterministic and B_k is a standard 1-dimensional Brownian motion, for $k = 1, 2$. Note that the interest rate processes r_1, r_2 are positively-valued under the Black-Karasinski model, which agrees with Condition 3.2.5.

Moreover, our risk-free interest rate process is not limited by the number of regimes in the model. In the above examples, the interest rate models change according to which regime the market is in. Suppose instead we wish to fit a interest rate model which changes according to how often the market has switched regimes over the last ten years. This could reflect the stability of the market; the more often the market has switched regimes in, say, the last ten years, the less stable is the market. In a less stable market, the risk-free interest rate may fluctuate more wildly than in a more stable market. We could fit such a model within our regime-switching model. However, we do not know a practical example to illustrate this.

The above example demonstrates the increased power of our model when modeling the risk-free interest rate process. Now we specify the stock price processes for our model. The price at time t of one unit holding in the nth stock will be denoted $S_n(t)$, with the convention that $S_n(0)$ is some positive constant, for each

$n = 1, \ldots, N$. The price process of the nth stock satisfies for each $n = 1, \ldots, N$, and for all $t \in [0, T]$,

$$\mathrm{d}S_n(t) = S_n(t)\left(b_n(t)\,\mathrm{d}t + \sum_{m=1}^{N} \sigma_{nm}(t)\,\mathrm{d}W^{(m)}(t)\right). \tag{3.2.22}$$

$b_n(t)$ is called the *mean rate of return process* of the nth 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 3.2.7. We assume that the entries of the mean rate of return process $\mathbf{b}(t) = \{b_n(t)\}_{n=1}^{N}$ and the entries of the volatility process $\boldsymbol{\sigma}(t) = \{\sigma_{nm}(t)\}_{n,m=1}^{N}$ are uniformly bounded, previsible, $\mathbb{R}$-valued processes on the set $\Omega \times [0, T]$.

Let us compare the market coefficients given in (3.2.22) to those found in (3.1.4), from Zhou and Yin [53]. We start with the mean rate of return process $\mathbf{b}$.

The mean rate of return process of the nth stock in (3.1.4) is of the form $b_n(t, \alpha(t))$, which is a function on the set $\Omega \times I$. The mean rate of return process $b_n(t)$ of the nth stock given in (3.2.22) is a function on the set $\Omega \times [0, T]$. It is clear that the comparison that we did above for the risk-free interest rate process will also hold for the mean rate of return process. As all our arguments about the risk-free interest rate process carry over exactly to the mean rate of return process, we do not repeat them here.

The same argument follows for the comparison of the volatility processes in our model and Zhou and Yin [53]. An example of a market which we could model is a hybrid of Merton's model, given by (2.1.2), and Zhou and Yin's [53] model, where the mean rate of return process is of the form $\mathbf{b}(t, S(t), \alpha(t_-))$ and the volatility process is of the form $\boldsymbol{\sigma}(t, S(t), \alpha(t_-))$. In other words, they depend on the time, the level of the stock price and the market regime.

Remark 3.2.8. Having compared the market coefficients of Zhou and Yin [53], given in (3.1.3) and (3.1.4), to the random market coefficients in Conditions 3.2.5 and 3.2.7, a natural question is to ask if we could extend or adapt their approach, which we summarized earlier, to encompass the MVO problem with portfolio constraints in our market model, where the market coefficients are genuine random processes. However, as we noted in Section 2.3, it is very challenging to apply a linear quadratic control method to problems with random market coefficients when there are portfolio constraints present. It is certainly not obvious how to do this. Indeed, as we noted earlier in connection with Hu and Zhou [22], this would make a very interesting research problem (which we leave to others more familiar with the mathematics of backward stochastic differential equations than ourselves!).

Market conditions and notations

Here we specify some conditions which are essential to the smooth operation of the market model. We also define the market price of risk.

Condition 3.2.9. There exists a constant $\kappa \in (0, \infty)$ such that

$$\mathbf{z}^\top \boldsymbol{\sigma}(\omega, t)\boldsymbol{\sigma}^\top(\omega, t)\mathbf{z} \geq \kappa \|\mathbf{z}\|^2 \qquad \forall(\mathbf{z}, \omega, t) \in \mathbb{R}^N \times \Omega \times [0, T], \tag{3.2.23}$$

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

Remark 3.2.10. Condition 3.2.9 is a very standard condition in portfolio optimization theory and, in particular, it is the natural analog of Condition 3.1.1. It implies that the matrices $\boldsymbol{\sigma}$ and $\boldsymbol{\sigma}^\top$ are non-singular and therefore invertible.

For, suppose that the matrix $\boldsymbol{\sigma}^\top$ is singular. Then there is a non-zero vector $y \in \mathbb{R}^N$ such that $\boldsymbol{\sigma}^\top y = 0$. Then $\|\boldsymbol{\sigma}^\top y\|^2 = 0$. This contradicts the positivity of the right-hand side of (3.2.23) for non-zero $y \in \mathbb{R}^N$. Since $\boldsymbol{\sigma}^\top$ is non-singular, then $\boldsymbol{\sigma}$ is also non-singular.

Remark 3.2.11. We will collectively call Conditions 3.2.2, 3.2.5, 3.2.7 and 3.2.9 the *market conditions*.

Definition 3.2.12. The *market price of risk* $\boldsymbol{\theta}$ is the mapping $\boldsymbol{\theta} : \Omega \times [0, T] \to \mathbb{R}^N$ given by

$$\boldsymbol{\theta}(\omega, t) := \boldsymbol{\sigma}^{-1}(\omega, t)(\mathbf{b}(\omega, t) - r(\omega, t)\mathbf{1}), \tag{3.2.24}$$

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

Remark 3.2.13. From Condition 3.2.5, Condition 3.2.7 and Remark 3.2.14, we see that $\boldsymbol{\theta} = \{\boldsymbol{\theta}(t) : t \in [0, T]\}$ is previsible and uniformly bounded on $\Omega \times [0, T]$.

Remark 3.2.14. In view of Condition 3.2.7 and Condition 3.2.9, there exists a constant $\kappa \in (0, \infty)$ such that

$$\max\{\|(\boldsymbol{\sigma}(\omega, t))^{-1}\mathbf{z}\|, \|\left(\boldsymbol{\sigma}^\top(\omega, t)\right)^{-1}\mathbf{z}\|\} \leq \frac{1}{\sqrt{\kappa}}\|\mathbf{z}\|, \quad \forall(\mathbf{z}, \omega, t) \in \mathbb{R}^N \times \Omega \times [0, T]. \tag{3.2.25}$$

The existence of such a constant $\kappa \in (0, \infty)$ is a standard result and the proof can be found in Karatzas, Lehoczky and Shreve [28].

Next we define some constant bounds which we use throughout the thesis. By Condition 3.2.7 and (3.2.25), there exists a constant $\kappa_{\boldsymbol{\sigma}} \in (0, \infty)$ such that

$$\max\{\|\boldsymbol{\sigma}(\omega, t)\mathbf{z}\|, \|\boldsymbol{\sigma}^\top(\omega, t)\mathbf{z}\|, \|(\boldsymbol{\sigma}(\omega, t))^{-1}\mathbf{z}\|, \|\left(\boldsymbol{\sigma}^\top(\omega, t)\right)^{-1}\mathbf{z}\|\} \leq \kappa_{\boldsymbol{\sigma}}\|\mathbf{z}\|, \tag{3.2.26}$$

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

By Condition 3.2.5, there exists a constant $\kappa_r \in (0, \infty)$ such that

$$|r(\omega, t)| \leq \kappa_r, \quad \forall(\omega, t) \in \Omega \times [0, T]. \tag{3.2.27}$$

Finally, from Remark 3.2.13, there exists a constant $\kappa_{\boldsymbol{\theta}} \in (0, \infty)$ such that

$$\|\boldsymbol{\theta}(\omega, t)\| \leq \kappa_{\boldsymbol{\theta}}, \quad \forall(\omega, t) \in \Omega \times [0, T]. \tag{3.2.28}$$

Definition 3.2.15. The processes $\{r(t)\}$, $\{\mathbf{b}(t)\}$, $\{\boldsymbol{\sigma}(t)\}$ and $\{\boldsymbol{\theta}(t)\}$ are called the *market coefficients* of the market model.

Remark 3.2.16. At each time t, we know the values of the market coefficients $r(\omega, t)$, $\mathbf{b}(\omega, t)$, $\boldsymbol{\sigma}(\omega, t)$ and $\boldsymbol{\theta}(\omega, t)$. The goal of portfolio optimization is to characterize, and, if possible, compute, the optimal portfolio in terms of these known quantities.

3.2.2 The mean-variance portfolio optimization problem

The investor

We consider an investor with an initial wealth $x_0 > 0$. The total wealth of an investor in the market at time 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 amount of wealth that the investor holds in the bank account at time t. We denote by $\pi_n(t)$ the amount of wealth that the investor holds in stock n at time t, for each $n = 1, \ldots, N$. Defining the vector

$$\boldsymbol{\pi}(t) := (\pi_1(t), \ldots, \pi_N(t))^\top, \tag{3.2.29}$$

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

$$X^{\boldsymbol{\pi}}(t) = \pi_0(t) + \boldsymbol{\pi}^\top(t)\mathbf{1}, \qquad \forall t \in [0,T], \tag{3.2.30}$$

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

Note that the value of $\pi_0(t)$ can be retrieved from (3.2.30) 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 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 3.2.17. A portfolio process $\{\boldsymbol{\pi}(t) : t \in [0,T]\}$ for the market model is a previsible process $\boldsymbol{\pi} : \Omega \times [0,T] \to \mathbb{R}^N$ such that $\int_0^T \|\boldsymbol{\pi}(t)\|^2 \, \mathrm{d}t < \infty$ a.s.

Zhou and Yin's portfolio process $\{\boldsymbol{u}(t)\}$, given by Definition 3.1.3, is therefore comparable to the portfolio process $\{\boldsymbol{\pi}(t)\}$.

Using (3.2.16), (3.2.22) and (3.2.24) and the self-financing assumption, (3.2.30) can be written in differential form as

$$\mathrm{d}X^{\boldsymbol{\pi}}(t) = \big(r(t)X^{\boldsymbol{\pi}}(t) + \boldsymbol{\pi}^\top(t)\boldsymbol{\sigma}(t)\boldsymbol{\theta}(t)\big)\,\mathrm{d}t + \boldsymbol{\pi}^\top(t)\boldsymbol{\sigma}(t)\,\mathrm{d}\mathbf{W}(t), \quad X^{\boldsymbol{\pi}}(0) = x_0. \tag{3.2.31}$$

We call (3.2.31) the *wealth equation*.

The *wealth process* $X^{\boldsymbol{\pi}} = \{X^{\boldsymbol{\pi}}(t) : t \in [0,T]\}$ is the unique (up to indistinguishability) solution of the wealth equation (3.2.31). It is a continuous, $\{\mathcal{F}_t\}$–adapted, $\mathbb{R}$-valued process given by

$$X^{\boldsymbol{\pi}}(t) = S_0(t)\left\{x_0 + \int_0^t S_0^{-1}(\tau)\boldsymbol{\pi}^\top(\tau)\boldsymbol{\sigma}(\tau)\boldsymbol{\theta}(\tau)\,\mathrm{d}\tau + \int_0^t S_0^{-1}(\tau)\boldsymbol{\pi}^\top(\tau)\boldsymbol{\sigma}(\tau)\,\mathrm{d}\mathbf{W}(\tau)\right\}. \tag{3.2.32}$$

We refer to $X^{\boldsymbol{\pi}}$ as the solution to the wealth equation (3.2.31) for the portfolio process $\boldsymbol{\pi}$.

Examining the right-hand side of (3.2.32), we see that the only parameter which is under the sole control of the investor is the portfolio process $\boldsymbol{\pi}$. All the other parameters are market-determined parameters and known to the investor.

Spaces of integrands

We will see in the next section that in order to solve the MVO problem, we must have $\mathrm{E}|X^{\boldsymbol{\pi}}(T)|^2 < \infty$. 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 processes.

However, the wealth process which solves the wealth equation (3.2.31) is a continuous process. For this reason, we define a subspace $\mathbb{A}$ of $\mathbb{B}$ whose members are continuous processes, in addition to being square-integrable. Potential solutions to the MVO problem will lie in the space $\mathbb{A}$.

We start by defining some appropriate L^2–spaces of integrands.

$$L_{21} := \left\{ f : \Omega \times [0,T] \to \mathbb{R} \big| \; f \in \mathcal{P}^\star \text{ and } \mathrm{E}\left(\int_0^T |f(t)| \, \mathrm{d}t \right)^2 < \infty \right\}. \tag{3.2.33}$$

$$L^2(\mathbf{W}) := \left\{ \mathbf{\Lambda} : \Omega \times [0,T] \to \mathbb{R}^N \big| \; \mathbf{\Lambda} \in \mathcal{P}^\star \text{ and } \mathrm{E} \int_0^T \|\mathbf{\Lambda}(t)\|^2 \, \mathrm{d}t < \infty \right\}. \tag{3.2.34}$$

$$\begin{aligned} L^2(\mathcal{Q}) := \Bigg\{ & \mathbf{\Gamma} = \{\Gamma_{ij}\}_{i,j=1}^D : \Omega \times [0,T] \to \mathbb{R}^{D\times D} \big| \Gamma_{ii} = 0 \; (\mathbb{P} \otimes \mathit{Leb})\text{-a.e.}, \forall i \in I, \\ & \Gamma_{ij} \in \mathcal{P}^\star \text{ and } \sum_{i,j=1}^D \mathrm{E} \int_0^T |\Gamma_{ij}(t)|^2 \, \mathrm{d}[\mathcal{Q}_{ij}](t) < \infty \; \forall \, i,j \in I, \, i \neq j \Bigg\}. \end{aligned} \tag{3.2.35}$$

Remark 3.2.18. From the definition of the space $L^2(\mathcal{Q})$, the reason for the particular form of the notation $\nu_{[\mathcal{Q}]}$-a.e. in (3.2.13) should now be clear.

Define the space $\mathbb{B}$ as

$$\mathbb{B} := \mathbb{R} \times L_{21} \times L^2(\mathbf{W}) \times L^2(\mathcal{Q}). \tag{3.2.36}$$

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

$$X(t) := X_0 + \int_0^t \dot{X}(\tau) \, \mathrm{d}\tau + \sum_{n=1}^N \int_0^t \Lambda_n^X(\tau) \, \mathrm{d}W_n(\tau) + \sum_{i,j=1}^D \int_0^t \Gamma_{ij}^X(\tau) \, \mathrm{d}\mathcal{Q}_{ij}(\tau), \tag{3.2.37}$$

for some $X_0 \in \mathbb{R}$, $\dot{X} \in L_{21}$, $\mathbf{\Lambda}^X := (\Lambda_1^X, \ldots, \Lambda_N^X)^\top \in L^2(\mathbf{W})$ and $\mathbf{\Gamma}^X := (\Gamma_{ij}^X)_{i,j=1}^D \in L^2(\mathcal{Q})$. We write $X \equiv (X_0, \dot{X}, \mathbf{\Lambda}^X, \mathbf{\Gamma}^X)$ to indicate that (3.2.37) holds for $X \in \mathbb{B}$ and we call the quadruple $(X_0, \dot{X}, \mathbf{\Lambda}^X, \mathbf{\Gamma}^X)$ the *components* of X.

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

$$\mathbb{A} := \{X \equiv (X_0, \dot{X}, \mathbf{\Lambda}^X, \mathbf{\Gamma}^X) \in \mathbb{B} \,|\, \mathbf{\Gamma}^X = \mathbf{0} \ \nu_{[\mathcal{Q}]}\text{-a.e.}\}. \tag{3.2.38}$$

Remark 3.2.19. The subspace $\mathbb{A}$ consists of all continuous processes in the space $\mathbb{B}$.

The proof of the next proposition can be found beneath Proposition A.1.1.

Proposition 3.2.20. *Suppose we have* $X \equiv (X_0, \dot{X}, \mathbf{\Lambda}^X, \mathbf{\Gamma}^X) \in \mathbb{B}$ *and* $Y \equiv (Y_0, \dot{Y}, \mathbf{\Lambda}^Y, \mathbf{\Gamma}^Y) \in \mathbb{B}$. *If for all* $t \in [0,T]$,

$$\begin{aligned} X_0 &+ \int_0^t \dot{X}(\tau)\,\mathrm{d}\tau + \int_0^t \left(\mathbf{\Lambda}^X\right)^\top(\tau)\,\mathrm{d}\mathbf{W}(\tau) + \sum_{i,j=1}^D \int_0^t \Gamma_{ij}^X(\tau)\,\mathrm{d}\mathcal{Q}_{ij}(\tau) \\ &= Y_0 + \int_0^t \dot{Y}(\tau)\,\mathrm{d}\tau + \int_0^t \left(\mathbf{\Lambda}^Y\right)^\top(\tau)\,\mathrm{d}\mathbf{W}(\tau) + \sum_{i,j=1}^D \int_0^t \Gamma_{ij}^Y(\tau)\,\mathrm{d}\mathcal{Q}_{ij}(\tau) \ \ \text{a.s.} \end{aligned} \tag{3.2.39}$$

then $X_0 = Y_0$, $\dot{X} = \dot{Y}$, $\mathbf{\Lambda}^X = \mathbf{\Lambda}^Y$ $(\mathbb{P} \otimes Leb)$*-a.e. and* $\mathbf{\Gamma}^X = \mathbf{\Gamma}^Y$ $\nu_{[\mathcal{Q}]}$*-a.e.*

Note that by Proposition 3.2.20, the representation of $X \in \mathbb{B}$ by the components $(X_0, \dot{X}, \mathbf{\Lambda}^X, \mathbf{\Gamma}^X) \in \mathbb{R} \times L_{21} \times L^2(\mathbf{W}) \times L^2(\mathcal{Q})$ is unique up to indistinguishability.

The following lemma shows that the members of $\mathbb{B}$ are square-integrable. The proof can be found under Lemma A.1.2.

Lemma 3.2.21. *For all* $X \equiv (X_0, \dot{X}, \mathbf{\Lambda}^X, \mathbf{\Gamma}^X) \in \mathbb{B}$,

$$E\left(\sup_{t\in[0,T]} |X(t)|^2\right) < \infty. \tag{3.2.40}$$

Remark 3.2.22. Observe that the wealth process $X^{\boldsymbol{\pi}}$ which solves the wealth equation (3.2.31) for a portfolio process $\boldsymbol{\pi}$ is a continuous process. However, we don't know if the wealth process $X^{\boldsymbol{\pi}}$ is square-integrable, since we don't know if $rX^{\boldsymbol{\pi}} + \boldsymbol{\pi}^\top\boldsymbol{\sigma}\boldsymbol{\theta} \in L_{21}$ and $\boldsymbol{\sigma}^\top\boldsymbol{\pi} \in L^2(\mathbf{W})$. However, as we shall see in the next proposition, whose proof can be found beneath Proposition A.1.3, if the portfolio process $\boldsymbol{\pi}$ is in $L^2(\mathbf{W})$ then $X^{\boldsymbol{\pi}} \in \mathbb{A}$.

Proposition 3.2.23. *For* $X^{\boldsymbol{\pi}}$ *which solves the wealth equation (3.2.31) for a portfolio process* $\boldsymbol{\pi}$*, we have*

$$X^{\boldsymbol{\pi}} \in \mathbb{A} \text{ if and only if } \boldsymbol{\pi} \in L^2(\mathbf{W}). \tag{3.2.41}$$

Remark 3.2.24. Suppose that $\boldsymbol{\pi} \in L^2(\mathbf{W})$. Then from the wealth equation (3.2.31) and Proposition 3.2.23, we see that

$$X^{\boldsymbol{\pi}} \equiv (x_0, rX^{\boldsymbol{\pi}} + \boldsymbol{\pi}^\top\boldsymbol{\sigma}\boldsymbol{\theta}, \boldsymbol{\sigma}^\top\boldsymbol{\pi}, \mathbf{0}) \in \mathbb{A}. \tag{3.2.42}$$

The investor's problem

As discussed in Section 2.2, we begin by defining an MVO problem without a terminal wealth constraint. This is to keep the focus on the application of a convex duality method, which is the essence of the research. However, in Section 4.9, we define and solve the MVO problem with a terminal wealth constraint.

Before we define the MVO problem, we outline a simple, motivating example. This is a typical example of the problem that we are attempting to solve. We will refer back to this motivating example as we are defining the MVO problem, to illustrate the chosen formulation of the problem.

At time 0, an investor has a fixed sum of money x_0, which is called the *initial wealth*. The investor would like to have d units of wealth at the end of a finite time horizon $[0, T]$. We call T the *terminal time* and we call the investor's wealth at time T the *terminal wealth*.

However, the investor realizes that the market returns are not guaranteed, so she can only expect to attain on average the d units of wealth. As a result, she decides to minimize the variance of her actual terminal wealth from the d units of wealth. In other words, she wishes to minimize the risk, where risk is measured by the variance of her actual terminal wealth from the d units of wealth.

She also requires that at no point over the finite time horizon $[0, T]$ will she have negative wealth in any stock. Such a requirement is called a *portfolio constraint*.

The investor's requirements can be summarized as follows: the investor is seeking to minimize her risk, subject to starting with an initial wealth of x_0 and meeting the portfolio constraints over the time horizon. Can she find a portfolio process so that, starting with an initial wealth of x_0 and ensuring that at all times the portfolio constraints are met, her risk is minimized?

Having outlined the motivating example, we next specify precisely the general MVO problem that we propose to solve. We start by defining the risk measure, the minimization of which is analogous to the investor minimizing her risk.

Definition 3.2.25. Define a *risk measure* J on the wealth process by

$$\begin{aligned} J : \Omega \times \mathbb{R} &\to \mathbb{R} \\ (\omega, x) &\mapsto J(\omega, x) := \frac{1}{2}a(\omega)x^2 + b(\omega)x + c(\omega), \end{aligned} \tag{3.2.43}$$

subject to the following three conditions:

Condition 3.2.26. a is $\mathcal{F}_T$–measurable random variable on $(\Omega, \mathcal{F}, \mathbb{P})$ and satisfies

$$0 < \inf_{\omega \in \Omega}\{a(\omega)\} \leq \sup_{\omega \in \Omega}\{a(\omega)\} < \infty, \tag{3.2.44}$$

Condition 3.2.27. b is an $\mathcal{F}_T$–measurable, square-integrable random variable on $(\Omega, \mathcal{F}, \mathbb{P})$.

Condition 3.2.28. c is an $\mathcal{F}_T$–measurable, integrable random variable on $(\Omega, \mathcal{F}, \mathbb{P})$.

Remark 3.2.29. If we set the random variables $a = 2$, $b = -2d$ and $c = d^2$, where $d > 0$ is a real number, (3.2.43) reduces to $J(x) = (x - d)^2$. Then taking the expectation of $J(X(T))$ we obtain

$$\mathrm{E}\left(J(X(T))\right) = \mathrm{E}\left(X(T) - d\right)^2, \tag{3.2.45}$$

which is the risk measure in the motivating example.

Remark 3.2.30. Consider a pension fund whose liability value at time T is represented by a random variable $L(T)$. Assume that $L(T)$ is $\{\mathcal{F}_T\}$–measurable and square-integrable. The pension fund's asset value at time t is represented by $X(t)$. The pension fund wishes to minimize the variance of the liability value $L(T)$ from the asset value $X(T)$ at time T. Setting the random variables $a = 2$, $b = -L(T)$ and $c = L(T)^2$, (3.2.43) reduces to $J(x) = (x - L(T))^2$. Then taking expectations of $J(X(T))$ we obtain

$$\mathrm{E}\left(J(X(T))\right) = \mathrm{E}\left(X(T) - L(T)\right)^2, \tag{3.2.46}$$

which is the required risk measure.

Next we define the portfolio constraint. The portfolio constraint is that the portfolio process we seek must always lie in some convex set K. We define the convex set K below.

Condition 3.2.31. We are given a closed, convex set $K \subset \mathbb{R}^N$ with $\mathbf{0} \in K$.

Remark 3.2.32. If we define the set $K := \{p \in \mathbb{R}^N | p_1 \geq 0, \ldots, p_N \geq 0\} \equiv [0, \infty)^N$, then this corresponds to an investor never going short in the stocks. This is the portfolio constraint for the investor in the motivating example at the start of this section.

Remark 3.2.33. Consider a pension fund which is not permitted to invest in the company which sponsors it. Labelling the sponsoring company's stock as stock N, then the pension fund's portfolio constraints correspond to the set $K := \{p \in \mathbb{R}^N | p_N = 0\}$. If the pension fund is, in addition, subject to a no-short-selling constraint, then the portfolio constraints correspond to the set $K := \{p \in \mathbb{R}^N | p_1 \geq 0, \ldots, p_{N-1} \geq 0, p_N = 0\}$.

More generally, defining the set $K := \{p \in \mathbb{R}^N | p_{M+1} = \cdots = p_N = 0\}$ for some $M \in \{1, \ldots, N-1\}$ corresponds to investment in an incomplete market, where only stocks 1 to M are available to the investor.

Defining the set $K := \{p \in \mathbb{R}^N | p_1 \geq 0, \ldots, p_M \geq 0,\ p_{M+1} = \cdots = p_N = 0\}$ for some $M \in \{1, \ldots, N-1\}$ corresponds to investment in an incomplete market by an investor subject to a no-short-selling constraint.

Remark 3.2.34. Zhou and Yin [53] deal only with no portfolio constraints, which corresponds to the set $K := \mathbb{R}^N$. Their model cannot incorporate any of the constraints given in the above remarks.

From Proposition 3.2.23, 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 define an *admissible portfolio* to be those portfolio processes which are members of $L^2(\mathbf{W})$ and which satisfy the portfolio constraints by belonging to the convex set K.

Definition 3.2.35. Define the set $\mathcal{A}$ of admissible portfolios as

$$\mathcal{A} := \{\boldsymbol{\pi} \in L^2(\mathbf{W}) \,\big|\, \boldsymbol{\pi} \in K \ (\mathbb{P} \otimes Leb)\text{-a.e.}\}, \tag{3.2.47}$$

Remark 3.2.36. The portfolio constraint is that the portfolio process $\bar{\boldsymbol{\pi}}$ we seek is an admissible portfolio process, that is $\bar{\boldsymbol{\pi}} \in \mathcal{A}$.

Definition 3.2.37. The value of the problem, denoted by $\mathcal{V}$, is defined as

$$\mathcal{V} := \inf_{\boldsymbol{\pi} \in \mathcal{A}} \{\mathrm{E}\left(J(X^{\boldsymbol{\pi}}(T))\right)\}, \tag{3.2.48}$$

where $X^{\boldsymbol{\pi}}$ is the solution to the wealth equation (3.2.31) corresponding to $\boldsymbol{\pi}$, the set of admissible portfolios $\mathcal{A}$ is given by Definition 3.2.35 and the risk measure J is given by Definition 3.2.25.

Problem 3.2.38. The MVO problem is to determine the existence and characterization of a portfolio process $\bar{\boldsymbol{\pi}} \in \mathcal{A}$ such that the value of the problem $\mathcal{V} = \mathrm{E}\left(J(X^{\bar{\boldsymbol{\pi}}}(T))\right)$.

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

Remark 3.2.39. If we can determine the existence of a portfolio process $\bar{\boldsymbol{\pi}} \in \mathcal{A}$ such that the value of the problem $\mathcal{V} = \mathrm{E}\left(J(X^{\bar{\boldsymbol{\pi}}}(T))\right)$ then the infimum in (3.2.48) is attained. It is not obvious that such a portfolio process exists. For example, consider the function $f(x) = \exp(-x)$ for all $x \in \mathbb{R}$. Then $\inf_{x \in \mathbb{R}} f(x) = 0$ but there is no $x \in \mathbb{R}$ for which $\exp(-x) = 0$.

Remark 3.2.40. Note that Problem 3.2.38 is non-trivial since, as we show in Lemma A.1.4, the value of the problem $\mathcal{V}$ is such that $-\infty < \mathcal{V} < \infty$.

Convention 3.2.41. $\inf\{\emptyset\} := +\infty$

Convention 3.2.42. $+\infty - \infty := +\infty$

3.2.3 A convex duality method

We use a convex duality method to solve Problem 3.2.38. Convex duality methods establish a connection between the original problem and another problem, called the dual problem. The hope is that the dual problem is easier to solve than the original problem. The convexity properties of the original problem are critical in establishing the connection between the original and dual problems. Using the solution to the dual problem, this connection may allow us to generate the solution to the original problem.

Mathematically, suppose that we are given a convex functional $\Phi : \mathcal{X} \to (-\infty, +\infty]$, where $\mathcal{X}$ is a vector space. The problem $(\mathcal{P})$ is to

$$(\mathcal{P}) \qquad \text{find } \bar{X} \in \mathcal{X} \text{ such that } \Phi(\bar{X}) = \inf_{X \in \mathcal{X}} \{\Phi(X)\}. \tag{3.2.49}$$

Convex duality theory allows us to construct another convex functional $\Psi : \mathcal{Y} \to (-\infty, +\infty]$ over another vector space $\mathcal{Y}$ such that

$$\Phi(X) + \Psi(Y) \geq 0, \qquad \forall (X, Y) \in \mathcal{X} \times \mathcal{Y}. \tag{3.2.50}$$

If for some $(\bar{X}, \bar{Y}) \in \mathcal{X} \times \mathcal{Y}$, we have

$$\Phi(\bar{X}) + \Psi(\bar{Y}) = 0 \tag{3.2.51}$$

then

$$\Phi(\bar{X}) = \inf_{X \in \mathcal{X}} \{\Phi(X)\} = \sup_{Y \in \mathcal{Y}} \{-\Psi(Y)\} = -\Psi(\bar{Y}). \tag{3.2.52}$$

From (3.2.52), we see that $\bar{X}$ solves the problem $(\mathcal{P})$. Thus if we can find a pair $(\bar{X}, \bar{Y}) \in \mathcal{X} \times \mathcal{Y}$ which satisfy (3.2.51) then $\bar{X}$ solves the original problem $(\mathcal{P})$.

How do these observations help us? Their helpfulness rests on being able to show two more things. The first thing we require is the existence of a solution $\bar{Y}$ to the dual problem $(\mathcal{D})$, which is to

$$(\mathcal{D}) \qquad \text{find } \bar{Y} \in \mathcal{Y} \text{ such that } -\Psi(\bar{Y}) = \sup_{Y \in \mathcal{Y}} \{-\Psi(Y)\} \equiv -\inf_{Y \in \mathcal{Y}} \{\Psi(Y)\}. \tag{3.2.53}$$

The second thing we require are conditions on $(\bar{X}, \bar{Y}) \in \mathcal{X} \times \mathcal{Y}$ which are equivalent to (3.2.51). We get these necessary conditions from the optimality of $\bar{Y}$ at (3.2.53). Sometimes these are called saddle-point conditions or Kuhn-Tucker conditions.

Armed with the existence of $\bar{Y}$, we use these necessary conditions to characterize the solution $\bar{X}$ to the original problem $(\mathcal{P})$ in terms of the solution $\bar{Y}$ to the dual problem $(\mathcal{D})$.

This is the convex duality method that we use to solve Problem 3.2.38. The stages involved in applying the convex duality method are outlined in Subsection 3.2.4. Note that the vector space $\mathbb{A}$ defined by (3.2.38) plays the same role as the vector space $\mathcal{X}$, and the vector space $\mathbb{B}$ defined by (3.2.36) plays the same role as the vector space $\mathcal{Y}$ above.

3.2.4 Steps required to apply the convex duality method

The steps involved in applying the convex duality method are as follows.

1. Re-statement of the MVO problem as the primal problem.
2. Define the dual problem.
3. Relationship between the primal and dual cost functionals.
4. Necessary and sufficient conditions to solve the primal problem.
5. Existence of a solution to the dual problem.
6. Construction of a candidate solution to the primal problem.
7. Check if the candidate solution solves the primal problem.

In this subsection, we give a broad overview of each step.

Step 1 Re-statement of the MVO problem as the primal problem

To enable us to use the machinery of convex duality theory to determine a solution to Problem 3.2.38, we first need to re-state it in a suitable form. To do this, we express the value of the problem, which is defined as

$$\mathcal{V} := \inf_{\boldsymbol{\pi} \in \mathcal{A}} \{ \mathrm{E} \left(J(X^{\boldsymbol{\pi}}(T)) \right) \}, \tag{3.2.54}$$

in the form

$$\mathcal{V} = \inf_{X \in \mathbb{A}} \{ \Phi(X) \}, \tag{3.2.55}$$

where Φ is a functional on the space $\mathbb{A}$ defined by (3.2.38). Essentially, we are embedding the value of the problem $\mathcal{V}$ in a larger problem.

The idea is to change the optimization from one over the set of portfolio processes $\boldsymbol{\pi} \in \mathcal{A}$ to one over the set of wealth processes $X^{\boldsymbol{\pi}}$. We do this by finding the wealth processes $X^{\boldsymbol{\pi}}$ in $\mathbb{A}$ which correspond to the admissible portfolios $\boldsymbol{\pi} \in \mathcal{A}$. We should obtain an expression of the form

$$\mathcal{V} := \inf_{\substack{X^{\boldsymbol{\pi}} \in \mathbb{A} \\ \text{for } \boldsymbol{\pi} \in \mathcal{A}}} \{ \mathrm{E} \left(J(X^{\boldsymbol{\pi}}(T)) \right) \}. \tag{3.2.56}$$

Now we want to express this as (3.2.55). We do this by defining a functional Φ on the space $\mathbb{A}$. For each process $X \in \mathbb{A}$, the functional Φ equals $\mathrm{E}(J(X(T)))$, the expected value of the risk measure, if there is a portfolio process $\boldsymbol{\pi} \in \mathcal{A}$ which corresponds to X, in the sense that $X = X^{\boldsymbol{\pi}}$, where $X^{\boldsymbol{\pi}}$ is the solution to the wealth equation (3.2.31) for $\boldsymbol{\pi} \in \mathcal{A}$. If there is no such portfolio process $\boldsymbol{\pi} \in \mathcal{A}$ corresponding to X, then the functional Φ equals infinity.

We call the functional Φ the primal cost functional. To construct the primal cost functional Φ, we define two penalty functions on the space $\mathbb{A}$. These penalty functions have value infinity if $X \in \mathbb{A}$ cannot satisfy the constraints on the solution to Problem 3.2.38. Otherwise they have value zero. By summing the penalty functions and adding the expected value of the risk measure J, we obtain the primal cost functional Φ.

The primal problem is then defined in terms of the primal cost functional Φ.

Problem 3.2.43. The primal problem is to find $\bar{X} \in \mathbb{A}$ such that

$$\Phi(\bar{X}) = \inf_{X \in \mathbb{A}} \{\Phi(X)\}. \tag{3.2.57}$$

Remark 3.2.44. If we can find $\bar{X} \in \mathbb{A}$ which solves the primal problem, then we can find the corresponding portfolio process $\bar{\boldsymbol{\pi}} \in \mathcal{A}$, which solves Problem 3.2.38.

Step 2 Define the dual problem

Next we construct the dual cost functional Ψ so that we can define the dual problem. We use the convex conjugates of the penalty functions to define the dual functional Ψ (see Definition C.17.1). We construct the convex conjugates of the risk measure and the penalty functions which were used to define the primal functional. Adding these together, we obtain the dual cost functional Ψ. We can then define the dual problem as follows.

Problem 3.2.45. The dual problem is to find $\bar{Y} \in \mathbb{B}$ such that

$$-\Psi(\bar{Y}) = \sup_{Y \in \mathbb{B}} \{-\Psi(Y)\} \equiv -\inf_{Y \in \mathbb{B}} \{\Psi(Y)\}. \tag{3.2.58}$$

Step 3 Relationship between the primal and dual cost functionals

In this step, we establish the relationship between the primal and dual functionals. Concretely, we wish to show that for all pairs $(X, Y) \in \mathbb{A} \times \mathbb{B}$ we have

$$\Phi(X) + \Psi(Y) \geq 0. \tag{3.2.59}$$

This weak duality principle is a critical step in establishing the connection between the solutions to the primal and dual problems. Without showing that (3.2.59) holds, we are unable to show that equality is achieved in (3.2.59) when $X \in \mathbb{A}$ solves the primal problem and $Y \in \mathbb{B}$ solves the dual problem.

The inequality (3.2.59) follows as a consequence of the conjugate duality between the primal and dual cost functionals and the application of a result from Bismut [3].

Step 4 Necessary and sufficient conditions to solve the primal problem

After showing that equality is achieved in (3.2.59) when $X \in \mathbb{A}$ solves the primal problem and $Y \in \mathbb{B}$ solves the dual problem, so that strong duality holds, we derive necessary and sufficient conditions for this to happen. That is, we derive necessary and sufficient conditions for a pair of elements $(X, Y) \in \mathbb{A} \times \mathbb{B}$ to satisfy $\Phi(X) + \Psi(Y) = 0$.

The idea is this. Suppose we know that $\bar{Y} \in \mathbb{B}$ solves the dual problem. We have another element $\bar{X} \in \mathbb{A}$ and we wish to know if it solves the primal problem. All we have to do is see if the pair $(\bar{X}, \bar{Y}) \in \mathbb{A} \times \mathbb{B}$ satisfy these necessary and sufficient conditions.

The necessary and sufficient conditions are derived from the penalty functions, which are used to construct the primal cost functional, and their convex conjugates, which are used to construct the dual cost functional.

Step 5 Existence of a solution to the dual problem

Here we consider only the dual problem. We show that a solution $\bar{Y} \in \mathbb{B}$ to the dual problem exists using Ekeland and Témam [13], Chapter II, Proposition 1.2, page 35.

Step 6 Construction of a candidate solution to the primal problem

Using the solution $\bar{Y} \in \mathbb{B}$ to the dual problem, which we know exists by Step 5, we construct a candidate solution $\bar{X} \in \mathbb{B}$ to the primal problem. Note that we cannot say yet that $\bar{X} \in \mathbb{A}$.

We use one of the necessary and sufficient conditions derived in Step 4 and a result about a solution to the wealth equation (3.2.31) (the solution to the primal problem must satisfy the wealth equation) to construct the candidate solution $\bar{X} \in \mathbb{B}$.

Step 7 Check if the candidate solution solves the primal problem

Finally, we check that the candidate solution solves the primal problem. This amounts to showing that $\bar{X} \in \mathbb{A}$ and checking that the pair $(\bar{X}, \bar{Y}) \in \mathbb{A} \times \mathbb{B}$ satisfies all the necessary and sufficient conditions given in Step 4.

Conclusion

Once we have shown Steps 1-7, we will have solved the primal problem and hence Problem 3.2.38.

3.2.5 Summary

In this chapter, we began by outlining Zhou and Yin [53]. They solved an MVO problem which is similar to the one which we solve, except for two critical differences. The MVO problem they solve has no portfolio constraints and the market coefficients they use are Markov-modulated. The MVO problem we outline in this chapter has convex portfolio constraints and random market coefficients. Adding portfolio constraints and random market coefficients makes the MVO problem much more challenging to solve.

Zhou and Yin [53] use a linear quadratic control method to solve their MVO problem and they find an explicit solution for the optimal portfolio. However, as we discussed in Section 2.3, the convex duality approach is a more appropriate method of solution for our MVO problem, due to the portfolio constraints and random market coefficients.

After outlining Zhou and Yin [53], we describe the market model that we use. As Zhou and Yin [53] do, we model the regime-switching using a continuous-time, finite state space Markov chain and we assume that the stock prices are driven by an N-dimensional Brownian motion. We detail the canonical martingales of the Markov chain and briefly stated their properties, the most important of which is their orthogonality to one another.

We assume that the N-dimensional Brownian motion is independent of the Markov chain. This is the same assumption made by Zhou and Yin [53]. From the N-dimensional Brownian motion and the Markov chain, we construct a filtration, which we will often refer to as the joint filtration.

The price processes of the bank account and the stocks in the market satisfy the usual stochastic differential equations. We emphasize that the parameters of these stochastic differential equations, the market coefficents, are random and adapted to the joint filtration. The market coefficents are uniformly bounded and, in addition, the risk-free interest rate process is nonnegative.

Next we detail the wealth equation (3.2.31), which is the stochastic differential equation that the wealth process of the investor satisfies uniquely. From this, we see that the solution to the wealth equation (3.2.31) is a continuous process.

We define some L^2-spaces of integrands, which we use to define a space $\mathbb{B}$ consisting of right-continuous, square-integrable semimartingales. In particular, we note that the members of $\mathbb{B}$ are not necessarily continuous processes, being right-continuous. This contrasts with the fact that the wealth process corresponding to the solution to the MVO problem is a continuous process. From this, we observe that the $L^2(\mathcal{Q})$-component of the wealth process corresponding to the solution to the MVO problem must equal zero. This leads us to define a subspace $\mathbb{A}$ of $\mathbb{B}$, which contains all the continuous processes of the space $\mathbb{B}$.

Finally, we set out the MVO problem to be solved. We seek to determine a portfolio process which not only satisfies the portfolio constraints, but also minimizes the expected value of a risk measure. The portfolio constraints are that the portfolio process must at all times lie in a specified closed, convex set. The risk measure is defined on the terminal wealth of the investor.

We also describe briefly the convex duality method that we use to solve the MVO problem. The main idea is that we construct a dual problem and show existence of the solution to the dual problem. From this, existence of the solution to the MVO problem will follow.

Chapter 4

The mean-variance portfolio optimization problem

The majority of this chapter is concerned with the solution to Problem 3.2.38, which is the MVO problem without a terminal wealth constraint. We apply the convex duality method outlined in Subsection 3.2.4. The steps required to apply the method and the section where the step is completed is shown below.

1. Re-statement of the MVO problem as the primal problem (Section 4.1).
2. Define the dual problem (Section 4.2).
3. Relationship between the primal and dual cost functionals (Section 4.3).
4. Necessary and sufficient conditions to solve the primal problem (Section 4.4).
5. Existence of a solution to the dual problem (Section 4.5).
6. Construction of a candidate solution to the primal problem (Section 4.6).
7. Check if the candidate solution solves the primal problem (Section 4.7).

The results of these steps are summarized in Section 4.8. The section culminates in Proposition 4.8.1, which gives the existence and characterization of the solution to Problem 3.2.38.

Finally, in Section 4.9, we solve the MVO problem with a terminal wealth constraint. Using a Lagrange multiplier technique, the solution to the MVO problem with a terminal wealth constraint is bootstrapped from the solution to the MVO problem without a terminal wealth constraint. The method is taken from Labbé and Heunis [33] and requires no modification as a consequence of the regime-switching. However, we include it for the sake of completeness.

4.1 Re-statement of the MVO problem as the primal problem

Our goal in this section is to re-formulate Problem 3.2.38 as the primal problem. To do this, we seek to re-write the value of the problem

$$\mathcal{V} \overset{(3.2.48)}{=} \inf_{\boldsymbol{\pi}\in\mathcal{A}} \{\mathrm{E}\left(J(X^{\boldsymbol{\pi}}(T))\right)\}, \tag{4.1.1}$$

in the following form

$$\mathcal{V} = \inf_{X\in\mathbb{A}} \{\Phi(X)\}. \tag{4.1.2}$$

for some functional Φ. Notice that the infimum in (4.1.2) is over the set of processes $X \in \mathbb{A}$, compared to the infimum in (4.1.1) which is over the set of admissible portfolios $\boldsymbol{\pi} \in \mathcal{A}$. Essentially, we seek to embed the value of the problem $\mathcal{V}$ in a larger problem. We do this by regarding the value of the problem $\mathcal{V}$ as an optimization over the set of processes $X \in \mathbb{A}$. The functional Φ is carefully constructed so that any $X \in \mathbb{A}$ which does not satisfy the constraints of the original problem results in $\Phi(X) = \infty$.

Define for each $X \equiv (X_0, \dot{X}, \boldsymbol{\Lambda}^X, \mathbf{0}) \in \mathbb{A}$,

$$\begin{aligned} U(X) := \big\{\boldsymbol{\pi} \in \mathcal{A} \,|\, \dot{X}(t) = r(t)X(t) + \boldsymbol{\pi}^\top(t)\boldsymbol{\sigma}(t)\boldsymbol{\theta}(t) \ (\mathbb{P}\otimes Leb)\text{-a.e.} \\ \text{and } \boldsymbol{\Lambda}^X(t) = \boldsymbol{\sigma}^\top(t)\boldsymbol{\pi}(t) \ (\mathbb{P}\otimes Leb)\text{-a.e.}\big\}. \end{aligned} \tag{4.1.3}$$

Lemma 4.1.1. *For each $\boldsymbol{\pi} \in \mathcal{A}$ and $X \equiv (X_0, \dot{X}, \boldsymbol{\Lambda}^X, \mathbf{0}) \in \mathbb{A}$, we have*

$$X(t) \equiv X^{\boldsymbol{\pi}}(t) \ \textit{a.s.}, \quad \forall t \in [0,T] \qquad \Rightarrow \qquad \begin{cases} X_0 = x_0 \\ U(X) \neq \emptyset. \end{cases} \tag{4.1.4}$$

Conversely, for each $X \equiv (X_0, \dot{X}, \boldsymbol{\Lambda}^X, \mathbf{0}) \in \mathbb{A}$,

$$\left.\begin{array}{l} X_0 = x_0 \\ U(X) \neq \emptyset \end{array}\right\} \qquad \Rightarrow \qquad X(t) \equiv X^{\boldsymbol{\pi}}(t) \ \textit{a.s.}, \quad \forall t \in [0,T], \tag{4.1.5}$$

for $\boldsymbol{\pi} \in \mathcal{A}$ given by $\boldsymbol{\pi}(\omega,t) := (\boldsymbol{\sigma}^\top(\omega,t))^{-1}\boldsymbol{\Lambda}^X(\omega,t)$ for all $(\omega,t) \in \Omega\times[0,T]$.

Proof. Fix $\boldsymbol{\pi} \in \mathcal{A}$. Then $\boldsymbol{\pi} \in L^2(\mathbf{W})$ and applying Proposition 3.2.23, we see that the corresponding wealth process $X^{\boldsymbol{\pi}}$ which solves the wealth equation (3.2.31) belongs to the space $\mathbb{A}$, so that

$$X^{\boldsymbol{\pi}} \equiv (x_0, rX^{\boldsymbol{\pi}} + \boldsymbol{\pi}^\top\boldsymbol{\sigma}\boldsymbol{\theta}, \boldsymbol{\sigma}^\top\boldsymbol{\pi}, \mathbf{0}) \in \mathbb{A}. \tag{4.1.6}$$

Now fix some $X \equiv (X_0, \dot{X}, \boldsymbol{\Lambda}^X, \mathbf{0}) \in \mathbb{A}$. If

$$X(t) = X^{\boldsymbol{\pi}}(t) \quad \text{a.s.}, \quad \forall t \in [0,T], \tag{4.1.7}$$

then from Proposition A.1.1 and (4.1.6), the following three relations hold:

$$X_0 = x_0; \tag{4.1.8}$$

$$\begin{aligned}\dot{X}(t) &= r(t)X^{\boldsymbol{\pi}}(t) + \boldsymbol{\pi}^\top(t)\boldsymbol{\sigma}(t)\boldsymbol{\theta}(t) \\ &\overset{(4.1.7)}{=} r(t)X(t) + \boldsymbol{\pi}^\top(t)\boldsymbol{\sigma}(t)\boldsymbol{\theta}(t) \ (\mathbb{P}\otimes Leb)\text{-a.e; and}\end{aligned} \tag{4.1.9}$$

$$\boldsymbol{\Lambda}^X(t) = \boldsymbol{\sigma}^\top(t)\boldsymbol{\pi}(t) \ (\mathbb{P}\otimes Leb)\text{-a.e.} \tag{4.1.10}$$

Together, (4.1.9), (4.1.10) and the assumption that $\boldsymbol{\pi} \in \mathcal{A}$ imply that $\boldsymbol{\pi} \in U(X)$, so that $U(X) \neq \emptyset$. This and (4.1.8) give (4.1.4).

For the converse, fix $X \equiv (X_0, \dot{X}, \boldsymbol{\Lambda}^X, \mathbf{0}) \in \mathbb{A}$ and assume that $X_0 = x_0$ and $U(X) \neq \emptyset$. By the definition of $U(X)$, given by (4.1.3), there exists $\boldsymbol{\pi} \in \mathcal{A}$ such that

$$\dot{X}(t) = r(t)X(t) + \boldsymbol{\pi}^\top(t)\boldsymbol{\sigma}(t)\boldsymbol{\theta}(t) \ (\mathbb{P}\otimes Leb)\text{-a.e; and} \tag{4.1.11}$$

$$\boldsymbol{\Lambda}^X(t) = \boldsymbol{\sigma}^\top(t)\boldsymbol{\pi}(t) \ (\mathbb{P}\otimes Leb)\text{-a.e.} \tag{4.1.12}$$

Then, using the assumption that $X_0 = x_0$, X solves the wealth equation (3.2.31) for the portfolio process $\boldsymbol{\pi}$. As the solution is unique up to indistinguishability, we get $X(t) \equiv X^{\boldsymbol{\pi}}(t)$ a.s. for all $t \in [0, T]$.

In particular, from (4.1.12), $\boldsymbol{\pi} \in \mathcal{A}$ satisfies $\boldsymbol{\pi}(t) = (\boldsymbol{\sigma}^\top(t))^{-1}\boldsymbol{\Lambda}^X(t)$ $(\mathbb{P}\otimes Leb)$-a.e. □

Using Lemma 4.1.1, we can rewrite the value of the problem as

$$\mathcal{V} \overset{(4.1.1)}{=} \inf_{\boldsymbol{\pi}\in\mathcal{A}} \{\mathrm{E}\left(J(X^{\boldsymbol{\pi}}(T))\right)\} \equiv \inf_{\substack{X\in\mathbb{A},\\ X_0=x_0,\\ U(X)\neq\emptyset}} \{\mathrm{E}\left(J(X(T))\right)\}. \tag{4.1.13}$$

Note in the above equation as we move from the left-hand infimum to the right-hand infimum, the wealth process $X^{\boldsymbol{\pi}}$ becomes the process X and the constraints change from constraints on the portfolio process $\boldsymbol{\pi} \in \mathcal{A}$ to constraints on the process $X \in \mathbb{A}$.

The next step is to eliminate the constraints under the infimum on the process $X \in \mathbb{A}$. We do this by defining penalty functions on the space $\mathbb{A}$ which take value zero when the constraints are satisfied and value infinity otherwise.

We begin by defining a penalty function l_0 to take care of the initial wealth constraint $X_0 = x_0$. It assigns value zero if its argument equals x_0 and value infinity otherwise.

Definition 4.1.2. The penalty function l_0 takes values in $\{0, \infty\}$ and is given by

$$l_0(x) := \begin{cases} 0 & \text{if } x = x_0 \\ \infty & \text{otherwise,} \end{cases} \qquad \forall x \in \mathbb{R}. \tag{4.1.14}$$

Next we define a penalty function l_1 which gives value zero to those $X \in \mathbb{A}$ which have $U(X) \neq \emptyset$ and value infinity to those $X \in \mathbb{A}$ which have $U(X) = \emptyset$.

Definition 4.1.3. The penalty function l_1 takes values in $\{0, \infty\}$ and is given by

$$l_1(\omega, t, x, \nu, \boldsymbol{\lambda}) := \begin{cases} 0 & \text{if } \nu = r(\omega, t)x + \boldsymbol{\lambda}^\top \boldsymbol{\theta}(\omega, t) \text{ and } (\boldsymbol{\sigma}^\top(\omega, t))^{-1}\boldsymbol{\lambda} \in K \\ \infty & \text{otherwise,} \end{cases} \tag{4.1.15}$$

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

Together, the penalty functions l_0 and l_1 eliminate the constraints on $X \in \mathbb{A}$ under the infimum in (4.1.13).

Next we define a function l_T equal to the risk measure J. This function will only contribute meaningfully to the value of the functional Φ if the penalty functions l_0 and l_1 both equal zero.

Definition 4.1.4. The function l_T takes values in $\mathbb{R}$ and is given by

$$l_T(\omega, x) := J(\omega, x) \overset{(3.2.43)}{=} \frac{1}{2}a(\omega)x^2 + b(\omega)x + c(\omega), \tag{4.1.16}$$

for all $(\omega, x) \in \Omega \times \mathbb{R}$.

Remark 4.1.5. In order to take expectations of these functions, we must show that the functions l_1 and $l_T(X(T))$ are measurable (the measurability of the function l_0 is immediate from its definition). We show this in Proposition A.2.1.

Definition 4.1.6. The *primal cost functional* $\Phi : \mathbb{A} \to (-\infty, \infty]$ is given by

$$\Phi(X) := l_0(X_0) + \mathrm{E} \int_0^T l_1(t, X(t), \dot{X}(t), \boldsymbol{\Lambda}^X(t))\, \mathrm{d}t + \mathrm{E}\left(l_T(X(T))\right), \tag{4.1.17}$$

for all $X \equiv (X_0, \dot{X}, \boldsymbol{\Lambda}^X, \mathbf{0}) \in \mathbb{A}$.

For any $X \in \mathbb{A}$, we see from the definitions of the penalty functions l_0 and l_1 that if the constraints in (4.1.13) are not satisfied then $\Phi(X) = +\infty$. On the other hand, if the constraints are satisfied then $\Phi(X) = \mathrm{E}\left(l_T(X(T))\right) = \mathrm{E}\left(J(X(T))\right)$. Thus we can rewrite the value of the problem as

$$\mathcal{V} = \inf_{X \in \mathbb{A}} \{\Phi(X)\}. \tag{4.1.18}$$

Problem 4.1.7. The primal problem is to find $\bar{X} \equiv (\bar{X}_0, \dot{\bar{X}}, \boldsymbol{\Lambda}^{\bar{X}}, \mathbf{0}) \in \mathbb{A}$ such that the infimum of the primal cost functional Φ is attained, in other words

$$\Phi(\bar{X}) = \inf_{X \in \mathbb{A}} \{\Phi(X)\}. \tag{4.1.19}$$

4.2 Define the dual problem

Having defined the primal problem, we define the dual problem. This is formed by taking the convex conjugates of the functions given by (4.1.14), (4.1.15) and (4.1.16) to obtain the dual functions. Upon summing and taking expectations, the *dual cost functional* is obtained.

Definition 4.2.1. The dual function of l_0 is a function m_0 which takes values in $\mathbb{R} \cup \{\pm\infty\}$ and is given by

$$m_0(y) := l_0^\star(y) := \sup_{x\in\mathbb{R}}\{xy - l_0(x)\}, \qquad \forall y \in \mathbb{R}. \tag{4.2.1}$$

Definition 4.2.2. The dual function of l_1 is a function m_1 which takes values in $\mathbb{R} \cup \{\pm\infty\}$ and is given by

$$\begin{aligned} m_1(\omega, t, y, s, \xi) &:= l_1^\star(\omega, t, y, s, \xi) \\ &:= \sup_{\substack{x,\nu\in\mathbb{R} \\ \boldsymbol{\lambda}\in\mathbb{R}^N}} \left\{xs + \nu y + \boldsymbol{\lambda}^\top \xi - l_1(\omega, t, x, \nu, \boldsymbol{\lambda})\right\}, \end{aligned} \tag{4.2.2}$$

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

Definition 4.2.3. The dual function of l_T is a function m_T which takes values in $\mathbb{R} \cup \{\pm\infty\}$ and is given by

$$m_T(\omega, y) := l_T^\star(\omega, -y) := \sup_{x\in\mathbb{R}} \{x(-y) - l_T(\omega, x)\}, \tag{4.2.3}$$

for all $(\omega, y) \in \Omega \times \mathbb{R}$.

Remark 4.2.4. The reason for using a negative y in the argument of $l_T^\star$ in (4.2.3) is to allow us to use a result derived from Bismut [3], Proposition I-1. This will become apparent in the proof of Lemma 4.3.2.

Remark 4.2.5. In order to take expectations of the dual functions, we must show that the dual functions m_1 and $m_T(X(T))$ are measurable (the measurability of m_0 is immediate from its definition). This is shown in Proposition A.2.2.

Definition 4.2.6. The dual cost functional $\Psi : \mathbb{B} \to (-\infty, \infty]$ is given by

$$\Psi(Y) := m_0(Y_0) + \mathrm{E}\int_0^T m_1(t, Y(t), \dot{Y}(t), \boldsymbol{\Lambda}^Y(t))\,\mathrm{d}t + \mathrm{E}\left(m_T(Y(T))\right), \tag{4.2.4}$$

for all $Y \equiv (Y_0, \dot{Y}, \boldsymbol{\Lambda}^Y, \boldsymbol{\Gamma}^Y) \in \mathbb{B}$.

Remark 4.2.7. The dual cost functional Ψ is convex on $\mathbb{B}$.

Problem 4.2.8. The dual problem is to find $\bar{Y} \in \mathbb{B}$ such that the infimum of the dual cost functional $\Psi(Y)$ is attained, in other words

$$\Psi(\bar{Y}) = \inf_{Y\in\mathbb{B}}\{\Psi(Y)\}. \tag{4.2.5}$$

4.3 Relationship between the primal and dual functionals

In this section, we show that the sum of the primal cost functional and the dual cost functional is at least zero. This is critical to retrieving the solution to the primal problem (Problem 4.1.7) from the solution to the dual problem (Problem 4.2.8).

We will need the following proposition, which is a minor adaptation of Bismut [3], Proposition I-1. The proof can be found under Proposition A.2.3.

Proposition 4.3.1. *For any* $X \equiv (X_0, \dot{X}, \mathbf{\Lambda}^X, \mathbf{\Gamma}^X) \in \mathbb{B}$ *and* $Y \equiv (Y_0, \dot{Y}, \mathbf{\Lambda}^Y, \mathbf{\Gamma}^Y) \in \mathbb{B}$*, the process* $\{\mathbb{M}(X,Y)(t) : t \in [0,T]\}$ *defined by*

$$\begin{aligned} \mathbb{M}(X,Y)(t) :=& X(t)Y(t) - X_0Y_0 - \int_0^t \left(\dot{X}(\tau)Y(\tau) + X(\tau)\dot{Y}(\tau)\right) \mathrm{d}\tau \\ & - \sum_{n=1}^{N} \int_0^t \Lambda_n^X(\tau)\Lambda_n^Y(\tau)\,\mathrm{d}\tau - \sum_{i,j=1}^{D} \int_0^t \Gamma_{ij}^X(\tau)\Gamma_{ij}^Y(\tau)\,\mathrm{d}[\mathcal{Q}_{ij}](\tau) \end{aligned} \tag{4.3.1}$$

is such that $\mathbb{M}(X,Y) \in \mathcal{M}_0(\{\mathcal{F}_t\}, \mathbb{P})$.

We use Proposition 4.3.1 in the proof of the next lemma. This lemma is then used in Proposition 4.3.3 to show that the weak duality principle holds between the primal and dual cost functionals.

Lemma 4.3.2. *For any* $X \equiv (X_0, \dot{X}, \mathbf{\Lambda}^X, \mathbf{0}) \in \mathbb{A}$ *and* $Y \equiv (Y_0, \dot{Y}, \mathbf{\Lambda}^Y, \mathbf{\Gamma}^Y) \in \mathbb{B}$*, for all* $t \in [0,T]$*, we have a.s. the relations*

$$l_0(X_0) + m_0(Y_0) \geq X_0Y_0, \tag{4.3.2}$$

$$\begin{aligned} & l_1(t, X(t), \dot{X}(t), \mathbf{\Lambda}^X(t)) + m_1(t, Y(t), \dot{Y}(t), \mathbf{\Lambda}^Y(t)) \\ & \qquad \geq X(t)\dot{Y}(t) + \dot{X}(t)Y(t) + \left(\mathbf{\Lambda}^X\right)^\top(t)\mathbf{\Lambda}^Y(t), \end{aligned} \tag{4.3.3}$$

$$l_T(X(T)) + m_T(Y(T)) \geq -X(T)Y(T), \tag{4.3.4}$$

and

$$\begin{aligned} & X_0Y_0 + E\int_0^T \left(X(t)\dot{Y}(t) + \dot{X}(t)Y(t) + \left(\mathbf{\Lambda}^X\right)^\top(t)\mathbf{\Lambda}^Y(t)\right) \mathrm{d}t - E(X(T)Y(T)) \\ & = 0. \end{aligned} \tag{4.3.5}$$

Proof. Fix $X \equiv (X_0, \dot{X}, \mathbf{\Lambda}^X, \mathbf{0}) \in \mathbb{A}$ and $Y \equiv (Y_0, \dot{Y}, \mathbf{\Lambda}^Y, \mathbf{\Gamma}^Y) \in \mathbb{B}$. From the definitions of the dual functions in (4.2.1) - (4.2.3), it is immediate that (4.3.2) - (4.3.4) hold.

From Proposition 4.3.1, the process $\mathbb{M}(X,Y) = \{\mathbb{M}(X,Y)(t) : t \in [0,T]\}$ defined by

$$\begin{aligned}\mathbb{M}(X,Y)(t) := X(t)Y(t) - X_0Y_0 - \int_0^t \left(\dot{X}(\tau)Y(\tau) + X(\tau)\dot{Y}(\tau)\right)\,\mathrm{d}\tau \\ - \int_0^t \left(\mathbf{\Lambda}^X\right)^\top(\tau)\mathbf{\Lambda}^Y(\tau)\,\mathrm{d}\tau,\end{aligned} \tag{4.3.6}$$

is such that $\mathbb{M}(X,Y) \in \mathcal{M}_0(\{\mathcal{F}_t\},\mathbb{P})$. Setting $t = T$ and taking expectations in (4.3.6), we find

$$\begin{aligned}\mathrm{E}\left(\mathbb{M}(X,Y)(T)\right) &= \mathrm{E}\left(X(T)Y(T)\right) - X_0Y_0 - \mathrm{E}\int_0^T \left(\dot{X}(\tau)Y(\tau) + X(\tau)\dot{Y}(\tau)\right)\,\mathrm{d}\tau \\ &\quad - \mathrm{E}\int_0^T \left(\mathbf{\Lambda}^X\right)^\top(\tau)\mathbf{\Lambda}^Y(\tau)\,\mathrm{d}\tau \\ &= 0,\end{aligned} \tag{4.3.7}$$

giving (4.3.5). □

The following "weak duality" result is absolutely essential for applying convex duality.

Proposition 4.3.3. *For any* $X \equiv (X_0, \dot{X}, \mathbf{\Lambda}^X, \mathbf{0}) \in \mathbb{A}$ *and* $Y \equiv (Y_0, \dot{Y}, \mathbf{\Lambda}^Y, \mathbf{\Gamma}^Y) \in \mathbb{B}$*, we have the relation*

$$\Phi(X) + \Psi(Y) \geq 0. \tag{4.3.8}$$

Hence

$$\inf_{X\in\mathbb{A}}\{\Phi(X)\} \geq \sup_{Y\in\mathbb{B}}\{-\Psi(Y)\}. \tag{4.3.9}$$

Proof. Fix $X \equiv (X_0, \dot{X}, \mathbf{\Lambda}^X, \mathbf{0}) \in \mathbb{A}$ and $Y \equiv (Y_0, \dot{Y}, \mathbf{\Lambda}^Y, \mathbf{\Gamma}^Y) \in \mathbb{B}$. Adding (4.1.17) and (4.2.4) and using Lemma 4.3.2, we obtain

$$\begin{aligned}\Phi(X) + \Psi(Y) &= l_0(X_0) + m_0(Y_0) \\ &\quad + \mathrm{E}\int_0^T \left(l_1(t, X(t), \dot{X}(t), \mathbf{\Lambda}^X(t)) + m_1(t, Y(t), \dot{Y}(t), \mathbf{\Lambda}^Y(t))\right)\,\mathrm{d}t \\ &\quad + \mathrm{E}\left(l_T(X(T)) + m_T(Y(T))\right) \\ &\overset{(4.3.2),(4.3.3),(4.3.4)}{\geq} X_0Y_0 \\ &\quad + \mathrm{E}\int_0^T \left(X(t)\dot{Y}(t) + \dot{X}(t)Y(t) + \left(\mathbf{\Lambda}^X\right)^\top(t)\mathbf{\Lambda}^Y(t)\right)\,\mathrm{d}t \\ &\quad - \mathrm{E}\left(X(T)Y(T)\right) \\ &\overset{(4.3.5)}{=} 0.\end{aligned} \tag{4.3.10}$$

Hence we get (4.3.8), from which (4.3.9) follows immediately. □

4.4 Necessary and sufficient conditions to solve the primal problem

Here we give necessary and sufficient conditions to solve the primal problem. We first show in Corollary 4.4.1 that the sum of the primal and dual cost functionals is zero for some pair $(\bar{X}, \bar{Y}) \in \mathbb{A} \times \mathbb{B}$ if and only if $\bar{X} \in \mathbb{A}$ solves the primal problem and $\bar{Y} \in \mathbb{B}$ solves the dual problem. Then in Proposition 4.4.8, we show that the sum of the primal and dual cost functionals equals zero if and only if certain "optimality" conditions hold between the solutions $\bar{X} \in \mathbb{A}$ and $\bar{Y} \in \mathbb{B}$.

Corollary 4.4.1. *For any pair $(\bar{X}, \bar{Y}) \in \mathbb{A} \times \mathbb{B}$, we have*

$$\Phi(\bar{X}) + \Psi(\bar{Y}) = 0 \tag{4.4.1}$$

if and only if

$$\Phi(\bar{X}) = \inf_{X \in \mathbb{A}} \{\Phi(X)\} = \sup_{Y \in \mathbb{B}} \{-\Psi(Y)\} = -\Psi(\bar{Y}). \tag{4.4.2}$$

Proof. That (4.4.2) implies (4.4.1) is immediate.

Suppose $(\bar{X}, \bar{Y}) \in \mathbb{A} \times \mathbb{B}$ are such that (4.4.1) holds. From Proposition 4.3.3, for all $X \in \mathbb{A}$, we have the inequality $\Phi(X) + \Psi(\bar{Y}) \geq 0$, that is

$$\begin{aligned} \Phi(X) &\geq -\Psi(\bar{Y}), \quad \forall X \in \mathbb{A} \\ \Rightarrow \inf_{X \in \mathbb{A}} \{\Phi(X)\} &\geq -\Psi(\bar{Y}) \overset{(4.4.1)}{=} \Phi(\bar{X}). \end{aligned} \tag{4.4.3}$$

However, by definition of the infimum, we must have equality, that is

$$\Phi(\bar{X}) = \inf_{X \in \mathbb{A}} \{\Phi(X)\}. \tag{4.4.4}$$

Then from the definition of the supremum and Proposition 4.3.3, we have

$$\sup_{Y \in \mathbb{B}} \{-\Psi(Y)\} \geq -\Psi(\bar{Y}) \overset{(4.4.1)}{=} \Phi(\bar{X}) \overset{(4.4.4)}{=} \inf_{X \in \mathbb{A}} \{\Phi(X)\} \overset{(4.3.9)}{\geq} \sup_{Y \in \mathbb{B}} \{-\Psi(Y)\}. \tag{4.4.5}$$

We must have equality in (4.4.5), whence we obtain (4.4.2). □

Remark 4.4.2. It is immediate from Corollary 4.4.1 that if we can find a pair $(\bar{X}, \bar{Y}) \in \mathbb{A} \times \mathbb{B}$ such that $\Phi(\bar{X}) + \Psi(\bar{Y}) = 0$, then $\bar{X} \in \mathbb{A}$ solves the primal problem (Problem 4.1.7) and $\bar{Y} \in \mathbb{B}$ solves the dual problem (Problem 4.2.8).

Corollary 4.4.1 is not descriptive enough for our purposes. In order to construct the primal solution $\bar{X} \in \mathbb{A}$ from the dual solution $\bar{Y} \in \mathbb{B}$ we need to be more explicit about how the components of each one relate to the other. Thus we arrive at the need for Proposition 4.4.8. However, we will first introduce some notation, which we will use to simplify the dual function m_1. We will also simplify the dual function m_T. We use these simplifications in Proposition 4.4.8.

Definition 4.4.3. For each $Y \equiv (Y_0, \dot{Y}, \mathbf{\Lambda}^Y, \mathbf{\Gamma}^Y) \in \mathbb{B}$, we define a process $\mathbf{\Theta}_Y = \{\mathbf{\Theta}_Y(t)\}$ by

$$\mathbf{\Theta}_Y(\omega, t) := -\boldsymbol{\sigma}(\omega, t)\left(\boldsymbol{\theta}(\omega, t)Y(\omega, t) + \mathbf{\Lambda}^Y(\omega, t)\right), \quad \forall(\omega, t) \in \Omega \times [0, T]. \tag{4.4.6}$$

We next define the *support function* of the convex set $-K$.

Definition 4.4.4. The support function δ of the convex set $-K$ is defined as

$$\delta(\mathbf{z}) := \sup_{\boldsymbol{\pi} \in K}\left\{-\boldsymbol{\pi}^\top \mathbf{z}\right\}, \quad \forall \mathbf{z} \in \mathbb{R}^N. \tag{4.4.7}$$

Remark 4.4.5. The support function δ has many nice properties that we will use. From Condition 3.2.31, $\mathbf{0} \in K$, which means that $\delta(\mathbf{z}) \geq 0$ for all $\mathbf{z} \in \mathbb{R}^N$. It is a lower semi-continuous, convex function, which is positively homogeneous, that is $\delta(\epsilon\mathbf{z}) = \epsilon\delta(\mathbf{z})$ for all $\epsilon \geq 0$, and subadditive, that is $\delta(\mathbf{z}_1 + \mathbf{z}_2) \leq \delta(\mathbf{z}_1) + \delta(\mathbf{z}_2)$ for all $\mathbf{z}_1, \mathbf{z}_2 \in \mathbb{R}^N$. These properties are given in Karatzas and Shreve [31], page 206.

Lemma 4.4.6. *For all* $(\omega, t, y, s, \xi) \in \Omega \times [0, T] \times \mathbb{R} \times \mathbb{R} \times \mathbb{R}^N$,

$$m_1(\omega, t, y, s, \xi) = \begin{cases} \delta\left(-\boldsymbol{\sigma}(\omega, t)\left(\boldsymbol{\theta}(\omega, t)y + \xi\right)\right) & \text{if } s + r(\omega, t)y = 0 \\ \infty & \text{otherwise,} \end{cases} \tag{4.4.8}$$

for the dual function m_1 *given by (4.2.2).*

Proof. Using the definition of l_1, given by (4.1.15), we can re-write (4.2.2) as

$$\begin{aligned} m_1(\omega, t, y, s, \xi) := \sup_{\substack{x, \nu \in \mathbb{R} \\ \boldsymbol{\lambda} \in \mathbb{R}^N}} \Big\{ & xs + \nu y + \boldsymbol{\lambda}^\top \xi \,|\, \nu = r(\omega, t)x + \boldsymbol{\lambda}^\top \boldsymbol{\theta}(\omega, t), \\ & \text{and } \left(\boldsymbol{\sigma}^\top(\omega, t)\right)^{-1} \boldsymbol{\lambda} \in K \Big\}. \end{aligned} \tag{4.4.9}$$

Defining $\boldsymbol{\pi}(\omega, t) := (\boldsymbol{\sigma}^\top(\omega, t))^{-1}\boldsymbol{\lambda}$, so that $\boldsymbol{\pi}(\omega, t) \in K$, we get $\boldsymbol{\lambda} = \boldsymbol{\sigma}^\top(\omega, t)\boldsymbol{\pi}(\omega, t)$. Substituting $\nu = r(\omega, t)x + \boldsymbol{\lambda}^\top\boldsymbol{\theta}(\omega, t)$ (the first constraint in the supremum above) and $\boldsymbol{\lambda} = \boldsymbol{\sigma}^\top(\omega, t)\boldsymbol{\pi}(\omega, t)$ into $xs + \nu y + \boldsymbol{\lambda}^\top\xi$, we obtain

$$\begin{aligned} m_1(\omega, t, y, s, \xi) &= \sup_{\substack{x \in \mathbb{R} \\ \boldsymbol{\pi}(\omega, t) \in K}} \left\{x\left(s + r(\omega, t)y\right) + \boldsymbol{\pi}^\top(\omega, t)\boldsymbol{\sigma}(\omega, t)\left(\boldsymbol{\theta}(\omega, t)y + \xi\right)\right\} \\ &\overset{(4.4.7)}{=} \begin{cases} \delta\left(-\boldsymbol{\sigma}(\omega, t)\left(\boldsymbol{\theta}(\omega, t)y + \xi\right)\right) & \text{if } s + r(\omega, t)y = 0 \\ \infty & \text{otherwise.} \end{cases} \end{aligned} \tag{4.4.10}$$

□

For $Y \equiv (Y_0, \dot{Y}, \mathbf{\Lambda}^Y, \mathbf{\Gamma}^Y) \in \mathbb{B}$, we can substitute the process $\mathbf{\Theta}_Y$, given by (4.4.6), into (4.4.8) to obtain a.s. for all $t \in [0, T]$,

$$m_1(t, Y(t), \dot{Y}(t), \mathbf{\Lambda}^Y(t)) = \begin{cases} \delta\left(\mathbf{\Theta}_Y(t)\right) & \text{if } \dot{Y}(t) + r(t)Y(t) = 0 \\ \infty & \text{otherwise.} \end{cases} \tag{4.4.11}$$

The next lemma gives an explicit form for the dual function m_T.

Lemma 4.4.7. *For all $(\omega, y) \in \Omega \times \mathbb{R}$, we have*

$$m_T(\omega, y) = \frac{(y + b(\omega))^2}{2a(\omega)} - c(\omega), \tag{4.4.12}$$

for the dual function m_T given by (4.2.3).

Proof. Substituting for l_T from (4.1.16) into (4.2.3), we obtain

$$\begin{aligned} m_T(\omega, y) &= \sup_{x \in \mathbb{R}} \left\{ x(-y) - \frac{1}{2}a(\omega)x^2 - b(\omega)x - c(\omega) \right\} \\ &= \sup_{x \in \mathbb{R}} \left\{ -\frac{a(\omega)}{2} \left(\left(x + \frac{y + b(\omega)}{a(\omega)} \right)^2 - \frac{(y + b(\omega))^2}{a^2(\omega)} \right) - c(\omega) \right\} \\ &= \sup_{x \in \mathbb{R}} \left\{ -\frac{a(\omega)}{2} \left(x + \frac{y + b(\omega)}{a(\omega)} \right)^2 \right\} + \frac{(y + b(\omega))^2}{2a(\omega)} - c(\omega). \end{aligned} \tag{4.4.13}$$

As $(x + \frac{y+b(\omega)}{a(\omega)})^2 \geq 0$ and, by the strict positivity of the random variable a assumed in Condition 3.2.26, $-\frac{a(\omega)}{2} < 0$, then the supremum equals zero. Thus we get

$$m_T(\omega, y) = \frac{(y + b(\omega))^2}{2a(\omega)} - c(\omega). \tag{4.4.14}$$

□

Having found simplifying expressions for the dual functions m_1 and m_T, we now state and prove Proposition 4.4.8.

Proposition 4.4.8. *For $\bar{X} \equiv (\bar{X}_0, \dot{\bar{X}}, \mathbf{\Lambda}^{\bar{X}}, \mathbf{0}) \in \mathbb{A}$ and $\bar{Y} \equiv (\bar{Y}_0, \dot{\bar{Y}}, \mathbf{\Lambda}^{\bar{Y}}, \mathbf{\Gamma}^{\bar{Y}}) \in \mathbb{B}$, we have*

$$\Phi(\bar{X}) + \Psi(\bar{Y}) = 0, \tag{4.4.15}$$

if and only if the following are satisfied:

$$\bar{X}_0 = x_0. \tag{4.4.16}$$

$$\bar{X}(T) = -\left(\frac{\bar{Y}(T) + b}{a} \right) \quad a.s. \tag{4.4.17}$$

$$\dot{\bar{Y}}(t) + r(t)\bar{Y}(t) = 0 \qquad (\mathbb{P} \otimes Leb)\text{-}a.e. \tag{4.4.18}$$

For

$$\bar{\boldsymbol{\pi}}(\omega, t) := (\boldsymbol{\sigma}^\top(\omega, t))^{-1} \mathbf{\Lambda}^{\bar{X}}(\omega, t), \quad \forall(\omega, t) \in \Omega \times [0, T], \tag{4.4.19}$$

we have $(\mathbb{P} \otimes Leb)$-a.e,

$$\delta(\mathbf{\Theta}_{\bar{Y}}(t)) + \bar{\boldsymbol{\pi}}^\top(t)\mathbf{\Theta}_{\bar{Y}}(t) = 0 \tag{4.4.20}$$

and

$$\bar{\boldsymbol{\pi}} \in U(\bar{X}), \tag{4.4.21}$$

for the set $U(\cdot)$ given by (4.1.3).

Proof. Suppose $\bar{X} \equiv (\bar{X}_0, \dot{\bar{X}}, \mathbf{\Lambda}^{\bar{X}}, \mathbf{0}) \in \mathbb{A}$ and $\bar{Y} \equiv (\bar{Y}_0, \dot{\bar{Y}}, \mathbf{\Lambda}^{\bar{Y}}, \mathbf{\Gamma}^{\bar{Y}}) \in \mathbb{B}$ satisfy $\Phi(\bar{X}) + \Psi(\bar{Y}) = 0$. As

$$\begin{aligned}
\Phi(\bar{X}) + \Psi(\bar{Y}) &\overset{(4.1.17),(4.2.4)}{=} l_0(\bar{X}_0) + m_0(Y_0) \\
&\quad + \mathrm{E} \int_0^T \left(l_1(t, \bar{X}(t), \dot{\bar{X}}(t), \mathbf{\Lambda}^{\bar{X}}(t)) + m_1(t, \bar{Y}(t), \dot{\bar{Y}}(t), \mathbf{\Lambda}^{\bar{Y}}(t)) \right) \mathrm{d}t \\
&\quad + \mathrm{E} \left(l_T(\bar{X}(T)) + m_T(\bar{Y}(T)) \right) \\
&\overset{(4.4.15)}{=} 0,
\end{aligned} \tag{4.4.22}$$

it is immediate from Lemma 4.3.2 that the following three relations must hold a.s. since otherwise we would not have the above equality with zero:

$$l_0(\bar{X}_0) + m_0(\bar{Y}_0) = \bar{X}_0 \bar{Y}_0. \tag{4.4.23}$$

$$\begin{aligned}
l_1(t, &\bar{X}(t), \dot{\bar{X}}(t), \mathbf{\Lambda}^{\bar{X}}(t)) + m_1(t, \bar{Y}(t), \dot{\bar{Y}}(t), \mathbf{\Lambda}^{\bar{Y}}(t)) \\
&= \bar{X}(t)\dot{\bar{Y}}(t) + \dot{\bar{X}}(t)\bar{Y}(t) + \left(\mathbf{\Lambda}^{\bar{X}}\right)^{\top}(t)\mathbf{\Lambda}^{\bar{Y}}(t), \quad \forall t \in [0, T].
\end{aligned} \tag{4.4.24}$$

$$l_T(\bar{X}(T)) + m_T(\bar{Y}(T)) = -\bar{X}(T)\bar{Y}(T). \tag{4.4.25}$$

First consider (4.4.23). From (4.1.14) and (4.2.1), we have

$$l_0(\bar{X}_0) + m_0(\bar{Y}_0) = \bar{X}_0 \bar{Y}_0 \quad \Leftrightarrow \quad \bar{X}_0 = x_0. \tag{4.4.26}$$

This gives (4.4.16).

Next consider (4.4.25).

$$\begin{aligned}
-\bar{X}(T)\bar{Y}(T) &\overset{(4.4.25)}{=} l_T(\bar{X}(T)) + m_T(\bar{Y}(T)) \\
&\overset{(4.1.16),(4.4.12)}{=} \frac{1}{2} a\bar{X}^2(T) + b\bar{X}(T) + c + \frac{(\bar{Y}(T) + b)^2}{2a} - c.
\end{aligned} \tag{4.4.27}$$

After some algebra, we find that the above equation holds if and only if $\bar{X}(T) = -\frac{\bar{Y}(T)+b}{a}$ a.s., giving (4.4.17).

Finally, consider (4.4.24). From (4.1.15), we have

$$l_1(t, \bar{X}(t), \dot{\bar{X}}(t), \mathbf{\Lambda}^{\bar{X}}(t)) = \begin{cases} 0 & \text{if } \dot{\bar{X}}(t) = r(t)\bar{X}(t) + (\mathbf{\Lambda}^{\bar{X}})^{\top}(t)\boldsymbol{\theta}(t) \\ & \quad \text{and } (\boldsymbol{\sigma}^{\top}(t))^{-1}\mathbf{\Lambda}^{\bar{X}}(t) \in K \\ \infty & \text{otherwise.} \end{cases} \tag{4.4.28}$$

Recalling the definition of $\bar{\boldsymbol{\pi}}$ in (4.4.19), we see that as $\mathbf{\Lambda}^{\bar{X}} \in L^2(\mathbf{W})$ and $\boldsymbol{\sigma}$ is uniformly bounded, we have $\bar{\boldsymbol{\pi}} \in L^2(\mathbf{W})$.

Adding (4.4.28) to (4.4.11), and substituting for $\mathbf{\Lambda}^{\bar{X}}$ from (4.4.19), we obtain

$$l_1(t, \bar{X}(t), \dot{\bar{X}}(t), \mathbf{\Lambda}^{\bar{X}}(t)) + m_1(t, \bar{Y}(t), \dot{\bar{Y}}(t), \mathbf{\Lambda}^{\bar{Y}}(t)) \\ = \begin{cases} \delta\left(\mathbf{\Theta}_Y(t)\right) & \text{if } \dot{\bar{X}}(t) = r(t)\bar{X}(t) + \bar{\boldsymbol{\pi}}^\top(t)\boldsymbol{\sigma}(t)\boldsymbol{\theta}(t), \\ & \bar{\boldsymbol{\pi}}(t) \in K \\ & \text{ and } \dot{\bar{Y}}(t) + r(t)\bar{Y}(t) = 0 \\ \infty & \text{otherwise.} \end{cases} \tag{4.4.29}$$

Comparing (4.4.29) to (4.4.24), we must have $(\mathbb{P} \otimes Leb)$-a.e.,

$$\delta(\mathbf{\Theta}_{\bar{Y}}(t)) = \bar{X}(t)\dot{\bar{Y}}(t) + \dot{\bar{X}}(t)\bar{Y}(t) + \left(\mathbf{\Lambda}^{\bar{X}}\right)^\top(t)\mathbf{\Lambda}^{\bar{Y}}(t), \tag{4.4.30}$$

$$\dot{\bar{X}}(t) = r(t)\bar{X}(t) + \bar{\boldsymbol{\pi}}^\top(t)\boldsymbol{\sigma}(t)\boldsymbol{\theta}(t), \tag{4.4.31}$$

$$\bar{\boldsymbol{\pi}}(t) \in K, \tag{4.4.32}$$

and

$$\dot{\bar{Y}}(t) + r(t)\bar{Y}(t) = 0. \tag{4.4.33}$$

(4.4.33) is the same as (4.4.18).

As we noted above, $\bar{\boldsymbol{\pi}} \in L^2(\mathbf{W})$. From this fact and (4.4.32), we get $\bar{\boldsymbol{\pi}} \in \mathcal{A}$. Furthermore, from (4.4.31) and the definition of $\bar{\boldsymbol{\pi}}$ in (4.4.19), it follows that $\bar{\boldsymbol{\pi}} \in U(\bar{X})$. This gives (4.4.21).

From (4.4.30), we get

$$\begin{aligned} \delta(\mathbf{\Theta}_{\bar{Y}}(t)) &\overset{(4.4.33)}{=} \bar{X}(t)\left(-r(t)\bar{Y}(t)\right) + \dot{\bar{X}}(t)\bar{Y}(t) + \left(\mathbf{\Lambda}^{\bar{X}}\right)^\top(t)\mathbf{\Lambda}^{\bar{Y}}(t) \\ &= \bar{Y}(t)\left(\dot{\bar{X}}(t) - r(t)\bar{X}(t)\right) + \left(\mathbf{\Lambda}^{\bar{X}}\right)^\top(t)\mathbf{\Lambda}^{\bar{Y}}(t) \\ &\overset{(4.4.31)}{=} \bar{Y}(t)\left(\left(\mathbf{\Lambda}^{\bar{X}}\right)^\top(t)\boldsymbol{\theta}(t)\right) + \left(\mathbf{\Lambda}^{\bar{X}}\right)^\top(t)\mathbf{\Lambda}^{\bar{Y}}(t) \\ &= \left(\mathbf{\Lambda}^{\bar{X}}\right)^\top(t)\left(\boldsymbol{\theta}(t)\bar{Y}(t) + \mathbf{\Lambda}^{\bar{Y}}(t)\right) \\ &\overset{(4.4.19)}{=} \bar{\boldsymbol{\pi}}^\top(t)\boldsymbol{\sigma}(t)\left(\boldsymbol{\theta}(t)\bar{Y}(t) + \mathbf{\Lambda}^{\bar{Y}}(t)\right) \\ &\overset{(4.4.6)}{=} -\bar{\boldsymbol{\pi}}^\top(t)\mathbf{\Theta}_{\bar{Y}}(t). \end{aligned} \tag{4.4.34}$$

Hence (4.4.20) holds.

Conversely, suppose (4.4.16) - (4.4.21) hold for $\bar{X} \equiv (\bar{X}_0, \dot{\bar{X}}, \mathbf{\Lambda}^{\bar{X}}, \mathbf{0}) \in \mathbb{A}$ and $\bar{Y} \equiv (\bar{Y}_0, \dot{\bar{Y}}, \mathbf{\Lambda}^{\bar{Y}}, \mathbf{\Gamma}^{\bar{Y}}) \in \mathbb{B}$.

From (4.4.16) and (4.4.26), (4.4.23) must hold.

From (4.3.4) of Lemma 4.3.2,

$$l_T(\bar{X}(T)) + m_T(\bar{Y}(T)) \geq -\bar{X}(T)\bar{Y}(T) \quad \text{a.s.} \tag{4.4.35}$$

Substituting from (4.4.17) and using (4.1.16) and (4.4.12), after some algebra we find that equality must hold in the above inequality. This gives (4.4.25).

Similarly, from (4.3.3) of Lemma 4.3.2, we have a.s. for all $t \in [0, T]$,

$$\begin{aligned} l_1(t, X(t), \dot{X}(t), \mathbf{\Lambda}^X(t)) + m_1(t, Y(t), \dot{Y}(t), \mathbf{\Lambda}^Y(t)) \\ \geq X(t)\dot{Y}(t) + \dot{X}(t)Y(t) + \left(\mathbf{\Lambda}^X\right)^\top(t)\mathbf{\Lambda}^Y(t). \end{aligned} \tag{4.4.36}$$

Using (4.4.29) and (4.4.18) - (4.4.21), we find

$$\delta(\mathbf{\Theta}_{\bar{Y}}(t)) \geq X(t)\dot{Y}(t) + \dot{X}(t)Y(t) + \left(\mathbf{\Lambda}^X\right)^\top(t)\mathbf{\Lambda}^Y(t). \tag{4.4.37}$$

By the same calculation as in (4.4.34) (so that we show the first line of (4.4.34) holds), we must have equality in the above inequality. This gives (4.4.24).

It follows from (4.4.23) - (4.4.25) that $\Phi(\bar{X}) + \Psi(\bar{Y}) = 0$. □

Remark 4.4.9. If we can find $\bar{Y} \in \mathbb{B}$ that solves the dual problem (Problem 4.2.8) then, given any $\bar{X} \in \mathbb{A}$ which, together with $\bar{Y} \in \mathbb{B}$, satisfies (4.4.16) - (4.4.21) of Proposition 4.4.8, it follows that $\Phi(\bar{X}) + \Psi(\bar{Y}) = 0$. Thus $\bar{X}$ solves the primal problem (Problem 4.1.7).

Before we do this, we must show that there exists $\bar{Y} \in \mathbb{B}$ which solves the dual problem. This is the focus of the next section.

4.5 Existence of a solution to the dual problem

In this section, we show that a solution to the dual problem exists. In other words, we show that there exists $\bar{Y} \in \mathbb{B}$ satisfying

$$\Psi(\bar{Y}) = \inf_{Y \in \mathbb{B}} \{\Psi(Y)\}, \tag{4.5.1}$$

for the dual cost functional Ψ given by (4.2.4), that is

$$\Psi(Y) = m_0(Y_0) + \mathrm{E}\int_0^T m_1(t, Y(t), \dot{Y}(t), \mathbf{\Lambda}^Y(t))\,\mathrm{d}t + \mathrm{E}\left(m_T(Y(T))\right), \tag{4.5.2}$$

for $Y \equiv (Y_0, \dot{Y}, \mathbf{\Lambda}^Y, \mathbf{\Gamma}^Y) \in \mathbb{B}$. We show the existence of $\bar{Y} \in \mathbb{B}$ indirectly, by constructing another functional $\tilde{\Psi}$ which is very closely related to the dual cost functional Ψ. We show that a solution to a problem involving $\tilde{\Psi}$ exists and, from this solution, we can easily find the solution $\bar{Y} \in \mathbb{B}$ of (4.5.1).

From (4.4.11) and (4.4.18) of Proposition 4.4.8, we note first that a solution $\bar{Y} \in \mathbb{B}$ to the dual problem must satisfy $\dot{\bar{Y}}(t) = -r(t)\bar{Y}(t)$ $(\mathbb{P} \otimes Leb)$-a.e. Motivated by this fact, we define a space $\mathbb{B}_1 \subset \mathbb{B}$ as

$$\mathbb{B}_1 := \left\{ Y \equiv (Y_0, \dot{Y}, \mathbf{\Lambda}^Y, \mathbf{\Gamma}^Y) \in \mathbb{B} \,|\, \dot{Y}(t) = -r(t)Y(t) \ (\mathbb{P} \otimes Leb)\text{-a.e.} \right\}. \tag{4.5.3}$$

We observe from (4.4.11) that for $Y \equiv (Y_0, \dot{Y}, \mathbf{\Lambda}^Y, \mathbf{\Gamma}^Y) \in \mathbb{B} - \mathbb{B}_1$, we have $m_1(t, Y(t), \dot{Y}(t), \mathbf{\Lambda}^Y(t)) = +\infty$ $(\mathbb{P} \otimes Leb)$-a.e. Substituting this m_1 into (4.5.2), we get $\Psi(Y) = +\infty$. Thus, without loss of generality, we can restrict our attention to the space $\mathbb{B}_1$ when solving (4.5.1) since

$$\Psi(\bar{Y}) = \inf_{Y \in \mathbb{B}} \{\Psi(Y)\} = \inf_{Y \in \mathbb{B}_1} \{\Psi(Y)\}. \tag{4.5.4}$$

Next we define a map Ξ, which explicitly shows the component parts of an element of $\mathbb{B}_1$, in terms of $\mathbb{R} \times L^2(\mathbf{W}) \times L^2(\mathcal{Q})$. We will define the functional $\tilde{\Psi}$ on these components.

Definition 4.5.1. For each $t \in [0, T]$, define

$$\beta(t) := \exp\left\{-\int_0^t r(\tau)\,\mathrm{d}\tau\right\}, \tag{4.5.5}$$

and for all $y \in \mathbb{R}$, $\boldsymbol{\lambda} \in L^2(\mathbf{W})$ and $\boldsymbol{\gamma} \equiv (\gamma_{ij})_{i,j=1}^D \in L^2(\mathcal{Q})$, define

$$\mathbb{J}(\boldsymbol{\lambda})(t) := \int_0^t \beta^{-1}(\tau)\boldsymbol{\lambda}^\top(\tau)\,\mathrm{d}\mathbf{W}(\tau), \tag{4.5.6}$$

$$\mathbb{I}(\boldsymbol{\gamma})(t) := \sum_{i,j=1}^D \int_0^t \beta^{-1}(\tau)\gamma_{ij}(\tau)\,\mathrm{d}\mathcal{Q}_{ij}(\tau), \tag{4.5.7}$$

and

$$\Xi(y, \boldsymbol{\lambda}, \boldsymbol{\gamma})(t) := \beta(t)\left(y + \mathbb{J}(\boldsymbol{\lambda})(t) + \mathbb{I}(\boldsymbol{\gamma})(t)\right). \tag{4.5.8}$$

Lemma 4.5.2. *Let $(y, \boldsymbol{\lambda}, \boldsymbol{\gamma}) \in \mathbb{R} \times L^2(\mathbf{W}) \times L^2(\mathcal{Q})$ and set $Y \equiv \Xi(y, \boldsymbol{\lambda}, \boldsymbol{\gamma})$. Then*

$$Y \equiv (y, -rY_-, \boldsymbol{\lambda}, \boldsymbol{\gamma}) \in \mathbb{B}_1, \tag{4.5.9}$$

so that, recalling (4.5.3), we have

$$Y_0 = y \in \mathbb{R}, \quad \dot{Y} = -rY_- \in L_{21}, \quad \mathbf{\Lambda}^Y = \boldsymbol{\lambda} \in L^2(\mathbf{W}) \quad \text{and} \quad \mathbf{\Gamma}^Y = \boldsymbol{\gamma} \in L^2(\mathcal{Q}). \tag{4.5.10}$$

Proof. Fix $(y, \boldsymbol{\lambda}, \boldsymbol{\gamma}) \in \mathbb{R} \times L^2(\mathbf{W}) \times L^2(\mathcal{Q})$. Setting $Y \equiv \Xi(y, \boldsymbol{\lambda}, \boldsymbol{\gamma})$, we have a.s. for all $t \in [0, T]$,

$$\begin{aligned} Y(t) &= \Xi(y, \boldsymbol{\lambda}, \boldsymbol{\gamma})(t) \\ &\overset{(4.5.8)}{=} \beta(t)\left(y + \mathbb{J}(\boldsymbol{\lambda})(t) + \mathbb{I}(\boldsymbol{\gamma})(t)\right). \end{aligned} \tag{4.5.11}$$

We begin by showing that Y is square-integrable. Expanding the last line of (4.5.11) using (4.5.6) and (4.5.7), squaring and using the fact that $\beta(t) \le 1$ a.s., we get

$$\begin{aligned} |Y(t)|^2 &\le \left| y + \int_0^t \beta^{-1}(\tau)\boldsymbol{\lambda}^\top(\tau)\,\mathrm{d}\mathbf{W}(\tau) + \sum_{i,j=1}^D \int_0^t \beta^{-1}(\tau)\gamma_{ij}(\tau)\,\mathrm{d}\mathcal{Q}_{ij}(\tau)\right|^2 \\ &\le 3|y|^2 + 3\left|\int_0^t \beta^{-1}(\tau)\boldsymbol{\lambda}^\top(\tau)\,\mathrm{d}\mathbf{W}(\tau)\right|^2 + 3\left|\sum_{i,j=1}^D \int_0^t \beta^{-1}(\tau)\gamma_{ij}(\tau)\,\mathrm{d}\mathcal{Q}_{ij}(\tau)\right|^2. \end{aligned} \tag{4.5.12}$$

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

$$\mathrm{E}\left(\sup_{t\in[0,T]}|Y(t)|^2\right) \le 3|y|^2 + 3\kappa_\beta^2 \mathrm{E}\left(\sup_{t\in[0,T]}\left|\int_0^t \boldsymbol{\lambda}^\top(\tau)\,\mathrm{d}\mathbf{W}(\tau)\right|^2\right) + 3\kappa_\beta^2\mathrm{E}\left(\sup_{t\in[0,T]}\left|\sum_{i,j=1}^D\int_0^t \gamma_{ij}(\tau)\,\mathrm{d}\mathcal{Q}_{ij}(\tau)\right|^2\right). \tag{4.5.13}$$

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

$$\begin{aligned}\mathrm{E}\left(\sup_{t\in[0,T]}|Y(t)|^2\right) &\overset{\text{Doob}}{\le} 3|y|^2 + 12\kappa_\beta^2\mathrm{E}\left|\int_0^T \boldsymbol{\lambda}^\top(\tau)\,\mathrm{d}\mathbf{W}(\tau)\right|^2 \\ &\qquad + 12\kappa_\beta^2\mathrm{E}\left|\sum_{i,j=1}^D\int_0^T \gamma_{ij}(\tau)\,\mathrm{d}\mathcal{Q}_{ij}(\tau)\right|^2 \\ &\overset{\text{Itô}}{=} 3|y|^2 + 12\kappa_\beta^2\mathrm{E}\int_0^T \|\boldsymbol{\lambda}(\tau)\|^2\,\mathrm{d}\tau \\ &\qquad + 12\kappa_\beta^2\mathrm{E}\sum_{i,j=1}^D\int_0^T |\gamma_{ij}(\tau)|^2\,\mathrm{d}[\mathcal{Q}_{ij}](\tau) \\ &< \infty.\end{aligned} \tag{4.5.14}$$

The finiteness comes from the facts that $\boldsymbol{\lambda} \in L^2(\mathbf{W})$ and $\boldsymbol{\gamma} \in L^2(\mathcal{Q})$. Thus we have shown that Y is square-integrable.

Using the integration-by-parts formula (Theorem C.14.1) to expand the last line of (4.5.11), we get

$$\begin{aligned}Y(t) &= y + \int_0^t \beta(\tau)\left(\beta^{-1}(\tau)\boldsymbol{\lambda}^\top(\tau)\,\mathrm{d}\mathbf{W}(\tau) + \sum_{i,j=1}^D \beta^{-1}(\tau)\gamma_{ij}(\tau)\,\mathrm{d}\mathcal{Q}_{ij}(\tau)\right) \\ &\qquad - \int_0^t r(\tau)Y(\tau_-)\,\mathrm{d}\tau \\ &= y - \int_0^t r(\tau)Y(\tau_-)\,\mathrm{d}\tau + \int_0^t \boldsymbol{\lambda}^\top(\tau)\,\mathrm{d}\mathbf{W}(\tau) + \sum_{i,j=1}^D\int_0^t \gamma_{ij}(\tau)\,\mathrm{d}\mathcal{Q}_{ij}(\tau).\end{aligned} \tag{4.5.15}$$

By the uniform boundedness of r and the square-integrability of Y, shown above, we have $rY_- \in L_{21}$. Then, as $(y,\boldsymbol{\lambda},\boldsymbol{\gamma}) \in \mathbb{R}\times L^2(\mathbf{W})\times L^2(\mathcal{Q})$, we get

$$Y \equiv (y, -rY_-, \boldsymbol{\lambda}, \boldsymbol{\gamma}) \in \mathbb{B}, \tag{4.5.16}$$

so that (4.5.10) holds. Finally, as $Y(t) \neq Y(t_-)$ only on a set of Lebesgue measure zero, we find that

$$\dot{Y}(t) = -r(t)Y(t_-) = -r(t)Y(t), \quad (\mathbb{P}\otimes \mathit{Leb})\text{-a.e.}, \tag{4.5.17}$$

so that from the definition of $\mathbb{B}_1$ in (4.5.3), we get $Y \in \mathbb{B}_1$. $\square$

Lemma 4.5.3. *The map* $\Xi : \mathbb{R} \times L^2(\mathbf{W}) \times L^2(\mathcal{Q}) \to \mathbb{B}_1$ *is a linear bijection.*

Proof. First we show that the map is linear by showing that it is additive and homogeneous.

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

To show that the map Ξ is bijective, we show that it is both injective and surjective.

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

$$\Xi(y^{(1)}, \boldsymbol{\lambda}^{(1)}, \boldsymbol{\gamma}^{(1)}) = \Xi(y^{(2)}, \boldsymbol{\lambda}^{(2)}, \boldsymbol{\gamma}^{(2)}) \tag{4.5.18}$$

implies $y^{(1)} = y^{(2)}$, $\boldsymbol{\lambda}^{(1)} = \boldsymbol{\lambda}^{(2)}$ $(\mathbb{P} \otimes Leb)$-a.e. and $\boldsymbol{\gamma}^{(1)} = \boldsymbol{\gamma}^{(2)}$ $\nu_{[\mathcal{Q}]}$-a.e.

Assuming the left-hand side of (4.5.18), from (4.5.8) we have for all $t \in [0, T]$

$$\beta(t)\left(y^{(1)} + \mathbb{J}(\boldsymbol{\lambda}^{(1)})(t) + \mathbb{I}(\boldsymbol{\gamma}^{(1)})(t)\right) = \beta(t)\left(y^{(2)} + \mathbb{J}(\boldsymbol{\lambda}^{(2)})(t) + \mathbb{I}(\boldsymbol{\gamma}^{(2)})(t)\right). \tag{4.5.19}$$

Thus,

$$y^{(1)} + \mathbb{J}(\boldsymbol{\lambda}^{(1)})(t) + \mathbb{I}(\boldsymbol{\gamma}^{(1)})(t) = y^{(2)} + \mathbb{J}(\boldsymbol{\lambda}^{(2)})(t) + \mathbb{I}(\boldsymbol{\gamma}^{(2)})(t). \tag{4.5.20}$$

From Proposition A.1.1 and the strict positivity of $\beta(t)$ (see (4.5.5)), we must then have $y^{(1)} = y^{(2)}$, $\boldsymbol{\lambda}^{(1)} = \boldsymbol{\lambda}^{(2)}$ $(\mathbb{P} \otimes Leb)$-a.e. and $\boldsymbol{\gamma}^{(1)} = \boldsymbol{\gamma}^{(2)}$ $\nu_{[\mathcal{Q}]}$-a.e. 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 $(y, \boldsymbol{\lambda}, \boldsymbol{\gamma}) \in \mathbb{R} \times L^2(\mathbf{W}) \times L^2(\mathcal{Q})$ such that $\Xi(y, \boldsymbol{\lambda}, \boldsymbol{\gamma}) = Y$. Fix $Y \equiv (Y_0, \dot{Y}, \boldsymbol{\Lambda}^Y, \boldsymbol{\Gamma}^Y) \in \mathbb{B}_1$. From the definition of $\mathbb{B}_1$ in (4.5.3), we see that Y has the particular integral form

$$Y(t) = Y_0 - \int_0^t r(\tau) Y(\tau)\, \mathrm{d}\tau + \int_0^t \left(\boldsymbol{\Lambda}^Y\right)^\top (\tau)\, \mathrm{d}\mathbf{W}(\tau) + \sum_{i,j=1}^{D} \int_0^t \Gamma_{ij}^Y(\tau)\, \mathrm{d}\mathcal{Q}_{ij}(\tau). \tag{4.5.21}$$

Now consider an arbitrary triple $(y, \boldsymbol{\lambda}, \boldsymbol{\gamma}) \in \mathbb{R} \times L^2(\mathbf{W}) \times L^2(\mathcal{Q})$. By Lemma 4.5.2, we can express $\Xi(y, \boldsymbol{\lambda}, \boldsymbol{\gamma})$ as

$$\begin{aligned} &\Xi(y, \boldsymbol{\lambda}, \boldsymbol{\gamma})(t) \\ &= y - \int_0^t r(\tau) \Xi(y, \boldsymbol{\lambda}, \boldsymbol{\gamma})(\tau_-)\, \mathrm{d}\tau + \int_0^t \boldsymbol{\lambda}^\top(\tau)\, \mathrm{d}\mathbf{W}(\tau) + \sum_{i,j=1}^{D} \int_0^t \gamma_{ij}(\tau)\, \mathrm{d}\mathcal{Q}_{ij}(\tau). \end{aligned} \tag{4.5.22}$$

Setting $Y \equiv \Xi(y, \boldsymbol{\lambda}, \boldsymbol{\gamma})$ and applying Proposition A.1.1, we must have $y = Y_0$, $\boldsymbol{\lambda} = \boldsymbol{\Lambda}^Y$ $(\mathbb{P} \otimes Leb)$-a.e. and $\boldsymbol{\gamma} = \boldsymbol{\Gamma}^Y$ $\nu_{[\mathcal{Q}]}$-a.e. Since $(y, \boldsymbol{\lambda}, \boldsymbol{\gamma}) = (Y_0, \boldsymbol{\Lambda}^Y, \boldsymbol{\Gamma}^Y) \in \mathbb{R} \times L^2(\mathbf{W}) \times L^2(\mathcal{Q})$, we have shown that the map Ξ is surjective. It follows that the map Ξ is bijective. $\square$

Now that we have defined the map Ξ, we define a functional $\tilde{\Psi}$ to act on it. We show that the dual problem is equivalent to a problem involving the functional $\tilde{\Psi}$. Then we show that a solution exists to the problem involving the functional $\tilde{\Psi}$. Thus a solution will exist to the dual problem.

Definition 4.5.4. Define the functional $\tilde{\Psi} : \mathbb{R} \times L^2(\mathbf{W}) \times L^2(\mathcal{Q}) \to (-\infty, +\infty]$ by

$$\tilde{\Psi}(y, \boldsymbol{\lambda}, \boldsymbol{\gamma}) := \Psi(\Xi(y, \boldsymbol{\lambda}, \boldsymbol{\gamma})), \qquad \forall (y, \boldsymbol{\lambda}, \boldsymbol{\gamma}) \in \mathbb{R} \times L^2(\mathbf{W}) \times L^2(\mathcal{Q}). \tag{4.5.23}$$

Ψ is the dual cost functional given by (4.2.4) and the map Ξ is given by (4.5.8).

Remark 4.5.5. By the bijectivity of the map Ξ, given by Lemma 4.5.3, we have

$$\begin{aligned} \inf_{(y,\boldsymbol{\lambda},\boldsymbol{\gamma}) \in \mathbb{R} \times L^2(\mathbf{W}) \times L^2(\mathcal{Q})} \left\{ \tilde{\Psi}(y, \boldsymbol{\lambda}, \boldsymbol{\gamma}) \right\} &\overset{(4.5.23)}{=} \inf_{(y,\boldsymbol{\lambda},\boldsymbol{\gamma}) \in \mathbb{R} \times L^2(\mathbf{W}) \times L^2(\mathcal{Q})} \left\{ \Psi(\Xi(y, \boldsymbol{\lambda}, \boldsymbol{\gamma}) \right\} \\ &= \inf_{Y \in \mathbb{B}_1} \left\{ \Psi(Y) \right\}. \end{aligned} \tag{4.5.24}$$

If we can show that there exists a triple $(\bar{y}, \bar{\boldsymbol{\lambda}}, \bar{\boldsymbol{\gamma}}) \in \mathbb{R} \times L^2(\mathbf{W}) \times L^2(\mathcal{Q})$ such that

$$\tilde{\Psi}(\bar{y}, \bar{\boldsymbol{\lambda}}, \bar{\boldsymbol{\gamma}}) = \inf_{(y,\boldsymbol{\lambda},\boldsymbol{\gamma}) \in \mathbb{R} \times L^2(\mathbf{W}) \times L^2(\mathcal{Q})} \{\tilde{\Psi}(y, \boldsymbol{\lambda}, \boldsymbol{\gamma})\}, \tag{4.5.25}$$

then $\bar{Y} := \Xi(\bar{y}, \bar{\boldsymbol{\lambda}}, \bar{\boldsymbol{\gamma}}) \in \mathbb{B}_1$ solves the dual problem (4.5.1).

To show that there exists a solution to (4.5.25), we use Ekeland and Témam [13], Chapter II, Proposition 1.2, page 35. We rework this proposition into our notation.

Proposition 4.5.6. *Suppose* $\mathbb{R} \times L^2(\mathbf{W}) \times L^2(\mathcal{Q})$ *is a reflexive Banach space with norm* $\|(\cdot, \cdot, \cdot)\|$. *We are given a functional* $\tilde{\Psi} : \mathbb{R} \times L^2(\mathbf{W}) \times L^2(\mathcal{Q}) \to \mathbb{R}$ *such that* $\tilde{\Psi}$ *is convex, lower semi-continuous and proper. Suppose we wish to find a solution to the problem*

$$\inf_{(y,\boldsymbol{\lambda},\boldsymbol{\gamma}) \in \mathbb{R} \times L^2(\mathbf{W}) \times L^2(\mathcal{Q})} \{\tilde{\Psi}(y, \boldsymbol{\lambda}, \boldsymbol{\gamma})\}, \tag{4.5.26}$$

that is to find $(\bar{y}, \bar{\boldsymbol{\lambda}}, \bar{\boldsymbol{\gamma}}) \in \mathbb{R} \times L^2(\mathbf{W}) \times L^2(\mathcal{Q})$ *such that*

$$\tilde{\Psi}(\bar{y}, \bar{\boldsymbol{\lambda}}, \bar{\boldsymbol{\gamma}}) = \inf_{(y,\boldsymbol{\lambda},\boldsymbol{\gamma}) \in \mathbb{R} \times L^2(\mathbf{W}) \times L^2(\mathcal{Q})} \{\tilde{\Psi}(y, \boldsymbol{\lambda}, \boldsymbol{\gamma})\}. \tag{4.5.27}$$

If the functional $\tilde{\Psi}$ *is coercive over* $\mathbb{R} \times L^2(\mathbf{W}) \times L^2(\mathcal{Q})$, *that is if*

$$\lim_{\|(y,\boldsymbol{\lambda},\boldsymbol{\gamma})\| \to \infty} \tilde{\Psi}(y, \boldsymbol{\lambda}, \boldsymbol{\gamma}) = +\infty, \tag{4.5.28}$$

then the problem has at least one solution.

In order to apply Proposition 4.5.6 and thus conclude that a solution to the dual problem does indeed exist, we must show that $\mathbb{R} \times L^2(\mathbf{W}) \times L^2(\mathcal{Q})$ is a reflexive Banach space and that the functional $\tilde{\Psi}$ defined by (4.5.23) is convex, lower semi-continuous, proper and coercive.

Remark 4.5.7. A *Banach space* is a complete, normed vector space. An example of a Banach space is the space $\mathbb{R}$ with the Euclidean norm $\|x\| = |x|$ for all $x \in \mathbb{R}$. Other examples are the L^2-spaces. In particular, the spaces $L^2(\mathbf{W})$ and $L^2(\mathcal{Q})$ are Banach spaces.

Define a norm on $L^2(\mathbf{W})$ by

$$\|\boldsymbol{\lambda}\|_{L^2(\mathbf{W})}^2 := \mathrm{E}\int_0^T \|\boldsymbol{\lambda}(t)\|^2 \,\mathrm{d}t, \qquad \forall \boldsymbol{\lambda} \in L^2(\mathbf{W}), \tag{4.5.29}$$

where $\|\cdot\|$ denotes Euclidean length.

Similarly, define a norm on $L^2(\mathcal{Q})$ by

$$\|\boldsymbol{\gamma}\|_{L^2(\mathcal{Q})}^2 := \mathrm{E}\sum_{i,j=1}^{D}\int_0^T \gamma_{ij}^2(t)\,\mathrm{d}[\mathcal{Q}_{ij}](t), \qquad \forall \boldsymbol{\gamma} \in L^2(\mathcal{Q}). \tag{4.5.30}$$

We can then define a norm $\|(\cdot,\cdot,\cdot)\|$ on $\mathbb{R} \times L^2(\mathbf{W}) \times L^2(\mathcal{Q})$ by

$$\|(y, \boldsymbol{\lambda}, \boldsymbol{\gamma})\|^2 := |y|^2 + \|\boldsymbol{\lambda}\|_{L^2(\mathbf{W})}^2 + \|\boldsymbol{\gamma}\|_{L^2(\mathcal{Q})}^2, \tag{4.5.31}$$

for all $(y, \boldsymbol{\lambda}, \boldsymbol{\gamma}) \in \mathbb{R} \times L^2(\mathbf{W}) \times L^2(\mathcal{Q})$.

Remark 4.5.8. A *Hilbert space* is a vector space V with an inner product $\langle\cdot,\cdot\rangle$ such that the norm defined by $\|f\| = \langle f, f\rangle$ turns V into a complete metric space. While every Hilbert space is a Banach space, not every Banach space is a Hilbert space. An example of a Hilbert space is the space $\mathbb{R}$ with the inner product $\langle x, y\rangle = xy$ for all $x, y \in \mathbb{R}$. Other examples are the L^2-spaces. In particular, we show in Lemma A.2.5 that the space $L^2(\mathcal{Q})$ is a Hilbert space. That the space $L^2(\mathbf{W})$ is a Hilbert space is well-known. As $\mathbb{R}$, $L^2(\mathbf{W})$ and $L^2(\mathcal{Q})$ are Hilbert spaces, it follows that $\mathbb{R} \times L^2(\mathbf{W}) \times L^2(\mathcal{Q})$ is also a Hilbert space, since it is the direct sum of the Hilbert spaces $\mathbb{R}$, $L^2(\mathbf{W})$ and $L^2(\mathcal{Q})$.

Remark 4.5.9. A Banach space V is called *reflexive* if it satisfies a property involving its dual and bidual space. The dual space $V^\star$ of a Banach space V is the set of all continuous linear functionals on V. The bidual space $V^{\star\star}$ is the set of all continuous linear functionals on $V^\star$. If the natural mapping between V and $V^{\star\star}$ is bijective, then V is called a reflexive Banach space. A well-known example of a reflexive Banach space is a Hilbert space (see Ekeland and Témam [13], page 34). It follows from Remark 4.5.8 that $\mathbb{R} \times L^2(\mathbf{W}) \times L^2(\mathcal{Q})$ is a reflexive Banach space.

Remark 4.5.10. In order to show that all the conditions of Proposition 4.5.6 hold, we simplify the expression for $\tilde{\Psi}$. To do this, we first simplify the dual cost functional Ψ and then, using this and (4.5.23), we find a suitable form for the functional $\tilde{\Psi}$.

From the definition of l_0 in (4.1.14), it is immediate that (4.2.1) simplifies to

$$m_0(y) = x_0 y, \qquad \forall y \in \mathbb{R}. \tag{4.5.32}$$

Using (4.5.32) and substituting for m_T from (4.4.12) into (4.2.4), we get

$$\Psi(Y) = x_0 Y_0 + \mathrm{E}\int_0^T m_1(t, Y(t), \dot{Y}(t), \mathbf{\Lambda}^Y(t))\,\mathrm{d}t + \mathrm{E}\left(\frac{(Y(T)+b)^2}{2a}\right) - \mathrm{E}c, \quad (4.5.33)$$

for all $Y \equiv (Y_0, \dot{Y}, \mathbf{\Lambda}^Y, \mathbf{\Gamma}^Y) \in \mathbb{B}$.

Substituting for m_1 from (4.4.11) into (4.5.33), we get

$$\Psi(Y) = x_0 Y_0 + \mathrm{E}\int_0^T \delta(\mathbf{\Theta}_Y(t))\,\mathrm{d}t + \mathrm{E}\left(\frac{(Y(T)+b)^2}{2a}\right) - \mathrm{E}c, \quad (4.5.34)$$

for all $Y \equiv (Y_0, \dot{Y}, \mathbf{\Lambda}^Y, \mathbf{\Gamma}^Y) \in \mathbb{B}_1$.

As (4.5.34) is for $Y \in \mathbb{B}_1$, using the bijectivity of the map Ξ, given by Lemma 4.5.3, we can find a unique triple $(y, \boldsymbol{\lambda}, \boldsymbol{\gamma}) \in \mathbb{R} \times L^2(\mathbf{W}) \times L^2(\mathcal{Q})$ such that $Y = \Xi(y, \boldsymbol{\lambda}, \boldsymbol{\gamma})$. From this and the definition of the functional $\tilde{\Psi}$ (see (4.5.23)), we obtain for all $Y \equiv \Xi(y, \boldsymbol{\lambda}, \boldsymbol{\gamma}) \in \mathbb{B}_1$,

$$\tilde{\Psi}(y, \boldsymbol{\lambda}, \boldsymbol{\gamma}) = x_0 y + \mathrm{E}\int_0^T \delta(\mathbf{\Theta}_Y(t))\ \mathrm{d}t + \mathrm{E}\left(\frac{(Y(T)+b)^2}{2a}\right) - \mathrm{E}c, \quad (4.5.35)$$

for $\mathbf{\Theta}_Y(t)$ defined by (4.4.6) and the support function δ given by (4.4.7). This is the form of the functional $\tilde{\Psi}$ that we use to show that the conditions of Proposition 4.5.6 hold.

Lemma 4.5.11. *The functional $\tilde{\Psi}$ is convex on $\mathbb{R} \times L^2(\mathbf{W}) \times L^2(\mathcal{Q})$.*

Proof. From (4.5.35), for any $(y, \boldsymbol{\lambda}, \boldsymbol{\gamma}) \in \mathbb{R} \times L^2(\mathbf{W}) \times L^2(\mathcal{Q})$,

$$\tilde{\Psi}(y, \boldsymbol{\lambda}, \boldsymbol{\gamma}) = x_0 y + \mathrm{E}\int_0^T \delta(\mathbf{\Theta}_Y(t))\,\mathrm{d}t + \mathrm{E}\left(\frac{(Y(T)+b)^2}{2a}\right) - \mathrm{E}c, \quad (4.5.36)$$

for $Y \equiv \Xi(y, \boldsymbol{\lambda}, \boldsymbol{\gamma}) \in \mathbb{B}_1$. To show that the functional $\tilde{\Psi}$ is convex on the space $\mathbb{R} \times L^2(\mathbf{W}) \times L^2(\mathcal{Q})$, we need to show that for all $\epsilon \in (0,1)$ and for all triples $(y^{(1)}, \boldsymbol{\lambda}^{(1)}, \boldsymbol{\gamma}^{(1)}), (y^{(2)}, \boldsymbol{\lambda}^{(2)}, \boldsymbol{\gamma}^{(2)}) \in \mathbb{R} \times L^2(\mathbf{W}) \times L^2(\mathcal{Q})$, setting $(y^{(3)}, \boldsymbol{\lambda}^{(3)}, \boldsymbol{\gamma}^{(3)}) = \epsilon(y^{(1)}, \boldsymbol{\lambda}^{(1)}, \boldsymbol{\gamma}^{(1)}) + (1-\epsilon)(y^{(2)}, \boldsymbol{\lambda}^{(2)}, \boldsymbol{\gamma}^{(2)})$ we have

$$\Psi(y^{(3)}, \boldsymbol{\lambda}^{(3)}, \boldsymbol{\gamma}^{(3)}) \le \epsilon\Psi(y^{(1)}, \boldsymbol{\lambda}^{(1)}, \boldsymbol{\gamma}^{(1)}) + (1-\epsilon)\Psi(y^{(2)}, \boldsymbol{\lambda}^{(2)}, \boldsymbol{\gamma}^{(2)}). \quad (4.5.37)$$

Trivially, the first and last terms on the right-hand side of (4.5.36) are convex. Thus we need only show that the second and third terms are convex.

Fix $(y^{(1)}, \boldsymbol{\lambda}^{(1)}, \boldsymbol{\gamma}^{(1)}), (y^{(2)}, \boldsymbol{\lambda}^{(2)}, \boldsymbol{\gamma}^{(2)}) \in \mathbb{R} \times L^2(\mathbf{W}) \times L^2(\mathcal{Q})$ and $\epsilon \in (0,1)$. Set

$$(y^{(3)}, \boldsymbol{\lambda}^{(3)}, \boldsymbol{\gamma}^{(3)}) := \epsilon(y^{(1)}, \boldsymbol{\lambda}^{(1)}, \boldsymbol{\gamma}^{(1)}) + (1-\epsilon)(y^{(2)}, \boldsymbol{\lambda}^{(2)}, \boldsymbol{\gamma}^{(2)}). \quad (4.5.38)$$

Then $(y^{(3)}, \boldsymbol{\lambda}^{(3)}, \boldsymbol{\gamma}^{(3)}) \in \mathbb{R} \times L^2(\mathbf{W}) \times L^2(\mathcal{Q})$. From Lemma 4.5.3, the map Ξ is linear and we see from (4.4.6) that $\boldsymbol{\Theta}_{\cdot}(t)$ is also a linear map. Then a.s. for all $t \in [0, T]$, we have

$$\boldsymbol{\Theta}_{\Xi(y^{(3)},\boldsymbol{\lambda}^{(3)},\boldsymbol{\gamma}^{(3)})}(t) = \epsilon \boldsymbol{\Theta}_{\Xi(y^{(1)},\boldsymbol{\lambda}^{(1)},\boldsymbol{\gamma}^{(1)})}(t) + (1-\epsilon)\boldsymbol{\Theta}_{\Xi(y^{(2)},\boldsymbol{\lambda}^{(2)},\boldsymbol{\gamma}^{(2)})}(t). \tag{4.5.39}$$

From the positive homogeneity and subadditivity of the support function δ (see Remark 4.4.5), we have for all $t \in [0, T]$,

$$\delta\left(\boldsymbol{\Theta}_{\Xi(y^{(3)},\boldsymbol{\lambda}^{(3)},\boldsymbol{\gamma}^{(3)})}(t)\right) \leq \epsilon\delta\left(\boldsymbol{\Theta}_{\Xi(y^{(1)},\boldsymbol{\lambda}^{(1)},\boldsymbol{\gamma}^{(1)})}(t)\right) + (1-\epsilon)\delta\left(\boldsymbol{\Theta}_{\Xi(y^{(2)},\boldsymbol{\lambda}^{(2)},\boldsymbol{\gamma}^{(2)})}(t)\right). \tag{4.5.40}$$

It follows that the second term on the right-hand side of (4.5.36) is convex on $\mathbb{R} \times L^2(\mathbf{W}) \times L^2(\mathcal{Q})$.

The convexity of the third term on the right-hand side of (4.5.36) follows from the linearity of the mapping $Y(T) = \Xi(y, \boldsymbol{\lambda}, \boldsymbol{\gamma})(T)$, the convexity of the quadratic function and the strict positivity of the random variable a (Condition 3.2.26). □

Lemma 4.5.12. *The functional $\tilde{\Psi}$ is proper, that is to say it is strictly greater than $-\infty$ and is not identically equal to $+\infty$ on $\mathbb{R} \times L^2(\mathbf{W}) \times L^2(\mathcal{Q})$.*

Proof. First we show that the functional $\tilde{\Psi}$ is strictly greater than $-\infty$.

From (4.5.35), for any $(y, \boldsymbol{\lambda}, \boldsymbol{\gamma}) \in \mathbb{R} \times L^2(\mathbf{W}) \times L^2(\mathcal{Q})$,

$$\tilde{\Psi}(y, \boldsymbol{\lambda}, \boldsymbol{\gamma}) = x_0 y + \mathrm{E}\int_0^T \delta(\boldsymbol{\Theta}_Y(t))\,\mathrm{d}t + \mathrm{E}\left(\frac{(Y(T)+b)^2}{2a}\right) - \mathrm{E}c, \tag{4.5.41}$$

for $Y \equiv \Xi(y, \boldsymbol{\lambda}, \boldsymbol{\gamma}) \in \mathbb{B}_1$.

From Remark 4.4.5, for the support function δ, we have that $\delta(\mathbf{x}) \geq 0$ for all $\mathbf{x} \in \mathbb{R}^N$. Applying this to (4.5.41), we find

$$\tilde{\Psi}(y, \boldsymbol{\lambda}, \boldsymbol{\gamma}) \geq x_0 y + \mathrm{E}\left(\frac{(Y(T)+b)^2}{2a}\right) - \mathrm{E}c \geq x_0 y - \mathrm{E}c > -\infty, \tag{4.5.42}$$

where the last inequality follows from $x_0, y, \mathrm{E}c \in \mathbb{R}$.

Now we show that there is an element in $\mathbb{R} \times L^2(\mathbf{W}) \times L^2(\mathcal{Q})$ for which the functional $\tilde{\Psi}$ is less than $+\infty$. Consider

$$\tilde{\Psi}(0, \mathbf{0}, \mathbf{0}) = \mathrm{E}\int_0^T \delta(\boldsymbol{\Theta}_{\mathbf{0}}(t))\,\mathrm{d}t + \mathrm{E}\left(\frac{b^2}{2a}\right) - \mathrm{E}c = \mathrm{E}\left(\frac{b^2}{2a}\right) - \mathrm{E}c < +\infty. \tag{4.5.43}$$

Hence there is an element in $\mathbb{R} \times L^2(\mathbf{W}) \times L^2(\mathcal{Q})$ for which the functional $\tilde{\Psi}$ is finite. So the functional $\tilde{\Psi}$ is proper. □

Lemma 4.5.13. *The functional $\tilde{\Psi}$ is lower semi-continuous on $\mathbb{R} \times L^2(\mathbf{W}) \times L^2(\mathcal{Q})$.*

Proof. Let $\{(y^{(m)}, \boldsymbol{\lambda}^{(m)}, \boldsymbol{\gamma}^{(m)})\}_{m\in\mathbb{N}}$ be a sequence in the space $\mathbb{R} \times L^2(\mathbf{W}) \times L^2(\mathcal{Q})$ converging in the norm to $(y, \boldsymbol{\lambda}, \boldsymbol{\gamma}) \in \mathbb{R} \times L^2(\mathbf{W}) \times L^2(\mathcal{Q})$, that is

$$\|(y^{(m)} - y, \boldsymbol{\lambda}^{(m)} - \boldsymbol{\lambda}, \boldsymbol{\gamma}^{(m)} - \boldsymbol{\gamma})\| \to 0 \quad \text{as } m \to \infty, \tag{4.5.44}$$

for the norm given by (4.5.31). We want to show that

$$\tilde{\Psi}(y, \boldsymbol{\lambda}, \boldsymbol{\gamma}) \leq \liminf_{m\to\infty} \tilde{\Psi}(y^{(m)}, \boldsymbol{\lambda}^{(m)}, \boldsymbol{\gamma}^{(m)}). \tag{4.5.45}$$

From (4.5.35),

$$\tilde{\Psi}(y, \boldsymbol{\lambda}, \boldsymbol{\gamma}) = x_0 y + \mathrm{E} \int_0^T \delta(\boldsymbol{\Theta}_Y(t))\, \mathrm{d}t + \mathrm{E}\left(\frac{(Y(T) + b)^2}{2a}\right) - \mathrm{E}c, \tag{4.5.46}$$

for $Y \equiv \Xi(y, \boldsymbol{\lambda}, \boldsymbol{\gamma}) \in \mathbb{B}_1$. Trivially, the first and last terms on the right-hand side of (4.5.46) are continuous, thus we need only consider the second and third terms on the right-hand side.

Consider the second term on the right-hand side of (4.5.46). Define

$$Y^{(m)} := \Xi(y^{(m)}, \boldsymbol{\lambda}^{(m)}, \boldsymbol{\gamma}^{(m)}) \in \mathbb{B}_1. \tag{4.5.47}$$

In view of the nonnegativity of the risk-free interest rate process $\{r(t)\}$, the process $\{\beta(t)\}$, given by (4.5.5), is uniformly bounded above by the constant 1. Then

$$\begin{aligned}
\mathrm{E}\int_0^T |Y^{(m)}(t) - Y(t)|^2\, \mathrm{d}t &= \mathrm{E}\int_0^T |\Xi(y^{(m)}, \boldsymbol{\lambda}^{(m)}, \boldsymbol{\gamma}^{(m)})(t) - \Xi(y, \boldsymbol{\lambda}, \boldsymbol{\gamma})(t)|^2\, \mathrm{d}t \\
&\overset{(4.5.8)}{=} \mathrm{E}\int_0^T \left|\beta(t)\left((y^{(m)} - y) + \mathbb{J}(\boldsymbol{\lambda}^{(m)} - \boldsymbol{\lambda})(t) + \mathbb{I}(\boldsymbol{\gamma}^{(m)} - \boldsymbol{\gamma})(t)\right)\right|^2 \mathrm{d}t \\
&\leq 3\mathrm{E}\int_0^T \left((y^{(m)} - y)^2 + \mathbb{J}^2(\boldsymbol{\lambda}^{(m)} - \boldsymbol{\lambda})(t) + \mathbb{I}^2(\boldsymbol{\gamma}^{(m)} - \boldsymbol{\gamma})(t)\right)\, \mathrm{d}t \\
&\leq 3T\, |y^{(m)} - y|^2 + 3T\, \mathrm{E}\left(\sup_{t\in[0,T]} \mathbb{J}^2(\boldsymbol{\lambda}^{(m)} - \boldsymbol{\lambda})(t)\right) \\
&\qquad + 3T\, \mathrm{E}\left(\sup_{t\in[0,T]} \mathbb{I}^2(\boldsymbol{\gamma}^{(m)} - \boldsymbol{\gamma})(t)\right),
\end{aligned} \tag{4.5.48}$$

Let $\kappa_\beta \geq 1$ be a uniform bound on $\{\beta^{-1}(t)\}$. Applying Doob's L^2-inequality, the Itô isometry and recalling the definitions of the stochastic integrals $\mathbb{J}$ and $\mathbb{I}$ from

(4.5.6) and (4.5.7), we obtain

$$
\begin{aligned}
&\mathrm{E}\int_0^T |Y^{(m)}(t) - Y(t)|^2 \,\mathrm{d}t \\
&\overset{(4.5.6),(4.5.7)}{\leq} 3T\,|y^{(m)} - y|^2 + 3\kappa_\beta^2\, T\,\mathrm{E}\left(\sup_{t\in[0,T]}\int_0^t \left(\boldsymbol{\lambda}^{(m)} - \boldsymbol{\lambda}\right)^\top(\tau)\,\mathrm{d}\mathbf{W}(\tau)\right)^2 \\
&\qquad + 3\kappa_\beta^2\, T\,\mathrm{E}\left(\sup_{t\in[0,T]}\sum_{i,j=1}^{D}\int_0^t \left(\gamma_{ij}^{(m)} - \gamma_{ij}\right)(\tau)\,\mathrm{d}\mathcal{Q}_{ij}(\tau)\right)^2 \\
&\overset{\text{Doob}}{\leq} 3T\,|y^{(m)} - y|^2 + 12\kappa_\beta^2\, T\,\mathrm{E}\left(\int_0^T \left(\boldsymbol{\lambda}^{(m)} - \boldsymbol{\lambda}\right)^\top(t)\,\mathrm{d}\mathbf{W}(t)\right)^2 \\
&\qquad + 12\kappa_\beta^2\, T\,\mathrm{E}\left(\sum_{i,j=1}^{D}\int_0^T \left(\gamma_{ij}^{(m)} - \gamma_{ij}\right)(t)\,\mathrm{d}\mathcal{Q}_{ij}(t)\right)^2 \\
&\overset{\text{Itô}}{=} 3T\,|y^{(m)} - y|^2 + 12\kappa_\beta^2\, T\,\mathrm{E}\int_0^T \|\left(\boldsymbol{\lambda}^{(m)} - \boldsymbol{\lambda}\right)(t)\|^2\,\mathrm{d}t \\
&\qquad + 12\kappa_\beta^2\, T\,\mathrm{E}\sum_{i,j=1}^{D}\int_0^T (\gamma_{ij}^{(m)} - \gamma_{ij})^2(t)\,\mathrm{d}[\mathcal{Q}_{ij}](t) \\
&\overset{(4.5.31)}{\leq} 12\kappa_\beta^2\,\|\left(y^{(m)} - y, \boldsymbol{\lambda}^{(m)} - \boldsymbol{\lambda}, \boldsymbol{\gamma}^{(m)} - \boldsymbol{\gamma}\right)\|^2 \\
&\overset{(4.5.44)}{\to} 0 \text{ as } m \to \infty.
\end{aligned}
\tag{4.5.49}
$$

Hence $Y^{(m)} \to Y$ as $m \to \infty$ in the space $L^2(\Omega \times [0,T], \mathcal{P}^\star, \mathbb{P} \otimes Leb)$.

Recall the definition of $\boldsymbol{\Theta}_Y(t)$ from (4.4.6). We don't know if the sequence $\{\mathrm{E}\int_0^T \delta(\boldsymbol{\Theta}_{Y^{(m)}}(t))\,\mathrm{d}t\}_{m\in\mathbb{N}}$ has a limit, but we do know that the liminf will certainly exist. Choose a subsequence $\{m_j\} \subset \mathbb{N}$ such that

$$
\lim_{m_j\to\infty}\mathrm{E}\int_0^T \delta(\boldsymbol{\Theta}_{Y^{(m_j)}}(t))\,\mathrm{d}t = \liminf_{m\to\infty}\mathrm{E}\int_0^T \delta(\boldsymbol{\Theta}_{Y^{(m)}}(t))\,\mathrm{d}t. \tag{4.5.50}
$$

Since $Y^{(m)} \to Y$ as $m \to \infty$ in the L^2-norm, we can extract from any subsequence of $\{Y^{(m)}\}_{m\in\mathbb{N}}$ a further subsequence which converges $(\mathbb{P} \otimes Leb)$-a.e. to Y. As by assumption, $\boldsymbol{\lambda}^{(m)} \to \boldsymbol{\lambda}$ as $m \to \infty$ in the L^2-norm, we can also extract from any subsequence of $\{\boldsymbol{\lambda}^{(m)}\}_{m\in\mathbb{N}}$ a further subsequence which converges $(\mathbb{P} \otimes Leb)$-a.e. to $\boldsymbol{\lambda}$. Hence, from the subsequence $\{m_j\} \subset \mathbb{N}$ we can extract a further subsequence $\{m_{j_k}\} \subset \mathbb{N}$ such that $Y^{(m_{j_k})} \to Y$ $(\mathbb{P} \otimes Leb)$-a.e. and $\boldsymbol{\lambda}^{(m_{j_k})} \to \boldsymbol{\lambda}$ $(\mathbb{P} \otimes Leb)$-a.e. as $m_{j_k} \to \infty$. Then for

$$
\boldsymbol{\Theta}_{Y^{(m_{j_k})}}(t) \overset{(4.4.6)}{=} -\boldsymbol{\sigma}(t)\left(\boldsymbol{\theta}(t)Y^{(m_{j_k})}(t) + \boldsymbol{\lambda}^{(m_{j_k})}(t)\right), \tag{4.5.51}
$$

it follows that $\boldsymbol{\Theta}_{Y^{(m_{j_k})}} \to \boldsymbol{\Theta}_Y$ $(\mathbb{P} \otimes Leb)$-a.e. as $m_{j_k} \to \infty$.

By Remark 4.4.5, for the support function δ, we have $\delta \geq 0$, so we can apply Fatou's Lemma (Theorem C.1.4) to the nonnegative measurable functions $\{\delta(\mathbf{\Theta}_{Y^{(m_{j_k})}}(t))\}$. Then, recalling from Remark 4.4.5 that the support function δ is itself a lower semi-continuous function, and the fact that any subsequence of a convergent sequence will converge to the same limit, we have

$$\begin{aligned}
\mathrm{E}\int_0^T \delta(\mathbf{\Theta}_Y(t))\,\mathrm{d}t &\overset{\delta \text{ l.s.c.}}{\leq} \mathrm{E}\int_0^T \liminf_{m_{j_k}\to\infty} \delta(\mathbf{\Theta}_{Y^{(m_{j_k})}}(t))\,\mathrm{d}t \\
&\overset{\text{Fatou}}{\leq} \liminf_{m_{j_k}\to\infty} \mathrm{E}\int_0^T \delta(\mathbf{\Theta}_{Y^{(m_{j_k})}}(t))\,\mathrm{d}t \\
&= \lim_{m_j\to\infty} \mathrm{E}\int_0^T \delta(\mathbf{\Theta}_{Y^{(m_j)}}(t))\,\mathrm{d}t \\
&\overset{(4.5.50)}{=} \liminf_{m\to\infty} \mathrm{E}\int_0^T \delta(\mathbf{\Theta}_{Y^{(m)}}(t))\,\mathrm{d}t.
\end{aligned} \tag{4.5.52}$$

Hence the second term on the right-hand side of (4.5.46) is lower semi-continuous.

For the third term on the right-hand side of (4.5.46), we need to show that

$$\mathrm{E}\left(\frac{(Y(T)+b)^2}{2a}\right) \leq \liminf_{m\to\infty} \mathrm{E}\left(\frac{(Y^{(m)}(T)+b)^2}{2a}\right). \tag{4.5.53}$$

By the nonnegativity of the risk-free interest rate process $\{r(t)\}$, the process $\{\beta(t)\}$ is uniformly bounded above by the constant 1. As before, let $\kappa_\beta \geq 1$ be a uniform bound on $\{\beta^{-1}(t)\}$, we have

$$\begin{aligned}
&\mathrm{E}|Y^{(m)}(T) - Y(T)|^2 \\
&\overset{(4.5.8)}{=} \mathrm{E}|\beta(T)\left((y^{(m)} - y) + \mathbb{J}(\boldsymbol{\lambda}^{(m)} - \boldsymbol{\lambda})(T) + \mathbb{I}(\boldsymbol{\gamma}^{(m)} - \boldsymbol{\gamma})(T)\right)|^2 \\
&\leq 3\mathrm{E}\left((y^{(m)} - y)^2 + \mathbb{J}^2(\boldsymbol{\lambda}^{(m)} - \boldsymbol{\lambda})(T) + \mathbb{I}^2(\boldsymbol{\gamma}^{(m)} - \boldsymbol{\gamma})(T)\right) \\
&\leq 3\mathrm{E}|y^{(m)} - y|^2 + 3\kappa_\beta^2 \mathrm{E}\left(\int_0^T (\boldsymbol{\lambda}^{(m)} - \boldsymbol{\lambda})^\top(t)\,\mathrm{d}\mathbf{W}(t)\right)^2 \\
&\qquad + 3\kappa_\beta^2 \mathrm{E}\sum_{i,j=1}^{D}\left(\int_0^T \left(\gamma_{ij}^{(m)} - \gamma_{ij}\right)(t)\,\mathrm{d}\mathcal{Q}_{ij}(t)\right)^2.
\end{aligned} \tag{4.5.54}$$

Applying the Itô isometry to the second and third terms of the right-most side of the above inequality, we obtain

$$\begin{aligned}
&\mathrm{E}|Y^{(m)}(T) - Y(T)|^2 \\
&\leq 3|y^{(m)} - y|^2 + 3\kappa_\beta^2 \mathrm{E}\int_0^T \|\left(\boldsymbol{\lambda}^{(m)} - \boldsymbol{\lambda}\right)(t)\|^2 \,\mathrm{d}t \\
&\qquad\qquad + 3\kappa_\beta^2 \mathrm{E}\sum_{i,j=1}^{D}\int_0^T (\gamma_{ij}^{(m)} - \gamma_{ij})^2(t)\,\mathrm{d}[\mathcal{Q}_{ij}](t) \qquad (4.5.55)\\
&\overset{(4.5.31)}{\leq} 3\kappa_\beta^2 \|\left(y^{(m)} - y, \boldsymbol{\lambda}^{(m)} - \boldsymbol{\lambda}, \boldsymbol{\gamma}^{(m)} - \boldsymbol{\gamma}\right)\|^2 \\
&\overset{(4.5.44)}{\rightarrow} 0 \text{ as } m \to \infty.
\end{aligned}$$

Hence $Y^{(m)}(T) \to Y(T)$ as $m \to \infty$ in the space $L^2(\Omega, \mathcal{F}_T, \mathbb{P})$. From Condition 3.2.27, the sequence of random variables $\{Y^{(m)}(T) + b\}_{m\in\mathbb{N}}$ converges to $Y(T) + b$ in the space $L^2(\Omega, \mathcal{F}_T, \mathbb{P})$. From the bounds on the random variable a given by Condition 3.2.26, the mapping

$$X \mapsto \mathrm{E}\left(\frac{X^2}{2a}\right) : L^2(\Omega, \mathcal{F}_T, \mathbb{P}) \to \mathbb{R} \qquad (4.5.56)$$

is continuous and this implies that

$$\lim_{m\to\infty} \mathrm{E}\left|\frac{\left(Y^{(m)}(T) + b\right)^2}{2a}\right| = \mathrm{E}\left|\frac{(Y(T) + b)^2}{2a}\right|. \qquad (4.5.57)$$

Hence (4.5.53) is established and we have shown that the functional $\tilde{\Psi}$ is lower semi-continuous. □

Lemma 4.5.14. *The functional* $\tilde{\Psi}$ *is coercive on* $\mathbb{R} \times L^2(\mathbf{W}) \times L^2(\mathcal{Q})$*, that is*

$$\lim_{\|(y,\boldsymbol{\lambda},\boldsymbol{\gamma})\|\to\infty} \tilde{\Psi}(y, \boldsymbol{\lambda}, \boldsymbol{\gamma}) = +\infty, \qquad (4.5.58)$$

for the norm $\|(\cdot,\cdot,\cdot)\|$ *given by (4.5.31).*

Proof. Fix $Y \equiv \Xi(y, \boldsymbol{\lambda}, \boldsymbol{\gamma}) \in \mathbb{B}_1$. From Remark 4.4.5, for the support function δ, we have $\delta \geq 0$. Applying this to (4.5.35), we get

$$\tilde{\Psi}(y, \boldsymbol{\lambda}, \boldsymbol{\gamma}) \geq x_0 y + \mathrm{E}\left(\frac{(Y(T) + b)^2}{2a}\right) - \mathrm{E}c. \qquad (4.5.59)$$

We show that $\mathrm{E}\left(\frac{(Y(T)+b)^2}{2a}\right) \to \infty$ as $\|(y, \boldsymbol{\lambda}, \boldsymbol{\gamma})\| \to \infty$. To simplify this task, we will show that we can replace the random variable b with zero. To do this, we first write the random variable b in the form of the map Ξ. Then we add the random variables $Y(T)$ and b using their expressions in terms of the map Ξ.

From Condition 3.2.27 and the uniform boundedness of the process $\{\beta^{-1}(t)\}$, the random variable $\beta^{-1}(T)b$ is $\mathcal{F}_T$-measurable and square-integrable. Then the process

$$\left\{\left(\mathrm{E}\left(\beta^{-1}(T)b|\mathcal{F}_t\right), \mathcal{F}_t\right) : t \in [0, T]\right\} \tag{4.5.60}$$

is a square-integrable martingale. Setting $\tilde{y} := \mathrm{E}\left(\beta^{-1}(T)b\right)$ and applying Theorem B.4.6 to the square-integrable martingale, there exists $\tilde{\boldsymbol{\lambda}} \in L^2(\mathbf{W})$ and $\tilde{\boldsymbol{\gamma}} = (\tilde{\gamma}_{ij})_{i,j=1}^{D} \in L^2(\mathcal{Q})$ such that

$$\mathrm{E}\left(\beta^{-1}(T)b|\mathcal{F}_T\right) = \tilde{y} + \int_0^T \tilde{\boldsymbol{\lambda}}^\top(t)\,\mathrm{d}\mathbf{W}(t) + \sum_{i,j=1}^{D} \int_0^T \tilde{\gamma}_{ij}(t)\,\mathrm{d}\mathcal{Q}_{ij}(t). \tag{4.5.61}$$

Multiplying the integrands by $\beta^{-1}(t)\beta(t)$ and using the $\mathcal{F}_T$-measurability of the random variable $\beta^{-1}(T)b$, we obtain upon rearranging the above equation,

$$\begin{aligned} b &= \beta(T)\left(\tilde{y} + \int_0^T \beta^{-1}(t)\beta(t)\tilde{\boldsymbol{\lambda}}^\top(t)\,\mathrm{d}W(t) + \sum_{i,j=1}^{D} \int_0^T \beta^{-1}(t)\beta(t)\tilde{\boldsymbol{\gamma}}_{ij}(t)\,\mathrm{d}\mathcal{Q}_{ij}(t)\right) \\ &\overset{(4.5.6),(4.5.7)}{=} \beta(T)\left(\tilde{y} + \mathbb{J}\left(\beta\tilde{\boldsymbol{\lambda}}\right)(T) + \mathbb{I}\left(\beta\tilde{\boldsymbol{\gamma}}\right)(T)\right) \\ &\overset{(4.5.8)}{=} \Xi(\tilde{y}, \beta\tilde{\boldsymbol{\lambda}}, \beta\tilde{\boldsymbol{\gamma}})(T). \end{aligned} \tag{4.5.62}$$

Since $Y(T) \equiv \Xi(y, \boldsymbol{\lambda}, \boldsymbol{\gamma})(T)$, we have from the linearity of the map Ξ that

$$\begin{aligned} Y(T) + b &\overset{(4.5.62)}{=} \Xi(y, \boldsymbol{\lambda}, \boldsymbol{\gamma})(T) + \Xi(\tilde{y}, \beta\tilde{\boldsymbol{\lambda}}, \beta\tilde{\boldsymbol{\gamma}})(T) \\ &= \Xi(y + \tilde{y}, \boldsymbol{\lambda} + \beta\tilde{\boldsymbol{\lambda}}, \boldsymbol{\gamma} + \beta\tilde{\boldsymbol{\gamma}})(T). \end{aligned} \tag{4.5.63}$$

Recall that we wish to show $\mathrm{E}\left(\frac{(Y(T)+b)^2}{2a}\right) \to \infty$ as $\|(y, \boldsymbol{\lambda}, \boldsymbol{\gamma})\| \to \infty$, that is from (4.5.63) we wish to show

$$\mathrm{E}\left(\frac{\left(\Xi(y + \tilde{y}, \boldsymbol{\lambda} + \beta\tilde{\boldsymbol{\lambda}}, \boldsymbol{\gamma} + \beta\tilde{\boldsymbol{\gamma}})(T)\right)^2}{2a}\right) \to \infty \quad \text{as} \quad \|(y, \boldsymbol{\lambda}, \boldsymbol{\gamma})\| \to \infty. \tag{4.5.64}$$

Since for fixed $(\tilde{y}, \beta\tilde{\boldsymbol{\lambda}}, \beta\tilde{\boldsymbol{\gamma}}) \in \mathbb{R} \times L^2(\mathbf{W}) \times L^2(\mathcal{Q})$, $\|(y + \tilde{y}, \boldsymbol{\lambda} + \beta\tilde{\boldsymbol{\lambda}}, \boldsymbol{\gamma} + \beta\tilde{\boldsymbol{\gamma}})\| \to \infty$ as $\|(y, \boldsymbol{\lambda}, \boldsymbol{\gamma})\| \to \infty$, we can, without loss of generality, replace the random variable b with zero.

Setting

$$\epsilon_0 := \inf_{\omega \in \Omega} \left\{\frac{1}{2a(\omega)}\right\}, \tag{4.5.65}$$

we see from the strict positivity of the random variable a (see Condition 3.2.26), that $\epsilon_0 > 0$.

Let $\kappa_r \in (0, \infty)$ satisfy (3.2.27), that is κ_r is a uniform upper bound on the risk-free interest rate process $\{r(t)\}$. Then

$$1 \geq \beta(t) \geq \exp\{-\kappa_r t\} \quad \text{a.s.}, \quad \forall t \in [0, T]. \tag{4.5.66}$$

Expanding $\mathrm{E}\left(\frac{1}{2a}(Y(T))^2\right)$, we get

$$\begin{aligned}
&\mathrm{E}\left(\frac{1}{2a}(Y(T))^2\right) \\
&= \mathrm{E}\left(\frac{1}{2a}\Xi(y, \boldsymbol{\lambda}, \boldsymbol{\gamma})^2(T)\right) \\
&\overset{(4.5.8)}{=} \mathrm{E}\left(\frac{\beta^2(T)}{2a}\left(y + \mathbb{J}(\boldsymbol{\lambda})(T) + \mathbb{I}(\boldsymbol{\gamma})(T)\right)^2\right) \\
&\overset{(4.5.65),(4.5.66)}{\geq} \epsilon_0 \exp\{-2\kappa_r T\} \mathrm{E}\left(y + \mathbb{J}(\boldsymbol{\lambda})(T) + \mathbb{I}(\boldsymbol{\gamma})(T)\right)^2 \\
&\overset{(4.5.6),(4.5.7)}{=} \epsilon_0 \exp\{-2\kappa_r T\} \\
&\qquad \cdot \mathrm{E}\left(y + \int_0^T \beta^{-1}(t)\boldsymbol{\lambda}^\top(t)\,\mathrm{d}\mathbf{W}(t) + \sum_{i,j=1}^{D}\int_0^T \beta^{-1}(t)\gamma_{ij}(t)\,\mathrm{d}\mathcal{Q}_{ij}(t)\right)^2.
\end{aligned} \tag{4.5.67}$$

Applying the Itô isometry to the last line above gives

$$\begin{aligned}
&\mathrm{E}\left(\frac{(Y(T))^2}{2a}\right) \\
&\geq \epsilon_0 \exp\{-2\kappa_r T\}\Bigg(|y|^2 + \mathrm{E}\int_0^T \beta^{-2}(t)\|\boldsymbol{\lambda}(t)\|^2\,\mathrm{d}t + \mathrm{E}\sum_{i,j=1}^{D}\int_0^T \beta^{-2}(t)\gamma_{ij}^2(t)\,\mathrm{d}[\mathcal{Q}_{ij}](t) \\
&\qquad + \mathrm{E}\sum_{n=1}^{N}\sum_{i,j=1}^{D}\int_0^T \beta^{-2}(t)\lambda_n(t)\gamma_{ij}(t)\,\mathrm{d}[W_n, \mathcal{Q}_{ij}](t)\Bigg).
\end{aligned} \tag{4.5.68}$$

By Lemma B.3.18, the last term on the last line above vanishes. Using the fact that for all $t \in [0, T]$, $\beta^{-1}(t) \geq 1$ a.s., we get

$$\begin{aligned}
&\mathrm{E}\left(\frac{(Y(T))^2}{2a}\right) \\
&\geq \epsilon_0 \exp\{-2\kappa_r T\}\left(|y|^2 + \mathrm{E}\int_0^T \|\boldsymbol{\lambda}(t)\|^2\,\mathrm{d}t + \mathrm{E}\sum_{i,j=1}^{D}\int_0^T \gamma_{ij}^2(t)\,\mathrm{d}[\mathcal{Q}_{ij}](t)\right) \\
&\overset{(4.5.29),(4.5.30)}{=} \epsilon_0 \exp\{-2\kappa_r T\}\left(|y|^2 + \|\boldsymbol{\lambda}\|^2_{L^2(\mathbf{W})} + \|\boldsymbol{\gamma}\|^2_{L^2(\mathcal{Q})}\right) \\
&\overset{(4.5.31)}{=} \epsilon_0 \exp\{-2\kappa_r T\}\|(y, \boldsymbol{\lambda}, \boldsymbol{\gamma})\|^2 \\
&\to \infty \quad \text{as } \|(y, \boldsymbol{\lambda}, \boldsymbol{\gamma})\| \to \infty.
\end{aligned} \tag{4.5.69}$$

This shows that the functional $\tilde{\Psi}$ is coercive on $\mathbb{R} \times L^2(\mathbf{W}) \times L^2(\mathcal{Q})$. □

The next proposition summarizes the results obtained in this section.

Proposition 4.5.15. *For the functional $\tilde{\Psi}$ given by (4.5.23), there exists a triple $(\bar{y}, \bar{\boldsymbol{\lambda}}, \bar{\boldsymbol{\gamma}}) \in \mathbb{R} \times L^2(\mathbf{W}) \times L^2(\mathcal{Q})$ such that*

$$\tilde{\Psi}(\bar{y}, \bar{\boldsymbol{\lambda}}, \bar{\boldsymbol{\gamma}}) = \inf_{(y,\boldsymbol{\lambda},\boldsymbol{\gamma}) \in \mathbb{R} \times L^2(\mathbf{W}) \times L^2(\mathcal{Q})} \{\tilde{\Psi}(y, \boldsymbol{\lambda}, \boldsymbol{\gamma})\}. \tag{4.5.70}$$

Furthermore, setting $\bar{Y} = \Xi(\bar{y}, \bar{\boldsymbol{\lambda}}, \bar{\boldsymbol{\gamma}})$, for the map Ξ defined by (4.5.8), we have

$$\Psi(\bar{Y}) = \inf_{Y \in \mathbb{B}} \{\Psi(Y)\}, \tag{4.5.71}$$

for the dual cost functional Ψ given by (4.2.4). In other words, there exists a solution to the dual problem (Problem 4.2.8).

Proof. We seek to apply Proposition 4.5.6. By Remark 4.5.9, $\mathbb{R} \times L^2(\mathbf{W}) \times L^2(\mathcal{Q})$ is a reflexive Banach space. From Lemma 4.5.11, Lemma 4.5.12, Lemma 4.5.13 and Lemma 4.5.14, the functional $\tilde{\Psi}$ is a convex, proper, lower semi-continuous, coercive function over $\mathbb{R} \times L^2(\mathbf{W}) \times L^2(\mathcal{Q})$. Hence by Proposition 4.5.6, there is at least one solution $(\bar{y}, \bar{\boldsymbol{\lambda}}, \bar{\boldsymbol{\gamma}}) \in \mathbb{R} \times L^2(\mathbf{W}) \times L^2(\mathcal{Q})$ to the problem

$$\inf_{(y,\boldsymbol{\lambda},\boldsymbol{\gamma}) \in \mathbb{R} \times L^2(\mathbf{W}) \times L^2(\mathcal{Q})} \{\tilde{\Psi}(y, \boldsymbol{\lambda}, \boldsymbol{\gamma})\}. \tag{4.5.72}$$

This gives (4.5.70).

Now set $\bar{Y} = \Xi(\bar{y}, \bar{\boldsymbol{\lambda}}, \bar{\boldsymbol{\gamma}})$. Then (4.5.71) follows from Remark 4.5.5. □

4.6 Construction of a candidate solution to the primal problem

In this section, we propose a candidate solution $\tilde{X}$ to the primal problem, which we show belongs to the space $\mathbb{B}$. Using the candidate solution $\tilde{X}$, we construct a candidate portfolio process $\tilde{\boldsymbol{\pi}}$. However, it is not until the next section that we show $\tilde{X} \in \mathbb{A}$ and $\tilde{X} = X^{\tilde{\boldsymbol{\pi}}}$.

We begin by using the definition of $\beta(t)$ in (4.5.5) to define the *state price density process*

$$H(t) := \beta(t) \exp\left(-\int_0^t \boldsymbol{\theta}^\top(\tau)\, \mathrm{d}\mathbf{W}(\tau) - \frac{1}{2}\int_0^t \|\boldsymbol{\theta}(\tau)\|^2\, \mathrm{d}\tau\right), \quad \forall t \in [0,T]. \tag{4.6.1}$$

Using the symbol $\bullet$ to indicate stochastic integration, so that

$$(-\boldsymbol{\theta} \bullet \mathbf{W})(t) := \int_0^t -\boldsymbol{\theta}^\top(\tau)\ \mathrm{d}\mathbf{W}(\tau), \tag{4.6.2}$$

and denoting by $\mathcal{E}(-\boldsymbol{\theta} \bullet \mathbf{W})$ the *Doléans-Dade exponential* of the continuous martingale $(-\boldsymbol{\theta} \bullet \mathbf{W})$, which, by Remark C.15.2, satisfies

$$\mathcal{E}(-\boldsymbol{\theta} \bullet \mathbf{W})(t) = \exp\left\{\int_0^t -\boldsymbol{\theta}^\top(\tau)\, \mathrm{d}\mathbf{W}(\tau) - \frac{1}{2}\int_0^t \|\boldsymbol{\theta}(\tau)\|^2\, \mathrm{d}\tau\right\}, \tag{4.6.3}$$

we can rewrite (4.6.1) in the more compact form

$$H(t) = \beta(t)\mathcal{E}(-\boldsymbol{\theta} \bullet \mathbf{W})(t), \quad \forall t \in [0, T]. \tag{4.6.4}$$

Remark 4.6.1. $\{H(t)\}$ is a continuous, strictly positive process.

Remark 4.6.2. We aim to show that the state price density process H has the following properties:

- for all $p \in \mathbb{R}$, $\mathrm{E}\left(\sup_{t\in[0,T]}|H(t)|^p\right) < \infty$ (Proposition 4.6.5);
- $H \equiv (1, -rH, -H\boldsymbol{\theta}, \mathbf{0}) \in \mathbb{A}$ (Proposition 4.6.6); and
- for all $\boldsymbol{\pi} \in L^2(\mathbf{W})$, $X^{\boldsymbol{\pi}}H \in \mathcal{M}(\{\mathcal{F}_t\}, \mathbb{P})$ where $X^{\boldsymbol{\pi}}$ is the solution to the wealth equation (3.2.31) for the portfolio process $\boldsymbol{\pi}$ (Proposition 4.6.7).

Of most importance is the last item, since this motivates the form of the candidate solution to the primal problem.

We begin by showing that, for any $p \in \mathbb{R}$, the local martingale $\mathcal{E}(-p\boldsymbol{\theta} \bullet \mathbf{W})$ is a square-integrable martingale. We then use this fact to show that $H \in \mathbb{A}$.

Proposition 4.6.3. *For any $p \in \mathbb{R}$, $\mathcal{E}(-p\boldsymbol{\theta} \bullet \mathbf{W}) \in \mathcal{M}^2(\{\mathcal{F}_t\}, \mathbb{P})$.*

Proof. 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 C.15.4). Consider

$$Z(t) := \int_0^t -p\boldsymbol{\theta}^\top(\tau)\, \mathrm{d}\mathbf{W}(\tau). \tag{4.6.5}$$

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

$$[Z](t) = p^2 \int_0^t \|\boldsymbol{\theta}(\tau)\|^2\, \mathrm{d}\tau \quad \text{a.s.}, \quad \forall t \in [0, T]. \tag{4.6.6}$$

Using the constant $\kappa_{\boldsymbol{\theta}} \in (0, \infty)$ which satisfies (3.2.28), we obtain the bound $[Z](t) \leq (p\kappa_{\boldsymbol{\theta}})^2\, t$ a.s. for each $t \in [0, T]$. Hence

$$\mathrm{E}\left(\exp\left\{\frac{1}{2}[Z](T)\right\}\right) \leq \exp\left\{\frac{1}{2}(p\kappa_{\boldsymbol{\theta}})^2 T\right\} < \infty. \tag{4.6.7}$$

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

Using Corollary C.15.3, we have

$$(\mathcal{E}(-p\boldsymbol{\theta}\bullet\mathbf{W})(t))^2 = \exp\left\{p^2\int_0^t \|\boldsymbol{\theta}(\tau)\|^2\,\mathrm{d}\tau\right\}\mathcal{E}(-2p\boldsymbol{\theta}\bullet\mathbf{W})(t), \tag{4.6.8}$$

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

$$\mathrm{E}|\,\mathcal{E}(-p\boldsymbol{\theta}\bullet\mathbf{W})(t)|^2 \le \exp\{(p\kappa_{\boldsymbol{\theta}})^2T\}\,\mathrm{E}|\,\mathcal{E}(-2p\boldsymbol{\theta}\bullet\mathbf{W})(t)| < \infty. \tag{4.6.9}$$

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

Lemma 4.6.4. *For all $p\in\mathbb{R}$ and for all $t\in[0,T]$,*

$$|\,\mathcal{E}(-\boldsymbol{\theta}\bullet\mathbf{W})(t)|^p = \exp\left\{-p\int_0^t \boldsymbol{\theta}^\top(\tau)\,\mathrm{d}\mathbf{W}(\tau) - \frac{1}{2}p\int_0^t\|\boldsymbol{\theta}(\tau)\|^2\,\mathrm{d}\tau\right\}. \tag{4.6.10}$$

Proof. Clearly, (4.6.10) holds for $p=0$. It also holds for $p=1$, upon recalling (4.6.3). Set

$$Z(t) := \mathcal{E}(-\boldsymbol{\theta}\bullet\mathbf{W})(t), \tag{4.6.11}$$

so that $Z(t)$ satisfies (by (C.15.1)),

$$Z(t) = 1 - \int_0^t Z(\tau)\boldsymbol{\theta}(\tau)\,\mathrm{d}\mathbf{W}(\tau), \tag{4.6.12}$$

and hence

$$[Z](t) = \int_0^t Z^2(\tau)\|\boldsymbol{\theta}(\tau)\|^2\,\mathrm{d}\tau. \tag{4.6.13}$$

Fix $p\in\mathbb{R}$, $p\neq 0$, $p\neq 1$ and set $f(x)=x^p$. Using Itô's Formula (Theorem C.14.2) to expand $f(Z(t))$, we obtain

$$\begin{aligned} Z^p(t) &= 1 + \int_0^t pZ^{p-1}(\tau)\,\mathrm{d}Z(\tau) + \frac{1}{2}\int_0^t p(p-1)Z^{p-2}(\tau)\,\mathrm{d}[Z](\tau) \\ &\overset{(4.6.12),(4.6.13)}{=} 1 - \int_0^t pZ^p(\tau)\boldsymbol{\theta}^\top(\tau)\,\mathrm{d}\mathbf{W}(\tau) + \frac{1}{2}\int_0^t p(p-1)Z^p(\tau)\|\boldsymbol{\theta}(\tau)\|^2\,\mathrm{d}\tau \\ &= 1 + \int_0^t Z^p\left(-p\boldsymbol{\theta}^\top(\tau)\,\mathrm{d}\mathbf{W}(\tau) + \frac{1}{2}p(p-1)\|\boldsymbol{\theta}(\tau)\|^2\,\mathrm{d}\tau\right). \end{aligned} \tag{4.6.14}$$

Then Z^p is an exponential semimartingale which satisfies (C.15.3), that is

$$\begin{aligned} Z^p(t) &= \exp\left\{-p\int_0^t\boldsymbol{\theta}^\top(\tau)\,\mathrm{d}\mathbf{W}(\tau) + \frac{1}{2}p(p-1)\int_0^t\|\boldsymbol{\theta}(\tau)\|^2\,\mathrm{d}\tau - \frac{1}{2}p^2\int_0^t\|\boldsymbol{\theta}(\tau)\|^2\,\mathrm{d}\tau\right\} \\ &= \exp\left\{-p\int_0^t\boldsymbol{\theta}^\top(\tau)\,\mathrm{d}\mathbf{W}(\tau) - \frac{1}{2}p\int_0^t\|\boldsymbol{\theta}(\tau)\|^2\,\mathrm{d}\tau\right\}, \end{aligned} \tag{4.6.15}$$

giving (4.6.10). □

Proposition 4.6.5. *For any $p \in \mathbb{R}$,*

$$E\left(\sup_{t\in[0,T]} |H(t)|^p\right) < \infty. \tag{4.6.16}$$

Proof. Fix $t \in [0,T]$. From the nonnegativity of the risk-free interest rate process $\{r(t)\}$, we have $\beta(t) \leq 1$ a.s. for all $t \in [0,T]$. Expanding $|H(t)|^p$ using Lemma 4.6.4, we get

$$\begin{aligned} |H(t)|^p &\overset{(4.6.4)}{=} |\beta(t)\mathcal{E}(-\boldsymbol{\theta} \bullet \mathbf{W})(t)|^p \\ &\overset{(4.6.2),(4.6.3)}{\leq} \exp\left\{-\int_0^t p\,\boldsymbol{\theta}^\top(\tau)\,\mathrm{d}\mathbf{W}(\tau) - \frac{1}{2}p\int_0^t \|\boldsymbol{\theta}(\tau)\|^2\,\mathrm{d}\tau\right\} \\ &= \exp\left\{\frac{1}{4}p(p-2)\int_0^t \|\boldsymbol{\theta}(\tau)\|^2\,\mathrm{d}\tau\right\}\left|\mathcal{E}(-\frac{p}{2}\boldsymbol{\theta} \bullet \mathbf{W})(t)\right|^2. \end{aligned} \tag{4.6.17}$$

Recall the constant $\kappa_{\boldsymbol{\theta}} \in (0,\infty)$ which satisfies (3.2.28). Taking the supremum over $t \in [0,T]$ in (4.6.17), we get

$$\sup_{t\in[0,T]} |H(t)|^p \leq \max\left[1, \exp\left\{\frac{1}{4}p(p-2)\kappa_{\boldsymbol{\theta}}^2 T\right\}\right] \sup_{t\in[0,T]}\left|\mathcal{E}(-\frac{p}{2}\boldsymbol{\theta} \bullet \mathbf{W})(t)\right|^2, \tag{4.6.18}$$

where the maximum will equal one if $p \in (0,2)$ and will otherwise equal $\exp\left\{\frac{1}{4}p(p-2)\kappa_{\boldsymbol{\theta}}^2 T\right\}$. Upon taking expectations, we get

$$\mathrm{E}\left(\sup_{t\in[0,T]} |H(t)|^p\right) \leq \max\left[1, \exp\left\{\frac{1}{4}p(p-2)\kappa_{\boldsymbol{\theta}}^2 T\right\}\right] \mathrm{E}\left(\sup_{t\in[0,T]}\left|\mathcal{E}(-\frac{p}{2}\boldsymbol{\theta} \bullet \mathbf{W})(t)\right|^2\right). \tag{4.6.19}$$

By Proposition 4.6.3, $\mathcal{E}(-\frac{p}{2}\boldsymbol{\theta} \bullet \mathbf{W})$ is a square-integrable martingale, so upon applying Doob's L^2-inequality, we get

$$\mathrm{E}\left(\sup_{t\in[0,T]} |H(t)|^p\right) < \infty. \tag{4.6.20}$$

□

Proposition 4.6.6. *Recalling (3.2.38), we have $H \equiv (1, -rH, -H\boldsymbol{\theta}, \mathbf{0}) \in \mathbb{A}$. That is,*

$$H_0 = 1 \in \mathbb{R}, \quad \dot{H} = -rH \in L_{21}, \quad \boldsymbol{\Lambda}^H = -H\boldsymbol{\theta} \in L^2(\mathbf{W}) \quad \textit{and} \quad \boldsymbol{\Gamma}^H = \mathbf{0} \in L^2(\mathcal{Q}). \tag{4.6.21}$$

Proof. Setting $p = 2$ in Proposition 4.6.5 shows that H is square-integrable. Expanding $H(t) = \beta(t)\mathcal{E}(-\boldsymbol{\theta} \bullet \mathbf{W})(t)$ using the integration-by-parts formula (see Theorem C.14.1), we get

$$\begin{aligned} H(t) &= \beta(t)\mathcal{E}(-\boldsymbol{\theta} \bullet \mathbf{W})(t) \\ &= 1 + \int_0^t \mathcal{E}(-\boldsymbol{\theta} \bullet \mathbf{W})(\tau)\,\mathrm{d}\beta(\tau) + \int_0^t \beta(\tau)\,\mathrm{d}\mathcal{E}(-\boldsymbol{\theta} \bullet \mathbf{W})(\tau) \\ &\quad + [\beta, \mathcal{E}(-\boldsymbol{\theta} \bullet \mathbf{W})](t). \end{aligned} \tag{4.6.22}$$

The square-bracket quadratic co-variation term $[\beta, \mathcal{E}(-\boldsymbol{\theta} \bullet \mathbf{W})](t) = 0$ a.s. since $\{\beta(t)\}$ is a continuous, finite variation process. As $\mathcal{E}(-\boldsymbol{\theta} \bullet \mathbf{W})$ is a Doléans-Dade exponential, it satisfies (C.15.1), that is

$$\begin{aligned}\mathcal{E}(-\boldsymbol{\theta} \bullet \mathbf{W})(t) &= 1 + \int_0^t \mathcal{E}(-\boldsymbol{\theta} \bullet \mathbf{W})(\tau)\ \mathrm{d}\left(-\boldsymbol{\theta} \bullet \mathbf{W}\right)(\tau) \\ &= 1 - \int_0^t \mathcal{E}(-\boldsymbol{\theta} \bullet \mathbf{W})(\tau)\, \boldsymbol{\theta}^\top(\tau)\, \mathrm{d}\mathbf{W}(\tau).\end{aligned} \tag{4.6.23}$$

From the definition of $\beta(t)$ in (4.5.5),

$$\mathrm{d}\beta(\tau) = -r(\tau)\beta(\tau)\, \mathrm{d}\tau. \tag{4.6.24}$$

Substituting these last two equations into (4.6.22), we find

$$\begin{aligned}H(t) &= 1 - \int_0^t r(\tau)\beta(\tau)\mathcal{E}(-\boldsymbol{\theta} \bullet \mathbf{W})(\tau)\ \mathrm{d}\tau - \int_0^t \beta(\tau)\mathcal{E}(-\boldsymbol{\theta} \bullet \mathbf{W})(\tau)\, \boldsymbol{\theta}^\top(\tau)\, \mathrm{d}\mathbf{W}(\tau) \\ &= 1 - \int_0^t r(\tau)H(\tau)\ \mathrm{d}\tau - \int_0^t H(\tau)\, \boldsymbol{\theta}^\top(\tau)\ \mathrm{d}\mathbf{W}(\tau).\end{aligned} \tag{4.6.25}$$

From the square-integrability of $\{H(t)\}$ and the constant bound $\kappa_r \in (0, \infty)$ on the risk-free interest rate process $\{r(t)\}$ which satisfies (3.2.27), we get $-rH \in L_{21}$. Similarly, from the square-integrability of $\{H(t)\}$ and the constant bound $\kappa_{\boldsymbol{\theta}} \in (0, \infty)$ on the market price of risk $\{\boldsymbol{\theta}(t)\}$ which satisfies (3.2.28), we obtain $-H\boldsymbol{\theta} \in L^2(\mathbf{W})$. Hence $H \equiv (1, -rH, -H\boldsymbol{\theta}, \mathbf{0}) \in \mathbb{A}$. □

The next proposition is one of the motivations for the form of the candidate solution, since any solution to the primal problem must be of the form $X^{\boldsymbol{\pi}}$ for some $\boldsymbol{\pi} \in L^2(\mathbf{W})$, where $X^{\boldsymbol{\pi}}$ is the solution to the wealth equation (3.2.31) for the portfolio process $\boldsymbol{\pi}$. The next proposition gives a useful property of any such process $X^{\boldsymbol{\pi}}$.

Proposition 4.6.7. *$X^{\boldsymbol{\pi}} H \in \mathcal{M}(\{\mathcal{F}_t\}, \mathbb{P})$ for every portfolio process $\boldsymbol{\pi} \in L^2(\mathbf{W})$.*

Proof. Fix $\boldsymbol{\pi} \in L^2(\mathbf{W})$. From the wealth equation (3.2.31) and recalling (3.2.38), we can write the wealth process $X^{\boldsymbol{\pi}}$ in component form as

$$X^{\boldsymbol{\pi}} \equiv (x_0, rX^{\boldsymbol{\pi}} + \boldsymbol{\pi}^\top \boldsymbol{\sigma}\boldsymbol{\theta}, \boldsymbol{\sigma}^\top \boldsymbol{\pi}, \mathbf{0}) \in \mathbb{A}. \tag{4.6.26}$$

From Proposition 4.6.6,

$$H \equiv (1, -rH, -H\boldsymbol{\theta}, \mathbf{0}) \in \mathbb{A}. \tag{4.6.27}$$

From Proposition 4.3.1, the process $\{\mathbb{M}(X^{\boldsymbol{\pi}}, H)(t) : t \in [0, T]\}$ defined by

$$\begin{aligned}\mathbb{M}(X^{\boldsymbol{\pi}}, H)(t) := X^{\boldsymbol{\pi}}(t)H(t) - x_0 &- \int_0^t \left\{r(\tau)X^{\boldsymbol{\pi}}(\tau) + \boldsymbol{\pi}^\top(\tau)\boldsymbol{\sigma}(\tau)\boldsymbol{\theta}(\tau)\right\} H(\tau)\ \mathrm{d}\tau \\ &- \int_0^t X^{\boldsymbol{\pi}}(\tau)\left\{-r(\tau)H(\tau)\right\}\ \mathrm{d}\tau - \int_0^t \left\{\boldsymbol{\pi}^\top(\tau)\boldsymbol{\sigma}(\tau)\right\}\left\{-H(\tau)\boldsymbol{\theta}(\tau)\right\}\ \mathrm{d}\tau\end{aligned} \tag{4.6.28}$$

is a martingale, null at the origin. Upon simplifying the above equation, we obtain $\mathbb{M}(X^{\boldsymbol{\pi}}, H)(t) = X^{\boldsymbol{\pi}}(t)H(t) - x_0$. Thus $X^{\boldsymbol{\pi}} H \in \mathcal{M}(\{\mathcal{F}_t\}, \mathbb{P})$ for any $\boldsymbol{\pi} \in L^2(\mathbf{W})$. $\square$

Next we propose the candidate solution $\tilde{X}$ to the primal problem (Problem 4.1.7). The candidate solution $\tilde{X}$ is motivated both by Proposition 4.6.7 and by (4.4.17), which is one of the necessary and sufficient conditions for a pair $(\bar{X}, \bar{Y}) \in \mathbb{A} \times \mathbb{B}$ to solve the primal and dual problems. The reasoning is as follows. Suppose that the pair $(\bar{X}, \bar{Y}) \in \mathbb{A} \times \mathbb{B}$ solves the primal and dual problems. Then there exists $\bar{\boldsymbol{\pi}} \in L^2(\mathbf{W})$ corresponding to $\bar{X} \in \mathbb{A}$, so that $\bar{X} \equiv X^{\bar{\boldsymbol{\pi}}}$. By Proposition 4.6.7, $X^{\bar{\boldsymbol{\pi}}} H = \bar{X} H$ is a martingale, so for all $t \in [0, T]$,

$$\bar{X}(t)H(t) = \mathrm{E}\left(\bar{X}(T)H(T) | \mathcal{F}_t\right) \quad \text{a.s.} \tag{4.6.29}$$

From (4.4.17), we have

$$\bar{X}(T) = -\left(\frac{\bar{Y}(T) + b}{a}\right) \quad \text{a.s.} \tag{4.6.30}$$

Using (4.6.30) to replace $\bar{X}(T)$ in (4.6.29), we obtain

$$\bar{X}(t)H(t) = \mathrm{E}\left(-\left(\frac{\bar{Y}(T) + b}{a}\right) H(T) \,\middle|\, \mathcal{F}_t\right) \quad \text{a.s.} \tag{4.6.31}$$

Guided by (4.6.31), we define a candidate solution $\tilde{X}$ to the primal problem as

$$\tilde{X}(t) := -\frac{1}{H(t)} \mathrm{E}\left(\left(\frac{\bar{Y}(T) + b}{a}\right) H(T) \,\middle|\, \mathcal{F}_t\right) \quad \text{a.s., } \forall t \in [0, T]. \tag{4.6.32}$$

Remark 4.6.8. The candidate solution $\tilde{X}$ is well-defined as $H > 0$ by Remark 4.6.1.

Remark 4.6.9. We aim to show that the candidate solution $\tilde{X}$ has the following properties:

- $\mathrm{E}\left(\sup_{t\in[0,T]} |\tilde{X}(t)|^2\right) < \infty$ (Proposition 4.6.10); and
- $\tilde{X} H \in \mathcal{M}^2_{\text{loc}}(\{\mathcal{F}_t\}, \mathbb{P})$ (Lemma 4.6.12).

If we can show the last item, then we can apply Theorem B.4.22, which is the martingale representation theorem for locally square-integrable martingales, to $\tilde{X} H$. Using Itô's Formula, we can then express the candidate solution $\tilde{X}$ as a stochastic integral equation. As ultimately we want to show that the candidate solution $\tilde{X}$ satisfies the wealth equation (3.2.31) for some portfolio process $\tilde{\boldsymbol{\pi}}$, we compare the stochastic integral equation for the candidate solution $\tilde{X}$ to the wealth equation in order to ascertain the form of the candidate portfolio process $\tilde{\boldsymbol{\pi}}$. Then we check that $\tilde{\boldsymbol{\pi}} \in L^2(\mathbf{W})$ and from this it will follow that $\tilde{X} \in \mathbb{B}$. However, we won't have shown yet that $\tilde{X} \in \mathbb{A}$ and $\tilde{X}$ is the solution to the wealth equation (3.2.31) for some portfolio process $\tilde{\boldsymbol{\pi}}$. That must wait until Section 4.7.

Proposition 4.6.10. *The candidate solution $\tilde{X}$, defined by (4.6.32), is such that*

$$E\left(\sup_{t\in[0,T]} |\tilde{X}(t)|^2\right) < \infty. \tag{4.6.33}$$

Proof. The following proof carries over directly from Labbé and Heunis [33]. We include it for completeness, as we wish to ensure that it continues to work for our larger filtration.

For ease of notation, define

$$D_T := \frac{\bar{Y}(T) + b}{a}. \tag{4.6.34}$$

By the strict positivity and finiteness of the random variable a, given by Condition 3.2.26, the square-integrability of the random variable b, given by Condition 3.2.27, and the existence of the solution $\bar{Y}$ to the dual problem, as demonstrated in Section 4.5, we can immediately conclude that

$$\mathrm{E}|D_T|^2 < \infty. \tag{4.6.35}$$

Using D_T to rewrite (4.6.32), we obtain

$$\tilde{X}(t) = -\mathrm{E}\left(D_T H(T) H^{-1}(t) \,\middle|\, \mathcal{F}_t\right) \quad \text{a.s.} \tag{4.6.36}$$

From Proposition 4.6.5 and Hölder's inequality, for any $p \in \mathbb{R}$, we have

$$\mathrm{E}|H(T)H^{-1}(t)|^p \le \left(\mathrm{E}|H(T)|^{2p}\right)^{\frac{1}{2}} \left(\mathrm{E}|H(t)|^{-2p}\right)^{\frac{1}{2}} < \infty. \tag{4.6.37}$$

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

$$\mathrm{E}\big|D_T H(T)H^{-1}(t)\big| \le \left(\mathrm{E}|D_T|^2\right)^{\frac{1}{2}} \left(\mathrm{E}|H(T)H^{-1}(t)|^2\right)^{\frac{1}{2}} < \infty. \tag{4.6.38}$$

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

Now fix real numbers $p \in (2,\infty)$ and $q \in (1,2)$ which together satisfy the equation $\frac{1}{p} + \frac{1}{q} = 1$. Applying Hölder's inequality for conditional expectations (Theorem C.3.1),

$$|\tilde{X}(t)| \le \left(\mathrm{E}\left(|H(T)H^{-1}(t)|^p \big| \mathcal{F}_t\right)\right)^{\frac{1}{p}} \left(\mathrm{E}\left(|D_T|^q \big| \mathcal{F}_t\right)\right)^{\frac{1}{q}} \quad \text{a.s.} \tag{4.6.39}$$

Raising all terms to the power of q, we obtain

$$|\tilde{X}(t)|^q \le \left(\mathrm{E}\left(|H(T)H^{-1}(t)|^p \big| \mathcal{F}_t\right)\right)^{\frac{q}{p}} \mathrm{E}\left(|D_T|^q \big| \mathcal{F}_t\right) \quad \text{a.s.} \tag{4.6.40}$$

Consider separately the terms composing the product on the right-hand side. We will show that each is finite almost surely.

Expanding $|H(t)|^p$ using Lemma 4.6.4 and recalling (4.6.3), we get

$$|H(t)|^p = \beta^p(t)\,\exp\left\{\frac{1}{2}p(p-1)\int_0^t \|\boldsymbol{\theta}(\tau)\|^2\,\mathrm{d}\tau\right\}\mathcal{E}(-p\boldsymbol{\theta}\bullet\mathbf{W})(t). \tag{4.6.41}$$

By the nonnegativity of the risk-free interest rate process $\{r(t)\}$, $\beta(t)\le 1$ a.s., so for all $t\in[0,T]$,

$$|H(t)|^p \le \exp\left\{\frac{1}{2}p(p-1)\int_0^t \|\boldsymbol{\theta}(\tau)\|^2\,\mathrm{d}\tau\right\}\mathcal{E}(-p\boldsymbol{\theta}\bullet\mathbf{W})(t). \tag{4.6.42}$$

Then, recalling the constant $\kappa_{\boldsymbol{\theta}}\in(0,\infty)$ which satisfies (3.2.28), we have

$$\begin{aligned}|H(T)|^p\,|H(t)|^{-p} &\le \exp\left\{\frac{1}{2}p(p-1)\int_0^T \|\boldsymbol{\theta}(\tau)\|^2\,\mathrm{d}\tau + \frac{1}{2}p(p+1)\int_0^t \|\boldsymbol{\theta}(\tau)\|^2\,\mathrm{d}\tau\right\}\\ &\qquad\mathcal{E}(-p\boldsymbol{\theta}\bullet\mathbf{W})(T)\,\mathcal{E}(p\boldsymbol{\theta}\bullet\mathbf{W})(t)\\ &\le \exp\left\{p^2\kappa_{\boldsymbol{\theta}}^2 T\right\}\,\mathcal{E}(-p\boldsymbol{\theta}\bullet\mathbf{W})(T)\,\mathcal{E}(p\boldsymbol{\theta}\bullet\mathbf{W})(t).\end{aligned} \tag{4.6.43}$$

From Proposition 4.6.3, $\mathcal{E}(-p\boldsymbol{\theta}\bullet\mathbf{W})$ is an $\{\mathcal{F}_t\}$–martingale. Using this fact and Corollary C.15.3, we obtain

$$\begin{aligned}\mathrm{E}\left(|H(T)|^p\,|H(t)|^{-p}\,\middle|\,\mathcal{F}_t\right) &\le \exp\left\{p^2\kappa_{\boldsymbol{\theta}}^2 T\right\}\,\mathcal{E}(-p\boldsymbol{\theta}\bullet\mathbf{W})(t)\,\mathcal{E}(p\boldsymbol{\theta}\bullet\mathbf{W})(t)\\ &= \exp\left\{p^2\kappa_{\boldsymbol{\theta}}^2 T\right\}\,\exp\{-p^2\int_0^t \|\boldsymbol{\theta}(\tau)\|^2\,\mathrm{d}\tau\}\\ &\le \exp\left\{p^2\kappa_{\boldsymbol{\theta}}^2 T\right\}.\end{aligned} \tag{4.6.44}$$

Substituting this bound into (4.6.40), we get

$$|\tilde{X}(t)|^q \le \exp\left\{pq\kappa_{\boldsymbol{\theta}}^2 T\right\}\,\mathrm{E}\left(|D_T|^q\middle|\mathcal{F}_t\right)\quad\text{a.s.} \tag{4.6.45}$$

We now show that $\mathrm{E}\left(|D_T|^q\middle|\mathcal{F}_t\right)$ is finite. Note first that as $q\in(1,2)$, we have

$$\mathrm{E}\left(|D_T|^q\right) < \mathrm{E}\left(|D_T|^2\right) \overset{(4.6.35)}{<} \infty. \tag{4.6.46}$$

Fix $q\in(1,2)$ and define

$$N(t) := \mathrm{E}\left(|D_T|^q\middle|\mathcal{F}_t\right). \tag{4.6.47}$$

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

$$p_1 := \frac{2}{q} \tag{4.6.48}$$

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

$$\begin{aligned} \mathrm{E}|N(t)|^{p_1} &\overset{(4.6.47)}{=} \mathrm{E}\Big| \mathrm{E}\left(|D_T|^q \middle| \mathcal{F}_t\right) \Big|^{p_1} \leq \mathrm{E}\Big| \mathrm{E}\left(|D_T|^{qp_1} \middle| \mathcal{F}_t\right) \Big| \\ &\overset{(4.6.48)}{=} \mathrm{E}\Big| \mathrm{E}\left(|D_T|^2 \middle| \mathcal{F}_t\right) \Big| = \mathrm{E}|D_T|^2 \overset{(4.6.35)}{<} \infty. \end{aligned} \tag{4.6.49}$$

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

$$\mathrm{E}\left(\sup_{t\in[0,T]} |N(t)|^{p_1}\right) \leq \left(\frac{p_1}{p_1 - 1}\right)^{p_1} \mathrm{E}|N(T)|^{p_1} \overset{(4.6.49)}{<} \infty. \tag{4.6.50}$$

Substituting $N(t) = \mathrm{E}\left(|D_T|^q \middle| \mathcal{F}_t\right)$ into (4.6.45), and raising both sides of the above inequality to the power $p_1 = \frac{2}{q}$, we get

$$|\tilde{X}(t)|^2 \leq \exp\left\{2p\kappa_{\boldsymbol{\theta}}^2 T\right\} N(t)^{p_1}. \tag{4.6.51}$$

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

$$\mathrm{E}\left(\sup_{t\in[0,T]} |\tilde{X}(t)|^2\right) \leq \exp\left\{2p\kappa_{\boldsymbol{\theta}}^2 T\right\} \mathrm{E}\left(\sup_{t\in[0,T]} |N(t)|^{p_1}\right) \overset{(4.6.50)}{<} \infty. \tag{4.6.52}$$

□

Next we show that $\tilde{X}H$ is a locally square-integrable martingale, so that we can apply a martingale representation theorem to express the candidate solution $\tilde{X}$ as a stochastic integral equation.

Remark 4.6.11. We do not know yet if there exists a portfolio process $\tilde{\boldsymbol{\pi}} \in L^2(\mathbf{W})$ for which $\tilde{X} = X^{\tilde{\boldsymbol{\pi}}}$ is the solution to the wealth equation (3.2.31). Thus we cannot use Proposition 4.6.7 to conclude that $\tilde{X}H \in \mathcal{M}(\{\mathcal{F}_t\}, \mathbb{P})$.

Lemma 4.6.12. *$\tilde{X}H \in \mathcal{M}^2_{loc}(\{\mathcal{F}_t\}, \mathbb{P})$.*

Proof. From Hölder's inequality and the square-integrability of $\frac{\bar{Y}(T)+b}{a}$ and $H(T)$,

$$\mathrm{E}\left| \left(\frac{\bar{Y}(T)+b}{a}\right) H(T)\right| \leq \left(\mathrm{E}\left| \left(\frac{\bar{Y}(T)+b}{a}\right)\right|^2\right)^{\frac{1}{2}} \left(\mathrm{E}|H(T)|^2\right)^{\frac{1}{2}} < \infty. \tag{4.6.53}$$

It follows that for all $t \in [0,T]$,

$$\mathrm{E}|\tilde{X}(t)H(t)| \overset{(4.6.32)}{=} \mathrm{E}\left(\left| \left(\frac{\bar{Y}(T)+b}{a}\right) H(T)\right| \,\middle|\, \mathcal{F}_t\right) < \infty. \tag{4.6.54}$$

Then from this integrability and the definition of $\tilde{X}H$ in terms of a conditional expectation, we have that

$$\tilde{X}H \in \mathcal{M}(\{\mathcal{F}_t\}, \mathbb{P}). \tag{4.6.55}$$

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

$$T^m := \inf\{t > 0 \,:\, |H(t)| > m\} \wedge T. \tag{4.6.56}$$

Then T^m is an $\{\mathcal{F}_t\}$–stopping time (by Proposition C.5.1) and $T^m \Uparrow T$ a.s. (see Definition C.7.1), 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 $t \in [0, T]$,

$$\mathrm{E}|\tilde{X}(t \wedge T^m)H(t \wedge T^m)|^2 \overset{(4.6.56)}{\leq} m^2 \mathrm{E}\left| \sup_{t\in[0,T]} \tilde{X}^2(t)\right| \overset{(4.6.33)}{<} \infty, \tag{4.6.57}$$

showing that $\tilde{X}H$ is locally square-integrable. Thus, from this and (4.6.55), $\tilde{X}H$ is a locally square-integrable martingale. As any martingale is also a local martingale, we have $\tilde{X}H \in \mathcal{M}^2_{\mathrm{loc}}(\{\mathcal{F}_t\}, \mathbb{P})$. □

From Lemma 4.6.12, $\tilde{X}H \in \mathcal{M}^2_{\mathrm{loc}}(\{\mathcal{F}_t\}, \mathbb{P})$ so we can apply the martingale representation theorem for locally square-integrable martingales, Theorem B.4.22, to $\tilde{X}H$ to find processes

$$\mathbf{\Lambda}^{\tilde{X}H} = \left(\Lambda_1^{\tilde{X}H}, \ldots, \Lambda_N^{\tilde{X}H}\right)^\top \in L^2_{\mathrm{loc}}(\mathbf{W}) \quad \text{and} \quad \mathbf{\Gamma}^{\tilde{X}H} = \left(\Gamma_{ij}^{\tilde{X}H}\right)_{i,j=1}^{D} \in L^2_{\mathrm{loc}}(\mathcal{Q}) \tag{4.6.58}$$

such that $\tilde{X}H$ has the representation for all $t \in [0, T]$,

$$\tilde{X}(t)H(t) = \tilde{X}(0)H(0) + \sum_{n=1}^{N} \int_0^t \Lambda_n^{\tilde{X}H}(\tau)\, \mathrm{d}W_n(\tau) + \sum_{i,j=1}^{D} \int_0^t \Gamma_{ij}^{\tilde{X}H}(\tau)\, \mathrm{d}\mathcal{Q}_{ij}(\tau) \quad \text{a.s.} \tag{4.6.59}$$

Remark 4.6.13. We use (4.6.59) to express the candidate solution $\tilde{X}$ as a stochastic integral equation by applying Itô's Formula. Ultimately we want to show that the candidate solution $\tilde{X}$ satisfies the wealth equation (3.2.31) for some portfolio process. For this reason, we compare the resulting equation to the wealth equation (3.2.31) in order to ascertain the form of the candidate portfolio process $\tilde{\boldsymbol{\pi}}$.

Begin by defining for all $t \in [0, T]$,

$$\xi(t) := \tilde{X}(t)H(t). \tag{4.6.60}$$

Applying the integration-by-parts formula (Theorem C.14.1) to the product $\tilde{X}(t) = H^{-1}(t)\xi(t)$ and using the continuity of H gives

$$\tilde{X}(t) = \tilde{X}(0) + \int_0^t H^{-1}(\tau)\, \mathrm{d}\xi(\tau) + \int_0^t \xi(\tau_-)\, \mathrm{d}H^{-1}(\tau) + [H^{-1}, \xi](t). \tag{4.6.61}$$

Expanding H^{-1} using Itô's Formula (Theorem C.14.2), we find

$$H^{-1}(t) = 1 - \int_0^t H^{-2}(\tau)\,\mathrm{d}H(\tau) + \int_0^t H^{-3}(\tau)\,\mathrm{d}[H,H](\tau). \tag{4.6.62}$$

From Proposition 4.6.6,

$$\mathrm{d}H(\tau) = -r(\tau)H(\tau)\ \mathrm{d}\tau - H(\tau)\,\boldsymbol{\theta}^\top(\tau)\ \mathrm{d}\mathbf{W}(\tau) \tag{4.6.63}$$

and

$$\mathrm{d}[H,H](\tau) = H^2(\tau)\|\boldsymbol{\theta}(\tau)\|^2\,\mathrm{d}\tau. \tag{4.6.64}$$

Substituting (4.6.63) and (4.6.64) into (4.6.62) gives (after some algebra),

$$H^{-1}(t) = 1 + \int_0^t H^{-1}(\tau)\left(r(\tau) + \|\boldsymbol{\theta}(\tau)\|^2\right)\ \mathrm{d}\tau + \int_0^t H^{-1}(\tau)\boldsymbol{\theta}^\top(\tau)\,\mathrm{d}\mathbf{W}(\tau). \tag{4.6.65}$$

Substituting into (4.6.61) using the above equation, as well as (4.6.59) and (4.6.60), and using the continuity of H, we find that

$$\begin{aligned}
\tilde{X}(t) &= \tilde{X}(0) + \int_0^t H^{-1}(\tau)\left(\left(\Lambda^{\tilde{X}H}\right)^\top(\tau)\ \mathrm{d}\mathbf{W}(\tau) + \sum_{i,j=1}^{D} \Gamma_{ij}^{\tilde{X}H}(\tau)\ \mathrm{d}\mathcal{Q}_{ij}(\tau)\right) \\
&+ \int_0^t \tilde{X}(\tau_-)H(\tau)\left(H^{-1}(\tau)\left(r(\tau) + \|\boldsymbol{\theta}(\tau)\|^2\right)\ \mathrm{d}\tau + H^{-1}(\tau)\boldsymbol{\theta}^\top(\tau)\,\mathrm{d}\mathbf{W}(\tau)\right) \\
&+ \Bigg[1 + \int_0^{\cdot} H^{-1}(\tau)\left(r(\tau) + \|\boldsymbol{\theta}(\tau)\|^2\right)\ \mathrm{d}\tau + \int_0^{\cdot} H^{-1}(\tau)\boldsymbol{\theta}^\top(\tau)\,\mathrm{d}\mathbf{W}(\tau), \\
&\qquad \tilde{X}(0) + \int_0^{\cdot} \left(\Lambda^{\tilde{X}H}\right)^\top(\tau)\ \mathrm{d}\mathbf{W}(\tau) + \sum_{i,j=1}^{D}\int_0^{\cdot} \Gamma_{ij}^{\tilde{X}H}(\tau)\ \mathrm{d}\mathcal{Q}_{ij}(\tau)\Bigg](t) \\
&= \tilde{X}(0) + \int_0^t H^{-1}(\tau)\left(\Lambda^{\tilde{X}H}\right)^\top(\tau)\ \mathrm{d}\mathbf{W}(\tau) + \sum_{i,j=1}^{D}\int_0^t H^{-1}(\tau)\Gamma_{ij}^{\tilde{X}H}(\tau)\ \mathrm{d}\mathcal{Q}_{ij}(\tau) \\
&+ \int_0^t \tilde{X}(\tau_-)r(\tau)\,\mathrm{d}\tau + \int_0^t \tilde{X}(\tau_-)\|\boldsymbol{\theta}(\tau)\|^2\,\mathrm{d}\tau + \int_0^t \tilde{X}(\tau_-)\boldsymbol{\theta}^\top(\tau)\,\mathrm{d}\mathbf{W}(\tau) \\
&+ \int_0^t H^{-1}(\tau)\left(\Lambda^{\tilde{X}H}\right)^\top(\tau)\boldsymbol{\theta}(\tau)\ \mathrm{d}\tau \\
&+ \sum_{i,j=1}^{D}\sum_{n=1}^{N}\int_0^t H^{-1}(\tau)\theta_n(\tau)\Gamma_{ij}^{\tilde{X}H}(\tau)\ \mathrm{d}[W_n, \mathcal{Q}_{ij}](\tau).
\end{aligned} \tag{4.6.66}$$

Applying Lemma B.3.18, the term one the last line above vanishes, leaving

$$\begin{aligned}\tilde{X}(t) &= \tilde{X}(0) \\ &\quad + \int_0^t \left(r(\tau)\tilde{X}(\tau_-) + \left(\frac{1}{H(\tau)} \left(\mathbf{\Lambda}^{\tilde{X}H}\right)^\top (\tau) + \tilde{X}(\tau_-)\boldsymbol{\theta}^\top(\tau) \right) \boldsymbol{\theta}(\tau) \right) \, \mathrm{d}\tau \\ &\quad + \int_0^t \left(\frac{1}{H(\tau)} \left(\mathbf{\Lambda}^{\tilde{X}H}\right)^\top (\tau) + \tilde{X}(\tau_-)\boldsymbol{\theta}^\top(\tau) \right) \, \mathrm{d}\mathbf{W}(\tau) \\ &\quad + \sum_{i,j=1}^{D} \int_0^t \frac{1}{H(\tau)} \Gamma_{ij}^{\tilde{X}H}(\tau) \, \mathrm{d}\mathcal{Q}_{ij}(\tau).\end{aligned} \tag{4.6.67}$$

Motivated by the comparison of (4.6.67) with the wealth equation (3.2.31), define for all $(\omega, t) \in \Omega \times [0, T]$, the *candidate portfolio process*

$$\tilde{\boldsymbol{\pi}}(\omega, t) := (\boldsymbol{\sigma}^\top(\omega, t))^{-1} \left(\frac{\mathbf{\Lambda}^{\tilde{X}H}(\omega, t)}{H(\omega, t)} + \tilde{X}(\omega, t_-)\boldsymbol{\theta}(\omega, t) \right). \tag{4.6.68}$$

Remark 4.6.14. We show that the candidate portfolio process $\tilde{\boldsymbol{\pi}} \in L^2(\mathbf{W})$ and from this it will follow that the candidate solution $\tilde{X} \in \mathbb{B}$. This is the subject of the remainder of the present section.

Showing that $\tilde{X} \in \mathbb{A}$, $\tilde{X} = X^{\tilde{\boldsymbol{\pi}}}$ is the solution to the wealth equation (3.2.31) for the candidate portfolio process $\tilde{\boldsymbol{\pi}}$ and that the candidate solution $\tilde{X}$ is the solution to the primal problem is done in Section 4.7.

Substituting the candidate portfolio process $\tilde{\boldsymbol{\pi}}$ into (4.6.67) gives

$$\begin{aligned}\tilde{X}(t) = \tilde{X}(0) &+ \int_0^t \left(r(\tau)\tilde{X}(\tau_-) + \tilde{\boldsymbol{\pi}}^\top(\tau)\boldsymbol{\sigma}(\tau)\boldsymbol{\theta}(\tau) \right) \, \mathrm{d}\tau + \int_0^t \tilde{\boldsymbol{\pi}}^\top(\tau)\boldsymbol{\sigma}(\tau) \, \mathrm{d}\mathbf{W}(\tau) \\ &+ \sum_{i,j=1}^{D} \int_0^t \frac{1}{H(\tau)} \Gamma_{ij}^{\tilde{X}H}(\tau) \, \mathrm{d}\mathcal{Q}_{ij}(\tau).\end{aligned} \tag{4.6.69}$$

We begin by showing that $\tilde{\boldsymbol{\pi}}$ is a portfolio process.

Lemma 4.6.15. $\tilde{\boldsymbol{\pi}} = \{\tilde{\boldsymbol{\pi}}(t) : t \in [0, T]\}$ *is a portfolio process, that is* $\tilde{\boldsymbol{\pi}}$ *is a previsible process satisfying*

$$\int_0^T \|\tilde{\boldsymbol{\pi}}(t)\|^2 \, \mathrm{d}t < \infty \quad a.s. \tag{4.6.70}$$

Proof. The previsibility of $\tilde{\boldsymbol{\pi}}$ is immediate from (4.6.68). It remains to show that $\int_0^T \|\tilde{\boldsymbol{\pi}}(t)\|^2 \, \mathrm{d}t < \infty$ a.s. Recall the constants $\kappa_{\boldsymbol{\sigma}} \in (0, \infty)$ and $\kappa_{\boldsymbol{\theta}} \in (0, \infty)$ from

(3.2.26) and (3.2.28), respectively. Then

$$\begin{aligned}\int_0^T \|\tilde{\boldsymbol{\pi}}(t)\|^2\, \mathrm{d}t &\overset{(4.6.68)}{=} \int_0^T \left\|(\boldsymbol{\sigma}^\top(t))^{-1}\left(\frac{\boldsymbol{\Lambda}^{\tilde{X}H}(t)}{H(t)} + \tilde{X}(t_-)\boldsymbol{\theta}(t)\right)\right\|^2 \mathrm{d}t \\ &\le 2\kappa_{\boldsymbol{\sigma}}^2 \int_0^T \left\|\frac{\boldsymbol{\Lambda}^{\tilde{X}H}(t)}{H(t)}\right\|^2 \mathrm{d}t + 2\kappa_{\boldsymbol{\sigma}}^2 \int_0^T \left\|\tilde{X}(t_-)\boldsymbol{\theta}(t)\right\|^2 \mathrm{d}t \\ &\le 2\kappa_{\boldsymbol{\sigma}}^2 \sup_{t\in[0,T]}\left\{H^{-2}(t)\right\} \int_0^T \left\|\boldsymbol{\Lambda}^{\tilde{X}H}(t)\right\|^2 \mathrm{d}t + 2\kappa_{\boldsymbol{\sigma}}^2\,\kappa_{\boldsymbol{\theta}}^2\, T \sup_{t\in[0,T]}\left|\tilde{X}(t)\right|^2.\end{aligned} \tag{4.6.71}$$

We show that the last line of the above inequality is finite.

Since H^{-2} is pathwise a continuous function on the 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, $\sup_{t\in[0,T]}\{H^{-2}(t)\}$ is finite a.s. We also have from (4.6.58) that $\boldsymbol{\Lambda}^{\tilde{X}H} \in L^2_{\mathrm{loc}}(\mathbf{W})$, 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 $\boldsymbol{\Lambda}^{\tilde{X}H}[0,S^m] \in L^2(\mathbf{W})$ for all $m \in \mathbb{N}$. Then for each $m \in \mathbb{N}$,

$$\mathrm{E}\int_0^{T\wedge S^m} \|\boldsymbol{\Lambda}^{\tilde{X}H}(t)\|^2\, \mathrm{d}t < \infty \quad \Rightarrow \quad \int_0^{T\wedge S^m} \|\boldsymbol{\Lambda}^{\tilde{X}H}(t)\|^2\, \mathrm{d}t < \infty \text{ a.s.} \tag{4.6.72}$$

By definition of $S^m \Uparrow T$ a.s. (recall Definition C.7.1) there exists $M(\omega) \in \mathbb{N}$ such that $S^m(\omega) = T$ for all $m \geq M(\omega)$ for all $\omega \in \Omega$. Then, letting $m \to \infty$ in (4.6.72), we obtain $\int_0^T \|\boldsymbol{\Lambda}^{\tilde{X}H}(t)\|^2\, \mathrm{d}t < \infty$ a.s. Hence, on the last line of (4.6.71), the first term is finite.

Proposition 4.6.10 implies that $\sup_{t\in[0,T]}\left|\tilde{X}(t)\right|^2 < \infty$ a.s. Hence, on the last line of (4.6.71), the second term is finite.

We conclude that $\int_0^T \|\tilde{\boldsymbol{\pi}}(t)\|^2\, \mathrm{d}t < \infty$ a.s. and so $\tilde{\boldsymbol{\pi}}$ is a portfolio process. □

Proposition 4.6.16.

$$\boldsymbol{\sigma}^\top\tilde{\boldsymbol{\pi}} \in L^2(\mathbf{W}) \quad \textit{and} \quad \frac{1}{H}\boldsymbol{\Gamma}^{\tilde{X}H} \in L^2(\mathcal{Q}). \tag{4.6.73}$$

Proof. For each $m \in \mathbb{N}$, let

$$R^m := \inf\left\{t > 0 \,:\, \int_0^t \|\tilde{\boldsymbol{\pi}}(\tau)\|^2\, \mathrm{d}\tau > m\right\} \wedge T. \tag{4.6.74}$$

Then R^m is an $\{\mathcal{F}_t\}$–stopping time (by Proposition C.5.1) and $R^m \Uparrow T$ a.s., since $\int_0^T \|\tilde{\boldsymbol{\pi}}(\tau)\|^2\, \mathrm{d}\tau < \infty$ a.s. by Lemma 4.6.15.

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

$$S^m := \inf\left\{t > 0 \,:\, |\tilde{X}(t_-)|^2 > m\right\} \wedge T. \tag{4.6.75}$$

Then S^m is an $\{\mathcal{F}_t\}$–stopping time since $\tilde{X}(t_-)$ is locally bounded and $S^m \Uparrow T$ a.s., since $\sup_{t\in[0,T]} \left|\tilde{X}(t)\right|^2 < \infty$ a.s. by Proposition 4.6.10.

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

$$T^m := \inf\left\{t > 0 \,:\, |H^{-1}(t)|^2 > m\right\} \wedge T. \tag{4.6.76}$$

Then T^m is an $\{\mathcal{F}_t\}$–stopping time (by Proposition C.5.1) and $T^m \Uparrow T$ a.s., since $\sup_{t\in[0,T]} \{H^{-2}(t)\}$ is finite a.s. by the pathwise continuity of H^{-2} on the compact interval $[0, T]$.

Since $\mathbf{\Gamma}^{\tilde{X}H} \in L^2_{\text{loc}}(\mathcal{Q})$ (by (4.6.58)), 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 $\mathbf{\Gamma}^{\tilde{X}H}[0, U^m] \in L^2(\mathcal{Q})$ for all $m \in \mathbb{N}$. Then for all $m \in \mathbb{N}$,

$$\mathrm{E} \sum_{i,j=1}^{D} \int_0^{T\wedge U^m} |\Gamma^{\tilde{X}H}_{ij}(t)|^2 \,\mathrm{d}[\mathcal{Q}_{ij}](t) < \infty. \tag{4.6.77}$$

Finally, define

$$V^m := R^m \wedge S^m \wedge T^m \wedge U^m. \tag{4.6.78}$$

Then V^m is an $\{\mathcal{F}_t\}$–stopping time and $V^m \Uparrow T$ a.s.

Applying the integration-by-parts formula (Theorem C.14.1) to (4.6.69), we can expand the mapping $t \mapsto \tilde{X}^2(t)$. Evaluating the expansion at time $t \wedge V^m$, gives for all $t \in [0, T]$,

$$\begin{aligned}
\tilde{X}^2(t\wedge V^m) &= \tilde{X}^2(0) + 2\int_0^{t\wedge V^m} \tilde{X}(\tau_-)\,\mathrm{d}\tilde{X}(\tau) + [\tilde{X}, \tilde{X}](t\wedge V^m) \\
&= \tilde{X}^2(0) + 2\int_0^{t\wedge V^m} \tilde{X}(\tau_-)\left(r(\tau)\tilde{X}(\tau_-) + \tilde{\boldsymbol{\pi}}^\top(\tau)\boldsymbol{\sigma}(\tau)\boldsymbol{\theta}(\tau)\right)\,\mathrm{d}\tau \\
&+ 2\int_0^{t\wedge V^m} \tilde{X}(\tau_-)\tilde{\boldsymbol{\pi}}^\top(\tau)\boldsymbol{\sigma}(\tau)\,\mathrm{d}\mathbf{W}(\tau) + 2\sum_{i,j=1}^{D}\int_0^{t\wedge V^m} \frac{\tilde{X}(\tau_-)}{H(\tau)}\,\Gamma^{\tilde{X}H}_{ij}(\tau)\,\mathrm{d}\mathcal{Q}_{ij}(\tau) \\
&+ \int_0^{t\wedge V^m} \|\boldsymbol{\sigma}^\top(\tau)\tilde{\boldsymbol{\pi}}(\tau)\|^2\,\mathrm{d}\tau + \sum_{i,j=1}^{D}\int_0^{t\wedge V^m} \frac{1}{H^2(\tau)}|\Gamma^{\tilde{X}H}_{ij}(\tau)|^2\,\mathrm{d}[\mathcal{Q}_{ij}](\tau).
\end{aligned} \tag{4.6.79}$$

We show that the last third and fourth-to-last terms above are square-integrable martingales. Recall the constant $\kappa_{\boldsymbol{\sigma}} \in (0, \infty)$ satisfying (3.2.26). For the fourth-to-last term, we have

$$
\begin{aligned}
&\mathrm{E}\int_0^{T\wedge V^m} \|\tilde{X}(\tau_-)\boldsymbol{\sigma}^\top(\tau)\tilde{\boldsymbol{\pi}}(\tau)\|^2 \,\mathrm{d}\tau \\
&\leq \kappa_{\boldsymbol{\sigma}}^2\,\mathrm{E}\left(\sup_{t\in[0,T]}|\tilde{X}(t)|^2 \int_0^{T\wedge V^m} \|\tilde{\boldsymbol{\pi}}(\tau)\|^2 \,\mathrm{d}\tau\right) \\
&\overset{(4.6.74)}{\leq} \kappa_{\boldsymbol{\sigma}}^2\, m\,\mathrm{E}\left(\sup_{t\in[0,T]}|\tilde{X}(t)|^2\right) \\
&\overset{\text{Propn 4.6.10}}{<} \infty.
\end{aligned}
\tag{4.6.80}
$$

Then for each $m \in \mathbb{N}$, the fourth-to-last term in (4.6.79) is a square-integrable martingale, which is clearly null at the origin, so that for all $t \in [0, T]$,

$$
\mathrm{E}\int_0^{t\wedge V^m} \tilde{X}(\tau_-)\boldsymbol{\sigma}^\top(\tau)\tilde{\boldsymbol{\pi}}(\tau)\,\mathrm{d}\mathbf{W}(\tau) = 0. \tag{4.6.81}
$$

For the third-to-last term in (4.6.79), we have

$$
\begin{aligned}
&\mathrm{E}\sum_{i,j=1}^{D}\int_0^{T\wedge V^m} \left|\frac{\tilde{X}(\tau_-)}{H(\tau)}\,\Gamma_{ij}^{\tilde{X}H}(\tau)\right|^2 \mathrm{d}[\mathcal{Q}_{ij}](\tau) \\
&= \mathrm{E}\sum_{i,j=1}^{D}\int_0^{T\wedge V^m} |H^{-2}(\tau)|\,|\tilde{X}(\tau_-)|^2\,|\Gamma_{ij}^{\tilde{X}H}(\tau)|^2 \,\mathrm{d}[\mathcal{Q}_{ij}](\tau) \\
&\overset{(4.6.75),(4.6.76)}{\leq} m^2\,\mathrm{E}\sum_{i,j=1}^{D}\int_0^{T\wedge V^m} |\Gamma_{ij}^{\tilde{X}H}(\tau)|^2 \,\mathrm{d}[\mathcal{Q}_{ij}](\tau) \\
&\overset{(4.6.77)}{<} \infty.
\end{aligned}
\tag{4.6.82}
$$

Then for each $m \in \mathbb{N}$, the third-to-last term in (4.6.79) is a square-integrable martingale, which is clearly null at the origin, so that for all $t \in [0, T]$,

$$
\sum_{i,j=1}^{D}\mathrm{E}\int_0^{t\wedge V^m} \frac{\tilde{X}(\tau_-)}{H(\tau)}\,\Gamma_{ij}^{\tilde{X}H}(\tau)\,\mathrm{d}\mathcal{Q}_{ij}(\tau) = 0 \tag{4.6.83}
$$

Hence the the sum of the third and fourth-to-last terms of (4.6.79) is a martingale, null at the origin, and thus has zero expectation for all $t \in [0, T]$. Evaluating (4.6.79) at time $t = T$, noting that $T \wedge V^m = V^m$, and taking expectations, we get

$$
\begin{aligned}
\mathrm{E}\tilde{X}^2(V^m) &= \mathrm{E}\tilde{X}^2(0) + \mathrm{E}\int_0^{V^m} 2\tilde{X}(\tau_-)\left(r(\tau)\tilde{X}(\tau_-) + \tilde{\boldsymbol{\pi}}^\top(\tau)\boldsymbol{\sigma}(\tau)\boldsymbol{\theta}(\tau)\right)\,\mathrm{d}\tau \\
&\quad + \mathrm{E}\int_0^{V^m} \|\boldsymbol{\sigma}^\top(\tau)\tilde{\boldsymbol{\pi}}(\tau)\|^2 \,\mathrm{d}\tau + \sum_{i,j=1}^{D}\mathrm{E}\int_0^{V^m} \frac{1}{H^2(\tau)}|\Gamma_{ij}^{\tilde{X}H}(\tau)|^2\,\mathrm{d}[\mathcal{Q}_{ij}](\tau).
\end{aligned}
\tag{4.6.84}
$$

From the nonnegativity of the risk-free interest rate process $\{r(t)\}$ and $\tilde{X}^2_-$, upon rearranging (4.6.84) we obtain the inequality

$$\begin{aligned} &\mathrm{E}\tilde{X}^2(V^m) - \mathrm{E}\int_0^{V^m} 2\tilde{X}(\tau_-)\tilde{\boldsymbol{\pi}}^\top(\tau)\boldsymbol{\sigma}(\tau)\boldsymbol{\theta}(\tau)\ \mathrm{d}\tau \\ &\geq \mathrm{E}\int_0^{V^m} \|\boldsymbol{\sigma}^\top(\tau)\tilde{\boldsymbol{\pi}}(\tau)\|^2\ \mathrm{d}\tau + \sum_{i,j=1}^{D} \mathrm{E}\int_0^{V^m} \frac{1}{H^2(\tau)}|\Gamma_{ij}^{\tilde{X}H}(\tau)|^2\,\mathrm{d}[\mathcal{Q}_{ij}](\tau). \end{aligned} \tag{4.6.85}$$

Now, for arbitrary $\boldsymbol{\nu}_1, \boldsymbol{\nu}_2 \in \mathbb{R}^N$, we have the inequality

$$\boldsymbol{\nu}_1^\top\boldsymbol{\nu}_2 \leq \frac{1}{2}\|\boldsymbol{\nu}_1\|^2 + \frac{1}{2}\|\boldsymbol{\nu}_2\|^2. \tag{4.6.86}$$

Setting $\boldsymbol{\nu}_1 = \boldsymbol{\sigma}^\top(\tau)\tilde{\boldsymbol{\pi}}(\tau)$ and $\boldsymbol{\nu}_2 = -2\tilde{X}(\tau_-)\boldsymbol{\theta}(\tau)$, we then get

$$-2\tilde{X}(\tau_-)\boldsymbol{\sigma}^\top(\tau)\tilde{\boldsymbol{\pi}}(\tau)\boldsymbol{\theta}(\tau) \leq \frac{1}{2}\|\boldsymbol{\sigma}^\top(\tau)\tilde{\boldsymbol{\pi}}(\tau)\|^2 + 2\tilde{X}^2(\tau_-)\,\|\boldsymbol{\theta}(\tau)\|^2. \tag{4.6.87}$$

Integrating, taking expectations and adding $\mathrm{E}\tilde{X}^2(V^m)$ to each side of the above inequality, we get

$$\begin{aligned} &\mathrm{E}\tilde{X}^2(V^m) - \mathrm{E}\int_0^{V^m} 2\tilde{X}(\tau_-)\boldsymbol{\sigma}^\top(\tau)\tilde{\boldsymbol{\pi}}(\tau)\boldsymbol{\theta}(\tau)\ \mathrm{d}\tau \\ &\leq \mathrm{E}\tilde{X}^2(V^m) + \frac{1}{2}\mathrm{E}\int_0^{V^m} \|\boldsymbol{\sigma}^\top(\tau)\tilde{\boldsymbol{\pi}}(\tau)\|^2\ \mathrm{d}\tau + 2\mathrm{E}\int_0^{V^m} \tilde{X}^2(\tau_-)\,\|\boldsymbol{\theta}(\tau)\|^2\ \mathrm{d}\tau. \end{aligned} \tag{4.6.88}$$

Combining (4.6.85) and (4.6.88), we get

$$\begin{aligned} &\mathrm{E}\int_0^{V^m} \|\boldsymbol{\sigma}^\top(\tau)\tilde{\boldsymbol{\pi}}(\tau)\|^2\ \mathrm{d}\tau + \sum_{i,j=1}^{D} \mathrm{E}\int_0^{V^m} \frac{1}{H^2(\tau)}|\Gamma_{ij}^{\tilde{X}H}(\tau)|^2\,\mathrm{d}[\mathcal{Q}_{ij}](\tau) \\ &\leq \mathrm{E}\tilde{X}^2(V^m) + \frac{1}{2}\mathrm{E}\int_0^{V^m} \|\boldsymbol{\sigma}^\top(\tau)\tilde{\boldsymbol{\pi}}(\tau)\|^2\ \mathrm{d}\tau + 2\mathrm{E}\int_0^{V^m} \tilde{X}^2(\tau_-)\,\|\boldsymbol{\theta}(\tau)\|^2\ \mathrm{d}\tau. \end{aligned} \tag{4.6.89}$$

Recall the constant $\kappa_{\boldsymbol{\theta}} \in (0, \infty)$ satisfying (3.2.28). Rearranging (4.6.89),

$$\begin{aligned} &\frac{1}{2}\mathrm{E}\int_0^{V^m} \|\boldsymbol{\sigma}^\top(\tau)\tilde{\boldsymbol{\pi}}(\tau)\|^2\ \mathrm{d}\tau + \sum_{i,j=1}^{D} \mathrm{E}\int_0^{V^m} \frac{1}{H^2(\tau)}|\Gamma_{ij}^{\tilde{X}H}(\tau)|^2\,\mathrm{d}[\mathcal{Q}_{ij}](\tau) \\ &\leq \left(1 + 2T\kappa_{\boldsymbol{\theta}}^2\right)\mathrm{E}\left(\sup_{t\in[0,T]} |\tilde{X}(t)|^2\right) \\ &\overset{\text{Propn 4.6.10}}{<} \infty. \end{aligned} \tag{4.6.90}$$

Since $V^m \Uparrow T$ a.s., upon letting $m \to \infty$ in (4.6.90), we obtain

$$\boldsymbol{\sigma}^\top\tilde{\boldsymbol{\pi}} \in L^2(\mathbf{W}) \quad \text{and} \quad \frac{1}{H}\boldsymbol{\Gamma}^{\tilde{X}H} \in L^2(\mathcal{Q}). \tag{4.6.91}$$

□

Corollary 4.6.17. $\tilde{\boldsymbol{\pi}} \in L^2(\mathbf{W})$.

Proof. This follows immediately from Proposition 4.6.16 and the uniform boundedness of $\boldsymbol{\sigma}$. □

Corollary 4.6.18. *Recalling (3.2.38), we have*

$$\tilde{X} \equiv (\tilde{X}(0), r\tilde{X}_- + \tilde{\boldsymbol{\pi}}^\top \boldsymbol{\sigma\theta}, \boldsymbol{\sigma}^\top \tilde{\boldsymbol{\pi}}, \frac{\boldsymbol{\Gamma}^{\tilde{X}H}}{H}) \in \mathbb{B}, \tag{4.6.92}$$

that is

$$\begin{aligned} &\tilde{X}_0 = \tilde{X}(0) \in \mathbb{R}, \quad \dot{\tilde{X}} = r\tilde{X}_- + \tilde{\boldsymbol{\pi}}^\top \boldsymbol{\sigma\theta} \in L_{21}, \quad \boldsymbol{\Lambda}^{\tilde{X}} = \boldsymbol{\sigma}^\top \tilde{\boldsymbol{\pi}} \in L^2(\mathbf{W}) \\ &\textit{and} \quad \boldsymbol{\Gamma}^{\tilde{X}} = \frac{\boldsymbol{\Gamma}^{\tilde{X}H}}{H} \in L^2(\mathcal{Q}). \end{aligned} \tag{4.6.93}$$

Proof. From Proposition 4.6.16, we have $\boldsymbol{\sigma}^\top \tilde{\boldsymbol{\pi}} \in L^2(\mathbf{W})$ and $\frac{1}{H}\boldsymbol{\Gamma}^{\tilde{X}H} \in L^2(\mathcal{Q})$. So we need only show that $r\tilde{X}_- + \tilde{\boldsymbol{\pi}}^\top \boldsymbol{\sigma\theta} \in L_{21}$. However, from Proposition 4.6.10, we have $\mathrm{E}\left(\sup_{t\in[0,T]} |\tilde{X}(t)|^2\right) < \infty$ and from Corollary 4.6.17, $\mathrm{E}\int_0^T \|\tilde{\boldsymbol{\pi}}(t)\|^2 \, \mathrm{d}t < \infty$. Then from these two facts and the uniform boundedness of r, $\boldsymbol{\sigma}$ and $\boldsymbol{\theta}$, we have $r\tilde{X}_- + \tilde{\boldsymbol{\pi}}^\top \boldsymbol{\sigma\theta} \in L_{21}$. □

Remark 4.6.19. We prove in Section 4.7 that $\boldsymbol{\Gamma}^{\tilde{X}H} = 0$ $\nu_{[\mathcal{Q}]}$-a.e. This implies that the candidate solution $\tilde{X}$ is a continuous process and will allow us to drop the dangling minus appended to the candidate solution $\tilde{X}$ in (4.6.92).

4.7 Check if the candidate solution solves the primal problem.

Having explored some of the properties of the candidate solution $\tilde{X}$, the next step is to verify if the candidate solution $\tilde{X}$ solves the primal problem (Problem 4.1.7). We show that for the candidate solution $\tilde{X}$, given by (4.6.32), we have $\tilde{X} \in \mathbb{A}$. Furthermore, we prove that the candidate solution $\tilde{X}$ and the solution to the dual problem $\bar{Y}$, which we know exists, satisfy (4.4.16) - (4.4.21) of Proposition 4.4.8. From this, we conclude that $\tilde{X} = X^{\tilde{\boldsymbol{\pi}}}$ for the candidate portfolio process $\tilde{\boldsymbol{\pi}}$.

We first examine the functional $\tilde{\Psi}$ given by Definition 4.5.4. Proposition 4.5.15 tells us that there exists $\bar{Y} \equiv \Xi(\bar{y}, \bar{\boldsymbol{\lambda}}, \bar{\boldsymbol{\gamma}}) \in \mathbb{B}_1$ satisfying

$$\tilde{\Psi}(\bar{y}, \bar{\boldsymbol{\lambda}}, \bar{\boldsymbol{\gamma}}) = \inf_{(y,\boldsymbol{\lambda},\boldsymbol{\gamma})\in\mathbb{R}\times L^2(\mathbf{W})\times L^2(\mathcal{Q})} \{\tilde{\Psi}(y, \boldsymbol{\lambda}, \boldsymbol{\gamma})\}. \tag{4.7.1}$$

It follows that

$$\tilde{\Psi}(\bar{y}, \bar{\boldsymbol{\lambda}}, \bar{\boldsymbol{\gamma}}) \leq \tilde{\Psi}(y, \boldsymbol{\lambda}, \boldsymbol{\gamma}), \quad \forall (y, \boldsymbol{\lambda}, \boldsymbol{\gamma}) \in \mathbb{R} \times L^2(\mathbf{W}) \times L^2(\mathcal{Q}). \tag{4.7.2}$$

We perform a variational analysis on (4.7.2). By doing so, we hope to find out some more information about the optimality of $\bar{Y}$. We do the variational analysis by perturbing each component $y, \boldsymbol{\lambda}$ and $\boldsymbol{\gamma}$ in turn, while holding the other two components fixed.

For $(y, \boldsymbol{\lambda}, \boldsymbol{\gamma}) \in \mathbb{R} \times L^2(\mathbf{W}) \times L^2(\mathcal{Q})$ and recalling (4.5.8), set

$$R := \Xi(y, \boldsymbol{\lambda}, \boldsymbol{\gamma}). \tag{4.7.3}$$

Define for all $\epsilon \in (0, \infty)$, a triple $(y^\epsilon, \boldsymbol{\lambda}^\epsilon, \boldsymbol{\gamma}^\epsilon) \in \mathbb{R} \times L^2(\mathbf{W}) \times L^2(\mathcal{Q})$ by

$$\begin{cases} y^\epsilon & := \bar{y} + \epsilon y \in \mathbb{R} \\ \boldsymbol{\lambda}^\epsilon & := \bar{\boldsymbol{\lambda}} + \epsilon \boldsymbol{\lambda} \in L^2(\mathbf{W}) \\ \boldsymbol{\gamma}^\epsilon & := \bar{\boldsymbol{\gamma}} + \epsilon \boldsymbol{\gamma} \in L^2(\mathcal{Q}), \end{cases} \tag{4.7.4}$$

and let

$$Y^\epsilon := \Xi(y^\epsilon, \boldsymbol{\lambda}^\epsilon, \boldsymbol{\gamma}^\epsilon). \tag{4.7.5}$$

Then, by the linearity and homogeneity of the map Ξ (shown by Lemma 4.5.3), we have

$$Y^\epsilon = \bar{Y} + \epsilon R \tag{4.7.6}$$

and by the linearity of $\Theta_\cdot(t)$ we have

$$\Theta_{Y^\epsilon}(t) \overset{(4.4.6)}{=} \Theta_{\bar{Y}}(t) + \epsilon \Theta_R(t). \tag{4.7.7}$$

By the minimality of the dual solution $\bar{Y}$ as shown by (4.7.1), we obtain from (4.7.2),

$$0 \leq \frac{\tilde{\Psi}(y^\epsilon, \boldsymbol{\lambda}^\epsilon, \boldsymbol{\gamma}^\epsilon) - \tilde{\Psi}(\bar{y}, \bar{\boldsymbol{\lambda}}, \bar{\boldsymbol{\gamma}})}{\epsilon}. \tag{4.7.8}$$

Using (4.5.35) to substitute for the functional $\tilde{\Psi}$, we get

$$\begin{aligned} 0 \leq & \frac{x_0(y^\epsilon - \bar{y})}{\epsilon} + \frac{\mathrm{E} \int_0^T \left(\delta\left(\Theta_{Y^\epsilon}(t)\right) - \delta\left(\Theta_{\bar{Y}}(t)\right)\right) \, \mathrm{d}t}{\epsilon} \\ & + \mathrm{E}\left(\frac{1}{\epsilon}\left(\frac{\left(Y^\epsilon(T) + b\right)^2}{2a} - \frac{\left(\bar{Y}(T) + b\right)^2}{2a}\right)\right). \end{aligned} \tag{4.7.9}$$

Substituting from (4.7.4), (4.7.6) and (4.7.7), letting $\epsilon \downarrow 0$ and some straightforward algebra gives,

$$\begin{aligned} 0 \leq x_0 y + & \lim_{\epsilon \downarrow 0} \left\{ \mathrm{E} \int_0^T \left(\frac{\delta\left(\Theta_{\bar{Y}}(t) + \epsilon \Theta_R(t)\right) - \delta\left(\Theta_{\bar{Y}}(t)\right)}{\epsilon} \right) \mathrm{d}t \right\} \\ & + \mathrm{E}\left(\frac{\left(\bar{Y}(T) + b\right) R(T)}{a} \right). \end{aligned} \tag{4.7.10}$$

Evaluating the candidate solution $\tilde{X}$ given by (4.6.32) at $t = T$, we obtain a.s.,

$$\tilde{X}(T) = -\left(\frac{\bar{Y}(T) + b}{a}\right). \tag{4.7.11}$$

Substituting this into the last term of (4.7.10), we get

$$0 \leq x_0 y + \lim_{\epsilon \downarrow 0} \left\{ \mathrm{E} \int_0^T \left(\frac{\delta\left(\Theta_{\bar{Y}}(t) + \epsilon \Theta_R(t)\right) - \delta\left(\Theta_{\bar{Y}}(t)\right)}{\epsilon} \right) \mathrm{d}t \right\} - \mathrm{E}\left(\tilde{X}(T) R(T) \right). \tag{4.7.12}$$

We seek to expand the last term on the right-hand side of (4.7.12). From (4.7.3) and Lemma 4.5.2, we find

$$R \equiv (y, -rR_-, \boldsymbol{\lambda}, \boldsymbol{\gamma}) \in \mathbb{B}_1. \tag{4.7.13}$$

From (4.4.6) applied to R, so that in particular $\boldsymbol{\Lambda}^R(t) := \boldsymbol{\lambda}(t)$, we get for all $t \in [0, T]$,

$$\boldsymbol{\Theta}_R(t) = -\boldsymbol{\sigma}(t)\left(\boldsymbol{\theta}(t) R(t) + \boldsymbol{\lambda}(t)\right). \tag{4.7.14}$$

Recall from Corollary 4.6.18 that

$$\tilde{X} \equiv (\tilde{X}(0), r\tilde{X}_- + \tilde{\boldsymbol{\pi}}^\top \boldsymbol{\sigma}\boldsymbol{\theta}, \boldsymbol{\sigma}^\top \tilde{\boldsymbol{\pi}}, \frac{1}{H}\boldsymbol{\Gamma}^{\tilde{X}H}) \in \mathbb{B}. \tag{4.7.15}$$

Applying Proposition 4.3.1 to $\tilde{X}$ and R, we find that for

$$\begin{aligned} \mathbb{M}(\tilde{X}, R)(t) :=& \tilde{X}(t) R(t) - y\tilde{X}(0) - \int_0^t \tilde{\boldsymbol{\pi}}^\top(\tau) \boldsymbol{\sigma}(\tau)\left(\boldsymbol{\theta}(\tau) R(\tau) + \boldsymbol{\lambda}(\tau)\right) \mathrm{d}\tau \\ & - \sum_{i,j=1}^{D} \int_0^t \frac{1}{H(\tau)} \Gamma_{ij}^{\tilde{X}H}(\tau) \gamma_{ij}(\tau)\, \mathrm{d}[\mathcal{Q}_{ij}](\tau), \end{aligned} \tag{4.7.16}$$

we have $\mathbb{M}(\tilde{X}, R) \in \mathcal{M}_0(\{\mathcal{F}_t\}, \mathbb{P})$. Taking expectations at $t = T$ in (4.7.16) and using the fact that $\mathrm{E}(\mathbb{M}(\tilde{X}, R)(T)) = 0$, we find

$$\begin{aligned} \mathrm{E}(\tilde{X}(T) R(T)) \overset{(4.7.14)}{=} & y\tilde{X}(0) - \mathrm{E}\int_0^T \tilde{\boldsymbol{\pi}}^\top(t) \Theta_R(t)\, \mathrm{d}t \\ & + \mathrm{E}\sum_{i,j=1}^{D} \int_0^T \frac{1}{H(t)} \Gamma_{ij}^{\tilde{X}H}(t) \gamma_{ij}(t)\, \mathrm{d}[\mathcal{Q}_{ij}](t). \end{aligned} \tag{4.7.17}$$

Substituting (4.7.17) into (4.7.12), we get

$$\begin{aligned} 0 \leq y(x_0 - \tilde{X}(0)) + \lim_{\epsilon \downarrow 0} & \left\{ \mathrm{E}\int_0^T \left(\frac{\delta\left(\Theta_{\bar{Y}}(t) + \epsilon \Theta_R(t)\right) - \delta\left(\Theta_{\bar{Y}}(t)\right)}{\epsilon} \right) \mathrm{d}t \right\} \\ & + \mathrm{E}\int_0^T \tilde{\boldsymbol{\pi}}^\top(t) \Theta_R(t)\, \mathrm{d}t - \mathrm{E}\sum_{i,j=1}^{D} \int_0^T \frac{1}{H(t)} \Gamma_{ij}^{\tilde{X}H}(t) \gamma_{ij}(t)\, \mathrm{d}[\mathcal{Q}_{ij}](t), \end{aligned} \tag{4.7.18}$$

This equation holds for all $(y, \boldsymbol{\lambda}, \boldsymbol{\gamma}) \in \mathbb{R} \times L^2(\mathbf{W}) \times L^2(\mathcal{Q})$, for $R = \Xi(y, \boldsymbol{\lambda}, \boldsymbol{\gamma})$.

In Proposition 4.7.2 below, we show that $\tilde{X} \in \mathbb{A}$. We also show that (4.4.16) and (4.4.17) of Proposition 4.4.8 hold. First, we need the following lemma in the proof of Proposition 4.7.2.

Lemma 4.7.1. *For each* $\boldsymbol{\rho} \in L^2(\mathbf{W})$ *and* $\boldsymbol{\gamma} = (\gamma_{ij})_{i,j=1}^D \in L^2(\mathcal{Q})$, *there exists* $\boldsymbol{\lambda} \in L^2(\mathbf{W})$ *such that for all* $t \in [0,T]$ *we have a.s.*

$$\boldsymbol{\lambda}(t) + \boldsymbol{\theta}(t) \int_0^t \boldsymbol{\lambda}^\top(\tau)\,\mathrm{d}\mathbf{W}(\tau) = \boldsymbol{\rho}(t) - \boldsymbol{\theta}(t) \sum_{i,j=1}^D \int_0^t \gamma_{ij}(\tau)\,\mathrm{d}\mathcal{Q}_{ij}(\tau). \tag{4.7.19}$$

$\boldsymbol{\lambda}$ *is unique in the sense that if there exists* $\bar{\boldsymbol{\lambda}} \in L^2(\mathbf{W})$ *such that (4.7.19) holds with* $\boldsymbol{\lambda}$ *replaced by* $\bar{\boldsymbol{\lambda}}$, *then* $\boldsymbol{\lambda} = \bar{\boldsymbol{\lambda}}$ $(\mathbb{P} \otimes Leb)$*-a.e.*

Proof. See the proof of Lemma A.2.6. □

Proposition 4.7.2. *For the candidate solution* $\tilde{X} \in \mathbb{B}$, *defined by (4.6.32) (recall Corollary 4.6.18) and the solution to the dual problem* $\bar{Y} \in \mathbb{B}$ *(recall Proposition 4.5.15), we have*

$$\tilde{X}(T) = -\left(\frac{\bar{Y}(T) + b}{a}\right). \tag{4.7.20}$$

$$\tilde{X}(0) = x_0. \tag{4.7.21}$$

$$\boldsymbol{\Gamma}^{\tilde{X}H} = 0 \quad \nu_{[\mathcal{Q}]}\text{-a.e. (hence } \tilde{X} \in \mathbb{A}\text{)}. \tag{4.7.22}$$

For the candidate portfolio process $\tilde{\boldsymbol{\pi}}$ *given by (4.6.68), we have* $(\mathbb{P} \otimes Leb)$*-a.e.,*

$$\tilde{\boldsymbol{\pi}}(t) \in K, \tag{4.7.23}$$

$$\dot{\tilde{X}}(t) = r(t)\tilde{X}(t) + \tilde{\boldsymbol{\pi}}^\top(t)\boldsymbol{\sigma}(t)\boldsymbol{\theta}(t) \tag{4.7.24}$$

and

$$\delta(\boldsymbol{\Theta}_{\bar{Y}}(t)) + \tilde{\boldsymbol{\pi}}^\top(t)\boldsymbol{\Theta}_{\bar{Y}}(t) = 0. \tag{4.7.25}$$

Proof. We show that the equations in the proposition hold as a series of claims. We rely heavily on (4.7.18) and Lemma 4.7.1.

Claim 4.7.3. $\tilde{X}(T) = -\left(\frac{\bar{Y}(T)+b}{a}\right)$ a.s.

This follows immediately from (4.6.32) upon setting $t = T$. Hence Claim 4.7.3 is shown and (4.7.20) holds.

Claim 4.7.4. $\tilde{X}(0) = x_0$.

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 4.7.1 to $(\boldsymbol{\rho}, \boldsymbol{\gamma}) := (-y\boldsymbol{\theta}, \mathbf{0}) \in L^2(\mathbf{W}) \times L^2(\mathcal{Q})$, there exists $\boldsymbol{\lambda}_y \in L^2(\mathbf{W})$ such that

$$\boldsymbol{\lambda}_y(t) + \boldsymbol{\theta}(t) \int_0^t \boldsymbol{\lambda}_y^\top(\tau)\,\mathrm{d}\mathbf{W}(\tau) = -y\boldsymbol{\theta}(t) \quad \text{a.s.} \tag{4.7.26}$$

For all $t \in [0,T]$, set

$$\bar{\boldsymbol{\lambda}}_y(t) := \beta(t)\boldsymbol{\lambda}_y(t), \tag{4.7.27}$$

where $\beta(t)$ is defined by (4.5.5). Then as β is uniformly bounded and $\boldsymbol{\lambda}_y \in L^2(\mathbf{W})$, we have $\bar{\boldsymbol{\lambda}}_y \in L^2(\mathbf{W})$. Substituting $\boldsymbol{\lambda}_y = \beta^{-1}\bar{\boldsymbol{\lambda}}_y$ into (4.7.26) and multiplying across by $\beta(t)$, we get

$$\bar{\boldsymbol{\lambda}}_y(t) + \beta(t)\boldsymbol{\theta}(t)\int_0^t \beta^{-1}(\tau)\bar{\boldsymbol{\lambda}}_y^\top(\tau)\,\mathrm{d}\mathbf{W}(\tau) = -y\beta(t)\boldsymbol{\theta}(t) \quad \text{a.s.} \tag{4.7.28}$$

Rearranging, this becomes

$$\bar{\boldsymbol{\lambda}}_y(t) = -\boldsymbol{\theta}(t)\left(y\beta(t) + \beta(t)\int_0^t \beta^{-1}(\tau)\bar{\boldsymbol{\lambda}}_y^\top(\tau)\,\mathrm{d}\mathbf{W}(\tau)\right) \quad \text{a.s.} \tag{4.7.29}$$

Set $R := \Xi(y, \bar{\boldsymbol{\lambda}}_y, \mathbf{0})$. From Lemma 4.5.2 applied to $(y, \boldsymbol{\lambda}, \boldsymbol{\gamma}) := (y, \bar{\boldsymbol{\lambda}}_y, \mathbf{0})$, we have

$$R \equiv (y, -rR, \bar{\boldsymbol{\lambda}}_y, \mathbf{0}) \in \mathbb{B}_1. \tag{4.7.30}$$

From (4.4.6) applied to R, which is given by (4.7.30), so that in particular $\boldsymbol{\Lambda}^R(t) := \bar{\boldsymbol{\lambda}}_y(t)$, we get for all $t \in [0,T]$,

$$\boldsymbol{\Theta}_R(t) = -\boldsymbol{\sigma}(t)\left(\boldsymbol{\theta}(t)R(t) + \bar{\boldsymbol{\lambda}}_y(t)\right) \quad \text{a.s.} \tag{4.7.31}$$

Expanding $R = \Xi(y, \bar{\boldsymbol{\lambda}}_y, \mathbf{0})$ using (4.5.6), (4.5.7) and (4.5.8),

$$R(t) = y\beta(t) + \beta(t)\int_0^t \beta^{-1}(\tau)\bar{\boldsymbol{\lambda}}_y^\top(\tau)\,\mathrm{d}\mathbf{W}(\tau) \text{ a.s.}, \quad \forall t \in [0,T]. \tag{4.7.32}$$

Using (4.7.32) to replace the term in brackets in (4.7.29) gives $\bar{\boldsymbol{\lambda}}_y(t) = -\boldsymbol{\theta}(t)R(t)$. Substituting this into (4.7.31), we get for all $t \in [0,T]$,

$$\Theta_R(t) = 0 \quad \text{a.s.} \tag{4.7.33}$$

Substituting $(y, \boldsymbol{\lambda}, \boldsymbol{\gamma}) := (y, \bar{\boldsymbol{\lambda}}_y, \mathbf{0})$ and $\Theta_R(t) = 0$ into (4.7.18), we obtain

$$0 \le y(x_0 - \tilde{X}(0)). \tag{4.7.34}$$

By the arbitrary choice of $y \in \mathbb{R}$, we must have that $\tilde{X}(0) = x_0$. So Claim 4.7.4 is shown and hence (4.7.21) holds.

We use Claim 4.7.4 to simplify (4.7.18) to

$$0 \le \lim_{\epsilon \downarrow 0}\left\{\mathrm{E}\int_0^T \left(\frac{\delta\left(\Theta_{\bar{Y}}(t) + \epsilon\Theta_R(t)\right) - \delta\left(\Theta_{\bar{Y}}(t)\right)}{\epsilon}\right)\mathrm{d}t\right\}$$
$$+ \mathrm{E}\int_0^T \tilde{\boldsymbol{\pi}}^\top(t)\Theta_R(t)\,\mathrm{d}t - \mathrm{E}\sum_{i,j=1}^{D}\int_0^T \frac{1}{H(t)}\Gamma_{ij}^{\tilde{X}H}(t)\gamma_{ij}(t)\,\mathrm{d}[\mathcal{Q}_{ij}](t), \tag{4.7.35}$$

which holds for all $(y, \boldsymbol{\lambda}, \boldsymbol{\gamma}) \in \mathbb{R} \times L^2(\mathbf{W}) \times L^2(\mathcal{Q})$, for $R = \Xi(y, \boldsymbol{\lambda}, \boldsymbol{\gamma})$.

Note from the properties of the support function δ, as outlined in Remark 4.4.5, that for $\epsilon > 0$,

$$\begin{aligned} \frac{\delta\left(\Theta_{\bar{Y}}(t) + \epsilon\Theta_R(t)\right) - \delta\left(\Theta_{\bar{Y}}(t)\right)}{\epsilon} &\leq \frac{\delta\left(\Theta_{\bar{Y}}(t)\right) + \epsilon\delta\left(\Theta_R(t)\right) - \delta\left(\Theta_{\bar{Y}}(t)\right)}{\epsilon} \\ &\leq \delta\left(\Theta_R(t)\right), \end{aligned} \tag{4.7.36}$$

for all $(y, \boldsymbol{\lambda}, \boldsymbol{\gamma}) \in \mathbb{R} \times L^2(\mathbf{W}) \times L^2(\mathcal{Q})$, for $R = \Xi(y, \boldsymbol{\lambda}, \boldsymbol{\gamma})$. Applying this inequality to (4.7.35), we get

$$\mathrm{E}\sum_{i,j=1}^{D}\int_0^T \frac{1}{H(t)}\Gamma_{ij}^{\tilde{X}H}(t)\gamma_{ij}(t)\,\mathrm{d}[\mathcal{Q}_{ij}](t) \leq \mathrm{E}\int_0^T \left(\delta\left(\Theta_R(t)\right) + \tilde{\boldsymbol{\pi}}^\top(t)\Theta_R(t)\right)\,\mathrm{d}t, \tag{4.7.37}$$

for all $(y, \boldsymbol{\lambda}, \boldsymbol{\gamma}) \in \mathbb{R} \times L^2(\mathbf{W}) \times L^2(\mathcal{Q})$, for $R = \Xi(y, \boldsymbol{\lambda}, \boldsymbol{\gamma})$.

Claim 4.7.5. For the candidate portfolio process $\tilde{\boldsymbol{\pi}}$ defined by (4.6.68), we have $\tilde{\boldsymbol{\pi}} \in K$ $(\mathbb{P} \otimes Leb)$-a.e.

Set $y := 0$ and $\boldsymbol{\gamma} := \mathbf{0}$ in (4.7.37) to obtain

$$\mathrm{E}\int_0^T \left(\delta\left(\Theta_R(t)\right) + \tilde{\boldsymbol{\pi}}^\top(t)\Theta_R(t)\right)\,\mathrm{d}t \geq 0, \tag{4.7.38}$$

for all $\boldsymbol{\lambda} \in L^2(\mathbf{W})$, for $R = \Xi(0, \boldsymbol{\lambda}, \mathbf{0})$.

Define

$$B := \{(\omega, t) \in \Omega \times [0, T] \,:\, \tilde{\boldsymbol{\pi}}(\omega, t) \in K\}. \tag{4.7.39}$$

From Lemma C.1.1 with $\mathbf{p} := \tilde{\boldsymbol{\pi}}$, there exists a previsible mapping $\boldsymbol{\nu}^{\tilde{\boldsymbol{\pi}}} : \Omega \times [0, T] \to \mathbb{R}^N$ such that $\|\boldsymbol{\nu}^{\tilde{\boldsymbol{\pi}}}(t)\| \leq 1$ $(\mathbb{P} \otimes Leb)$-a.e., $|\,\delta(\boldsymbol{\nu}^{\tilde{\boldsymbol{\pi}}}(t))| \leq 1$ $(\mathbb{P} \otimes Leb)$-a.e. and

$$\begin{cases} \delta(\boldsymbol{\nu}^{\tilde{\boldsymbol{\pi}}}(t)) + \tilde{\boldsymbol{\pi}}^\top(t)\boldsymbol{\nu}^{\tilde{\boldsymbol{\pi}}}(t) = 0 & (\mathbb{P} \otimes Leb)\text{-a.e. on } B \\ \delta(\boldsymbol{\nu}^{\tilde{\boldsymbol{\pi}}}(t)) + \tilde{\boldsymbol{\pi}}^\top(t)\boldsymbol{\nu}^{\tilde{\boldsymbol{\pi}}}(t) < 0 & (\mathbb{P} \otimes Leb)\text{-a.e. on } \Omega \times [0, T] \setminus B. \end{cases} \tag{4.7.40}$$

Suppose

$$(\mathbb{P} \otimes Leb)(\Omega \times [0, T] \setminus B) > 0. \tag{4.7.41}$$

Then it follows from (4.7.40) that

$$\mathrm{E}\int_0^T \left(\delta\left(\boldsymbol{\nu}^{\tilde{\boldsymbol{\pi}}}(t)\right) + \tilde{\boldsymbol{\pi}}^\top(t)\boldsymbol{\nu}^{\tilde{\boldsymbol{\pi}}}(t)\right)\,\mathrm{d}t < 0. \tag{4.7.42}$$

Set $\boldsymbol{\rho}(t) := -\beta^{-1}(t)\boldsymbol{\sigma}^{-1}(t)\boldsymbol{\nu}^{\tilde{\boldsymbol{\pi}}}(t)$, where $\beta(t)$ is defined by (4.5.5). By the boundedness of β, $\boldsymbol{\sigma}$ and $\boldsymbol{\nu}^{\tilde{\boldsymbol{\pi}}}$, we have $\boldsymbol{\rho} \in L^2(\mathbf{W})$.

Applying Lemma 4.7.1 to $(\boldsymbol{\rho}, \boldsymbol{\gamma}) := (-\beta^{-1}\boldsymbol{\sigma}^{-1}\boldsymbol{\nu}^{\tilde{\boldsymbol{\pi}}}, \mathbf{0}) \in L^2(\mathbf{W}) \times L^2(\mathcal{Q})$, there exists $\xi \in L^2(\mathbf{W})$ such that for all $t \in [0, T]$,

$$\xi(t) + \boldsymbol{\theta}(t)\int_0^t \xi^\top(\tau)\,\mathrm{d}\mathbf{W}(\tau) = -\beta^{-1}(t)\boldsymbol{\sigma}^{-1}(t)\boldsymbol{\nu}^{\tilde{\boldsymbol{\pi}}}(t) \quad \text{a.s.} \tag{4.7.43}$$

For all $t \in [0, T]$, define

$$\bar{\xi}(t) := \beta(t)\xi(t), \tag{4.7.44}$$

where $\beta(t)$ is defined by (4.5.5). Then as β is uniformly bounded and $\xi \in L^2(\mathbf{W})$, we have $\bar{\xi} \in L^2(\mathbf{W})$. Substituting $\xi(t) = \beta^{-1}(t)\bar{\xi}(t)$ into (4.7.43), multiplying across by $\boldsymbol{\sigma}(t)\beta(t)$ and rearranging, we get

$$\boldsymbol{\nu}^{\tilde{\boldsymbol{\pi}}}(t) = -\boldsymbol{\sigma}(t)\left(\bar{\xi}(t) + \beta(t)\boldsymbol{\theta}(t)\int_0^t \beta^{-1}(\tau)\bar{\xi}^\top(\tau)\,\mathrm{d}\mathbf{W}(\tau)\right) \quad \text{a.s.} \tag{4.7.45}$$

Set $R_0 := \Xi(0, \bar{\xi}, \mathbf{0})$. From Lemma 4.5.2 applied to $(y, \boldsymbol{\lambda}, \boldsymbol{\gamma}) := (0, \bar{\xi}, \mathbf{0})$, we have

$$R_0 \equiv (0, -rR_{0\,-}, \bar{\xi}, \mathbf{0}) \in \mathbb{B}_1. \tag{4.7.46}$$

From (4.4.6) applied to R_0, which is given by (4.7.46), so that in particular $\boldsymbol{\Lambda}^{R_0}(t) := \bar{\xi}(t)$, we get for all $t \in [0, T]$,

$$\boldsymbol{\Theta}_{R_0}(t) = -\boldsymbol{\sigma}(t)\left(\boldsymbol{\theta}(t)R_0(t) + \bar{\xi}(t)\right) \quad \text{a.s.} \tag{4.7.47}$$

Expanding $R_0 = \Xi(0, \bar{\xi}, \mathbf{0})$ using (4.5.6), (4.5.7) and (4.5.8), we have a.s.,

$$R_0(t) = \beta(t)\int_0^t \beta^{-1}(\tau)\bar{\xi}^\top(\tau)\,\mathrm{d}\mathbf{W}(\tau), \quad \forall t \in [0, T]. \tag{4.7.48}$$

Then substituting (4.7.48) into (4.7.45), we get

$$\boldsymbol{\nu}^{\tilde{\boldsymbol{\pi}}}(t) = -\boldsymbol{\sigma}(t)\left(\bar{\xi}(t) + \boldsymbol{\theta}(t)R_0(t)\right) \overset{(4.7.47)}{=} \Theta_{R_0}(t) \quad \text{a.s.} \tag{4.7.49}$$

Substituting $\boldsymbol{\nu}^{\tilde{\boldsymbol{\pi}}}(t) = \Theta_{R_0}(t)$ into (4.7.42), we get

$$\mathrm{E}\int_0^T \left(\delta\left(\Theta_{R_0}(t)\right) + \tilde{\boldsymbol{\pi}}^\top(t)\Theta_{R_0}(t)\right)\,\mathrm{d}t < 0, \tag{4.7.50}$$

which contradicts (4.7.38). Hence (4.7.41) cannot hold so we must have,

$$(\mathbb{P} \otimes Leb)(\Omega \times [0, T] \setminus B) = 0, \tag{4.7.51}$$

which means $(\mathbb{P} \otimes Leb)(B) = 1$. By the definition of B in (4.7.39), it follows immediately that $\tilde{\boldsymbol{\pi}} \in K$ $(\mathbb{P} \otimes Leb)$-a.e. and Claim 4.7.5 is shown. Hence (4.7.23) holds for the candidate portfolio process $\tilde{\boldsymbol{\pi}}$ defined by (4.6.68).

Claim 4.7.6. $\boldsymbol{\Gamma}^{\tilde{X}H} \equiv \mathbf{0}$ $\nu_{[\mathcal{Q}]}$-a.e.

It follows from (4.7.51) and (4.7.40) that $(\mathbb{P} \otimes Leb)$-a.e.,

$$\delta(\boldsymbol{\nu}^{\tilde{\boldsymbol{\pi}}}(t)) + \tilde{\boldsymbol{\pi}}^\top(t)\boldsymbol{\nu}^{\tilde{\boldsymbol{\pi}}}(t) = 0. \tag{4.7.52}$$

Applying Lemma 4.7.1 to $(\boldsymbol{\rho}, \boldsymbol{\gamma}) := (-\beta^{-1}\boldsymbol{\sigma}^{-1}\boldsymbol{\nu}^{\tilde{\boldsymbol{\pi}}}, \beta^{-1}\frac{1}{H}\boldsymbol{\Gamma}^{\tilde{X}H}) \in L^2(\mathbf{W}) \times L^2(\mathcal{Q})$, where β is defined by (4.5.5), there exists $\boldsymbol{\eta} \in L^2(\mathbf{W})$ such that

$$\begin{aligned}&\boldsymbol{\eta}(t) + \boldsymbol{\theta}(t) \int_0^t \boldsymbol{\eta}^\top(\tau)\,\mathrm{d}\mathbf{W}(\tau) \\ &\quad = -\beta^{-1}(t)\boldsymbol{\sigma}^{-1}(t)\boldsymbol{\nu}^{\tilde{\boldsymbol{\pi}}}(t) - \boldsymbol{\theta}(t) \sum_{i,j=1}^{D} \int_0^t \beta^{-1}(\tau)\frac{1}{H(\tau)}\Gamma_{ij}^{\tilde{X}H}(\tau)\,\mathrm{d}\mathcal{Q}_{ij}(\tau) \quad \text{a.s.}\end{aligned} \tag{4.7.53}$$

For all $t \in [0, T]$, define

$$\bar{\boldsymbol{\eta}}(t) := \beta(t)\boldsymbol{\eta}(t), \tag{4.7.54}$$

where $\beta(t)$ is defined by (4.5.5). Then as 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 (4.7.53), multiplying across by $\boldsymbol{\sigma}(t)\beta(t)$ and rearranging, we get

$$\begin{aligned}\boldsymbol{\nu}^{\tilde{\boldsymbol{\pi}}}(t) = -\boldsymbol{\sigma}(t)\Bigg(&\bar{\boldsymbol{\eta}}(t) + \beta(t)\boldsymbol{\theta}(t)\int_0^t \beta^{-1}(\tau)\bar{\boldsymbol{\eta}}^\top(\tau)\,\mathrm{d}\mathbf{W}(\tau) \\ &+ \beta(t)\boldsymbol{\theta}(t)\sum_{i,j=1}^{D}\int_0^t \beta^{-1}(\tau)\frac{1}{H(\tau)}\Gamma_{ij}^{\tilde{X}H}(\tau)\,\mathrm{d}\mathcal{Q}_{ij}(\tau)\Bigg) \quad \text{a.s.}\end{aligned} \tag{4.7.55}$$

Set $R_1 := \Xi(0, \bar{\boldsymbol{\eta}}, \frac{1}{H}\boldsymbol{\Gamma}^{\tilde{X}H})$. From Lemma 4.5.2 applied to the triple $(y, \boldsymbol{\lambda}, \boldsymbol{\gamma}) := (0, \bar{\boldsymbol{\eta}}, \frac{1}{H}\boldsymbol{\Gamma}^{\tilde{X}H})$, we have

$$R_1 \equiv (0, -rR_{1-}, \bar{\boldsymbol{\eta}}, \frac{1}{H}\boldsymbol{\Gamma}^{\tilde{X}H}) \in \mathbb{B}_1. \tag{4.7.56}$$

From (4.4.6) applied to R_1, so that in particular $\boldsymbol{\Lambda}^{R_1}(t) := \bar{\boldsymbol{\eta}}(t)$, we get for all $t \in [0, T]$,

$$\boldsymbol{\Theta}_{R_1}(t) = -\boldsymbol{\sigma}(t)\left(\boldsymbol{\theta}(t)R_1(t) + \bar{\boldsymbol{\eta}}(t)\right) \quad \text{a.s.} \tag{4.7.57}$$

Expanding $R_1 = \Xi(0, \bar{\boldsymbol{\eta}}, \frac{1}{H}\boldsymbol{\Gamma}^{\tilde{X}H})$ using (4.5.6), (4.5.7) and (4.5.8), we have for all $t \in [0, T]$,

$$R_1(t) = \beta(t)\int_0^t \beta^{-1}(\tau)\bar{\boldsymbol{\eta}}^\top(\tau)\,\mathrm{d}\mathbf{W}(\tau) + \beta(t)\sum_{i,j=1}^{D}\int_0^t \beta^{-1}(\tau)\frac{1}{H(\tau)}\Gamma_{ij}^{\tilde{X}H}(\tau)\,\mathrm{d}\mathcal{Q}_{ij}(\tau). \tag{4.7.58}$$

Substituting (4.7.58) into (4.7.55), we get

$$\boldsymbol{\nu}^{\tilde{\boldsymbol{\pi}}}(t) = -\boldsymbol{\sigma}(t)\left(\bar{\boldsymbol{\eta}}(t) + \boldsymbol{\theta}(t)R_1(t)\right) \overset{(4.7.57)}{=} \Theta_{R_1}(t) \quad \text{a.s.} \tag{4.7.59}$$

Then substituting $\boldsymbol{\nu}^{\tilde{\boldsymbol{\pi}}}(t) = \Theta_{R_1}(t)$ into (4.7.52), we obtain

$$\delta\left(\Theta_{R_1}(t)\right) + \tilde{\boldsymbol{\pi}}^\top(t)\Theta_{R_1}(t) = 0. \tag{4.7.60}$$

Substituting $(y, \boldsymbol{\lambda}, \boldsymbol{\gamma}) := (0, \bar{\boldsymbol{\eta}}, \frac{1}{H}\boldsymbol{\Gamma}^{\tilde{X}H})$ and $R := R_1 = \Xi(0, \bar{\boldsymbol{\eta}}, \frac{1}{H}\boldsymbol{\Gamma}^{\tilde{X}H})$ into (4.7.37) and using (4.7.60), we find

$$\mathrm{E}\sum_{i,j=1}^{D}\int_0^T \frac{1}{H^2(t)}|\Gamma_{ij}^{\tilde{X}H}(t)|^2\,\mathrm{d}[\mathcal{Q}_{ij}](t) \leq 0. \tag{4.7.61}$$

However, the left-hand side of (4.7.61) is the sum of nonnegative terms, so we must have equality in (4.7.61), that is

$$\mathrm{E}\sum_{i,j=1}^{D}\int_0^T \frac{1}{H^2(t)}|\Gamma_{ij}^{\tilde{X}H}(t)|^2\,\mathrm{d}[\mathcal{Q}_{ij}](t) = 0. \tag{4.7.62}$$

By virtue of $\boldsymbol{\Gamma}^{\tilde{X}H} \in L^2(\mathcal{Q})$, $\Gamma_{ii}^{\tilde{X}H} = 0$ $(\mathbb{P}\otimes Leb)$-a.e. for each $i = 1, \dots, D$. Hence,

$$\frac{1}{H}\boldsymbol{\Gamma}^{\tilde{X}H} = 0 \quad \nu_{[\mathcal{Q}]}\text{-a.e.}, \tag{4.7.63}$$

and by the positivity of H,

$$\boldsymbol{\Gamma}^{\tilde{X}H} = 0 \quad \nu_{[\mathcal{Q}]}\text{-a.e.} \tag{4.7.64}$$

Hence Claim 4.7.6 is shown and (4.7.22) holds.

Claim 4.7.7. $\dot{\tilde{X}}(t) = r(t)\tilde{X}(t) + \tilde{\boldsymbol{\pi}}^\top(t)\boldsymbol{\sigma}(t)\boldsymbol{\theta}(t)$ $(\mathbb{P}\otimes Leb)$-a.e.

From (4.6.92),

$$\tilde{X} \equiv (\tilde{X}(0), r\tilde{X}_- + \tilde{\boldsymbol{\pi}}^\top\boldsymbol{\sigma}\boldsymbol{\theta}, \boldsymbol{\sigma}^\top\tilde{\boldsymbol{\pi}}, \frac{\boldsymbol{\Gamma}^{\tilde{X}H}}{H}) \in \mathbb{B}, \tag{4.7.65}$$

and upon applying Claim 4.7.6, this becomes $\tilde{X} \equiv (\tilde{X}(0), r\tilde{X}_- + \tilde{\boldsymbol{\pi}}^\top\boldsymbol{\sigma}\boldsymbol{\theta}, \boldsymbol{\sigma}^\top\tilde{\boldsymbol{\pi}}, \mathbf{0}) \in \mathbb{B}$. However, it is clear from the nullity of the $L^2(\mathcal{Q})$-component that $\tilde{X}$ is a continuous process. Hence we can drop the dangling minus appended to $\tilde{X}$ and simply write

$$\tilde{X} \equiv (\tilde{X}(0), r\tilde{X} + \tilde{\boldsymbol{\pi}}^\top\boldsymbol{\sigma}\boldsymbol{\theta}, \boldsymbol{\sigma}^\top\tilde{\boldsymbol{\pi}}, \mathbf{0}) \in \mathbb{B}. \tag{4.7.66}$$

Comparing (4.7.66) to the general form $\tilde{X} = (\tilde{X}(0), \dot{\tilde{X}}, \boldsymbol{\Lambda}^{\tilde{X}}, \boldsymbol{\Gamma}^{\tilde{X}})$, we have

$$\dot{\tilde{X}}(t) = r(t)\tilde{X}(t) + \tilde{\boldsymbol{\pi}}^\top(t)\boldsymbol{\sigma}(t)\boldsymbol{\theta}(t) \quad (\mathbb{P}\otimes Leb)\text{-a.e.}, \tag{4.7.67}$$

proving Claim 4.7.7 and showing that (4.7.24) holds.

Claim 4.7.8. $\delta\left(\Theta_{\bar{Y}}(t)\right) + \tilde{\boldsymbol{\pi}}^\top(t)\Theta_{\bar{Y}}(t) = 0$ $(\mathbb{P}\otimes Leb)$-a.e.

From Claim 4.7.6, $\boldsymbol{\Gamma}^{\tilde{X}H} \equiv \mathbf{0}$ $\nu_{[\mathcal{Q}]}$-a.e. Then (4.7.35) simplifies to

$$0 \leq \lim_{\epsilon\downarrow 0}\left\{\mathrm{E}\int_0^T \left(\frac{\delta\left(\Theta_{\bar{Y}}(t) + \epsilon\Theta_R(t)\right) - \delta\left(\Theta_{\bar{Y}}(t)\right)}{\epsilon}\right)\mathrm{d}t\right\} + \mathrm{E}\int_0^T \tilde{\boldsymbol{\pi}}^\top(t)\Theta_R(t)\,\mathrm{d}t, \tag{4.7.68}$$

for all $(y, \boldsymbol{\lambda}, \boldsymbol{\gamma}) \in \mathbb{R} \times L^2(\mathbf{W}) \times L^2(\mathcal{Q})$, for $R = \Xi(y, \boldsymbol{\lambda}, \boldsymbol{\gamma})$.

Recall that $\bar{Y} = \Xi(\bar{Y}_0, \mathbf{\Lambda}^{\bar{Y}}, \mathbf{\Gamma}^{\bar{Y}})$ is the solution to the dual problem. Set

$$R_2 := -\bar{Y} = \Xi(-\bar{Y}_0, -\mathbf{\Lambda}^{\bar{Y}}, -\mathbf{\Gamma}^{\bar{Y}}). \tag{4.7.69}$$

From Lemma 4.5.2 applied to $(y, \boldsymbol{\lambda}, \boldsymbol{\gamma}) := (-\bar{Y}_0, -\mathbf{\Lambda}^{\bar{Y}}, -\mathbf{\Gamma}^{\bar{Y}})$, we have

$$R_2 \equiv (-\bar{Y}_0, -rR_{2\,-}, -\mathbf{\Lambda}^{\bar{Y}}, -\mathbf{\Gamma}^{\bar{Y}}) \in \mathbb{B}_1. \tag{4.7.70}$$

From (4.4.6) applied to R_2, so that in particular $R_2(t) = -\bar{Y}(t)$ and $\mathbf{\Lambda}^{R_2}(t) := -\mathbf{\Lambda}^{\bar{Y}}(t)$, we get for all $t \in [0, T]$,

$$\begin{aligned}\mathbf{\Theta}_{R_2}(t) &= -\boldsymbol{\sigma}(t)\left(-\boldsymbol{\theta}(t)\bar{Y}(t) - \mathbf{\Lambda}^{\bar{Y}}(t)\right)\\ &= -\left(-\boldsymbol{\sigma}(t)\left(\boldsymbol{\theta}(t)\bar{Y}(t) + \mathbf{\Lambda}^{\bar{Y}}(t)\right)\right)\\ &= \mathbf{\Theta}_{\bar{Y}(t)} \quad \text{a.s.}\end{aligned} \tag{4.7.71}$$

Fix $\epsilon \in (0,1)$. Then using the positive homogeneity of the support function δ, as noted in Remark 4.4.5, we get

$$\delta\left(\Theta_{\bar{Y}}(t) + \epsilon\Theta_{R_2}(t)\right) \overset{(4.7.71)}{=} \delta\left(\Theta_{\bar{Y}}(t) - \epsilon\Theta_{\bar{Y}}(t)\right) = (1-\epsilon)\delta\left(\Theta_{\bar{Y}}(t)\right), \tag{4.7.72}$$

and hence

$$\frac{\delta\left(\Theta_{\bar{Y}}(t) + \epsilon\Theta_{R_2}(t)\right) - \delta\left(\Theta_{\bar{Y}}(t)\right)}{\epsilon} = \frac{(1-\epsilon)\delta\left(\Theta_{\bar{Y}}(t)\right) - \delta\left(\Theta_{\bar{Y}}(t)\right)}{\epsilon} = -\delta\left(\Theta_{\bar{Y}}(t)\right). \tag{4.7.73}$$

Substituting (4.7.71) and (4.7.73) into (4.7.68), we get

$$\mathrm{E}\int_0^T \left(-\delta\left(\Theta_{\bar{Y}}(t)\right) + \tilde{\boldsymbol{\pi}}^\top(t)\left(-\Theta_{\bar{Y}}(t)\right)\right)\,\mathrm{d}t \geq 0, \tag{4.7.74}$$

that is

$$\mathrm{E}\int_0^T \left(\delta\left(\Theta_{\bar{Y}}(t)\right) + \tilde{\boldsymbol{\pi}}^\top(t)\Theta_{\bar{Y}}(t)\right)\,\mathrm{d}t \leq 0. \tag{4.7.75}$$

Now as $\tilde{\boldsymbol{\pi}} \in K$ $(\mathbb{P} \otimes \mathit{Leb})$-a.e., we have for all $t \in [0, T]$,

$$\delta\left(\Theta_{\bar{Y}}(t)\right) = \sup_{\boldsymbol{\pi} \in K}\left\{-\boldsymbol{\pi}^\top\Theta_{\bar{Y}}(t)\right\} \geq -\tilde{\boldsymbol{\pi}}^\top(t)\Theta_{\bar{Y}}(t). \tag{4.7.76}$$

Combining (4.7.75) and (4.7.76) we obtain equality, that is

$$\delta\left(\Theta_{\bar{Y}}(t)\right) + \tilde{\boldsymbol{\pi}}^\top(t)\Theta_{\bar{Y}}(t) = 0 \quad (\mathbb{P} \otimes \mathit{Leb})\text{-a.e.} \tag{4.7.77}$$

This proves Claim 4.7.8 and hence (4.7.25) holds.

□

Remark 4.7.9. From Corollary 4.6.18, we have

$$\tilde{X} \equiv (\tilde{X}(0), r\tilde{X}_- + \tilde{\boldsymbol{\pi}}^\top \boldsymbol{\sigma\theta}, \boldsymbol{\sigma}^\top \tilde{\boldsymbol{\pi}}, \frac{\boldsymbol{\Gamma}^{\tilde{X}H}}{H}) \in \mathbb{B}. \tag{4.7.78}$$

Using (4.7.22) of Proposition 4.7.2, this becomes

$$\tilde{X} \equiv (\tilde{X}(0), r\tilde{X} + \tilde{\boldsymbol{\pi}}^\top \boldsymbol{\sigma\theta}, \boldsymbol{\sigma}^\top \tilde{\boldsymbol{\pi}}, \mathbf{0}) \in \mathbb{A}. \tag{4.7.79}$$

Hence the candidate solution $\tilde{X}$ is a continuous process. Then we can also drop the dangling minus sign in the definition of the candidate portfolio process, which is defined by (4.6.68), to get

$$\tilde{\boldsymbol{\pi}}(t) = (\boldsymbol{\sigma}^\top(t))^{-1} \left(\frac{\boldsymbol{\Lambda}^{\tilde{X}H}(t)}{H(t)} + \tilde{X}(t)\boldsymbol{\theta}(t) \right). \tag{4.7.80}$$

Moreover, from (4.7.21) of Proposition 4.7.2, $\tilde{X}(0) = x_0$. This initial condition, the form of the candidate solution $\tilde{X}$ given by (4.7.79) and the uniqueness of the solution of the wealth equation (3.2.31) gives

$$\tilde{X} = X^{\tilde{\boldsymbol{\pi}}} \quad (\mathbb{P} \otimes Leb)\text{-a.e.}, \tag{4.7.81}$$

that is the candidate solution $\tilde{X}$ solves the wealth equation (3.2.31) for the portfolio process $\tilde{\boldsymbol{\pi}}$.

Corollary 4.7.10. *For the candidate solution $\tilde{X} \in \mathbb{A}$, defined by (4.6.32) and expressed in component form as (4.7.79), and the solution to the dual problem $\bar{Y} \in \mathbb{B}$ (recall Proposition 4.5.15), we have*

$$\Phi(\tilde{X}) + \Psi(\bar{Y}) = 0. \tag{4.7.82}$$

Proof. To prove this, we show that (4.4.16) - (4.4.21) of Proposition 4.4.8 are satisfied.

First note the appearance of the candidate portfolio process $\tilde{\boldsymbol{\pi}}$ in (4.7.79). Recalling (3.2.37), this gives

$$\tilde{\boldsymbol{\pi}}(t) = (\boldsymbol{\sigma}^\top(t))^{-1} \boldsymbol{\Lambda}^{\tilde{X}}(t), \quad (\mathbb{P} \otimes Leb)\text{-a.e.} \tag{4.7.83}$$

This agrees with (4.4.19). Recall from Corollary 4.6.17 that $\tilde{\boldsymbol{\pi}} \in L^2(\mathbf{W})$. Together with (4.7.23) of Proposition 4.7.2, this gives $\tilde{\boldsymbol{\pi}} \in \mathcal{A}$ (recall Definition 3.2.35). Then this, (4.7.83) and (4.7.24) of Proposition 4.7.2 show that $\tilde{\boldsymbol{\pi}} \in U(\tilde{X})$, so that (4.4.21) of Proposition 4.4.8 holds.

Equations (4.4.16), (4.4.17) and (4.4.20) are shown by (4.7.21), (4.7.20) and (4.7.25) of Proposition 4.7.2, respectively.

It remains to show that (4.4.18) of Proposition 4.4.8 holds. From Proposition 4.5.15, the solution to the dual problem is $\bar{Y} \equiv \Xi(\bar{y}, \bar{\boldsymbol{\lambda}}, \bar{\boldsymbol{\gamma}})$ for some $(\bar{y}, \bar{\boldsymbol{\lambda}}, \bar{\boldsymbol{\gamma}}) \in \mathbb{R} \times L^2(\mathbf{W}) \times L^2(\mathcal{Q})$. Applying Lemma 4.5.2 to $\bar{Y}$, we find

$$\bar{Y} = (\bar{y}, -r\bar{Y}_-, \bar{\boldsymbol{\lambda}}, \bar{\boldsymbol{\gamma}}) \in \mathbb{B}_1. \tag{4.7.84}$$

Recalling (3.2.37), this gives $\dot{\bar{Y}}(t) = -r(t)\bar{Y}(t_-)$ $(\mathbb{P} \otimes Leb)$-a.e. Since $\bar{Y}(t) \neq \bar{Y}(t_-)$ only on a set of Lebesgue measure zero, we find $\dot{\bar{Y}}(t) = -r(t)\bar{Y}(t)$ $(\mathbb{P} \otimes Leb)$-a.e. and hence (4.4.18) of Proposition 4.4.8 is satisfied.

We can now use Proposition 4.4.8 to conclude that (4.7.82) holds. □

Remark 4.7.11. Corollary 4.7.10 tells us that the candidate solution $\tilde{X} \in \mathbb{A}$ is the solution to the primal problem (Problem 4.1.7). For, immediately from it and Corollary 4.4.1, we have

$$\Phi(\tilde{X}) = \inf_{X \in \mathbb{A}} \{\Phi(X)\}. \tag{4.7.85}$$

Furthermore, recalling the penalty function l_0, defined by (4.1.14), and the penalty function l_1, defined by (4.1.15), we have a.s. for all $t \in [0, T]$,

$$l_0(\tilde{X}_0) = 0 \quad \text{and} \quad l_1(t, \tilde{X}(t)\dot{\tilde{X}}(t), \mathbf{\Lambda}^{\tilde{X}}(t)) = 0. \tag{4.7.86}$$

It follows from the definition of the primal cost functional Φ given by (4.1.17) that

$$\Phi(\tilde{X}) = \mathrm{E}(l_T(\tilde{X}(T))) \overset{(4.1.16)}{=} \mathrm{E}(J(\tilde{X}(T))) \overset{(4.7.81)}{=} \mathrm{E}(J(X^{\tilde{\boldsymbol{\pi}}}(T))). \tag{4.7.87}$$

Now we just recall from (3.2.48) that the value of the problem $\mathcal{V}$ satisfies

$$\mathcal{V} = \inf_{\boldsymbol{\pi} \in \mathcal{A}} \mathrm{E}\left(J(X^{\boldsymbol{\pi}}(T))\right) \overset{(4.1.18)}{=} \inf_{X \in \mathbb{A}} \{\Phi(X)\}, \tag{4.7.88}$$

and combining this with (4.7.85) and (4.7.87), we get

$$\mathcal{V} = \inf_{\boldsymbol{\pi} \in \mathcal{A}} \mathrm{E}\left(J(X^{\boldsymbol{\pi}}(T))\right) = \mathrm{E}(J(X^{\tilde{\boldsymbol{\pi}}}(T))). \tag{4.7.89}$$

That is, the portfolio process $\tilde{\boldsymbol{\pi}}$, which is given by (4.7.80), solves Problem 3.2.38.

4.8 Summary

By applying a convex duality method, we have shown the existence of and characterized the solution to Problem 3.2.38, which is a problem of finding a portfolio process which minimizes the expected value of a quadratic risk measure subject to the portfolio process being an admissible portfolio.

We began by re-expressing Problem 3.2.38 as the primal problem (Problem 4.1.7), changing the minimization from one over the set of admissible portfolios to one over $\mathbb{A}$, which is a set of continuous, square-integrable processes. The constraints on the problem were coded as penalty functions. These penalty functions and the risk measure were used to construct the primal cost functional. The primal problem was to find an element of $\mathbb{A}$ which minimized the primal cost functional.

Next, we constructed the dual cost functional from the risk measure and the penalty functions by taking their convex conjugates. We defined the dual problem as one of finding an element of $\mathbb{B}$, which is a set of right-continuous, square-integrable processes, which minimized the dual cost functional.

From the convex conjugates of the penalty functions and the risk measure, we obtained necessary and sufficient conditions for a pair of elements to solve the primal problem and the dual problem.

After showing that a solution to the dual problem existed, we used some of the conditions and the solution to the dual problem to construct a candidate solution to the primal problem. To verify that the candidate solution did indeed solve the primal problem, we used the necessary and sufficient conditions. This allowed us to show the existence of and to characterize the solution to the primal problem. In turn, this showed the existence of and characterized the solution to Problem 3.2.38.

The results of this chapter are summarized by the following proposition.

Proposition 4.8.1. *Suppose that the market conditions (see Remark 3.2.11), Conditions 3.2.26 - 3.2.28 and Condition 3.2.31 are satisfied. Define the functional*

$$\tilde{\Psi}(y, \boldsymbol{\lambda}, \boldsymbol{\gamma}) := x_0 y + E\int_0^T \delta(\boldsymbol{\Theta}_Y(t))\ \mathrm{d}t + E\left(\frac{(Y(T)+b)^2}{2a}\right) - Ec, \tag{4.8.1}$$

for all $(y, \boldsymbol{\lambda}, \boldsymbol{\gamma}) \in \mathbb{R} \times L^2(\mathbf{W}) \times L^2(\mathcal{Q})$*, where* $Y := \Xi(y, \boldsymbol{\lambda}, \boldsymbol{\gamma})$*.*

Then there exists a triple $(\bar{y}, \bar{\boldsymbol{\lambda}}, \bar{\boldsymbol{\gamma}}) \in \mathbb{R} \times L^2(\mathbf{W}) \times L^2(\mathcal{Q})$ *which minimizes the functional* $\tilde{\Psi}$ *over* $\mathbb{R} \times L^2(\mathbf{W}) \times L^2(\mathcal{Q})$*, that is*

$$\mathcal{V} = -\tilde{\Psi}(\bar{y}, \bar{\boldsymbol{\lambda}}, \bar{\boldsymbol{\gamma}}) = -\inf_{(y,\boldsymbol{\lambda},\boldsymbol{\gamma}) \in \mathbb{R}\times L^2(\mathbf{W})\times L^2(\mathcal{Q})} \tilde{\Psi}(y, \boldsymbol{\lambda}, \boldsymbol{\gamma}). \tag{4.8.2}$$

Defining $\bar{Y} := \Xi(\bar{y}, \bar{\boldsymbol{\lambda}}, \bar{\boldsymbol{\gamma}})$*, for all* $t \in [0, T]$*, set*

$$\bar{X}(t) := -\frac{1}{H(t)} E\left(\left(\frac{\bar{Y}(T)+b}{a}\right) H(T) \,\middle|\, \mathcal{F}_t\right) \quad a.s., \tag{4.8.3}$$

for $\{H(t)\}$ *the state price density process given by (4.6.1) and set*

$$\bar{\boldsymbol{\pi}}(t) := (\boldsymbol{\sigma}^\top(t))^{-1} \left(\frac{\boldsymbol{\Lambda}^{\bar{X}H}(t)}{H(t)} + \bar{X}(t)\boldsymbol{\theta}(t)\right), \tag{4.8.4}$$

where $\boldsymbol{\Lambda}^{\bar{X}H} \in L^2_{loc}(\mathbf{W})$ *is given by Theorem B.4.22, a martingale representation theorem, applied to* $\bar{X}H$ *and which is consequently* $(\mathbb{P} \otimes Leb)$*-a.e. unique.*

Then $\bar{\boldsymbol{\pi}} \in \mathcal{A}$ *and* $\bar{X}(t) = X^{\bar{\boldsymbol{\pi}}}(t)$ $(\mathbb{P} \otimes Leb)$*-a.e., for* $X^{\bar{\boldsymbol{\pi}}}$ *the solution to the wealth equation (3.2.31) for the portfolio process* $\bar{\boldsymbol{\pi}}$*.*

In particular, the portfolio process $\bar{\boldsymbol{\pi}} \in \mathcal{A}$ *solves Problem 3.2.38, so that*

$$\mathcal{V} = E\big(J(X^{\bar{\boldsymbol{\pi}}}(T))\big) = \inf_{\boldsymbol{\pi} \in \mathcal{A}} E\left(J(X^{\boldsymbol{\pi}}(T))\right). \tag{4.8.5}$$

4.9 The fully constrained problem

In this chapter, we have already solved Problem 3.2.38, which we refer to in this section as the *partially-constrained optimization problem*. In this section, we solve the *fully-constrained optimization problem*, which is the partially-constrained optimization problem with the addition of another constraint. We assume that the market conditions (see Remark 3.2.11) are satisfied. The method of solution is taken from Labbé and Heunis [33] and, as it turns out, it follows through without any modifications. However, we include it for the sake of completeness.

To illustrate the fully-constrained optimization problem, we will extend the motivating example given in Section 3.2.2. As before, at time 0 an investor has an initial wealth of x_0 units and a no-short-selling portfolio constraint over the time horizon $[0, T]$. The investor would like to have d units of wealth at the end of the finite time horizon $[0, T]$. In the partially-constrained optimization problem, the investor decides to minimize the variance of her actual terminal wealth from the d units of wealth. For the fully-constrained optimization problem, we assume that the investor seeks in addition to attain, on average, d units of wealth at time T. This is called a *terminal wealth constraint*.

The investor's requirements can be summarized as follows: the investor is seeking to minimize her risk, subject to starting with an initial wealth of x_0 units, meeting the portfolio constraints over the time horizon and with an expected wealth at time T of d units. Can she find a portfolio process which will minimize her risk, subject to satisfying the initial wealth constraint, the portfolio constraints and the terminal wealth constraint?

Having outlined the motivating example, we next specify precisely the general MVO problem that we propose to solve. We start by defining the risk measure, which is analogous to the investor minimizing her risk.

Definition 4.9.1. Define a risk measure $\hat{J}$ on the wealth process by

$$\begin{aligned} &\hat{J} : \Omega \times \mathbb{R} \to \mathbb{R} \\ &(\omega, x) \mapsto \hat{J}(\omega, x) := \frac{1}{2}a(\omega)x^2 + b_0(\omega)x + c_0(\omega), \end{aligned} \tag{4.9.1}$$

subject to the following three conditions:

Condition 4.9.2. a is an $\mathcal{F}_T$–measurable random variable on $(\Omega, \mathcal{F}, \mathbb{P})$ and satisfies

$$0 < \inf_{\omega\in\Omega}\{a(\omega)\} \le \sup_{\omega\in\Omega}\{a(\omega)\} < \infty, \tag{4.9.2}$$

Condition 4.9.3. b_0 is an $\mathcal{F}_T$–measurable, square-integrable random variable on $(\Omega, \mathcal{F}, \mathbb{P})$.

Condition 4.9.4. c_0 is an $\mathcal{F}_T$–measurable, integrable random variable on $(\Omega, \mathcal{F}, \mathbb{P})$.

The portfolio constraint is the same as for the partially-constrained optimization problem, that is the portfolio process we seek must always lie in some closed, convex set.

Condition 4.9.5. We are given a closed, convex set K with $\mathbf{0} \in K$.

Remark 4.9.6. Conditions 4.9.2-4.9.5 are identical to Conditions 3.2.26 - 3.2.28 and Condition 3.2.31.

The set of admissible portfolios $\mathcal{A}$ is the same as the set in Definition 3.2.35, that is

$$\mathcal{A} := \{\boldsymbol{\pi} \in L^2(\mathbf{W}) \,\big|\, \boldsymbol{\pi} \in K \ (\mathbb{P} \otimes Leb)\text{-a.e.}\}. \tag{4.9.3}$$

Remark 4.9.7. As in the partially-constrained portfolio optimization problem, the portfolio constraint is that the portfolio process $\bar{\boldsymbol{\pi}}$ we seek is an admissible portfolio process, that is $\bar{\boldsymbol{\pi}} \in \mathcal{A}$.

Condition 4.9.8. Let b_1 be an $\mathcal{F}_T$–measurable, square-integrable random variable on $(\Omega, \mathcal{F}, \mathbb{P})$ and suppose we are given $d \in \mathbb{R}$.

Definition 4.9.9. Define a real-valued mapping G on the set $\mathcal{A}$ of admissible portfolios as

$$G(\boldsymbol{\pi}) := \mathrm{E}\left(b_1 X^{\boldsymbol{\pi}}(T)\right), \quad \forall \boldsymbol{\pi} \in \mathcal{A}, \tag{4.9.4}$$

for $X^{\boldsymbol{\pi}}$ the solution to the wealth equation (3.2.31) for the portfolio process $\boldsymbol{\pi} \in \mathcal{A}$.

Remark 4.9.10. The terminal wealth constraint is that the expected value of the terminal wealth multiplied by the random variable b_1 equals the real number d.

We also need to impose the *constraint qualification*. This is essential in order to apply the chosen method to solve the fully-constrained optimization problem.

Condition 4.9.11. The random variable b_1 and the convex set K are such that the interval $\{G(\boldsymbol{\pi}) : \boldsymbol{\pi} \in \mathcal{A}\}$ has non-empty interior which includes the number $d \in \mathbb{R}$.

Remark 4.9.12. An immediate question concerning Condition 4.9.11 is to ask if it is a reasonable assumption. Labbé and Heunis [33] use the following argument. Let us suppose that there is at least one portfolio process $\tilde{\boldsymbol{\pi}} \in \mathcal{A}$ such that $\mathrm{E}(b_1 X^{\tilde{\boldsymbol{\pi}}}(T)) > \mathrm{E}(b_1 x_0 S_0(T))$. It also follows from the strictness of the inequality that $\tilde{\boldsymbol{\pi}} \neq \mathbf{0}$. (If no such portfolio process exists then the largest expected terminal wealth is attained by investing the initial wealth x_0 entirely in the (risk-free) bank account and so the problem of portfolio optimization would be trivial.) From (3.2.32), for any $\boldsymbol{\pi}^{(1)}, \boldsymbol{\pi}^{(2)} \in \mathcal{A}$ and $\epsilon \in [0, 1]$,

$$X^{\epsilon \boldsymbol{\pi}^{(1)} + (1-\epsilon)\boldsymbol{\pi}^{(2)}}(t) = \epsilon X^{\boldsymbol{\pi}^{(1)}}(t) + (1-\epsilon) X^{\boldsymbol{\pi}^{(2)}}(t). \tag{4.9.5}$$

It follows from the above equation and the convexity of K in $\mathbb{R}^N$ that

$$\mathcal{R} := \{\mathrm{E}(b_1 X^{\boldsymbol{\pi}}(T)) : \boldsymbol{\pi} \in \mathcal{A}\} \tag{4.9.6}$$

is convex in $\mathbb{R}$ and hence an interval. As $\mathbf{0} \in K$, implying that $\mathbf{0} \in \mathcal{A}$, then $\mathcal{R}$ will contain the interval $[\mathrm{E}(b_1 x_0 S_0(T)), \mathrm{E}(b_1 X^{\tilde{\boldsymbol{\pi}}}(T))]$, which has non-empty interior.

Hence the interior of $\mathcal{R}$, which is given by

$$\mathring{\mathcal{R}} := \left(\inf_{\boldsymbol{\pi} \in \mathcal{A}} \{\mathrm{E}(b_1 X^{\boldsymbol{\pi}}(T))\}, \sup_{\boldsymbol{\pi} \in \mathcal{A}} \{\mathrm{E}(b_1 X^{\boldsymbol{\pi}}(T))\} \right), \tag{4.9.7}$$

is non-empty.

Provided we choose $d \in \mathring{\mathcal{R}}$, then Condition 4.9.11 will hold.

Remark 4.9.13. If we set the random variable $b_1 = 1$ then $G(\boldsymbol{\pi}) = \mathrm{E}\left(X^{\boldsymbol{\pi}}(T)\right)$. Let $d > 0$. The terminal wealth constraint in the motivating example at the start of this section is then given by

$$G(\boldsymbol{\pi}) = \mathrm{E}\left(X^{\boldsymbol{\pi}}(T)\right) = d. \tag{4.9.8}$$

Setting $a = 2$, $b_0 = -2d$ and $c_0 = d^2$, then (4.9.1) reduces to $\hat{J}(x) = (x - d)^2$. Taking the expectation of $\hat{J}(X(T))$ we obtain

$$\mathrm{E}\left(\hat{J}(X(T))\right) = \mathrm{E}\left(X(T) - d\right)^2, \tag{4.9.9}$$

which is the variance of the terminal wealth and is the risk measure in the motivating example. The problem of selecting a portfolio process in the set of admissible portfolios $\mathcal{A}$ which minimizes the variance of the terminal wealth subject to the terminal wealth constraint $\mathrm{E}\left(X^{\boldsymbol{\pi}}(T)\right) = d$ is often called *constrained mean-variance portfolio selection.*

Definition 4.9.14. We denote the value of the fully-constrained optimization problem by $\hat{\mathcal{V}}$ which is defined as

$$\hat{\mathcal{V}} := \inf_{\substack{\boldsymbol{\pi} \in \mathcal{A} \\ G(\boldsymbol{\pi}) = d}} \left\{ \mathrm{E}\left(\hat{J}(X^{\boldsymbol{\pi}}(T))\right) \right\}, \tag{4.9.10}$$

where $X^{\boldsymbol{\pi}}$ is the solution to the wealth equation (3.2.31) corresponding to the portfolio process $\boldsymbol{\pi}$, the set of admissible portfolios $\mathcal{A}$ is given by (4.9.3), the map G is given by Definition 4.9.9, the number d is given by Condition 4.9.8 and the risk measure $\hat{J}$ is given by Definition 4.9.1.

Remark 4.9.15. Compare the value $\hat{\mathcal{V}}$ of the fully-constrained optimization problem with the value $\mathcal{V}$ of the partially-constrained optimization problem, given by Definition 3.2.37. The main difference is the addition of the terminal wealth constraint $G(\boldsymbol{\pi}) = d$. For the fully-constrained optimization problem, we have also imposed the constraint qualification, given by Condition 4.9.11. This extra condition was not required to solve the partially-constrained optimization problem.

Problem 4.9.16. The fully-constrained optimization problem is to determine the existence of a portfolio process $\hat{\boldsymbol{\pi}} \in \mathcal{A}$ such that

$$G(\hat{\boldsymbol{\pi}}) = d \quad \text{and} \quad \hat{\mathcal{V}} = \mathrm{E}\left(\hat{J}\left(X^{\hat{\boldsymbol{\pi}}}(T)\right)\right). \tag{4.9.11}$$

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

Next we define the Lagrangian function for Problem 4.9.16.

$$\mathcal{L}(\mu;\boldsymbol{\pi}) := \mathrm{E}\left(\hat{J}\left(X^{\boldsymbol{\pi}}(T)\right)\right) + \mu\left(G(\boldsymbol{\pi}) - d\right), \quad \forall \boldsymbol{\pi} \in L^2(\mathbf{W}), \quad \forall \mu \in \mathbb{R}. \tag{4.9.12}$$

We want to apply Theorem C.17.4, so we need to check that the conditions of the theorem hold. The vector space we are applying Theorem C.17.4 to is $L^2(\mathbf{W})$. So we need to check that

1. The set of admissible portfolios $\mathcal{A}$ is a convex subset of $L^2(\mathbf{W})$.

   This follows directly from the convexity of K. Fix $\boldsymbol{\pi}_1, \boldsymbol{\pi}_2 \in \mathcal{A}$ and $\epsilon \in [0,1]$. Then $\epsilon\boldsymbol{\pi}_1 + (1-\epsilon)\boldsymbol{\pi}_2 \in L^2(\mathbf{W})$ and from the convexity of K in $\mathbb{R}^N$, we have for all $(\omega, t) \in \Omega \times [0,T]$,

   $$\epsilon\,\boldsymbol{\pi}_1(\omega,t) + (1-\epsilon)\boldsymbol{\pi}_2(\omega,t) \in K. \tag{4.9.13}$$

   Hence $\epsilon\boldsymbol{\pi}_1 + (1-\epsilon)\boldsymbol{\pi}_2 \in \mathcal{A}$ and so $\mathcal{A}$ is a convex subset of $L^2(\mathbf{W})$.

2. The mapping $\boldsymbol{\pi} \mapsto \mathrm{E}\left(\hat{J}\left(X^{\boldsymbol{\pi}}(T)\right)\right) : \mathcal{A} \to \mathbb{R}$ is convex.

   This follows from the risk measure $\hat{J}$ being a quadratic function of $X^{\boldsymbol{\pi}}$, and therefore convex, and (4.9.5).

3. The map $G : L^2(\mathbf{W}) \to \mathbb{R}$ is a linear operator.

   The linearity is immediate from (4.9.5) and the definition of G as a linear function of $X^{\boldsymbol{\pi}}$.

4. The number d is in the interior of $\{G(\boldsymbol{\pi}) : \boldsymbol{\pi} \in \mathcal{A}\}$.

   This is assumed by the constraint qualification (Condition 4.9.11).

Hence we can apply Theorem C.17.4 to see that there exists a Lagrange multiplier $\bar{\mu} \in \mathbb{R}$ such that

$$\inf_{\boldsymbol{\pi}\in\mathcal{A}}\{\mathcal{L}(\bar{\mu};\boldsymbol{\pi})\} = \sup_{\mu\in\mathbb{R}}\inf_{\boldsymbol{\pi}\in\mathcal{A}}\{\mathcal{L}(\mu;\boldsymbol{\pi})\} = \inf_{\substack{\boldsymbol{\pi}\in\mathcal{A}\\ G(\boldsymbol{\pi})=d}}\left\{\mathrm{E}\left(\hat{J}\left(X^{\boldsymbol{\pi}}(T)\right)\right)\right\} \overset{(4.9.10)}{=} \hat{\mathcal{V}}. \tag{4.9.14}$$

The idea is to show that, for each $\mu \in \mathbb{R}$, there exists a portfolio process $\bar{\boldsymbol{\pi}}(\mu) := \{\bar{\boldsymbol{\pi}}(\mu;t) : t \in [0,T]\}$ such that

$$\mathcal{L}(\mu;\bar{\boldsymbol{\pi}}(\mu)) = \inf_{\boldsymbol{\pi}\in\mathcal{A}}\{\mathcal{L}(\mu;\boldsymbol{\pi})\}. \tag{4.9.15}$$

Then, from (4.9.14),

$$\mathcal{L}(\bar{\mu};\bar{\boldsymbol{\pi}}(\bar{\mu})) = \sup_{\mu\in\mathbb{R}}\{\mathcal{L}(\mu;\bar{\boldsymbol{\pi}}(\mu))\} = \hat{\mathcal{V}}. \tag{4.9.16}$$

Finally, we check that the portfolio process $\bar{\boldsymbol{\pi}}(\bar{\mu}) := \{\bar{\boldsymbol{\pi}}(\bar{\mu};t) : t \in [0,T]\}$ solves Problem 4.9.16. We use a variational analysis on the optimality of $\bar{\mu}$ to show this.

We aim to use Proposition 4.8.1, which solves the partially-constrained optimization problem, to show that there exists for each $\mu \in \mathbb{R}$ a portfolio process $\bar{\boldsymbol{\pi}}(\mu) := \{\bar{\boldsymbol{\pi}}(\mu;t) : t \in [0,T]\}$ such that (4.9.15) holds. We re-write the Lagrangian $\mathcal{L}(\mu;\boldsymbol{\pi})$ in such a way that we can make use of the solution to the partially-constrained optimization problem. We begin by expanding the Lagrangian $\mathcal{L}(\mu;\boldsymbol{\pi})$ for a fixed $\mu \in \mathbb{R}$.

$$\begin{aligned}
\mathcal{L}(\mu;\boldsymbol{\pi}) &\overset{(4.9.12)}{=} \mathrm{E}\left(\hat{J}\left(X^{\boldsymbol{\pi}}(T)\right)\right) + \mu\left(G(\boldsymbol{\pi}) - d\right) \\
&\overset{(4.9.1),(4.9.4)}{=} \mathrm{E}\left(\frac{1}{2}a\left(X^{\boldsymbol{\pi}}(T)\right)^2 + b_0 X^{\boldsymbol{\pi}}(T) + c_0\right) + \mu\left(\mathrm{E}\left(b_1 X^{\boldsymbol{\pi}}(T)\right) - d\right) \\
&= \mathrm{E}\left(\frac{1}{2}a\left(X^{\boldsymbol{\pi}}(T)\right)^2 + \left(b_0 + \mu b_1\right) X^{\boldsymbol{\pi}}(T) + c_0 - \mu d\right).
\end{aligned} \tag{4.9.17}$$

Set

$$b_\mu := b_0 + \mu b_1 \quad \text{and} \quad c_\mu := c_0 - \mu d, \tag{4.9.18}$$

Remark 4.9.17. From Condition 4.9.3 and Condition 4.9.8, b_μ is an $\mathcal{F}_T$–measurable, square-integrable random variable. From Condition 4.9.4 and Condition 4.9.8, c_μ is an $\mathcal{F}_T$–measurable, integrable random variable.

Define for each $(\omega, x) \in \Omega \times [0,T]$,

$$J_\mu(\omega, x) := \frac{1}{2}a(\omega)x^2 + b_\mu(\omega)x + c_\mu(\omega). \tag{4.9.19}$$

Remark 4.9.18. Compare (4.9.19) to (3.2.43). We see that J_μ is the same function as J, with b replaced by b_μ and c replaced by c_μ. Furthermore, Conditions 3.2.26, 3.2.27 and 3.2.28 are satisfied by the random variables a, b_μ and c_μ, respectively, by Condition 4.9.2 and Remark 4.9.17.

Then from (4.9.17), (4.9.18) and (4.9.19), we have

$$\mathcal{L}(\mu;\boldsymbol{\pi}) = \mathrm{E}\left(J_\mu\left(X^{\boldsymbol{\pi}}(T)\right)\right). \tag{4.9.20}$$

Hence (4.9.15) can be written as

$$\mathrm{E}\left(J_\mu\left(X^{\bar{\boldsymbol{\pi}}(\mu)}(T)\right)\right) = \inf_{\boldsymbol{\pi} \in \mathcal{A}}\{\mathrm{E}\left(J_\mu\left(X^{\boldsymbol{\pi}}(T)\right)\right)\}. \tag{4.9.21}$$

Remark 4.9.19. We want to determine the existence of a portfolio process $\bar{\boldsymbol{\pi}}(\mu) := \{\bar{\boldsymbol{\pi}}(\mu;t) : t \in [0,T]\}$ such that (4.9.21) holds. Since this is exactly Problem 3.2.38 and all the necessary conditions hold, we can apply Proposition 4.8.1.

Define the functional

$$\tilde{\Psi}(\mu; y, \boldsymbol{\lambda}, \boldsymbol{\gamma}) := x_0 y + \mathrm{E}\int_0^T \delta(\boldsymbol{\Theta}_Y(t))\ \mathrm{d}t + \mathrm{E}\left(\frac{(Y(T)+b_\mu)^2}{2a}\right) - \mathrm{E}c_\mu, \tag{4.9.22}$$

for all $(y, \boldsymbol{\lambda}, \boldsymbol{\gamma}) \in \mathbb{R} \times L^2(\mathbf{W}) \times L^2(\mathcal{Q})$, where $Y := \Xi(y, \boldsymbol{\lambda}, \boldsymbol{\gamma})$.

Then, by Proposition 4.8.1, there exists a triple $(\bar{y}(\mu), \bar{\boldsymbol{\lambda}}(\mu), \bar{\boldsymbol{\gamma}}(\mu)) \in \mathbb{R}\times L^2(\mathbf{W})\times L^2(\mathcal{Q})$ which minimizes the functional $\tilde{\Psi}(\mu; \cdot, \cdot, \cdot)$ over $\mathbb{R} \times L^2(\mathbf{W}) \times L^2(\mathcal{Q})$.

Define

$$\bar{Y}(\mu; t) := \Xi(\bar{y}(\mu), \bar{\boldsymbol{\lambda}}(\mu), \bar{\boldsymbol{\gamma}}(\mu))(t), \tag{4.9.23}$$

$$\bar{X}(\mu; t) := -\frac{1}{H(t)}\mathrm{E}\left(\left(\frac{\bar{Y}(\mu; T)+b_\mu}{a}\right) H(T) \,\middle|\, \mathcal{F}_t\right), \tag{4.9.24}$$

and define the process $\bar{\boldsymbol{\pi}}(\mu) = \{\bar{\boldsymbol{\pi}}(\mu; t) : t \in [0, T]\}$ by

$$\bar{\boldsymbol{\pi}}(\mu; t) := (\boldsymbol{\sigma}^\top(t))^{-1}\left(\frac{\boldsymbol{\Lambda}^{\bar{X}(\mu)H}(\mu; t)}{H(t)} + \bar{X}(\mu; t)\boldsymbol{\theta}(t)\right), \tag{4.9.25}$$

where $\boldsymbol{\Lambda}^{\bar{X}(\mu)H}(\mu; \cdot) \in L^2_{\mathrm{loc}}(\mathbf{W})$ is given by Theorem B.4.22, a martingale representation theorem, applied to $\bar{X}(\mu; \cdot)H$ and which is consequently $(\mathbb{P} \otimes \mathit{Leb})$-a.e. unique.

Then $\bar{\boldsymbol{\pi}}(\mu) \in \mathcal{A}$ and $\bar{X}(\mu; t) = X^{\bar{\boldsymbol{\pi}}(\mu)}(t)$ $(\mathbb{P} \otimes \mathit{Leb})$-a.e., for $X^{\bar{\boldsymbol{\pi}}(\mu)}$ the solution to the wealth equation (3.2.31) for the portfolio process $\bar{\boldsymbol{\pi}}(\mu)$.

In particular, from the equivalence of (4.8.2) and (4.8.5), we get

$$\begin{aligned}\mathrm{E}\left(J_\mu(X^{\bar{\boldsymbol{\pi}}(\mu)}(T))\right) &= \inf_{\boldsymbol{\pi}\in\mathcal{A}}\{\mathrm{E}\left(J_\mu(X^{\boldsymbol{\pi}}(T))\right)\} \\ &= -\inf_{(y,\boldsymbol{\lambda},\boldsymbol{\gamma})\in\mathbb{R}\times L^2(\mathbf{W})\times L^2(\mathcal{Q})}\left\{\tilde{\Psi}(\mu; y, \boldsymbol{\lambda}, \boldsymbol{\gamma})\right\} \\ &= -\tilde{\Psi}(\mu; \bar{y}(\mu), \bar{\boldsymbol{\lambda}}(\mu), \bar{\boldsymbol{\gamma}}(\mu)).\end{aligned} \tag{4.9.26}$$

Hence, substituting from (4.9.20), we obtain

$$\begin{aligned}\mathcal{L}(\mu; \bar{\boldsymbol{\pi}}(\mu)) &= \inf_{\boldsymbol{\pi}\in\mathcal{A}}\{\mathcal{L}(\mu; \boldsymbol{\pi})\} \\ &= -\inf_{(y,\boldsymbol{\lambda},\boldsymbol{\gamma})\in\mathbb{R}\times L^2(\mathbf{W})\times L^2(\mathcal{Q})}\left\{\tilde{\Psi}(\mu; y, \boldsymbol{\lambda}, \boldsymbol{\gamma})\right\} = -\tilde{\Psi}(\mu; \bar{y}(\mu), \bar{\boldsymbol{\lambda}}(\mu), \bar{\boldsymbol{\gamma}}(\mu)).\end{aligned} \tag{4.9.27}$$

As the Lagrange multiplier $\bar{\mu} \in \mathbb{R}$ exists by Theorem C.17.4, clearly Proposition 4.8.1 will also hold with $\mu = \bar{\mu}$. So $\bar{\boldsymbol{\pi}}(\bar{\mu}) \in \mathcal{A}$. In order to establish that $\bar{\boldsymbol{\pi}}(\bar{\mu})$ solves Problem 4.9.16, it remains to show

$$G(\bar{\boldsymbol{\pi}}(\bar{\mu})) = d \quad \text{and} \quad \mathrm{E}\left(\hat{J}\left(X^{\bar{\boldsymbol{\pi}}(\bar{\mu})}(T)\right)\right) = \hat{\mathcal{V}}. \tag{4.9.28}$$

However, since

$$\hat{\mathcal{V}} \overset{(4.9.16)}{=} \mathcal{L}(\bar{\mu};\bar{\boldsymbol{\pi}}(\bar{\mu})) \overset{(4.9.12)}{=} \mathrm{E}\left(\hat{J}\left(X^{\bar{\boldsymbol{\pi}}(\bar{\mu})}(T)\right)\right) + \bar{\mu}\left(G(\bar{\boldsymbol{\pi}}(\bar{\mu})) - d\right), \tag{4.9.29}$$

it is enough to show that $G(\bar{\boldsymbol{\pi}}(\bar{\mu})) = d$. We show this using a variational analysis on the functional $\tilde{\Psi}$.

First observe that

$$\begin{aligned}\hat{\mathcal{V}} &\overset{(4.9.16)}{=} \sup_{\mu\in\mathbb{R}}\{\mathcal{L}(\mu;\bar{\boldsymbol{\pi}}(\mu))\} \\ &\overset{(4.9.27)}{=} \sup_{\mu\in\mathbb{R}}\left\{-\inf_{(y,\boldsymbol{\lambda},\boldsymbol{\gamma})\in\mathbb{R}\times L^2(\mathbf{W})\times L^2(\mathcal{Q})}\left\{\tilde{\Psi}(\mu;y,\boldsymbol{\lambda},\boldsymbol{\gamma})\right\}\right\} \\ &= -\inf_{(\mu,y,\boldsymbol{\lambda},\boldsymbol{\gamma})\in\mathbb{R}\times\mathbb{R}\times L^2(\mathbf{W})\times L^2(\mathcal{Q})}\left\{\tilde{\Psi}(\mu;y,\boldsymbol{\lambda},\boldsymbol{\gamma})\right\},\end{aligned} \tag{4.9.30}$$

and also that

$$\hat{\mathcal{V}} \overset{(4.9.16)}{=} \mathcal{L}(\bar{\mu};\bar{\boldsymbol{\pi}}(\bar{\mu})) \overset{(4.9.27)}{=} -\tilde{\Psi}(\bar{\mu};\bar{y}(\bar{\mu}),\bar{\boldsymbol{\lambda}}(\bar{\mu}),\bar{\boldsymbol{\gamma}}(\bar{\mu})). \tag{4.9.31}$$

Equating (4.9.30) and (4.9.31), we get

$$\tilde{\Psi}(\bar{\mu};\bar{y}(\bar{\mu}),\bar{\boldsymbol{\lambda}}(\bar{\mu}),\bar{\boldsymbol{\gamma}}(\bar{\mu})) = \inf_{(\mu,y,\boldsymbol{\lambda},\boldsymbol{\gamma})\in\mathbb{R}\times\mathbb{R}\times L^2(\mathbf{W})\times L^2(\mathcal{Q})}\left\{\tilde{\Psi}(\mu;y,\boldsymbol{\lambda},\boldsymbol{\gamma})\right\}. \tag{4.9.32}$$

For $\epsilon \in (0,\infty)$ and $\rho \in \mathbb{R}$, set

$$\mu^\epsilon = \bar{\mu} + \epsilon\rho. \tag{4.9.33}$$

Then from (4.9.32), it is true that

$$0 \leq \frac{\tilde{\Psi}(\mu^\epsilon;\bar{y}(\bar{\mu}),\bar{\boldsymbol{\lambda}}(\bar{\mu}),\bar{\boldsymbol{\gamma}}(\bar{\mu})) - \tilde{\Psi}(\bar{\mu};\bar{y}(\bar{\mu}),\bar{\boldsymbol{\lambda}}(\bar{\mu}),\bar{\boldsymbol{\gamma}}(\bar{\mu}))}{\epsilon}. \tag{4.9.34}$$

Using the definition of the functional $\tilde{\Psi}$ given by (4.9.22) and substituting for b_{μ^ϵ} and c_{μ^ϵ} from (4.9.18), this becomes after some straightforward algebra,

$$0 \leq \rho\mathrm{E}\left(\frac{b_1\left(\bar{Y}(\bar{\mu};T) + b_{\bar{\mu}}\right)}{a}\right) + \rho d + \epsilon\rho^2\mathrm{E}\left(\frac{b_1^2}{2a}\right). \tag{4.9.35}$$

Letting $\epsilon \downarrow 0$, we get

$$0 \leq \rho\left(\mathrm{E}\left(\frac{b_1\left(\bar{Y}(\bar{\mu};T) + b_{\bar{\mu}}\right)}{a}\right) + d\right). \tag{4.9.36}$$

Since this holds for all $\rho \in \mathbb{R}$, we must have equality, that is

$$\mathrm{E}\left(\frac{b_1\left(\bar{Y}(\bar{\mu};T) + b_{\bar{\mu}}\right)}{a}\right) + d = 0. \tag{4.9.37}$$

From (4.9.24), we find

$$b_1 \bar{X}(\bar{\mu};T) = -\frac{b_1\left(\bar{Y}(\bar{\mu};T)+b_{\bar{\mu}}\right)}{a}, \tag{4.9.38}$$

and upon taking expectations,

$$\mathrm{E}\left(b_1\bar{X}(\bar{\mu};T)\right) = -\mathrm{E}\left(\frac{b_1\left(\bar{Y}(\bar{\mu};T)+b_{\bar{\mu}}\right)}{a}\right) \overset{(4.9.37)}{=} d. \tag{4.9.39}$$

We know from Proposition 4.8.1 that $\bar{X}(\bar{\mu};t) = X^{\bar{\boldsymbol{\pi}}(\bar{\mu})}(t)$ $(\mathbb{P}\otimes Leb)$-a.e., for $X^{\bar{\boldsymbol{\pi}}(\bar{\mu})}$ the solution to the wealth equation (3.2.31) for the portfolio process $\bar{\boldsymbol{\pi}}(\bar{\mu})$. Hence

$$G(\bar{\boldsymbol{\pi}}(\bar{\mu})) \overset{(4.9.4)}{=} \mathrm{E}\left(b_1 X^{\bar{\boldsymbol{\pi}}(\bar{\mu})}(T)\right) = \mathrm{E}\left(b_1\bar{X}(\bar{\mu};T)\right) \overset{(4.9.39)}{=} d, \tag{4.9.40}$$

and so the terminal wealth constraint is satisfied for the portfolio process $\bar{\boldsymbol{\pi}}(\bar{\mu})$. Hence we have solved Problem 4.9.16, which is the MVO problem with portfolio constraints and a terminal wealth constraint.

The results of this section are summarized by the following proposition.

Proposition 4.9.20. *Suppose that the market conditions (see Remark 3.2.11) Conditions 4.9.2-4.9.5, Condition 4.9.8 and the constraint qualification (Condition 4.9.11) are satisfied. For each $\mu \in \mathbb{R}$, define the functional*

$$\tilde{\Psi}(\mu;y,\boldsymbol{\lambda},\boldsymbol{\gamma}) := x_0 y + E\int_0^T \delta(\boldsymbol{\Theta}_Y(t))\,\mathrm{d}t + E\left(\frac{(Y(T)+b_\mu)^2}{2a}\right) - Ec_\mu, \tag{4.9.41}$$

for all $(y,\boldsymbol{\lambda},\boldsymbol{\gamma}) \in \mathbb{R}\times L^2(\mathbf{W})\times L^2(\mathcal{Q})$, where $Y := \Xi(y,\boldsymbol{\lambda},\boldsymbol{\gamma})$ and b_μ and c_μ are given by (4.9.18).

Then there exists a triple $(\bar{y}(\mu),\bar{\boldsymbol{\lambda}}(\mu),\bar{\boldsymbol{\gamma}}(\mu)) \in \mathbb{R}\times L^2(\mathbf{W})\times L^2(\mathcal{Q})$ which minimizes the functional $\tilde{\Psi}(\mu;\cdot,\cdot,\cdot)$ over $\mathbb{R}\times L^2(\mathbf{W})\times L^2(\mathcal{Q})$, that is

$$\tilde{\Psi}(\mu;\bar{y}(\mu),\bar{\boldsymbol{\lambda}}(\mu),\bar{\boldsymbol{\gamma}}(\mu)) = \inf_{(y,\boldsymbol{\lambda},\boldsymbol{\gamma})\in\mathbb{R}\times L^2(\mathbf{W})\times L^2(\mathcal{Q})}\left\{\tilde{\Psi}(\mu;y,\boldsymbol{\lambda},\boldsymbol{\gamma})\right\}. \tag{4.9.42}$$

Moreover, there exists some $\bar{\mu}\in\mathbb{R}$ which minimizes $\tilde{\Psi}(\mu;\cdot,\cdot,\cdot)$ over $\mathbb{R}$, so that

$$\tilde{\Psi}(\bar{\mu};\bar{y}(\bar{\mu}),\bar{\boldsymbol{\lambda}}(\bar{\mu}),\bar{\boldsymbol{\gamma}}(\bar{\mu})) = \inf_{(\mu,y,\boldsymbol{\lambda},\boldsymbol{\gamma})\in\mathbb{R}\times\mathbb{R}\times L^2(\mathbf{W})\times L^2(\mathcal{Q})}\left\{\tilde{\Psi}(\mu;y,\boldsymbol{\lambda},\boldsymbol{\gamma})\right\}. \tag{4.9.43}$$

Defining $\bar{Y}(\bar{\mu};t) := \Xi(\bar{y}(\bar{\mu}),\bar{\boldsymbol{\lambda}}(\bar{\mu}),\bar{\boldsymbol{\gamma}}(\bar{\mu}))(t)$, set

$$\bar{X}(\bar{\mu};t) := -\frac{1}{H(t)}E\left(\left(\frac{\bar{Y}(\bar{\mu};T)+b_{\bar{\mu}}}{a}\right)H(T)\,\middle|\,\mathcal{F}_t\right), \tag{4.9.44}$$

for $\{H(t)\}$ *the state price density process given by (4.6.1) and define the process* $\bar{\boldsymbol{\pi}}(\bar{\mu}) = \{\bar{\boldsymbol{\pi}}(\bar{\mu}; t) : t \in [0, T]\}$ *by*

$$\bar{\boldsymbol{\pi}}(\bar{\mu}; t) := (\boldsymbol{\sigma}^\top(t))^{-1} \left(\frac{\boldsymbol{\Lambda}^{\bar{X}(\bar{\mu})H}(\bar{\mu}; t)}{H(t)} + \bar{X}(\bar{\mu}; t)\boldsymbol{\theta}(t) \right), \tag{4.9.45}$$

where $\boldsymbol{\Lambda}^{\bar{X}(\bar{\mu})H}(\bar{\mu}; \cdot) \in L^2_{loc}(\mathbf{W})$ *is given by Theorem B.4.22, a martingale representation theorem, applied to* $\bar{X}(\bar{\mu}; \cdot)H$ *and which is consequently* $(\mathbb{P} \otimes Leb)$*-a.e. unique.*

Then $\bar{\boldsymbol{\pi}}(\bar{\mu}) \in \mathcal{A}$ *and* $\bar{X}(\bar{\mu}; t) = X^{\bar{\boldsymbol{\pi}}(\bar{\mu})}(t)$ $(\mathbb{P} \otimes Leb)$*-a.e., for* $X^{\bar{\boldsymbol{\pi}}(\bar{\mu})}$ *the solution to the wealth equation (3.2.31) for the portfolio process* $\bar{\boldsymbol{\pi}}(\bar{\mu})$*.*

In particular, the portfolio process $\bar{\boldsymbol{\pi}}(\bar{\mu}) \in \mathcal{A}$ *solves Problem 4.9.16, that is*

$$G(\bar{\boldsymbol{\pi}}(\bar{\mu})) = d \quad \text{and} \quad \hat{\mathcal{V}} = E\left(\hat{J}\left(X^{\bar{\boldsymbol{\pi}}(\bar{\mu})}(T)\right)\right). \tag{4.9.46}$$

4.10 Conclusion

In this chapter, we have showed existence of and characterized the solution to an MVO problem with convex portfolio constraints and a terminal wealth constraint in a regime-switching model. The existence and characterization of the MVO problem without a terminal wealth constraint is given by Proposition 4.8.1. The existence and characterization of the MVO problem with a terminal wealth constraint is given by Proposition 4.9.20. In particular, we note that the extension of Proposition 4.8.1 to include the terminal wealth constraint, resulting in Proposition 4.9.20, was not affected by the regime-switching model.

Chapter 5

Conclusion and suggestions for further work

The goal of the thesis is to establish the existence and characterization of the solution to a mean-variance portfolio optimization with portfolio constraints. The market model in which the problem is set undergoes regime-switching. Unlike most regime-switching models where the market parameters are Markov-modulated, we allow the market parameters to be truly random processes. This is a significant generalization of the regime-switching models, since it allows for stochasticity within the market regimes.

Our approach to the problem is based on a convex duality method in Labbé and Heunis [33]. The method proves to be highly successful. There are two key elements to its success.

The first is that it systematically synthesizes the dual problem from the primal problem. The constraints on the MVO problem are encoded in the primal cost functional as penalty functions. The dual problem is obtained as the sum of the convex conjugates of these penalty functions and the value of the MVO problem.

Secondly, as a result of the synthesis, we find necessary and sufficient conditions on the solutions to the dual and primal problems. These conditions strongly suggest how the candidate solution to the primal problem should be constructed.

The effect of the regime-switching model indirectly impacts the method through the martingale representation theorems proved in Appendix B. As a consequence of this, both the solution to the dual problem and the candidate solution to the primal problem have a martingale part which is the sum of a stochastic integral with respect to the Brownian motion and a stochastic integral with respect to the canonical martingales of the Markov chain. We show that the martingale part of the candidate solution to the primal problem consists only of the stochastic integral with respect to the Brownian motion.

The martingale representation theorems proved in Appendix B are of interest in their own right. We found martingale representation theorems in the literature for filtrations generated by general jump processes, and one for a filtration generated by both a Brownian motion and a standard Poisson process. However, we did not find a martingale representation theorem for our particular case. Although there are very general martingale representation theorems in the literature (for example Jacod and Shiryaev [25], Section III.4c), it is not obvious how to extract our desired martingale representation theorem from them.

The work in this thesis demonstrates the power of the convex duality method. We show the existence of a solution to the MVO problem with portfolio constraints in a regime-switching model and, furthermore, we characterize it. This suggests that more complicated problems, such as the utility maximization problem in a regime-switching model, can be tackled using convex duality.

However, the preeminence of the convex duality method would be more assured if we could use it to find an actual solution to an MVO problem, especially one involving portfolio constraints. Zhou and Yin [53] found an explicit solution to the MVO problem without portfolio constraints, in a regime-switching model with Markov-modulated coefficients. We have not yet shown an explicit solution in our regime-switching model. However, this is certainly an interesting area in which to do further research.

Other extensions to the MVO problem would be to include transaction costs as well as portfolio insurance on the terminal wealth. This latter is called a "state constraint" and such constraints, in conjunction with the regular portfolio constraints we have considered in this thesis, make the problem very challenging.

Another topic of interest would be to extend the martingale representation theorem from locally square-integrable martingales to local martingales. Indeed, this would probably be required in order to apply the convex duality method to a utility maximization problem in a regime-switching model.

Appendix A

Supplementary results

A.1 Complement to Chapter 3

Proposition A.1.1. *Suppose* $(X_0, \dot{X}, \mathbf{\Lambda}^X, \mathbf{\Gamma}^X) \in \mathbb{B}$ *and* $(Y_0, \dot{Y}, \mathbf{\Lambda}^Y, \mathbf{\Gamma}^Y) \in \mathbb{B}$*. If for all* $t \in [0, T]$*,*

$$\begin{aligned} &X_0 + \int_0^t \dot{X}(\tau)\,\mathrm{d}\tau + \int_0^t \left(\mathbf{\Lambda}^X\right)^\top(\tau)\,\mathrm{d}\mathbf{W}(\tau) + \sum_{i,j=1}^{D} \int_0^t \Gamma_{ij}^X(\tau)\,\mathrm{d}\mathcal{Q}_{ij}(\tau) \\ &= Y_0 + \int_0^t \dot{Y}(\tau)\,\mathrm{d}\tau + \int_0^t \left(\mathbf{\Lambda}^Y\right)^\top(\tau)\,\mathrm{d}\mathbf{W}(\tau) + \sum_{i,j=1}^{D} \int_0^t \Gamma_{ij}^Y(\tau)\,\mathrm{d}\mathcal{Q}_{ij}(\tau) \quad a.s. \end{aligned} \tag{A.1.1}$$

then $X_0 = Y_0$*,* $\dot{X} = \dot{Y}$*,* $\mathbf{\Lambda}^X = \mathbf{\Lambda}^Y$ *(*$\mathbb{P} \otimes$ *Leb)-a.e. and* $\mathbf{\Gamma}^X = \mathbf{\Gamma}^Y$ $\nu_{[\mathcal{Q}]}$*-a.e.*

Proof. Set $t = 0$ in (A.1.1) to obtain immediately $X_0 = Y_0$. Then we have a.s.

$$\int_0^t (\mathbf{\Lambda}^X - \mathbf{\Lambda}^Y)^\top(\tau)\,\mathrm{d}\mathbf{W}(\tau) = \int_0^t (\dot{Y}(\tau) - \dot{X}(\tau))\,\mathrm{d}\tau + \sum_{i,j=1}^{D} \int_0^t (\Gamma_{ij}^Y - \Gamma_{ij}^X)(\tau)\,\mathrm{d}\mathcal{Q}_{ij}(\tau), \tag{A.1.2}$$

for all $t \in [0, T]$. The right-hand side of (A.1.2) is the sum of a 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.1.2) 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 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 [46], Theorem IV.30.4). Thus from (A.1.2) we obtain

$$\int_0^t (\mathbf{\Lambda}^X - \mathbf{\Lambda}^Y)^\top(\tau)\,\mathrm{d}\mathbf{W}(\tau) = 0 \text{ a.s.}, \qquad \forall t \in [0, T]. \tag{A.1.3}$$

Evaluating at time $t = T$, squaring, taking expectations and using the Itô isometry, we get from (A.1.3),

$$0 = \mathrm{E}\left(\int_0^T (\mathbf{\Lambda}^X - \mathbf{\Lambda}^Y)^\top(\tau)\,\mathrm{d}\mathbf{W}(\tau)\right)^2 = \mathrm{E}\int_0^T \|\left(\mathbf{\Lambda}^X - \mathbf{\Lambda}^Y\right)(\tau)\|^2\,\mathrm{d}\tau. \tag{A.1.4}$$

This implies that $\mathbf{\Lambda}^X = \mathbf{\Lambda}^Y$ $(\mathbb{P}\otimes Leb)$-a.e. From (A.1.2) and (A.1.3) we then obtain

$$\int_0^t (\dot{X}(\tau) - \dot{Y}(\tau))\,\mathrm{d}\tau = \sum_{i,j=1}^D \int_0^t (\Gamma_{ij}^Y - \Gamma_{ij}^X)(\tau)\,\mathrm{d}\mathcal{Q}_{ij}(\tau) \text{ a.s.} \tag{A.1.5}$$

The left-hand side of (A.1.5) 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

$$\sum_{i,j=1}^D \int_0^t (\Gamma_{ij}^Y - \Gamma_{ij}^X)(\tau)\,\mathrm{d}\mathcal{Q}_{ij}(\tau) = 0 \text{ a.s.}, \qquad \forall t \in [0, T]. \tag{A.1.6}$$

Evaluating at time $t = T$, squaring, taking expectations and using the Itô isometry, we get from (A.1.6),

$$0 = \mathrm{E}\left(\sum_{i,j=1}^D \int_0^t (\Gamma_{ij}^Y - \Gamma_{ij}^X)(\tau)\,\mathrm{d}\mathcal{Q}_{ij}(\tau)\right)^2 = \sum_{i,j=1}^D \mathrm{E}\int_0^T |\left(\Gamma_{ij}^Y - \Gamma_{ij}^X\right)(\tau)|^2\,\mathrm{d}[\mathcal{Q}_{ij}](\tau) \tag{A.1.7}$$

This implies that $\Gamma_{ij}^Y = \Gamma_{ij}^X$ $\nu_{[\mathcal{Q}_{ij}]}$-a.e for each $i, j = 1, \ldots, D$, $i \neq j$. By virtue of $\mathbf{\Gamma}^X, \mathbf{\Gamma}^Y \in L^2(\mathcal{Q})$, we have $\Gamma_{ii}^Y = \Gamma_{ii}^X = 0$ $(\mathbb{P}\otimes Leb)$-a.e. for each $i = 1, \ldots, D$. Hence, $\mathbf{\Gamma}^X = \mathbf{\Gamma}^Y$ $\nu_{[\mathcal{Q}]}$-a.e.

Finally, from (A.1.5) and (A.1.6), we are left with

$$\int_0^t (\dot{X}(\tau) - \dot{Y}(\tau))\,\mathrm{d}\tau = 0, \tag{A.1.8}$$

so we must have $\dot{X}(t) = \dot{Y}(t)$ for all $t \in [0, T]$, in other words $\dot{X} = \dot{Y}$. □

Lemma A.1.2. *For all $X \equiv (X_0, \dot{X}, \mathbf{\Lambda}^X, \mathbf{\Gamma}^X) \in \mathbb{B}$,*

$$E\left(\sup_{t\in[0,T]} |X(t)|^2\right) < \infty. \tag{A.1.9}$$

Proof. From (3.2.37), we have a.s.

$$X(t) = X_0 + \int_0^t \dot{X}(\tau)\,\mathrm{d}\tau + \sum_{n=1}^N \int_0^t \Lambda_n^X(\tau)\,\mathrm{d}W_n(\tau) + \sum_{i,j=1}^D \int_0^t \Gamma_{ij}^X(\tau)\,\mathrm{d}\mathcal{Q}_{ij}(\tau). \tag{A.1.10}$$

Then squaring both sides and using the identity

$$(z_1 + z_2 + z_3 + z_4)^2 \leq 4z_1^2 + 4z_2^2 + 4z_3^2 + 4z_4^2, \quad \forall z_1, z_2, z_3, z_4 \in \mathbb{R}, \tag{A.1.11}$$

and $|\int_0^t \dot{X}(\tau)\,\mathrm{d}\tau| \leq \int_0^t |\dot{X}(\tau)|\,\mathrm{d}\tau$, we get

$$\begin{aligned} |X(t)|^2 \leq 4|X_0|^2 + 4\left(\int_0^t |\dot{X}(\tau)|\,\mathrm{d}\tau\right)^2 + 4\left|\sum_{n=1}^{N}\int_0^t \Lambda_n^X(\tau)\,\mathrm{d}W_n(\tau)\right|^2 \\ + 4\left|\sum_{i,j=1}^{D}\int_0^t \Gamma_{ij}^X(\tau)\,\mathrm{d}\mathcal{Q}_{ij}(\tau)\right|^2. \end{aligned} \tag{A.1.12}$$

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

$$\begin{aligned} &\mathrm{E}\left(\sup_{t\in[0,T]} |X(t)|^2\right) \\ &\leq 4|X_0|^2 + 4\mathrm{E}\left(\int_0^T |\dot{X}(\tau)|\,\mathrm{d}\tau\right)^2 + 4\mathrm{E}\left(\sup_{t\in[0,T]}\left|\sum_{n=1}^{N}\int_0^t \Lambda_n^X(\tau)\,\mathrm{d}W_n(\tau)\right|^2\right) \\ &\quad + 4\mathrm{E}\left(\sup_{t\in[0,T]}\left|\sum_{i,j=1}^{D}\int_0^t \Gamma_{ij}^X(\tau)\,\mathrm{d}\mathcal{Q}_{ij}(\tau)\right|^2\right). \end{aligned} \tag{A.1.13}$$

Applying Doob's L^2-inequality to the last two terms on the right-hand side, both of which are martingales since $\boldsymbol{\Lambda}^X \in L^2(\mathbf{W})$ and $\boldsymbol{\Gamma}^X \in L^2(\mathbf{W})$, followed by the Itô isometry we get

$$\begin{aligned} &\mathrm{E}\left(\sup_{t\in[0,T]} |X(t)|^2\right) \\ &\leq 4|X_0|^2 + 4\mathrm{E}\left(\int_0^T |\dot{X}(\tau)|\,\mathrm{d}\tau\right)^2 + 16\mathrm{E}\sum_{n=1}^{N}\int_0^T |\Lambda_n^X(\tau)|^2\,\mathrm{d}\tau \\ &\quad + 16\mathrm{E}\sum_{i,j=1}^{D}\int_0^T |\Gamma_{ij}^X(\tau)|^2\,\mathrm{d}[\mathcal{Q}_{ij}](\tau). \end{aligned} \tag{A.1.14}$$

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

$$\mathrm{E}\left(\sup_{t\in[0,T]} |X(t)|^2\right) < \infty. \tag{A.1.15}$$

□

Proposition A.1.3. *For the wealth process* $\{X^{\boldsymbol{\pi}}(t) : t \in [0, T]\}$ *which is the solution to the wealth equation (3.2.31), we have*

$$X^{\boldsymbol{\pi}} \in \mathbb{A} \text{ if and only if } \boldsymbol{\pi} \in L^2(\mathbf{W}). \tag{A.1.16}$$

Proof. First we show that $X^{\boldsymbol{\pi}} \in \mathbb{A}$ implies that $\boldsymbol{\pi} \in L^2(\mathbf{W})$. From the wealth equation (3.2.31) for the portfolio process $\boldsymbol{\pi}$, we have

$$X^{\boldsymbol{\pi}} \equiv (x_0, rX^{\boldsymbol{\pi}} + \boldsymbol{\pi}^\top \boldsymbol{\sigma}\boldsymbol{\theta}, \boldsymbol{\sigma}^\top \boldsymbol{\pi}) \in \mathbb{R} \times L_{21} \times L^2(\mathbf{W}). \tag{A.1.17}$$

Therefore, $\boldsymbol{\sigma}^\top \boldsymbol{\pi} \in L^2(\mathbf{W})$. From the uniform boundedness of $\boldsymbol{\sigma}$ (Condition 3.2.9), it follows that $\boldsymbol{\pi} \in L^2(\mathbf{W})$.

Next we show that $\boldsymbol{\pi} \in L^2(\mathbf{W})$ implies that $X^{\boldsymbol{\pi}} \in \mathbb{A}$.

First note that by the uniform boundedness of r assumed in Condition 3.2.5, the bank account price process $S_0(t) = \exp\{\int_0^t r(\tau)\, \mathrm{d}\tau\}$, given by (3.2.16), 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 (3.2.32), we get

$$|X^{\boldsymbol{\pi}}(t)| \leq \kappa_{S_0} \left\{ x_0 + \int_0^t \boldsymbol{\pi}^\top(\tau)\boldsymbol{\sigma}(\tau)\boldsymbol{\theta}(\tau)\, \mathrm{d}\tau + \int_0^t \boldsymbol{\pi}^\top(\tau)\boldsymbol{\sigma}(\tau)\, \mathrm{d}\mathbf{W}(\tau) \right\}. \tag{A.1.18}$$

Then using the identity

$$|z_1 + z_2 + z_3|^2 \leq 3z_1^2 + 3z_2^2 + 3z_3^2, \quad \forall z_1, z_2, z_3 \in \mathbb{R} \tag{A.1.19}$$

and $|\int f\, \mathrm{d}\mu|^2 \leq \int |f|^2\, \mathrm{d}\mu$, we get from (A.1.18) that

$$\begin{aligned} \mathrm{E}\left| \sup_{t \in [0,T]} X^{\boldsymbol{\pi}}(t) \right|^2 \leq 3\kappa_{S_0}^2 \Bigg\{ & x_0^2 + \mathrm{E}\left(\sup_{t\in[0,T]} \int_0^t \left| \boldsymbol{\pi}^\top(\tau)\boldsymbol{\sigma}(\tau)\boldsymbol{\theta}(\tau) \right|^2 \mathrm{d}\tau \right) \\ & + \mathrm{E}\left(\sup_{t\in[0,T]} \left| \int_0^t \boldsymbol{\pi}^\top(\tau)\boldsymbol{\sigma}(\tau)\, \mathrm{d}\mathbf{W}(\tau) \right|^2 \right) \Bigg\}. \end{aligned} \tag{A.1.20}$$

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.1.20) occurs at $t = T$, we get

$$\begin{aligned} & \mathrm{E}\left| \sup_{t\in[0,T]} X^{\boldsymbol{\pi}}(t) \right|^2 \\ & \overset{\text{Doob}}{\leq} 3\kappa_{S_0}^2 \left\{ x_0^2 + \mathrm{E}\int_0^T \left| \boldsymbol{\pi}^\top(\tau)\boldsymbol{\sigma}(\tau)\boldsymbol{\theta}(\tau) \right|^2 \mathrm{d}\tau + 4\mathrm{E}\left(\int_0^T \boldsymbol{\pi}^\top(\tau)\boldsymbol{\sigma}(\tau)\, \mathrm{d}\mathbf{W}(\tau) \right)^2 \right\} \\ & \overset{\text{Itô}}{=} 3\kappa_{S_0}^2 \left\{ x_0^2 + \mathrm{E}\int_0^T \left| \boldsymbol{\pi}^\top(\tau)\boldsymbol{\sigma}(\tau)\boldsymbol{\theta}(\tau) \right|^2 \mathrm{d}\tau + 4\mathrm{E}\int_0^T \left\| \boldsymbol{\sigma}^\top(\tau)\boldsymbol{\pi}(\tau) \right\|^2 \mathrm{d}\tau \right\}. \end{aligned} \tag{A.1.21}$$

Using the bounds in (3.2.26) and (3.2.28), and the assumption that $\boldsymbol{\pi} \in L^2(\mathbf{W})$, we obtain

$$\mathrm{E}\int_0^T |\boldsymbol{\pi}^\top(t)\boldsymbol{\sigma}(t)\boldsymbol{\theta}(t)|^2\, \mathrm{d}t \leq \kappa_{\boldsymbol{\sigma}}^2\, \kappa_{\boldsymbol{\theta}}^2\, \mathrm{E}\int_0^T \|\boldsymbol{\pi}(t)\|^2\, \mathrm{d}t < \infty \tag{A.1.22}$$

and

$$\mathrm{E}\int_0^T \|\boldsymbol{\sigma}^\top(t)\boldsymbol{\pi}(t)\|^2\,\mathrm{d}t \le \kappa_{\boldsymbol{\sigma}}^2\,\mathrm{E}\int_0^T \|\boldsymbol{\pi}(t)\|^2\,\mathrm{d}t < \infty. \tag{A.1.23}$$

Applying these bounds to (A.1.20), we get $\mathrm{E}\left(\sup_{t\in[0,T]}|X^{\boldsymbol{\pi}}(t)|^2\right) < \infty$. It follows from this square-integrability, (A.1.22) and the uniform boundedness of r that $rX^{\boldsymbol{\pi}} + \boldsymbol{\pi}^\top\boldsymbol{\sigma\theta} \in L_{21}$. From (A.1.23), we have $\boldsymbol{\sigma}^\top\boldsymbol{\pi} \in L^2(\mathbf{W})$. Thus $X^{\boldsymbol{\pi}} \equiv (x_0, rX^{\boldsymbol{\pi}} + \boldsymbol{\pi}^\top\boldsymbol{\sigma\theta}, \boldsymbol{\pi}^\top\boldsymbol{\sigma}) \in \mathbb{A}$. □

Lemma A.1.4. *The value of the problem, given by (3.2.48), is such that* $-\infty < \mathcal{V} < \infty$.

Proof. First we show that $\mathcal{V} < \infty$. Since $\mathbf{0} \in K$ then $\mathcal{A} \neq \emptyset$. Choose $\boldsymbol{\pi} \in \mathcal{A}$. Then from Proposition A.1.3, the solution $X^{\boldsymbol{\pi}}$ of the wealth equation (3.2.31) for the portfolio process $\boldsymbol{\pi}$ is such that $X^{\boldsymbol{\pi}} \in \mathbb{A}$. Using the bounds on the random variable a (see Condition 3.2.26), the square-integrability of $X^{\boldsymbol{\pi}}$ (Lemma A.1.2) and Hölder's inequality, we find,

$$\begin{aligned}
\mathrm{E}\left(J(X^{\boldsymbol{\pi}}(T))\right) &\overset{(3.2.43)}{=} \frac{1}{2}\mathrm{E}\left(a\left(X^{\boldsymbol{\pi}}(T)\right)^2\right) + \mathrm{E}\left(bX^{\boldsymbol{\pi}}(T)\right) + \mathrm{E}c \\
&\le \frac{1}{2}\sup_{\omega\in\Omega}\{a(\omega)\}\,\mathrm{E}\left(\left(X^{\boldsymbol{\pi}}(T)\right)^2\right) + \left(\mathrm{E}b^2\right)^{\frac{1}{2}}\left(\mathrm{E}\left(X^{\boldsymbol{\pi}}(T)^2\right)\right)^{\frac{1}{2}} + \mathrm{E}c \\
&< \infty.
\end{aligned} \tag{A.1.24}$$

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

To show that $\mathcal{V} > -\infty$, we show that $\mathrm{E}\left(J(X^{\boldsymbol{\pi}}(T))\right)$ is bounded from below. Using the strict positivity of the random variable a (see Condition 3.2.26), we get

$$\begin{aligned}
\mathrm{E}\left(J(X^{\boldsymbol{\pi}}(T))\right) &= \frac{1}{2}\mathrm{E}\left(a\left(X^{\boldsymbol{\pi}}(T)\right)^2 + 2bX^{\boldsymbol{\pi}}(T)\right) + \mathrm{E}c \\
&= \frac{1}{2}\mathrm{E}\left(a\left(X^{\boldsymbol{\pi}}(T) + \frac{b}{a}\right)^2 - \left(\frac{b^2}{a}\right)\right) + \mathrm{E}c \\
&\ge \frac{1}{2}\inf_{\omega\in\Omega}\{a(\omega)\}\,\mathrm{E}\left(\left(X^{\boldsymbol{\pi}}(T) + \frac{b}{a}\right)^2\right) - \mathrm{E}\left(\frac{b^2}{2a}\right) + \mathrm{E}c \\
&\ge -\mathrm{E}\left(\frac{b^2}{2a}\right) + \mathrm{E}c.
\end{aligned} \tag{A.1.25}$$

Taking the infimum over $\boldsymbol{\pi} \in \mathcal{A}$, we have that $\mathcal{V} \ge -\mathrm{E}\left(\frac{b^2}{2a}\right) + \mathrm{E}c > -\infty$. □

A.2 Complement to Chapter 4

Proposition A.2.1. *The penalty function* l_1*, given by (4.1.15), is* $\mathcal{P}^\star \times \mathcal{B}(\mathbb{R}^{N+2})$*–measurable, where* $\mathcal{B}(\mathbb{R}^{N+2})$ *is the Borel* σ*–algebra on the set* $\mathbb{R}^{N+2}$*. The function* $l_T(X(T))$*, given by (4.1.16) is* $\mathcal{F}_T$*–measurable.*

Proof. We want to show that the penalty function l_1, given by (4.1.15), is $\mathcal{P}^\star \times \mathcal{B}(\mathbb{R}^{N+2})$–measurable. Now $l_1 = 0 \cdot \chi[A \cap B] + \infty \cdot (1 - \chi[A \cap B])$, where χ is the zero-one indicator function and

$$A := \left\{(\omega, t, x, \nu, \boldsymbol{\lambda}) \in \Omega \times [0, T] \times \mathbb{R} \times \mathbb{R} \times \mathbb{R}^N \,\middle|\, \nu = r(\omega, t)x + \boldsymbol{\lambda}^\top \boldsymbol{\theta}(\omega, t)\right\} \tag{A.2.1}$$

and

$$B := \left\{(\omega, t, x, \nu, \boldsymbol{\lambda}) \in \Omega \times [0, T] \times \mathbb{R} \times \mathbb{R} \times \mathbb{R}^N \,\middle|\, (\boldsymbol{\sigma}^\top)^{-1}(\omega, t)\boldsymbol{\lambda} \in K\right\}. \tag{A.2.2}$$

The set A is $\mathcal{P}^\star \times \mathcal{B}(\mathbb{R}^{N+2})$–measurable since r and $\boldsymbol{\theta}$ are $\mathcal{P}^\star$–measurable. The set B is $\mathcal{P}^\star \times \mathcal{B}(\mathbb{R}^{N+2})$–measurable since $\boldsymbol{\sigma}$ is $\mathcal{P}^\star$-measurable and the set K is closed (by Condition 3.2.31). It follows immediately that l_1 is $\mathcal{P}^\star \times \mathcal{B}(\mathbb{R}^{N+2})$-measurable.

Evaluating the function l_T, which is given by (4.1.16), at $X(T)$, we have,

$$l_T(X(T)) = \frac{1}{2}aX(T)^2 + bX(T) + c. \tag{A.2.3}$$

From Conditions 3.2.26, 3.2.27 and 3.2.28, $l_T(X(T))$ is a combination of the sum and product of $\mathcal{F}_T$–measurable functions, and hence is itself $\mathcal{F}_T$–measurable. □

Proposition A.2.2. *The dual function m_1, given by (4.2.2), is $\mathcal{P}^\star \times \mathcal{B}(\mathbb{R}^{N+2})$–measurable, where $\mathcal{B}(\mathbb{R}^{N+2})$ is the Borel σ–algebra on the set $\mathbb{R}^{N+2}$. The dual function $m_T(X(T))$, given by (4.2.3), is $\mathcal{F}_T$–measurable.*

Proof. To show that m_1 is $\mathcal{P}^\star \times \mathcal{B}(\mathbb{R}^{N+2})$–measurable, we use (4.4.11). This gives

$$m_1(\omega, t, y, s, \boldsymbol{\xi}) = \begin{cases} \delta\left(-\boldsymbol{\sigma}(\omega, t)(\boldsymbol{\theta}(\omega, t)y + \boldsymbol{\xi})\right) & \text{if } s + r(\omega, t)y = 0 \\ \infty & \text{otherwise,} \end{cases} \tag{A.2.4}$$

for all $(\omega, t, y, s, \boldsymbol{\xi}) \in \Omega \times [0, T] \times \mathbb{R} \times \mathbb{R} \times \mathbb{R}^N$. By Remark 4.4.5, the support function δ is a lower semi-continuous function and thus measurable. Combining this with the $\mathcal{P}^\star$–measurability of r, $\boldsymbol{\sigma}$ and $\boldsymbol{\theta}$, we conclude that m_1 is $\mathcal{P}^\star \times \mathcal{B}(\mathbb{R}^{N+2})$–measurable.

The $\mathcal{F}_T$–measurability of $m_T(X(T))$ follows from (4.4.12); it is a simple combination of $\mathcal{F}_T$–measurable functions. □

The following proposition is a minor adaptation of Bismut [3], Proposition I-1.

Proposition A.2.3. *For any $X \equiv (X_0, \dot{X}, \boldsymbol{\Lambda}^X, \boldsymbol{\Gamma}^X) \in \mathbb{B}$ and $Y \equiv (Y_0, \dot{Y}, \boldsymbol{\Lambda}^Y, \boldsymbol{\Gamma}^Y) \in \mathbb{B}$, the process $\{\mathbb{M}(X, Y)(t) : t \in [0, T]\}$ defined by*

$$\begin{aligned} \mathbb{M}(X, Y)(t) :=& X(t)Y(t) - X_0Y_0 - \int_0^t \left(\dot{X}(\tau)Y(\tau) + X(\tau)\dot{Y}(\tau)\right) \mathrm{d}\tau \\ &- \sum_{n=1}^{N} \int_0^t \Lambda_n^X(\tau)\Lambda_n^Y(\tau)\,\mathrm{d}\tau - \sum_{i,j=1}^{D} \int_0^t \Gamma_{ij}^X(\tau)\Gamma_{ij}^Y(\tau)\,\mathrm{d}[\mathcal{Q}_{ij}](\tau) \end{aligned} \tag{A.2.5}$$

is such that $\mathbb{M}(X, Y) \in \mathcal{M}_0(\{\mathcal{F}_t\}, \mathbb{P})$.

Proof. Fix $X \equiv (X_0, \dot{X}, \mathbf{\Lambda}^X, \mathbf{\Gamma}^X) \in \mathbb{B}$ and $Y \equiv (Y_0, \dot{Y}, \mathbf{\Lambda}^Y, \mathbf{\Gamma}^Y) \in \mathbb{B}$. Then from Bismut [3], Proposition I-1, we have that the process $\{\mathbb{N}(X,Y)(t) : t \in [0,T]\}$ defined by

$$\begin{aligned}\mathbb{N}(X,Y)(t) =& X(t)Y(t) - X_0Y_0 - \int_0^t \left(\dot{X}(\tau)Y(\tau) + X(\tau)\dot{Y}(\tau)\right) \mathrm{d}\tau \\ &- \sum_{n=1}^{N} \int_0^t \Lambda_n^X(\tau)\Lambda_n^Y(\tau)\,\mathrm{d}\tau - \sum_{i,j=1}^{D} \int_0^t \Gamma_{ij}^X(\tau)\Gamma_{ij}^Y(\tau)\,\mathrm{d}\langle \mathcal{Q}_{ij}\rangle(\tau)\end{aligned} \tag{A.2.6}$$

is such that $\mathbb{N}(X,Y) \in \mathcal{M}_0(\{\mathcal{F}_t\}, \mathbb{P})$. Subtracting (A.2.5) from (A.2.6), we get

$$\begin{aligned}\mathbb{N}(X,Y)(t) - \mathbb{M}(X,Y)(t) =& \sum_{i,j=1}^{D} \int_0^t \Gamma_{ij}^X(\tau)\Gamma_{ij}^Y(\tau)\,\mathrm{d}[\mathcal{Q}_{ij}](\tau) \\ &- \sum_{i,j=1}^{D} \int_0^t \Gamma_{ij}^X(\tau)\Gamma_{ij}^Y(\tau)\,\mathrm{d}\langle \mathcal{Q}_{ij}\rangle(\tau).\end{aligned} \tag{A.2.7}$$

We show that the right-hand side of (A.2.7) is in $\mathcal{M}_0(\{\mathcal{F}_t\}, \mathbb{P})$.

Set

$$A^X(t) := \sum_{i,j=1}^{D} \int_0^t \Gamma_{ij}^X(\tau)\ \mathrm{d}\mathcal{Q}_{ij}(\tau) \tag{A.2.8}$$

and

$$A^Y(t) := \sum_{i,j=1}^{D} \int_0^t \Gamma_{ij}^Y(\tau)\ \mathrm{d}\mathcal{Q}_{ij}(\tau). \tag{A.2.9}$$

Since $\mathbf{\Gamma}^X, \mathbf{\Gamma}^Y \in L^2(\mathcal{Q})$, we have $A^X, A^Y \in \mathcal{M}_0^2(\{\mathcal{F}_t\}, \mathbb{P})$, so their angle-bracket quadratic co-variation process exists. For all $t \in [0,T]$, we have

$$\begin{aligned}[A^X, A^Y](t) - \langle A^X, A^Y\rangle(t) =& \sum_{i,j=1}^{D} \int_0^t \Gamma_{ij}^X(\tau)\Gamma_{ij}^Y(\tau)\,\mathrm{d}[\mathcal{Q}_{ij}](\tau) \\ &- \sum_{i,j=1}^{D} \int_0^t \Gamma_{ij}^X(\tau)\Gamma_{ij}^Y(\tau)\,\mathrm{d}\langle \mathcal{Q}_{ij}\rangle(\tau).\end{aligned} \tag{A.2.10}$$

From Theorem C.9.1,

$$A^XA^Y - \langle A^X, A^Y\rangle \in \mathcal{M}_0(\{\mathcal{F}_t\}, \mathbb{P}), \tag{A.2.11}$$

and from Theorem C.9.4

$$A^XA^Y - [A^X, A^Y] \in \mathcal{M}_0(\{\mathcal{F}_t\}, \mathbb{P}). \tag{A.2.12}$$

Subtracting (A.2.12) from (A.2.11), we get

$$[A^X, A^Y] - \langle A^X, A^Y\rangle \in \mathcal{M}_0(\{\mathcal{F}_t\}, \mathbb{P}). \tag{A.2.13}$$

Then the right-hand side of (A.2.10) is also a martingale and hence, from (A.2.7), $\mathbb{N}(X,Y) - \mathbb{M}(X,Y) \in \mathcal{M}_0(\{\mathcal{F}_t\}, \mathbb{P})$. As $\mathbb{N}(X,Y) \in \mathcal{M}_0(\{\mathcal{F}_t\}, \mathbb{P})$, it follows that $\mathbb{M}(X,Y) \in \mathcal{M}_0(\{\mathcal{F}_t\}, \mathbb{P})$. □

The next lemma is called Gronwall's inequality and the proof can be found in Revuz and Yor [43], Appendix, Section 1. It is used in the proof of Lemma A.2.6.

Lemma A.2.4. Gronwall's inequality *If ϕ is a positive, locally bounded and Borel-measurable function on $\mathbb{R}$ such that for some constants a, b, and for every $t \in [0, \infty)$, we have*

$$\phi(t) \le a + b \int_0^t \phi(s)\,\mathrm{d}s, \tag{A.2.14}$$

then

$$\phi(t) \le a \exp\{bt\}. \tag{A.2.15}$$

Lemma A.2.5. *The space $L^2(\mathcal{Q})$ is a Hilbert space with inner product*

$$\langle \mathbf{\Gamma}^1, \mathbf{\Gamma}^2 \rangle_{L^2(\mathcal{Q})} := \sum_{\substack{i,j=1 \\ i \neq j}}^{D} E \int_0^T \Gamma^1_{ij}(t) \Gamma^2_{ij}(t)\,\mathrm{d}[\mathcal{Q}_{ij}](t), \tag{A.2.16}$$

for all $\mathbf{\Gamma}^1, \mathbf{\Gamma}^2 \in L^2(\mathcal{Q})$.

Proof. Define for all $i, j = 1, \ldots, D$, $i \neq j$,

$$L^2(\mathcal{Q}_{ij}) := \left\{ \Gamma : \Omega \times [0,T] \to \mathbb{R} \middle| \Gamma \in \mathcal{P}^\star \text{ and } \mathrm{E} \int_0^T |\Gamma(t)|^2\,\mathrm{d}[\mathcal{Q}_{ij}](t) < \infty \right\}, \tag{A.2.17}$$

and define an inner product $\langle\ ,\ \rangle_{L^2(\mathcal{Q}_{ij})}$ on $L^2(\mathcal{Q}_{ij})$ by the recipe

$$\langle \Gamma^1, \Gamma^2 \rangle_{L^2(\mathcal{Q}_{ij})} := \mathrm{E} \int_0^T \Gamma^1(t) \Gamma^2(t)\,\mathrm{d}[\mathcal{Q}_{ij}](t), \tag{A.2.18}$$

for all $\Gamma^1, \Gamma^2 \in L^2(\mathcal{Q}_{ij})$. Recalling the positive finite measure $\nu_{[\mathcal{Q}_{ij}]}$ defined by (3.2.12), from Jacod and Shiryaev [25], Subsection 2(b), page 48, we have

$$L^2(\mathcal{Q}_{ij}) \equiv L^2(\Omega \times [0,T], \mathcal{P}^\star, \nu_{[\mathcal{Q}_{ij}]}). \tag{A.2.19}$$

Since L^2-spaces are Hilbert spaces, then $L^2(\Omega \times [0,T], \mathcal{P}^\star, \nu_{[\mathcal{Q}_{ij}]})$ is a Hilbert space with inner product

$$\langle \Gamma^1, \Gamma^2 \rangle_{ij} := \int_{\Omega \times [0,T]} \Gamma^1 \Gamma^2\,\mathrm{d}\nu_{[\mathcal{Q}_{ij}]}, \tag{A.2.20}$$

for all $\Gamma^1, \Gamma^2 \in L^2(\Omega \times [0,T], \mathcal{P}^\star, \nu_{[\mathcal{Q}_{ij}]})$. From the definition of $\nu_{[\mathcal{Q}_{ij}]}$, given by (3.2.12), we get

$$\int_{\Omega \times [0,T]} \Gamma^1 \Gamma^2\,\mathrm{d}\nu_{[\mathcal{Q}_{ij}]} = \mathrm{E} \int_0^T \Gamma^1(t) \Gamma^2(t)\,\mathrm{d}[\mathcal{Q}_{ij}](t). \tag{A.2.21}$$

Then the two inner products $\langle \cdot, \cdot \rangle_{L^2(\mathcal{Q}_{ij})}$ and $\langle \cdot, \cdot \rangle_{ij}$ are equal, meaning that $L^2(\mathcal{Q}_{ij})$ is a Hilbert space with the inner product $\langle \cdot, \cdot \rangle_{L^2(\mathcal{Q}_{ij})}$.

Note that for the space $L^2(\mathcal{Q})$, which is defined by (3.2.35), we have

$$L^2(\mathcal{Q}) = \bigoplus_{\substack{i,j=1\\i\neq j}}^{D} L^2(\mathcal{Q}_{ij}). \tag{A.2.22}$$

Define an inner product on the space $L^2(\mathcal{Q})$ as

$$\langle \mathbf{\Gamma}^1, \mathbf{\Gamma}^2 \rangle_{L^2(\mathcal{Q})} := \sum_{\substack{i,j=1\\i\neq j}}^{D} \langle \Gamma^1_{ij}, \Gamma^2_{ij} \rangle_{L^2(\mathcal{Q}_{ij})}, \tag{A.2.23}$$

for all $\mathbf{\Gamma}^1, \mathbf{\Gamma}^2 \in L^2(\mathcal{Q})$. Then as $L^2(\mathcal{Q})$ is the direct sum of a finite number of Hilbert spaces, it is also a Hilbert space with the inner product defined above. Substituting into (A.2.23) from (A.2.18), we obtain (A.2.16). □

Lemma A.2.6. *For each $\boldsymbol{\rho} \in L^2(\mathbf{W})$ and $\boldsymbol{\gamma} = (\gamma_{ij})_{i,j=1}^{D} \in L^2(\mathcal{Q})$, there exists $\boldsymbol{\lambda} \in L^2(\mathbf{W})$ such that for all $t \in [0,T]$ we have a.s.*

$$\boldsymbol{\lambda}(t) + \boldsymbol{\theta}(t) \int_0^t \boldsymbol{\lambda}^\top(\tau)\,\mathrm{d}\mathbf{W}(\tau) = \boldsymbol{\rho}(t) - \boldsymbol{\theta}(t) \sum_{i,j=1}^{D} \int_0^t \gamma_{ij}(\tau)\,\mathrm{d}\mathcal{Q}_{ij}(\tau). \tag{A.2.24}$$

$\boldsymbol{\lambda}$ is unique in the sense that if there exists $\bar{\boldsymbol{\lambda}} \in L^2(\mathbf{W})$ such that (A.2.24) holds with $\boldsymbol{\lambda}$ replaced by $\bar{\boldsymbol{\lambda}}$, then $\boldsymbol{\lambda} = \bar{\boldsymbol{\lambda}}$ $(\mathbb{P} \otimes Leb)$-a.e.

Proof. This proof is adapted from the proof of Labbé [32], Lemma 4.4.21.

We construct a sequence $\{\boldsymbol{\lambda}^{(m)}\}_{m\in\mathbb{N}_0}$, which we show is a sequence of elements in the Banach space $L^2(\mathbf{W})$. Next we show that this sequence is a Cauchy sequence. By the completeness of $L^2(\mathbf{W})$, this Cauchy sequence must converge in the $L^2(\mathbf{W})$-norm to a limit in $L^2(\mathbf{W})$. We denote this limit by $\boldsymbol{\lambda}$ and show that it satisfies (A.2.24). Finally, we show that it is unique to within indistinguishability.

Define inductively

$$\boldsymbol{\lambda}^{(0)}(t) := \boldsymbol{\rho}(t) \tag{A.2.25}$$

$$\boldsymbol{\lambda}^{(m+1)}(t) := \boldsymbol{\rho}(t) - \boldsymbol{\theta}(t) \int_0^t \left(\boldsymbol{\lambda}^{(m)}\right)^\top(\tau)\,\mathrm{d}\mathbf{W}(\tau) - \boldsymbol{\theta}(t) \sum_{i,j=1}^{D} \int_0^t \gamma_{ij}(\tau)\,\mathrm{d}\mathcal{Q}_{ij}(\tau). \tag{A.2.26}$$

We show that the sequence $\{\boldsymbol{\lambda}^{(m)}\}_{m\in\mathbb{N}_0}$ lies in $L^2(\mathbf{W})$.

By assumption, $\boldsymbol{\rho} \in L^2(\mathbf{W})$ so we have immediately that $\boldsymbol{\lambda}^{(0)} \in L^2(\mathbf{W})$. Suppose that $\boldsymbol{\lambda}^{(m)} \in L^2(\mathbf{W})$ for some $m \in \mathbb{N}_0$. Let $\kappa_{\boldsymbol{\theta}} \in (0,\infty)$ satisfy (3.2.28), so that $\kappa_{\boldsymbol{\theta}}$ is a uniform bound on $\|\boldsymbol{\theta}\|$. Then

$$
\begin{aligned}
&\mathrm{E}\int_0^T \left\|\boldsymbol{\lambda}^{(m+1)}(t)\right\|^2 \,\mathrm{d}t \\
&\overset{(A.2.26)}{=} \mathrm{E}\int_0^T \left\|\boldsymbol{\rho}(t) - \boldsymbol{\theta}(t)\int_0^t \left(\boldsymbol{\lambda}^{(m)}\right)^\top(\tau)\,\mathrm{d}\mathbf{W}(\tau) - \boldsymbol{\theta}(t)\sum_{i,j=1}^{D}\int_0^t \gamma_{ij}(\tau)\,\mathrm{d}\mathcal{Q}_{ij}(\tau)\right\|^2 \mathrm{d}t \\
&\leq 3\mathrm{E}\int_0^T \|\boldsymbol{\rho}(t)\|^2 \,\mathrm{d}t + 3\kappa_{\boldsymbol{\theta}}^2 \mathrm{E}\int_0^T \left|\int_0^t \left(\boldsymbol{\lambda}^{(m)}\right)^\top(\tau)\,\mathrm{d}\mathbf{W}(\tau)\right|^2 \mathrm{d}t \\
&\qquad + 3\kappa_{\boldsymbol{\theta}}^2\, \mathrm{E}\int_0^T \left|\sum_{i,j=1}^{D}\int_0^t \gamma_{ij}(\tau)\,\mathrm{d}\mathcal{Q}_{ij}(\tau)\right|^2 \mathrm{d}t.
\end{aligned}
\tag{A.2.27}
$$

Since $\boldsymbol{\rho} \in L^2(\mathbf{W})$ then the first term on the last line of (A.2.27) is finite. It remains to show that the second and third terms on the last line are also finite.

Using Doob's L^2-inequality and the Itô isometry, we get

$$
\begin{aligned}
\mathrm{E}\int_0^T \left|\int_0^t \left(\boldsymbol{\lambda}^{(m)}\right)^\top(\tau)\,\mathrm{d}\mathbf{W}(\tau)\right|^2 \mathrm{d}t &\leq T\mathrm{E}\left(\sup_{s\in[0,T]}\left|\int_0^s \left(\boldsymbol{\lambda}^{(m)}\right)^\top(\tau)\,\mathrm{d}\mathbf{W}(\tau)\right|^2\right) \\
&\overset{\text{Doob}}{\leq} 4T\mathrm{E}\left|\int_0^T \left(\boldsymbol{\lambda}^{(m)}\right)^\top(\tau)\,\mathrm{d}\mathbf{W}(\tau)\right|^2 \\
&\overset{\text{Itô}}{=} 4T\mathrm{E}\int_0^T \|\boldsymbol{\lambda}^{(m)}(\tau)\|^2\,\mathrm{d}\tau \\
&= 4T\,\left\|\boldsymbol{\lambda}^{(m)}\right\|_{L^2(\mathbf{W})}^2 < \infty.
\end{aligned}
\tag{A.2.28}
$$

Again, using Doob's L^2-inequality and the Itô isometry, we get

$$
\begin{aligned}
\mathrm{E}\int_0^T \left|\sum_{i,j=1}^{D}\int_0^t \gamma_{ij}(\tau)\,\mathrm{d}\mathcal{Q}_{ij}(\tau)\right|^2 \mathrm{d}t &\leq T\mathrm{E}\left(\sup_{s\in[0,T]}\left|\sum_{i,j=1}^{D}\int_0^s \gamma_{ij}(\tau)\,\mathrm{d}\mathcal{Q}_{ij}(\tau)\right|^2\right) \\
&\overset{\text{Doob}}{\leq} 4T\mathrm{E}\left|\sum_{i,j=1}^{D}\int_0^T \gamma_{ij}(\tau)\,\mathrm{d}\mathcal{Q}_{ij}(\tau)\right|^2 \\
&\overset{\text{Itô}}{=} 4T\sum_{i,j=1}^{D}\mathrm{E}\int_0^T |\gamma_{ij}(\tau)|^2\,\mathrm{d}[\mathcal{Q}_{ij}](\tau) \\
&= 4T\,\|\boldsymbol{\gamma}\|_{L^2(\mathcal{Q})}^2 < \infty.
\end{aligned}
\tag{A.2.29}
$$

From (A.2.28) and (A.2.29), the last line of (A.2.27) is finite, hence $\boldsymbol{\lambda}^{(m+1)} \in L^2(\mathbf{W})$. Then by induction, $\boldsymbol{\lambda}^{(m)} \in L^2(\mathbf{W})$ for all $m \in \mathbb{N}_0$.

Now we show that the sequence $\{\boldsymbol{\lambda}^{(m)}\}_{m\in\mathbb{N}_0}$ is a Cauchy sequence in $L^2(\mathbf{W})$. To do this, we use induction to upper bound $\|\boldsymbol{\lambda}^{(m+1)} - \boldsymbol{\lambda}^{(m)}\|^2_{L^2(\mathbf{W})}$ by a constant which depends on m. Then we show that the norm $\|\boldsymbol{\lambda}^{(m_1)} - \boldsymbol{\lambda}^{(m_2)}\|^2_{L^2(\mathbf{W})}$ tends to zero as $m_1, m_2 \to \infty$.

First observe that from the Itô isometry that

$$\begin{aligned}
&\mathrm{E}\|\boldsymbol{\lambda}^{(1)}(t) - \boldsymbol{\lambda}^{(0)}(t)\|^2 \\
&\overset{(A.2.25),(A.2.26)}{=} \mathrm{E}\left\| -\boldsymbol{\theta}(t)\int_0^t \boldsymbol{\rho}^\top(\tau)\,\mathrm{d}\mathbf{W}(\tau) - \boldsymbol{\theta}(t)\sum_{i,j=1}^{D}\int_0^t \gamma_{ij}(\tau)\,\mathrm{d}\mathcal{Q}_{ij}(\tau)\right\|^2 \\
&\le 2\kappa_{\boldsymbol{\theta}}^2 \mathrm{E}\left|\int_0^t \boldsymbol{\rho}^\top(\tau)\,\mathrm{d}\mathbf{W}(\tau)\right|^2 + 2\kappa_{\boldsymbol{\theta}}^2 \mathrm{E}\left|\sum_{i,j=1}^{D}\int_0^t \gamma_{ij}(\tau)\,\mathrm{d}\mathcal{Q}_{ij}(\tau)\right|^2 \qquad \text{(A.2.30)} \\
&\overset{\text{Itô}}{=} 2\kappa_{\boldsymbol{\theta}}^2 \mathrm{E}\int_0^t \|\boldsymbol{\rho}(\tau)\|^2\,\mathrm{d}\tau + 2\kappa_{\boldsymbol{\theta}}^2 \sum_{i,j=1}^{D} \mathrm{E}\int_0^t |\gamma_{ij}(\tau)|^2\,\mathrm{d}[\mathcal{Q}_{ij}](\tau) \\
&\le 2\kappa_{\boldsymbol{\theta}}^2 \|\boldsymbol{\rho}\|^2_{L^2(\mathbf{W})} + 2\kappa_{\boldsymbol{\theta}}^2 \|\boldsymbol{\gamma}\|^2_{L^2(\mathcal{Q})}.
\end{aligned}$$

Then we can choose a real number $\kappa_1 > 0$ which satisfies

$$\sup_{t\in[0,T]} \mathrm{E}\|\boldsymbol{\lambda}^{(1)}(t) - \boldsymbol{\lambda}^{(0)}(t)\|^2 \le \kappa_1 < \infty. \qquad \text{(A.2.31)}$$

Claim A.2.7.

$$\mathrm{E}\|\boldsymbol{\lambda}^{(m+1)}(t) - \boldsymbol{\lambda}^{(m)}(t)\|^2 \le \frac{\kappa_1(\kappa_{\boldsymbol{\theta}}^2 t)^m}{m!}, \quad \forall t\in[0,T] \quad \forall m\in\mathbb{N}_0. \qquad \text{(A.2.32)}$$

To show that (A.2.32) holds, we proceed by induction. (A.2.32) holds for $m = 0$, from (A.2.31). Suppose that (A.2.32) holds for some $m \in \mathbb{N}_0$. First note that we can obtain the following inequality from the recurrence relation (A.2.26),

$$\begin{aligned}
\|\boldsymbol{\lambda}^{(m+2)}(t) - \boldsymbol{\lambda}^{(m+1)}(t)\|^2 &= \left\| -\boldsymbol{\theta}(t)\int_0^t \left(\boldsymbol{\lambda}^{(m+1)}(\tau) - \boldsymbol{\lambda}^{(m)}(\tau)\right)^\top \mathrm{d}\mathbf{W}(\tau)\right\|^2 \\
&\le \kappa_{\boldsymbol{\theta}}^2 \left|\int_0^t \left(\boldsymbol{\lambda}^{(m+1)}(\tau) - \boldsymbol{\lambda}^{(m)}(\tau)\right)^\top \mathrm{d}\mathbf{W}(\tau)\right|^2.
\end{aligned} \qquad \text{(A.2.33)}$$

Taking expectations, and using the Itô isometry and Fubini's theorem to interchange the expectation and the integral, we get

$$
\begin{aligned}
\mathrm{E}\|\boldsymbol{\lambda}^{(m+2)}(t)-\boldsymbol{\lambda}^{(m+1)}(t)\|^2 &\leq \kappa_{\boldsymbol{\theta}}^2 \mathrm{E}\left|\int_0^t \left(\boldsymbol{\lambda}^{(m+1)}(\tau)-\boldsymbol{\lambda}^{(m)}(\tau)\right)^\top \mathrm{d}\mathbf{W}(\tau)\right|^2 \\
&\overset{\text{Itô}}{=} \kappa_{\boldsymbol{\theta}}^2\, \mathrm{E}\int_0^t \left\|\boldsymbol{\lambda}^{(m+1)}(\tau)-\boldsymbol{\lambda}^{(m)}(\tau)\right\|^2 \mathrm{d}\tau \\
&\overset{\text{Fubini}}{=} \kappa_{\boldsymbol{\theta}}^2 \int_0^t \mathrm{E}\left\|\boldsymbol{\lambda}^{(m+1)}(\tau)-\boldsymbol{\lambda}^{(m)}(\tau)\right\|^2 \mathrm{d}\tau \\
&\overset{(A.2.32)}{\leq} \kappa_{\boldsymbol{\theta}}^2 \int_0^t \frac{\kappa_1(\kappa_{\boldsymbol{\theta}}^2\tau)^m}{m!}\,\mathrm{d}\tau \\
&= \frac{\kappa_1(\kappa_{\boldsymbol{\theta}}^2 t)^{m+1}}{(m+1)!}.
\end{aligned} \tag{A.2.34}
$$

Therefore, (A.2.32) holds by induction and Claim A.2.7 is shown.

Now using Fubini's theorem to interchange the expectation and the integral,

$$
\begin{aligned}
\left\|\boldsymbol{\lambda}^{(m+1)}-\boldsymbol{\lambda}^{(m)}\right\|_{L^2(\mathbf{W})}^2 &= \mathrm{E}\int_0^T \left\|\boldsymbol{\lambda}^{(m+1)}(t)-\boldsymbol{\lambda}^{(m)}(t)\right\|^2 \mathrm{d}t \\
&\overset{\text{Fubini}}{=} \int_0^T \mathrm{E}\left\|\boldsymbol{\lambda}^{(m+1)}(t)-\boldsymbol{\lambda}^{(m)}(t)\right\|^2 \mathrm{d}t \\
&\overset{(A.2.32)}{\leq} \int_0^T \frac{\kappa_1(\kappa_{\boldsymbol{\theta}}^2 t)^m}{m!}\,\mathrm{d}t = \frac{\kappa_1\kappa_{\boldsymbol{\theta}}^{2m}\,T^{m+1}}{(m+1)!}.
\end{aligned} \tag{A.2.35}
$$

For $m_1, m_2 \in \mathbb{N}_0$ with $m_1 < m_2$ by the triangle inequality for norms, we have

$$
\left\|\boldsymbol{\lambda}^{(m_2)}-\boldsymbol{\lambda}^{(m_1)}\right\|_{L^2(\mathbf{W})}^2 = \left\|\sum_{i=m_1}^{m_2-1}\left(\boldsymbol{\lambda}^{(i+1)}-\boldsymbol{\lambda}^{(i)}\right)\right\|_{L^2(\mathbf{W})}^2 \leq \left(\sum_{i=m_1}^{m_2-1}\left\|\boldsymbol{\lambda}^{(i+1)}-\boldsymbol{\lambda}^{(i)}\right\|_{L^2(\mathbf{W})}\right)^2. \tag{A.2.36}
$$

Substituting from (A.2.35), we get

$$
\begin{aligned}
\left\|\boldsymbol{\lambda}^{(m_2)}-\boldsymbol{\lambda}^{(m_1)}\right\|_{L^2(\mathbf{W})}^2 &\overset{(A.2.35)}{\leq} \sum_{i=m_1}^{m_2-1}\frac{\kappa_1\kappa_{\boldsymbol{\theta}}^{2i}\,T^{i+1}}{(i+1)!} \\
&= \left(\frac{\kappa_1}{\kappa_{\boldsymbol{\theta}}^2}\right)\sum_{i=m_1}^{m_2-1}\frac{(\kappa_{\boldsymbol{\theta}}^2 T)^{i+1}}{(i+1)!} \\
&\to 0 \text{ as } m_1, m_2 \to \infty.
\end{aligned} \tag{A.2.37}
$$

It follows that $\{\boldsymbol{\lambda}^{(m)}\}_{m\in\mathbb{N}_0}$ is a Cauchy sequence in $L^2(\mathbf{W})$. Therefore, by the completeness of $L^2(\mathbf{W})$, this sequence tends to a limit $\bar{\boldsymbol{\lambda}} \in L^2(\mathbf{W})$, so that

$$
\lim_{m\to\infty}\left\|\boldsymbol{\lambda}^{(m)}-\bar{\boldsymbol{\lambda}}\right\|_{L^2(\mathbf{W})}^2 = 0. \tag{A.2.38}
$$

Claim A.2.8. The limit $\bar{\boldsymbol{\lambda}} \in L^2(\mathbf{W})$ of the sequence $\{\boldsymbol{\lambda}^{(m)}\}_{m\in\mathbb{N}} \subset L^2(\mathbf{W})$ satisfies

$$\bar{\boldsymbol{\lambda}}(t) + \boldsymbol{\theta}(t)\int_0^t \bar{\boldsymbol{\lambda}}^\top(\tau)\,\mathrm{d}\mathbf{W}(\tau) = \boldsymbol{\rho}(t) - \boldsymbol{\theta}(t)\sum_{i,j=1}^{D}\int_0^t \gamma_{ij}(\tau)\,\mathrm{d}\mathcal{Q}_{ij}(\tau). \tag{A.2.39}$$

First consider

$$\begin{aligned}
&\mathrm{E}\int_0^T \left\| \theta(t)\int_0^t \left(\boldsymbol{\lambda}^{(m)} - \bar{\boldsymbol{\lambda}}\right)^\top(\tau)\,\mathrm{d}\mathbf{W}(\tau)\right\|^2 \mathrm{d}t \\
&\le T\kappa_{\boldsymbol{\theta}}^2 \mathrm{E}\left(\sup_{s\in[0,T]}\left|\int_0^s \left(\boldsymbol{\lambda}^{(m)} - \bar{\boldsymbol{\lambda}}\right)^\top(\tau)\,\mathrm{d}\mathbf{W}(\tau)\right|^2\right) \\
&\overset{\text{Doob}}{\le} 4T\kappa_{\boldsymbol{\theta}}^2 \mathrm{E}\left|\int_0^T \left(\boldsymbol{\lambda}^{(m)} - \bar{\boldsymbol{\lambda}}\right)^\top(\tau)\,\mathrm{d}\mathbf{W}(\tau)\right|^2 \\
&\overset{\text{Itô}}{=} 4T\kappa_{\boldsymbol{\theta}}^2 \mathrm{E}\int_0^T \|\boldsymbol{\lambda}^{(m)}(\tau) - \bar{\boldsymbol{\lambda}}(\tau)\|^2\,\mathrm{d}\tau \\
&= 4T\,\kappa_{\boldsymbol{\theta}}^2\,\left\|\boldsymbol{\lambda}^{(m)} - \bar{\boldsymbol{\lambda}}\right\|_{L^2(\mathbf{W})}^2 \\
&\overset{(A.2.38)}{\to} 0 \quad \text{as } m \to \infty.
\end{aligned} \tag{A.2.40}$$

It follows that

$$\begin{aligned}
&\left\|\boldsymbol{\lambda}^{(m+1)} - \bar{\boldsymbol{\lambda}} + \boldsymbol{\theta}(\cdot)\int_0^\cdot \left(\boldsymbol{\lambda}^{(m)} - \bar{\boldsymbol{\lambda}}\right)^\top(\tau)\,\mathrm{d}\mathbf{W}(\tau)\right\|_{L^2(\mathbf{W})}^2 \\
&\le 2\|\boldsymbol{\lambda}^{(m+1)} - \bar{\boldsymbol{\lambda}}\|_{L^2(\mathbf{W})}^2 + 2\left\|\boldsymbol{\theta}(\cdot)\int_0^\cdot \left(\boldsymbol{\lambda}^{(m)} - \bar{\boldsymbol{\lambda}}\right)^\top(\tau)\,\mathrm{d}\mathbf{W}(\tau)\right\|_{L^2(\mathbf{W})}^2 \\
&\overset{(A.2.38),(A.2.40)}{\to} 0 \quad \text{as } m \to \infty.
\end{aligned} \tag{A.2.41}$$

That is, in the $L^2(\mathbf{W})$-norm, we have

$$\boldsymbol{\lambda}^{(m+1)} + \boldsymbol{\theta}(\cdot)\int_0^\cdot \left(\boldsymbol{\lambda}^{(m)}\right)^\top(\tau)\,\mathrm{d}\mathbf{W}(\tau) \to \bar{\boldsymbol{\lambda}} + \boldsymbol{\theta}(\cdot)\int_0^\cdot \left(\bar{\boldsymbol{\lambda}}\right)^\top(\tau)\,\mathrm{d}\mathbf{W}(\tau) \quad \text{as } m \to \infty. \tag{A.2.42}$$

However, from (A.2.26), we have

$$\boldsymbol{\lambda}^{(m+1)}(t) + \boldsymbol{\theta}(t)\int_0^t \left(\boldsymbol{\lambda}^{(m)}\right)^\top(\tau)\,\mathrm{d}\mathbf{W}(\tau) = \boldsymbol{\rho}(t) - \boldsymbol{\theta}(t)\sum_{i,j=1}^{D}\int_0^t \gamma_{ij}(\tau)\,\mathrm{d}\mathcal{Q}_{ij}(\tau), \tag{A.2.43}$$

so then comparing (A.2.43) to (A.2.42), we must have $(\mathbb{P}\otimes \mathit{Leb})$-a.e.

$$\bar{\boldsymbol{\lambda}}(t) + \boldsymbol{\theta}(t)\int_0^t \bar{\boldsymbol{\lambda}}^\top(\tau)\,\mathrm{d}\mathbf{W}(\tau) = \boldsymbol{\rho}(t) - \boldsymbol{\theta}(t)\sum_{i,j=1}^{D}\int_0^t \gamma_{ij}(\tau)\,\mathrm{d}\mathcal{Q}_{ij}(\tau). \tag{A.2.44}$$

and Claim A.2.8 is shown.

Lastly, we show that the limit $\bar{\boldsymbol{\lambda}}$ is unique to within indistinguishability. Suppose $\bar{\boldsymbol{\lambda}}, \hat{\boldsymbol{\lambda}} \in L^2(\mathbf{W})$ both satisfy (A.2.24). Then

$$\bar{\boldsymbol{\lambda}}(T) - \hat{\boldsymbol{\lambda}}(T) = -\boldsymbol{\theta}(T) \int_0^T \left(\bar{\boldsymbol{\lambda}} - \hat{\boldsymbol{\lambda}}\right)^\top (\tau) \, \mathrm{d}\mathbf{W}(\tau). \tag{A.2.45}$$

Squaring and taking expectations, upon applying the Itô isometry, followed by Fubini's theorem, we get

$$\begin{aligned} &\mathrm{E}\|\bar{\boldsymbol{\lambda}}(T) - \hat{\boldsymbol{\lambda}}(T)\|^2 = \mathrm{E}\Big\|-\boldsymbol{\theta}(T) \int_0^T \left(\bar{\boldsymbol{\lambda}} - \hat{\boldsymbol{\lambda}}\right)^\top (\tau) \, \mathrm{d}\mathbf{W}(\tau)\Big\|^2 \\ &\leq \kappa_{\boldsymbol{\theta}}^2 \mathrm{E}\left| \int_0^T \left(\bar{\boldsymbol{\lambda}} - \hat{\boldsymbol{\lambda}}\right)^\top (\tau) \, \mathrm{d}\mathbf{W}(\tau)\right|^2 \\ &\overset{\text{Itô}}{=} \kappa_{\boldsymbol{\theta}}^2 \mathrm{E} \int_0^T \|\bar{\boldsymbol{\lambda}}(\tau) - \hat{\boldsymbol{\lambda}}(\tau)\|^2 \, \mathrm{d}\tau \\ &\overset{\text{Fubini}}{=} \kappa_{\boldsymbol{\theta}}^2 \int_0^T \mathrm{E}\|\bar{\boldsymbol{\lambda}}(\tau) - \hat{\boldsymbol{\lambda}}(\tau)\|^2 \, \mathrm{d}\tau. \end{aligned} \tag{A.2.46}$$

Applying Gronwall's inequality (Lemma A.2.4) to the positive function $\mathrm{E}\|\bar{\boldsymbol{\lambda}}(\cdot) - \hat{\boldsymbol{\lambda}}(\cdot)\|^2$ gives

$$\mathrm{E}\|\bar{\boldsymbol{\lambda}}(T) - \hat{\boldsymbol{\lambda}}(T)\|^2 \leq 0 \exp\{\kappa_{\boldsymbol{\theta}}^2 T\} = 0. \tag{A.2.47}$$

Hence, $\mathrm{E}\|\bar{\boldsymbol{\lambda}}(T) - \hat{\boldsymbol{\lambda}}(T)\|^2 = 0$ and so $\bar{\boldsymbol{\lambda}} = \hat{\boldsymbol{\lambda}}$ $(\mathbb{P} \otimes Leb)$-a.e. □

Appendix B

A martingale representation theorem

The goal of this rather lengthy appendix is to establish a martingale representation theorem for the joint filtration of a vector Brownian motion and a finite state space Markov chain. Specifically, we would like to show that a locally square-integrable martingale relative to this filtration can be written as the sum of two stochastic integrals; one with the Brownian motion as the integrator and the other with the canonical martingales of the Markov chain as the integrator. This is needed in particular to apply convex duality (see especially (4.6.59)).

We are aware that this result can be obtained, although not without considerable effort, from the general martingale representation theorem established in Jacod and Shiryayev [25], Chapter III (see in particular Theorem III.4.29). However, in this appendix, we have chosen to establish this result by generalizing an argument from Wong and Hajek [51] instead. The advantage of doing things this way is that the resulting proof, although quite lengthy, nevertheless relies only on elementary tools at every step.

The highlight of this appendix is Theorem B.4.6, which is the martingale representation theorem for square-integrable martingales. All the results leading up to it are in support of proving it. In Theorem B.4.22, we extend Theorem B.4.6 to locally square-integrable martingales by use of a standard argument.

The proof of Theorem B.4.6 requires showing that the N-dimensional standard Brownian motion remains an N-dimensional standard Brownian motion and the Markov chain remains a Markov chain under some specified change of measure. To demonstrate this, we use the properties of the *martingale problems* for the Brownian motion and the Markov chain. These martingale problems and their properties are set out in Section B.2. The martingale problem for the Markov chain requires some well-known results, such as Dynkin's Formula, which are set out in Section B.1. We also require a theorem from Ethier and Kurtz [15], which shows that the Brownian motion and the Markov chain remain independent under the specified change of measure.

In Section B.3, we determine the set of canonical martingales of the Markov chain. These are square-integrable martingales, one for each pair of states in the (finite) state space of the Markov chain. These canonical martingales are strongly orthogonal to each other and we are able to find explicit expressions for both their square-bracket and angle-bracket quadratic variation processes. These canonical martingales are used as integrators when expressing martingales as stochastic integrals.

The material in Section B.3 is primarily motivated by Rogers and Williams [46], Section IV.20. This provided the basic formulation of the canonical martingales, through the construction of their square-bracket and angle-bracket quadratic variation processes. From Boel, Varaiya and Wong [6], we discerned the basic properties of the canonical martingales, such as their square-integrability and orthogonality.

In Section B.4, we prove Theorem B.4.6, which is the martingale representation theorem for square-integrable martingales. The proof follows an argument adapted from Wong and Hajek [51]. They prove a martingale representation theorem for martingales with respect to a filtration generated by a Brownian motion and a standard Poisson process. The proof for Wong and Hajek [51] uses the Watanabe theorem of the characterization of a standard Poisson process. However, proving a similar result for the canonical martingales of the Markov chain did not appear feasible to us. The compensator of the standard Poisson process is deterministic (it equals the time t) and it is on this deterministic nature that the proof of the Watanabe theorem, or any extension of it, is hinged. In comparison, the compensator of the canonical martingales is non-deterministic (it depends on how long the Markov chain has spent in a particular state up to time t). To bypass this difficulty, we use martingale problems as a substitute for the Watanabe characterization theorem.

Before Theorem B.4.6, we prove Lemma B.4.1, which is used to construct the integrands for *uniformly bounded* martingales. Theorem B.4.6 begins by proving the martingale representation theorem for uniformly bounded martingales, using not only Lemma B.4.1 but also the martingale problems examined in Section B.4.6. Extending the martingale representation result for uniformly bounded martingales to square-integrable martingales is then relatively straightforward.

Finally in Section B.4, we extend the martingale representation theorem to *locally square-integrable martingales* (Theorem B.4.22). We also show that the orthogonal complement of $L^2(\mathcal{Q})$ is $L^2(\mathbf{W})$ and, conversely, the orthogonal complement of $L^2(\mathbf{W})$ is $L^2(\mathcal{Q})$.

B.1 The finite state space Markov chain

In this section, we define the continuous time, finite state space Markov chain and set out some results concerning it. The most important result for us is the (well-known) Dynkin's formula (Theorem B.1.6). We being by specifying the basic properties of the Markov chain.

Let $\alpha = \{\alpha(t) : t \in [0,T]\}$ be a continuous-time Markov chain defined on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and which takes values in a finite set I where

$$I = \{1, 2, \ldots, D\}. \tag{B.1.1}$$

We suppose that the Markov chain α starts in an initial state $i_0 \in I$, so that

$$\alpha(0) = i_0 \quad \text{a.s.} \tag{B.1.2}$$

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

$$q_{ij} \geq 0, \quad \forall i \neq j \quad \text{and} \quad -q_{ii} = \sum_{j \neq i} q_{ij}. \tag{B.1.3}$$

Remark B.1.1. From Rogers and Williams [46], 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,

$$\int_0^T f(\alpha(s_-)) \, \mathrm{d}s = \int_0^T f(\alpha(s)) \, \mathrm{d}s \tag{B.1.4}$$

for any function f on I.

Definition B.1.2. The *Markov transition function* $\{P_t\}$ *on* I is defined as

$$P_t := \exp\{tQ\}, \qquad \forall t \in [0,T]. \tag{B.1.5}$$

In particular, P_0 is the $D \times D$ identity matrix.

From Ethier and Kurtz [15], Chapter 4, we have the following.

Remark B.1.3. The *Markov property* of the Markov chain α is that for all functions f on I, for all $0 \leq s \leq t \leq T$,

$$\mathrm{E}\left(f(\alpha(t)) \,|\, \mathcal{F}_s\right) = (P_{t-s} f)(\alpha(s)), \tag{B.1.6}$$

where $(P_t f)(j) = \sum_{k \in I} P_t(j,k) f(k)$.

Proposition B.1.4.

$$\frac{\mathrm{d}}{\mathrm{d}t}(P_t) = QP_t = P_t Q. \tag{B.1.7}$$

Proof. See Norris [40], Theorem 2.1.1. □

The following proposition is a simple consequence of Proposition B.1.4.

Proposition B.1.5. *If f is a function on I and $t \in [0,T]$ then*

$$\frac{\mathrm{d}}{\mathrm{d}t}(P_t f) = QP_t f, \tag{B.1.8}$$

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.

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\left(P_t f\right)(j)\right) = \sum_{k\in I}\frac{\mathrm{d}}{\mathrm{d}t}\left(P_t(j,k)f(k)\right) \overset{(B.1.7)}{=} \sum_{k\in I}\left(QP_t\right)(j,k)f(k) = \left(QP_t f\right)(j). \tag{B.1.9}$$

□

We will now prove *Dynkin's formula.*

Theorem B.1.6. Dynkin's formula *If f is a function on I then*

$$f(\alpha(t)) - f(\alpha(0)) - \int_0^t (Qf)(\alpha(s))\ \mathrm{d}s \in \mathcal{M}_0((\Omega,\mathcal{F},\mathbb{P}),\{\mathcal{F}_t\}). \tag{B.1.10}$$

Proof. Fix a function $f : I \to \mathbb{R}$ and define

$$M(t) := f(\alpha(t)) - f(\alpha(0)) - \int_0^t (Qf)(\alpha(s))\ \mathrm{d}s. \tag{B.1.11}$$

By the boundedness of the function f, $M(t)$ is integrable for all $t \in [0,T]$.

We show that the martingale property holds for M. First note that upon integrating (B.1.8), we obtain for all $t \in [0,T]$ and each $i \in I$,

$$(P_t f)(i) - f(i) = \int_0^t (QP_u f)(i)\ \mathrm{d}u, \tag{B.1.12}$$

For $0 \le s \le t \le T$, we have from the definition of M in (B.1.11),

$$\begin{aligned}
\mathrm{E}\big(M(t) - M(s)\,|\,\mathcal{F}_s\big) &= \mathrm{E}\left(f(\alpha(t)) - f(\alpha(s))\,|\,\mathcal{F}_s\right) - \mathrm{E}\left(\int_s^t \left(Qf\right)\left(\alpha(u)\right)\ \mathrm{d}u\,\middle|\,\mathcal{F}_s\right)\\
&\overset{(B.1.6)}{=} (P_{t-s}f)(\alpha(s)) - f(\alpha(s)) - \mathrm{E}\left(\int_s^t \left(Qf\right)\left(\alpha(u)\right)\ \mathrm{d}u\,\middle|\,\mathcal{F}_s\right)\\
&\overset{(B.1.12)}{=} \int_0^{t-s}\left(QP_u f\right)\left(\alpha(s)\right)\ \mathrm{d}u - \mathrm{E}\left(\int_s^t \left(Qf\right)\left(\alpha(u)\right)\ \mathrm{d}u\,\middle|\,\mathcal{F}_s\right).
\end{aligned} \tag{B.1.13}$$

Applying Fubini's theorem for conditional expectations (see Ethier and Kurtz [15], Chapter 2, Proposition 4.6 and Remark 4.7), we obtain

$$\begin{aligned}\mathrm{E}\big(M(t)-M(s)\,|\,\mathcal{F}_s\big) &= \int_0^{t-s} (QP_u f)\,(\alpha(s))\;\mathrm{d}u - \int_s^t \mathrm{E}\left((Qf)(\alpha(u))\,|\,\mathcal{F}_s\right)\;\mathrm{d}u \\ &\overset{(B.1.6)}{=} \int_0^{t-s} (QP_u f)\,(\alpha(s))\;\mathrm{d}u - \int_s^t (P_{u-s}Qf)\,(\alpha(s))\;\mathrm{d}u \\ &\overset{(B.1.7)}{=} \int_0^{t-s} (QP_u f)\,(\alpha(s))\;\mathrm{d}u - \int_s^t (QP_{u-s}f)\,(\alpha(s))\;\mathrm{d}u \\ &= \int_0^{t-s} (QP_u f)\,(\alpha(s))\;\mathrm{d}u - \int_0^{t-s} (QP_u f)\,(\alpha(s))\;\mathrm{d}u \\ &= 0.\end{aligned} \tag{B.1.14}$$

Thus $M \in \mathcal{M}_0((\Omega,\mathcal{F},\mathbb{P}),\{\mathcal{F}_t\})$. □

B.2 Martingale problems

This section examines the properties of the *martingale problems* associated with the Brownian motion and the Markov chain. We begin by defining in Subsection B.2.1 a general martingale problem which involves a general set, an initial condition and a family of processes which depend both on a stochastic process and on functions on the general set. The solution to the martingale problem is a triple consisting of a probability space, a filtration and the particular stochastic process X which makes each member of the family of processes a martingale. The stochastic process X must also satisfy the initial condition. We then state two properties of a martingale problem: the properties of uniqueness and well-posedness. Showing that these properties hold for the martingale problems associated with the Brownian motion and the Markov chain will be critical in proving Theorem B.4.6, the martingale representation theorem for square-integrable martingales.

Subsection B.2.2 examines the martingale problem for the Markov chain and Subsection B.2.3 examines the martingale problem for the Brownian motion. These martingale problems are then used in Subsection B.2.4 to construct the joint martingale problem for both the Markov chain and the Brownian motion.

B.2.1 The general martingale problem

Here we define a general martingale problem and its solution. We also define some important properties of a martingale problem.

Definition B.2.1. Let (E,τ) be a metric space, that is a set E with an associated metric τ. Denote

$$B(E) := \{f : E \to \mathbb{R}\,|\, f \text{ is bounded and Borel-measurable}\}. \tag{B.2.1}$$

Suppose that we are given the following

- a subset $A \subset B(E) \times B(E)$;
- an initial state x_0;
- a filtered probability space $((\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}}), \{\tilde{\mathcal{F}}_t\})$ consisting of a probability space $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}})$ and a filtration $\{\tilde{\mathcal{F}}_t\}$ on $\tilde{\mathcal{F}}$; and
- a càdlàg, E-valued stochastic process $X = \{X(t) : t \in [0, T]\}$ on the probability space $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}})$ which is adapted to the filtration $\{\tilde{\mathcal{F}}_t\}$.

Define for all $\mathbf{f} \equiv (f, g) \in A$,

$$M^{\mathbf{f}}(X)(t) := f(X(t)) - f(X(0)) - \int_0^t g(X(\tau))\, \mathrm{d}\tau, \quad \forall t \in [0, T]. \tag{B.2.2}$$

Then the triplet $((\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}}), \{\tilde{\mathcal{F}}_t\}, X)$ is said to *solve the martingale problem for* (A, x_0) when

- $\tilde{\mathbb{P}}[X(0) = x_0] = 1$; and
- for all $\mathbf{f} \equiv (f, g) \in A$, $M^{\mathbf{f}}(X) \in \mathcal{M}_0((\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}}), \{\tilde{\mathcal{F}}_t\})$.

Remark B.2.2. Note that a process X is $\{\tilde{\mathcal{F}}_t\}$-adapted if and only if $\mathcal{F}_t^X \subset \tilde{\mathcal{F}}_t$ for all $t \in [0, T]$.

Remark B.2.3. Rather than specifying an initial state x_0, we could have specified an initial distribution $\mu_0 \in \mathscr{P}(E)$, the family of Borel probability measures on E. Any solution $((\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}}), \{\tilde{\mathcal{F}}_t\}, X)$ of the martingale problem for (A, μ_0) would then have to satisfy $\tilde{\mathbb{P}}[X^{-1}(0)] = \mu_0$, rather than $\tilde{\mathbb{P}}[X(0) = x_0] = 1$. However, as we will only be concerned with martingale problems with a fixed initial state, it is notationally more convenient to use the less general formulation.

We will also be interested in the following two properties of martingale problems.

Definition B.2.4. The martingale problem for (A, x_0) has the property of *uniqueness* when, for any two solutions $((\hat{\Omega}, \hat{\mathcal{F}}, \hat{\mathbb{P}}), \{\hat{\mathcal{F}}_t\}, X)$ and $((\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}}), \{\tilde{\mathcal{F}}_t\}, Y)$, it necessarily follows that X and Y have the same finite-dimensional distributions.

Definition B.2.5. The martingale problem for (A, x_0) is said to be *well-posed* when

- there exists a solution; and
- it has the property of uniqueness.

B.2.2 The martingale problem for the Markov chain

Here we examine the martingale problem for a Markov chain with generator Q, the generator Q having the properties given by (B.1.3). We do not consider only the Markov chain α, but allow for the possibility of other Markov chains which may live on different probability spaces to the Markov chain α. However, we assume that all these Markov chains have the same generator Q. We show that the martingale problem is well-posed, in the sense of Definition B.2.5.

Many of the arguments in this subsection are from Rogers and Williams [46]. However, we have shown them here because the proofs in Rogers and Williams [46] are for the filtration generated only by the Markov chain, whereas we have a more general filtration and we want to check that the arguments follow through. As it turns out they do, by reason of the solution to the martingale problem consisting of our more general filtration.

Recall the finite state space I of the Markov chain from (B.1.1) and the assumption that the Markov chain starts in state i_0. The metric space for the martingale problem associated with the Markov chain is the state space I equipped with the *discrete metric* τ,

$$\tau(i,j) := \begin{cases} 1 & \text{if } i \neq j \\ 0 & \text{if } i = j, \end{cases} \qquad \forall i,j \in I. \tag{B.2.3}$$

We define

$$A^Q := \{(f, Qf) \mid f : I \to \mathbb{R} \text{ is a function}\} \subset B(I) \times B(I), \tag{B.2.4}$$

for $B(I)$ given by (B.2.1).

Lemma B.2.6. $((\Omega, \mathcal{F}, \mathbb{P}), \{\mathcal{F}_t\}, \alpha)$ *is a solution to the martingale problem for* (A^Q, i_0).

Proof. By assumption, the Markov chain α starts in state i_0, thus $\mathbb{P}[\alpha(0) = i_0] = 1$.

We also know from Dynkin's formula (Theorem B.1.6) that for any function f on the set I, $M^{\mathbf{f}}(\alpha) \in \mathcal{M}_0((\Omega, \mathcal{F}, \mathbb{P}), \{\mathcal{F}_t\})$ for $\mathbf{f} = (f, Qf)$ and $M^{\mathbf{f}}$ given by (B.2.2). □

Proposition B.2.7. *If* $((\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}}), \{\tilde{\mathcal{F}}_t\}, X)$ *is a solution of the martingale problem for* (A^Q, i_0) *then* X *has the Markov property with respect to the filtration* $\{\tilde{\mathcal{F}}_t\}$*, that is for all* $0 \le s \le t \le T$,

$$E_{\tilde{\mathbb{P}}}\left(f\left(X(t)\right) \mid \tilde{\mathcal{F}}_s\right) = (P_{t-s}f)(X(s)) \quad \tilde{\mathbb{P}}\text{-}a.s. \tag{B.2.5}$$

Proof. The proof is taken from Rogers and Williams [46].

Suppose that $((\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}}), \{\tilde{\mathcal{F}}_t\}, X)$ is a solution of the martingale problem for (A^Q, i_0). Then by Definition B.2.1, X is a càdlàg, I-valued process on the probability space $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}})$ which is adapted to the filtration $\{\tilde{\mathcal{F}}_t\}$. Moreover, for all $\mathbf{f} \equiv (f, Qf) \in A^Q$, setting

$$M^{\mathbf{f}}(X)(t) := f(X(t)) - f(X(0)) - \int_0^t (Qf)(X(u))\,\mathrm{d}u, \quad \forall t \in [0,T], \tag{B.2.6}$$

we have $M^{\mathbf{f}}(X) \in \mathcal{M}_0((\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}}), \{\tilde{\mathcal{F}}_t\})$. Fix a function f on I and define for all $s \in [0,t]$,

$$N(s) := \left(P_{t-s} f\right)(X(s)) \tag{B.2.7}$$

We first show that for each $t \in [0,T]$,

$$N := \{N(s) : 0 \le s \le t\} \in \mathcal{M}((\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}}), \{\tilde{\mathcal{F}}_t\}). \tag{B.2.8}$$

Set

$$\phi(s,j) := \left(P_{t-s} f\right)(j), \quad \forall s \in [0,t] \quad \forall j \in I. \tag{B.2.9}$$

Then $N(s) = \phi(s, X(s))$. Differentiating ϕ for each fixed $j \in I$,

$$\frac{\partial}{\partial s}\phi(s,j) = \frac{\partial}{\partial s}\left(\left(P_{t-s} f\right)(j)\right) = \frac{\partial}{\partial s}\sum_{k\in I} P_{t-s}(j,k) f(k) = \sum_{k \in I}\frac{\partial}{\partial s} P_{t-s}(j,k) f(k). \tag{B.2.10}$$

As $P_{t-s} = \exp\{(t-s)Q\}$ by (B.1.5),

$$\begin{aligned}\frac{\partial}{\partial s}\phi(s,j) &= \sum_{k\in I}\left(-QP_{t-s}\right)(j,k) f(k) = -\sum_{k\in I} Q(j,k)\left(P_{t-s} f\right)(k)\\ &\overset{(B.2.9)}{=} -\sum_{k\in I} Q(j,k)\phi(s,k) = -\left(Q\phi\right)(s,j).\end{aligned} \tag{B.2.11}$$

Therefore for every $j \in I$ and $0 \le s \le t$,

$$\frac{\partial}{\partial s}\phi(s,j) + \left(Q\phi\right)(s,j) = 0. \tag{B.2.12}$$

Since

$$\phi(s, X(s)) = \sum_{j\in I}\phi(s,j)\chi\left[X(s) = j\right], \tag{B.2.13}$$

we can write ϕ as

$$\phi(s,j) = h(s) g(j), \tag{B.2.14}$$

for a real-valued function h on $[0,t]$ with continuous first derivative and a real-valued function g on I. We can then use the integration-by-parts formula (Theorem C.14.1) to expand $\phi(s, X(s))$.

$$\begin{aligned}\phi(s,X(s)) &= h(s) g(X(s))\\ &= h(0)g(X(0)) + \int_0^s g(X(u))\,\mathrm{d}h(u) + \int_0^s h(u)\,\mathrm{d}g(X(u)) + [h(\cdot), g(X(\cdot))](s).\end{aligned} \tag{B.2.15}$$

Since h is non-random, the quadratic co-variation term is zero. As $\mathbf{g} \equiv (g, Qg) \in A^Q$, we have $M^{\mathbf{g}}(X) \in \mathcal{M}_0((\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}}), \{\tilde{\mathcal{F}}_t\})$ and hence from (B.2.6),

$$\mathrm{d}M^{\mathbf{g}}(X)(u) = \mathrm{d}g(X(u)) - (Qg)(X(u))\,\mathrm{d}u. \tag{B.2.16}$$

Writing $\dot{h} = \frac{\mathrm{d}h}{\mathrm{d}u}$ and substituting (B.2.16) into (B.2.15),

$$\begin{aligned}
\phi(s, X(s)) &= h(0)g(X(0)) + \int_0^s g(X(u))\dot{h}\,\mathrm{d}u + \int_0^s h(u)(Qg)(X(u))\,\mathrm{d}u \\
&\quad + \int_0^s h(u)\,\mathrm{d}M^{\mathbf{g}}(X)(u) \\
&\overset{(B.2.14)}{=} \phi(0, X(0)) + \int_0^s \frac{\partial}{\partial u}\phi(u, X(u))\,\mathrm{d}u \\
&\quad + \int_0^s \sum_{k\in I} Q(X(u), k)g(k)h(u)\,\mathrm{d}u + \int_0^s h(u)\,\mathrm{d}M^{\mathbf{g}}(X)(u) \\
&\overset{(B.2.14)}{=} \phi(0, X(0)) + \int_0^s \frac{\partial}{\partial u}\phi(u, X(u))\,\mathrm{d}u \\
&\quad + \int_0^s \sum_{k\in I} Q(X(u), k)\phi(u, k)\,\mathrm{d}u + \int_0^s h(u)\,\mathrm{d}M^{\mathbf{g}}(X)(u) \\
&= \phi(0, X(0)) + \int_0^s \left(\frac{\partial}{\partial u}\phi(u, X(u)) + (Q\phi)(u, X(u))\right)\mathrm{d}u \\
&\quad + \int_0^s h(u)\,\mathrm{d}M^{\mathbf{g}}(X)(u) \\
&\overset{(B.2.12)}{=} \phi(0, X(0)) + \int_0^s h(u)\,\mathrm{d}M^{\mathbf{g}}(X)(u).
\end{aligned} \tag{B.2.17}$$

As $M^{\mathbf{g}}(X)$ is a martingale then from (B.2.17), $\phi(s, X(s)) - \phi(0, X(0))$ is a local martingale, and hence $N(s) = \phi(s, X(s))$ is also a local martingale. From (B.2.7), we see that N is uniformly bounded by $\max_{k\in I}|f(k)| < \infty$ and hence N is a martingale. Thus for the fixed $t \in [0, T]$ and fixed function f on I, we have for all $0 \le s \le t$,

$$\begin{aligned}
\mathrm{E}_{\tilde{\mathbb{P}}}\left(N(t)\,|\,\tilde{\mathcal{F}}_s\right) = N(s) \quad &\overset{(B.2.7)}{\Rightarrow} \quad \mathrm{E}_{\tilde{\mathbb{P}}}\left((P_0 f)(X(t))\,|\,\tilde{\mathcal{F}}_s\right) = (P_{t-s}f)(X(s)) \\
&\Rightarrow \quad \mathrm{E}_{\tilde{\mathbb{P}}}\left(f(X(t))\,|\,\tilde{\mathcal{F}}_s\right) = (P_{t-s}f)(X(s)).
\end{aligned} \tag{B.2.18}$$

As the choice of t and f was arbitrary, then the Markov property holds for all $0 \le s \le t \le T$ and all functions f on I.

□

Theorem B.2.8. *The martingale problem for (A^Q, i_0) is well-posed.*

Proof. Recall the definition of well-posedness from Definition B.2.5. By Lemma B.2.6, $((\Omega, \mathcal{F}, \mathbb{P}), \{\mathcal{F}_t\}, \alpha)$ solves the martingale problem for (A^Q, i_0), thus we have existence of a solution.

To show uniqueness, we will again use an argument from Rogers and Williams [46]. Let $((\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}}), \{\tilde{\mathcal{F}}_t\}, X)$ and $((\hat{\Omega}, \hat{\mathcal{F}}, \hat{\mathbb{P}}), \{\hat{\mathcal{F}}_t\}, Y)$ be solutions of the martingale problem for (A^Q, i_0). We will show that these solutions have the same finite-dimensional distributions. From Proposition B.2.7, both X and Y have the Markov property, that is for $0 \leq s \leq t \leq T$,

$$\mathrm{E}_{\tilde{\mathbb{P}}}\left(f\left(X(t)\right) \mid \tilde{\mathcal{F}}_s\right) = (P_{t-s}f)(X(s)) \quad \tilde{\mathbb{P}}\text{-a.s.,} \tag{B.2.19}$$

and

$$\mathrm{E}_{\hat{\mathbb{P}}}\left(f\left(Y(t)\right) \mid \hat{\mathcal{F}}_s\right) = (P_{t-s}f)(Y(s)) \quad \hat{\mathbb{P}}\text{-a.s.} \tag{B.2.20}$$

Let $p(t, i, j)$ be the (i, j)th entry of the Markov transition function P_t. Consider first the solution $((\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}}), \{\tilde{\mathcal{F}}_t\}, X)$. Set $f = \chi[j]$ in (B.2.19) to get,

$$\tilde{\mathbb{P}}\left[X(t) = j \mid \tilde{\mathcal{F}}_s\right] = \mathrm{E}_{\tilde{\mathbb{P}}}\left(\chi[X(t) = j] \mid \tilde{\mathcal{F}}_s\right) = (P_{t-s}\chi[j])(X(s)) = p(t-s, X(s), j) \tag{B.2.21}$$

$\tilde{\mathbb{P}}$-a.s. It follows that for any $j_1, j_2 \in I$ and $0 \leq s \leq t_1 \leq t_2 \leq T$,

$$\begin{aligned} &\tilde{\mathbb{P}}\left[X(t_1) = j_1,\ X(t_2) = j_2 \mid \tilde{\mathcal{F}}_s\right] \\ &= \tilde{\mathbb{P}}\left[X(t_2) = j_2 \mid \tilde{\mathcal{F}}_s, X(t_1) = j_1\right] \tilde{\mathbb{P}}\left[X(t_1) = j_1 \mid \tilde{\mathcal{F}}_s\right] \\ &= \tilde{\mathbb{P}}\left[X(t_2) = j_2 \mid X(t_1) = j_1\right] \tilde{\mathbb{P}}\left[X(t_1) = j_1 \mid \tilde{\mathcal{F}}_s\right] \\ &= p(t_2 - t_1, j_1, j_2)\, p(t_1 - s, X(s), j_1) \end{aligned} \tag{B.2.22}$$

and, more generally, fixing any $n \in \mathbb{N}$, setting $j_m \in I$ for $m = 1, \ldots, n$ and $s = t_0 < t_1 < \cdots < t_n \leq T$,

$$\tilde{\mathbb{P}}\left[X(t_1) = j_1, \ldots, X(t_n) = j_n \mid \tilde{\mathcal{F}}_s\right] = p(t_1 - s, X(s), j_1) \prod_{m=2}^{n} p(t_m - t_{m-1}, j_{m-1}, j_m). \tag{B.2.23}$$

Now set $s = 0$. As $((\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}}), \{\tilde{\mathcal{F}}_t\}, X)$ solves the martingale problem for (A^Q, i_0) then $X(0) = i_0$ $\tilde{\mathbb{P}}$-a.s. Then with $j_0 := i_0$, we obtain from (B.2.23) the finite-dimensional distributions of the process X,

$$\tilde{\mathbb{P}}\left[X(t_1) = j_1, \ldots, X(t_n) = j_n\right] = \prod_{m=1}^{n} p(t_m - t_{m-1}, j_{m-1}, j_m). \tag{B.2.24}$$

Upon repeating this argument for the solution $((\hat{\Omega}, \hat{\mathcal{F}}, \hat{\mathbb{P}}), \{\hat{\mathcal{F}}_t\}, Y)$, it is clear that we obtain the finite-dimensional distributions of the process Y,

$$\hat{\mathbb{P}}\left[Y(t_1) = j_1, \ldots, Y(t_n) = j_n\right] = \prod_{m=1}^{n} p(t_m - t_{m-1}, j_{m-1}, j_m). \tag{B.2.25}$$

From (B.2.24) and (B.2.25), we see that the processes X and Y have the same finite-dimensional distributions. Thus the martingale problem for (A^Q, i_0) has the property of uniqueness. □

Remark B.2.9. Demonstrating that $((\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}}), \{\tilde{\mathcal{F}}_t\}, X)$ is a solution for the martingale problem for (A^Q, i_0) involves showing that for all functions f on I, the process $M^{\mathbf{f}}(X)$ defined by (B.2.2) is a martingale. However, we will see shortly that we can simplify this task; we can find a set of processes which, if we can show are martingales, imply that for all functions f on I, setting $\mathbf{f} = (f, Qf)$ the process $M^{\mathbf{f}}(X)$ is a martingale. We shall see later that this set of processes are the canonical martingales of the Markov chain.

B.2.3 The martingale problem for the Brownian motion

We present here the martingale problem for the Brownian motion, which we demonstrate is well-posed. The metric space is the set $\mathbb{R}^N$ with Euclidean distance as metric. Denoting

$$\mathcal{C}_0^2(\mathbb{R}^N) := \{f : \mathbb{R}^N \to \mathbb{R} \mid f \text{ is a continuous function with compact support which has continuous derivatives up to order } 2\}, \tag{B.2.26}$$

we then define

$$A^\nabla := \{(a + f, \nabla f) \mid a \in \mathbb{R} \text{ and } f \in \mathcal{C}_0^2(\mathbb{R}^N)\}, \tag{B.2.27}$$

for $\nabla f := \frac{1}{2}\sum_{n=1}^N \frac{\partial^2 f}{\partial x_n^2}$.

Remark B.2.10. The *support* of a function f is the closure of the set of points on which f is non-zero. A function f has *compact support* if its support is a compact set. A property of continuous functions with compact support with continuous derivatives up to order $K \in \mathbb{N}$ is that each derivative up to order K is a bounded function. Thus $A^\nabla \subset B(\mathbb{R}^N) \times B(\mathbb{R}^N)$, for $B(\mathbb{R}^N)$ given by (B.2.1).

Remark B.2.11. Clearly $\nabla(a + f) = \nabla f$, so in the definition of A^∇ we write only ∇f instead of $\nabla(a + f)$. Including $a \in \mathbb{R}$ in the definition of A^∇ is to allow $(1, 0) \in A^\nabla$, a technical condition, which we will show later, that is required to show that uniqueness holds for the joint martingale problem.

Theorem B.2.12. $\mathbf{B} = \{\mathbf{B}(t) \equiv (B_1(t), \dots, B_N(t))^\top : t \in [0, T]\}$ *is an N-dimensional standard Brownian motion on* $((\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}}), \{\tilde{\mathcal{F}}_t\})$ *if and only if* $((\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}}), \{\tilde{\mathcal{F}}_t\}, \mathbf{B})$ *is a solution to the martingale problem for* $(A^\nabla, \mathbf{0})$.

Proof. Assume that $\mathbf{B} = \{\mathbf{B}(t) \equiv (B_1(t), \dots, B_N(t))^\top : t \in [0, T]\}$ is an N-dimensional standard Brownian motion on $((\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}}), \{\tilde{\mathcal{F}}_t\})$. Then applying Itô's

Formula (Theorem C.14.2), which holds for every continuous function $f : \mathbb{R}^N \to \mathbb{R}$ which has continuous second-order derivatives,

$$f(\mathbf{B}(t)) = f(\mathbf{B}(0)) + \sum_{n=1}^{N} \int_0^t \frac{\partial f}{\partial x_n}(\mathbf{B}(\tau))\ \mathrm{d}B_n(\tau) + \frac{1}{2} \sum_{n=1}^{N} \int_0^t \frac{\partial^2 f}{\partial x_n^2}(\mathbf{B}(\tau))\ \mathrm{d}\tau \quad \text{(B.2.28)}$$

Rearranging, this becomes

$$f(\mathbf{B}(t)) - f(\mathbf{B}(0)) - \int_0^t \nabla f(\mathbf{B}(\tau))\ \mathrm{d}\tau = \sum_{n=1}^{N} \int_0^t \frac{\partial f}{\partial x_n}(\mathbf{B}(\tau))\ \mathrm{d}B_n(\tau). \quad \text{(B.2.29)}$$

As f has compact support, its derivatives are bounded functions. Thus the right-hand side of the above equation is a (continuous) martingale and hence so is the left-hand side. Immediately, we have that $((\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}}), \{\tilde{\mathcal{F}}_t\}, \mathbf{B})$ is a solution to the martingale problem for $(A^\nabla, \mathbf{0})$.

To show the reverse direction, we use a standard argument (for example, see Karatzas and Shreve [30], Remark 5.4.12). Suppose that $((\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}}), \{\tilde{\mathcal{F}}_t\}, \mathbf{B})$ is a solution to the martingale problem for $(A^\nabla, \mathbf{0})$. Then for every $\mathbf{f} \equiv (a + f, \nabla f) \in A^\nabla$, defining the process

$$M^{\mathbf{f}}(\mathbf{B})(t) := f(\mathbf{B}(t)) - f(\mathbf{B}(0)) - \int_0^t \frac{1}{2} \nabla f(\mathbf{B}(s))\ \mathrm{d}s, \quad \forall t \in [0, T], \quad \text{(B.2.30)}$$

we have $M^{\mathbf{f}}(\mathbf{B}) \in \mathcal{M}_0^c((\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}}), \{\tilde{\mathcal{F}}_t\})$.

Fix $j \in \{1, \ldots, N\}$ and set $f_j(\mathbf{x}) = x_j$. Then for $n = 1, \ldots, N$, using δ in this proof only to represent the Dirac delta function, we have

$$\frac{\partial f_j}{\partial x_n}(\mathbf{x}) = \delta_{jn}; \quad \frac{\partial^2 f_j}{\partial x_n^2}(\mathbf{x}) = 0. \quad \text{(B.2.31)}$$

Note that f_j does not have compact support. Choose a sequence of functions $\{g_j^{(m)}\}_{m \in \mathbb{N}} \subset \mathcal{C}_0^2(\mathbb{R}^N)$ such that $g_j^{(m)}(\mathbf{x}) = f_j(\mathbf{x})$ for $\|\mathbf{x}\| \leq m$ for all $m \in \mathbb{N}$. Fix any $a \in \mathbb{R}$. Then for $\mathbf{g}_j^{(m)} \equiv (a + g_j^{(m)}, \nabla g_j^{(m)}) \in A^\nabla$, we have $M^{\mathbf{g}_j^{(m)}}(\mathbf{B}) \in \mathcal{M}_0^c((\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}}), \{\tilde{\mathcal{F}}_t\})$.

Define the increasing sequence of $\{\tilde{\mathcal{F}}_t\}$-stopping times

$$T^m := \inf\{t \geq 0 \, : \, \|\mathbf{B}(t)\| \geq m\} \wedge T, \quad \forall m \in \mathbb{N}, \quad \text{(B.2.32)}$$

for which $T^m \Uparrow T$ as $m \to \infty$. Then $M^{\mathbf{g}_j^{(m)}}(\mathbf{B})(\cdot \wedge T^m) \in \mathcal{M}_0^c((\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}}), \{\tilde{\mathcal{F}}_t\})$ and for $\mathbf{f}_j \equiv (a + f_j, \nabla f_j)$,

$$M^{\mathbf{g}_j^{(m)}}(\mathbf{B})(t \wedge T^m) = M^{\mathbf{f}_j}(\mathbf{B})(t \wedge T^m) \quad \tilde{\mathbb{P}}\text{-a.s.}, \quad \forall t \in [0, T], \quad \text{(B.2.33)}$$

that is $M^{\mathbf{f}_j}(\mathbf{B})(\cdot \wedge T^m) \in \mathcal{M}_0^c((\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}}), \{\tilde{\mathcal{F}}_t\})$. With the localizing sequence $\{T^m\}_{m \in \mathbb{N}}$ of $\{\tilde{\mathcal{F}}_t\}$-stopping times, $M^{\mathbf{f}_j}(\mathbf{B})$ satisfies the definition of a local martingale, that is

$$M^{\mathbf{f}_j}(\mathbf{B}) \in \mathcal{M}_{0,\mathrm{loc}}^c((\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}}), \{\tilde{\mathcal{F}}_t\}). \quad \text{(B.2.34)}$$

From (B.2.30) and the definition of $\mathbf{f}_j$, we obtain

$$M^{\mathbf{f}_j}(\mathbf{B})(t) = f_j(\mathbf{B}(t)) - f_j(\mathbf{B}(0)) - \int_0^t \nabla f_j(\mathbf{B}(\tau))\ \mathrm{d}\tau = B_j(t) - B_j(0), \quad \text{(B.2.35)}$$

which we have just seen is a continuous local martingale. This holds for every $j \in \{1, \ldots, N\}$.

Fix $j, k \in \{1, \ldots, N\}$ and set $f_{jk}(\mathbf{x}) = x_j x_k$. Then for $n = 1, \ldots, N$,

$$\frac{\partial f_{jk}}{\partial x_n}(\mathbf{x}) = \delta_{jn} x_k + \delta_{kn} x_j; \quad \frac{\partial^2 f_{jk}}{\partial x_n^2}(\mathbf{x}) = 2\delta_{jn}\delta_{kn}. \quad \text{(B.2.36)}$$

Note that f_{jk} does not have compact support. Following a similar argument as above, choose a sequence of functions $\{g_{jk}^{(m)}\}_{m\in\mathbb{N}} \subset \mathcal{C}_0^2(\mathbb{R}^N)$ such that $g_{jk}^{(m)}(\mathbf{x}) = f_{jk}(\mathbf{x})$ for $\|\mathbf{x}\| \leq m$ for all $m \in \mathbb{N}$. Fix any $a \in \mathbb{R}$. Then for $\mathbf{g}_{jk}^{(m)} \equiv (a + g_{jk}^{(m)}, \nabla g_{jk}^{(m)}) \in A^{\nabla}$, we have $M^{\mathbf{g}_{jk}^{(m)}}(\mathbf{B}) \in \mathcal{M}_0^c((\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}}), \{\tilde{\mathcal{F}}_t\})$.

For the sequence $\{T^m\}_{m\in\mathbb{N}}$ of $\{\tilde{\mathcal{F}}_t\}$-stopping times defined in (B.2.32), we have $M^{\mathbf{g}_{jk}^{(m)}}(\mathbf{B})(\cdot \wedge T^m) \in \mathcal{M}_0^c((\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}}), \{\tilde{\mathcal{F}}_t\})$ and for $\mathbf{f}_{jk} \equiv (a + f_{jk}, \nabla f_{jk})$,

$$M^{\mathbf{g}_{jk}^{(m)}}(\mathbf{B})(t \wedge T^m) = M^{\mathbf{f}_{jk}}(\mathbf{B})(t \wedge T^m) \quad \tilde{\mathbb{P}}\text{-a.s.}, \quad \forall t \in [0, T], \quad \text{(B.2.37)}$$

that is $M^{\mathbf{f}_{jk}}(\mathbf{B})(\cdot \wedge T^m) \in \mathcal{M}_0^c((\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}}), \{\tilde{\mathcal{F}}_t\})$. With the localizing sequence $\{T^m\}_{m\in\mathbb{N}}$ of $\{\tilde{\mathcal{F}}_t\}$-stopping times, $M^{\mathbf{f}_{jk}}(\mathbf{B})$ satisfies the definition of a local martingale, that is

$$M^{\mathbf{f}_{jk}}(\mathbf{B}) \in \mathcal{M}_{0,\mathrm{loc}}^c((\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}}), \{\tilde{\mathcal{F}}_t\}). \quad \text{(B.2.38)}$$

From (B.2.30) and the definition of $\mathbf{f}_{jk}$, we obtain

$$\begin{aligned} M^{\mathbf{f}_{jk}}(\mathbf{B})(t) &= f_{jk}(\mathbf{B}(t)) - f_{jk}(\mathbf{B}(0)) - \int_0^t \nabla f_{jk}(\mathbf{B}(\tau))\ \mathrm{d}\tau \\ &= B_j(t)B_k(t) - B_j(0)B_k(0) - t\delta_{jk}. \end{aligned} \quad \text{(B.2.39)}$$

which we have just seen is a continuous local martingale. This holds for every $j, k \in \{1, \ldots, N\}$.

From (B.2.35) and (B.2.39), the quadratic co-variation process of the local martingales $\{B_j(t) - B_j(0)\}$ and $\{B_k(t) - B_k(0)\}$ is

$$[B_j, B_k](t) = t\delta_{jk}. \quad \text{(B.2.40)}$$

Then applying Lévy's Theorem (Theorem C.14.4), $\mathbf{B}$ is an N-dimensional standard Brownian motion on $((\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}}), \{\tilde{\mathcal{F}}_t\})$.

□

Corollary B.2.13. $((\Omega, \mathcal{F}, \mathbb{P}), \{\mathcal{F}_t\}, \mathbf{W})$ *is a solution to the martingale problem for* $(A^{\nabla}, \mathbf{0})$.

Proof. This follows immediately from Theorem B.2.12, as $\mathbf{W}$ is a N-dimensional standard Brownian motion on $((\Omega, \mathcal{F}, \mathbb{P}), \{\mathcal{F}_t\})$. □

Theorem B.2.14. *The martingale problem for* $(A^{\nabla}, \mathbf{0})$ *is well-posed.*

Proof. Recall the definition of well-posedness from Definition B.2.5. From Corollary B.2.13, we have existence of a solution.

For uniqueness, suppose that $((\hat{\Omega}, \hat{\mathcal{F}}, \hat{\mathbb{P}}), \{\hat{\mathcal{F}}_t\}, X)$ and $((\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}}), \{\tilde{\mathcal{F}}_t\}, Y)$ are solutions to the martingale problem for $(A^{\nabla}, \mathbf{0})$. From Theorem B.2.12, X is an N-dimensional standard Brownian motion on $((\hat{\Omega}, \hat{\mathcal{F}}, \hat{\mathbb{P}}), \{\hat{\mathcal{F}}_t\})$ and Y is an N-dimensional standard Brownian motion on $((\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}}), \{\tilde{\mathcal{F}}_t\})$. Since the finite-dimensional distributions of a Brownian motion do not depend on the underlying stochastic basis, it follows that X and Y have the same finite-dimensional distributions. □

B.2.4 The joint martingale problem

Finally in this section, we introduce the joint martingale problem for the Markov chain and the Brownian motion. Using the well-posedness of the martingale problem for the Markov chain and the well-posedness of the martingale problem for the Brownian motion, we show that the joint martingale problem is well-posed. This result is part of a theorem from Ethier and Kurtz [15]. The remaining part of the theorem, which is critical to proving the required martingale representation theorem, tells us that any two stochastic processes comprising part of the solution to the joint martingale problem are independent.

Recall the definitions of A^Q and A^{∇} from (B.2.4) and (B.2.27), respectively.

Define

$$A^{Q,\nabla} := \{(f_1 f_2, (Qf_1)\, f_2 + f_1 \nabla f_2) \,|\, (f_1, Qf_1) \in A^Q,\, (f_2, \nabla f_2) \in A^{\nabla}\}, \tag{B.2.41}$$

where

$$(f_1 f_2)(j, \mathbf{x}) := f_1(j) f_2(\mathbf{x}), \quad \forall (j, \mathbf{x}) \in I \times \mathbb{R}^N, \tag{B.2.42}$$

and

$$((Qf_1)\, f_2 + f_1 \nabla f_2)(j, \mathbf{x}) := (Qf_1)\,(j) f_2(\mathbf{x}) + f_1(j) \nabla f_2(\mathbf{x}), \quad \forall (j, \mathbf{x}) \in I \times \mathbb{R}^N. \tag{B.2.43}$$

It is then clear from $A^Q \subset B(I) \times B(I)$ and $A^{\nabla} \subset B(\mathbb{R}^N) \times B(\mathbb{R}^N)$ that

$$A^{Q,\nabla} \subset B(I \times \mathbb{R}^N) \times B(I \times \mathbb{R}^N). \tag{B.2.44}$$

Proposition B.2.15. *If* $((\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}}), \{\tilde{\mathcal{F}}_t\}, X)$ *is a solution to the martingale problem for* (A^Q, i_0) *and* $((\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}}), \{\tilde{\mathcal{F}}_t\}, \mathbf{Y})$ *is a solution to the martingale problem for* $(A^{\nabla}, \mathbf{0})$ *then* $((\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}}), \{\tilde{\mathcal{F}}_t\}, (X, \mathbf{Y}))$ *is a solution to the martingale problem for* $(A^{Q,\nabla}, (i_0, \mathbf{0}))$.

Proof. As $((\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}}), \{\tilde{\mathcal{F}}_t\}, X)$ is a solution to the martingale problem for (A^Q, i_0), we have

$$\tilde{\mathbb{P}}[X(0) = i_0] = 1 \tag{B.2.45}$$

and for all $\mathbf{f}_1 \equiv (f_1, Qf_1) \in A^Q$, defining

$$M^{\mathbf{f}_1}(X)(t) := f_1(X(t)) - f_1(X(0)) - \int_0^t (Qf_1)(X(s))\, \mathrm{d}s, \quad \forall t \in [0,T], \tag{B.2.46}$$

we have $M^{\mathbf{f}_1}(X) \in \mathcal{M}_0((\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}}), \{\tilde{\mathcal{F}}_t\})$.

As $((\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}}), \{\tilde{\mathcal{F}}_t\}, \mathbf{Y})$ is a solution to the martingale problem for $(A^\nabla, \mathbf{0})$, we have

$$\tilde{\mathbb{P}}[\mathbf{Y}(0) = \mathbf{0}] = 1 \tag{B.2.47}$$

and for all $\mathbf{f}_2 \equiv (f_2, \nabla f_2) \in A^\nabla$, defining

$$M^{\mathbf{f}_2}(\mathbf{Y})(t) := f_2(\mathbf{Y}(t)) - f_2(\mathbf{Y}(0)) - \int_0^t \nabla f_2(\mathbf{Y}(s))\, \mathrm{d}s, \quad \forall t \in [0,T], \tag{B.2.48}$$

we have $M^{\mathbf{f}_2}(\mathbf{Y}) \in \mathcal{M}_0((\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}}), \{\tilde{\mathcal{F}}_t\})$.

From (B.2.45), the Markov chain X starts in a fixed initial state i_0, in other words $X(0) = i_0$ $\tilde{\mathbb{P}}$-a.s. From (B.2.47), the Brownian motion $\mathbf{Y}$ starts at zero, in other words $\mathbf{Y}(0) = \mathbf{0}$ $\tilde{\mathbb{P}}$-a.s. As these events are trivially independent, we have

$$\tilde{\mathbb{P}}[(X, \mathbf{Y})(0) = (i_0, \mathbf{0})] = \tilde{\mathbb{P}}[X(0) = i_0, \mathbf{Y}(0) = \mathbf{0}] = \tilde{\mathbb{P}}[X(0) = i_0]\, \tilde{\mathbb{P}}[\mathbf{Y}(0) = \mathbf{0}] = 1, \tag{B.2.49}$$

so the initial condition for the martingale problem for $(A^{Q,\nabla}, (i_0, \mathbf{0}))$ holds.

Now fix $\mathbf{f} \equiv (f, g) \in A^{Q,\nabla}$. From the definition of $A^{Q,\nabla}$ in (B.2.41), $f = f_1 f_2$ and $g = (Qf_1) f_2 + f_1 \nabla f_2$ for some $(f_1, Qf_1) \in A^Q$ and $(f_2, \nabla f_2) \in A^\nabla$. Upon rearranging (B.2.46), we have

$$f_1(X(t)) = f_1(X(0)) + M^{\mathbf{f}_1}(X)(t) + \int_0^t (Qf_1)(X(s))\, \mathrm{d}s, \tag{B.2.50}$$

we see that

$$f_1(X(\cdot)) \in \mathcal{SM}((\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}}), \{\tilde{\mathcal{F}}_t\}). \tag{B.2.51}$$

As the Markov chain X has generator Q, it makes a finitely many jumps in the finite time interval $[0, T]$ (see Remark B.1.1). Then X is a finite variation process and hence $f_1(X(\cdot))$ is also a finite variation process. Then by Theorem C.13.4, $f_1(X(\cdot))$ is a *purely discontinuous semimartingale*. Furthermore, by Theorem C.13.6, the square-bracket co-variation process of $f_1(X(\cdot))$ and any semimartingale Z is given by

$$[Z, f_1(X(\cdot))](t) = \sum_{0<s\leq t} \triangle Z(s) \triangle f_1(X(s)). \tag{B.2.52}$$

Similarly, upon rearranging (B.2.48) we have

$$f_2(\mathbf{Y}(t)) = f_2(\mathbf{Y}(0)) + M^{\mathbf{f}_2}(\mathbf{Y})(t) + \int_0^t \nabla f_2(\mathbf{Y}(s))\ \mathrm{d}s, \tag{B.2.53}$$

and we see from the above equation that

$$f_2(\mathbf{Y}(\cdot)) \in \mathcal{SM}((\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}}), \{\tilde{\mathcal{F}}_t\}). \tag{B.2.54}$$

In particular, as both $\mathbf{Y}$ and f_2 are continuous then so is $f_2(\mathbf{Y}(\cdot))$. Hence we have $\triangle f_2(\mathbf{Y}(t)) = 0$ for all $t \in [0, T]$ and, consequently, from (B.2.52), we get

$$[f_1(X(\cdot)),\ f_2(\mathbf{Y}(\cdot))](t) = 0, \quad \forall t \in [0, T]. \tag{B.2.55}$$

As $f_1(X(\cdot))$ and $f_2(\mathbf{Y}(\cdot))$ are semimartingales on the same filtered probability space $((\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}}), \{\tilde{\mathcal{F}}_t\})$, we can apply the integration-by-parts formula (Theorem C.14.1) to expand $f(X(t), \mathbf{Y}(t))$.

$$\begin{aligned} f(X(t), \mathbf{Y}(t)) &\overset{(B.2.42)}{=} f_1(X(t))\, f_2(\mathbf{Y}(t)) \\ &= f_1(X(0))\, f_2(\mathbf{Y}(0)) + \int_0^t f_1(X(s_-))\ \mathrm{d}f_2(\mathbf{Y}(s)) \\ &\quad + \int_0^t f_2(\mathbf{Y}(s))\ \mathrm{d}f_1(X(s)) + [f_1(X(\cdot)),\ f_2(\mathbf{Y}(\cdot))](t) \\ &\overset{(B.2.55)}{=} f_1(X(0))\, f_2(\mathbf{Y}(0)) + \int_0^t f_1(X(s_-))\ \mathrm{d}f_2(\mathbf{Y}(s)) \\ &\quad + \int_0^t f_2(\mathbf{Y}(s))\ \mathrm{d}f_1(X(s)) \\ &\overset{(B.2.50),(B.2.53)}{=} f_1(X(0))\, f_2(\mathbf{Y}(0)) + \int_0^t f_1(X(s_-))\ \mathrm{d}M^{\mathbf{f}_2}(\mathbf{Y})(s) \\ &\quad + \int_0^t f_1(X(s_-))\, \nabla f_2(\mathbf{Y}(s))\ \mathrm{d}s + \int_0^t f_2(\mathbf{Y}(s))\ \mathrm{d}M^{\mathbf{f}_1}(X)(s) \\ &\quad + \int_0^t f_2(\mathbf{Y}(s))\ (Qf_1)\,(X(s))\ \mathrm{d}s. \end{aligned} \tag{B.2.56}$$

From the boundedness of f_1 and f_2 we have

$$\int_0^t f_1(X(s_-))\ \mathrm{d}M^{\mathbf{f}_2}(\mathbf{Y})(s) + \int_0^t f_2(\mathbf{Y}(s))\ \mathrm{d}M^{\mathbf{f}_1}(X)(s) \in \mathcal{M}_0((\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}}), \{\tilde{\mathcal{F}}_t\}). \tag{B.2.57}$$

From (B.2.42) and (B.2.43),

$$\begin{aligned} &f_1(X(t))\, f_2(\mathbf{Y}(t)) - f_1(X(0))\, f_2(\mathbf{Y}(0)) \\ &\qquad - \int_0^t [f_1(X(s))\, \nabla f_2(\mathbf{Y}(s)) + f_2(\mathbf{Y}(s))\ (Qf_1)\,(X(s))]\ \mathrm{d}s \\ &= f(X(t), \mathbf{Y}(t)) - f(X(0), \mathbf{Y}(0)) - \int_0^t g(X(s), \mathbf{Y}(s))\ \mathrm{d}s \overset{(B.2.2)}{=} M^{\mathbf{f}}((X, \mathbf{Y}))(t). \end{aligned} \tag{B.2.58}$$

Combining (B.2.56), (B.2.57) and (B.2.58), $M^{\mathbf{f}}((X,\mathbf{Y})) \in \mathcal{M}_0((\tilde{\Omega},\tilde{\mathcal{F}},\tilde{\mathbb{P}}),\{\tilde{\mathcal{F}}_t\})$. By the arbitrary choice of $\mathbf{f} \equiv (f,g) \in A^{Q,\nabla}$, this means that $((\tilde{\Omega},\tilde{\mathcal{F}},\tilde{\mathbb{P}}),\{\tilde{\mathcal{F}}_t\},(X,\mathbf{Y}))$ is a solution to the martingale problem for $(A^{Q,\nabla},(i_0,\mathbf{0}))$. □

Corollary B.2.16. $((\Omega,\mathcal{F},\mathbb{P}),\{\mathcal{F}_t\},(\alpha,\mathbf{W}))$ *is a solution to the martingale problem for* $(A^{Q,\nabla},(i_0,\mathbf{0}))$.

Proof. From Lemma B.2.6, $((\Omega,\mathcal{F},\mathbb{P}),\{\mathcal{F}_t\},\alpha)$ solves the martingale problem for (A^Q,i_0). From Lemma B.2.13 $((\Omega,\mathcal{F},\mathbb{P}),\{\mathcal{F}_t\},\mathbf{W})$ is a solution to the martingale problem for $(A^\nabla,\mathbf{0})$. Then applying Proposition B.2.15, $((\Omega,\mathcal{F},\mathbb{P}),\{\mathcal{F}_t\},(\alpha,\mathbf{W}))$ is a solution to the martingale problem for $(A^{Q,\nabla},(i_0,\mathbf{0}))$. □

We need the following technical lemma in order to prove Proposition B.2.20.

Lemma B.2.17. $(1,0) \in A^Q$ *and* $(1,0) \in A^\nabla$.

Proof. Consider first the Markov chain case. Recall the definition of A^Q from (B.2.4). For each $j \in I$, set $f(j) = 1$. The generator Q is conservative, so for each $i \in I$ we have

$$\sum_{j=1}^{D} q_{ij} = 0. \tag{B.2.59}$$

The ith row of Qf is

$$(Qf)_i = \sum_{j=1}^{D} q_{ij} f(j) = \sum_{j=1}^{D} q_{ij} = 0. \tag{B.2.60}$$

Thus $Qf = \mathbf{0}$, in other words it is a column vector consisting of zeroes. Hence $(1,0) \in A^Q$.

For the Brownian motion case, recall the definition of A^∇ from (B.2.27). Set $a = 1 \in \mathbb{R}$ and for each $\mathbf{x} \in \mathbb{R}^N$, set $f(\mathbf{x}) = 0$. The support of $f = 0$ is the empty set, which is a compact set. Also, $\nabla f(\mathbf{x}) = 0$. Then $(a+f,\nabla f) = (1,0) \in A^\nabla$. □

The next theorem is from Ethier and Kurtz [15], Chapter 4, Theorem 10.1. This theorem is used to show in the two propositions following it that the joint martingale problem is well-posed and that the stochastic processes comprising the solution are independent.

The martingale problem for A in Theorem B.2.19 has the same meaning as in Definition B.2.1, with the difference that no initial distribution is specified.

Remark B.2.18. A metric space (E,τ) is *separable* if it contains a countably dense set, that is a set with a countable number of elements whose closure is the entire space (E,τ).

Theorem B.2.19. *Let (E_1, τ_1) and (E_2, τ_2) be complete, separable metric spaces. For $i = 1, 2$, let $A_i \subset B(E_i) \times B(E_i)$ and $(1, 0) \in A_i$, and suppose that uniqueness holds for the martingale problem for A_i. Then uniqueness holds for the martingale problem for A given by*

$$A = \{(f_1 f_2, g_1 f_2 + f_1 g_2 \, : \, (f_1, g_1) \in A_1, \, (f_2, g_2) \in A_2\}. \tag{B.2.61}$$

In particular, if $X = (X_1, X_2)$ is a solution of the martingale problem for A and $X_1(0)$ and $X_2(0)$ are independent, then X_1 and X_2 are independent.

Proposition B.2.20. *The martingale problem for $(A^{Q,\nabla}, (i_0, \mathbf{0}))$ is well-posed.*

Proof. Recall the definition of well-posedness from Definition B.2.5. From Corollary B.2.16, we have existence of a solution.

From Theorem B.2.8, the martingale problem for (A^Q, i_0) is well-posed and thus has the property of uniqueness. From Theorem B.2.14, the martingale problem for $(A^\nabla, \mathbf{0})$ is well-posed and thus also has the property of uniqueness. From Lemma B.2.17, we have $(1, 0) \in A^Q$ and $(1, 0) \in A^\nabla$. We can then apply Theorem B.2.19 once we show that the assumptions of the theorem hold. These are that the underlying metric spaces are complete and separable.

For the Brownian motion, the metric space is the set $\mathbb{R}^N$ equipped with Euclidean distance as metric. This is well-known to be a complete, separable metric space.

For the Markov chain, the metric space is the state space I of the Markov chain equipped with the discrete metric τ, as in (B.2.3). Then (I, τ) is complete, since the terms of any Cauchy sequence $\{i_n\}$ in (I, τ) will be the same after some index, and it is also separable, since I is a countably dense subset of itself.

Hence we can apply Theorem B.2.19, to see that the martingale problem for $(A^{Q,\nabla}, (i_0, \mathbf{0}))$ has the property of uniqueness. □

We also have the following result, which is a consequence of the last part of Theorem B.2.19.

Proposition B.2.21. *If $((\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}}), \{\tilde{\mathcal{F}}_t\}, (X, \mathbf{Y}))$ is a solution to the martingale problem for $(A^{Q,\nabla}, (i_0, \mathbf{0}))$, then X and $\mathbf{Y}$ are independent under the probability measure $\tilde{\mathbb{P}}$.*

Proof. We have already checked in the proof of Proposition B.2.20 that the assumptions of Theorem B.2.19 hold. So we need only show that $\alpha(0)$ and $\mathbf{W}(0)$ are independent. However, $\alpha(0)$ and $\mathbf{W}(0)$ assume constant values with probability one and are thus trivially independent under the probability measure $\tilde{\mathbb{P}}$. Hence we can apply Theorem B.2.19 to see that X and $\mathbf{Y}$ are independent under the measure $\tilde{\mathbb{P}}$ and hence the proposition is proved. □

B.3 The canonical martingales of the Markov chain

In this section, we determine the canonical martingales of the Markov chain. These are square-integrable martingales, one for each pair of states in the state space of the Markov chain. They are strongly orthogonal to each other and we are able to find explicit expressions for both their square-bracket and angle-bracket quadratic variation processes.

Motivated by Rogers and Williams [46], we also prove Lemma B.3.28, which makes verifying a potential solution to the martingale problem for a Markov chain much easier. Essentially, if we can show that the potential solution satisfies the initial condition requirement of the martingale problem and that a specific set of processes are martingales, then the lemma tells us that the potential solution is indeed a solution. We will see that these specific set of processes are the canonical martingales of the Markov chain.

The material in Section B.3 is primarily motivated by Rogers and Williams [46], Section IV.20. This provided the basic formulation of the canonical martingales, through the construction of their square-bracket and angle-bracket quadratic variation processes. From Boel, Varaiya and Wong [6], we discerned the basic properties of the canonical martingales, such as their square-integrability and orthogonality.

We begin by defining a process R_{ij} which counts the number of jumps between state i and state j. As we see later in this section, this process is the square-bracket quadratic variation process of the corresponding canonical martingale of the Markov chain.

Definition B.3.1. For each $i, j = 1, \ldots, D$, $i \neq j$, define a mapping $R_{ij} : \Omega \times [0, T] \to \mathbb{N}_0$ by

$$R_{ij}(\omega, t) := \sum_{0 < s \leq t} \chi[\alpha(s_-) = i](\omega)\, \chi[\alpha(s) = j](\omega), \tag{B.3.1}$$

and for each $i = 1, \ldots, D$, set

$$R_{ii}(\omega, t) := 0, \qquad \forall (\omega, t) \in \Omega \times [0, T]. \tag{B.3.2}$$

Remark B.3.2. For $i \neq j$, $R_{ij}(t)$ counts the number of jumps between distinct states i and j up to time t.

Remark B.3.3. $R_{ij} = \{R_{ij}(t) : t \in [0, T]\}$ is an $\{\mathcal{F}_t\}$-adapted, non-decreasing, càdlàg process which is null at the origin.

Next we define a process $\tilde{R}_{ij}$ which, as we see later in this section, is the angle-bracket quadratic variation process of the corresponding canonical martingale, as well as being the compensator of R_{ij}.

Definition B.3.4. For each $i, j = 1, \ldots, D$, $i \neq j$, define a mapping $\tilde{R}_{ij} : \Omega \times [0, T] \to [0, \infty)$ by

$$\tilde{R}_{ij}(\omega, t) := q_{ij} \int_0^t \chi[\alpha(s) = i](\omega)\ \mathrm{d}s, \tag{B.3.3}$$

and for each $i = 1, \dots, D$, set

$$\tilde{R}_{ii}(\omega, t) := 0, \qquad \forall (\omega, t) \in \Omega \times [0, T]. \tag{B.3.4}$$

Remark B.3.5. For $i \neq j$, $\tilde{R}_{ij}(t)/q_{ij}$ measures the time that the Markov chain α spends in state i up to time t.

Remark B.3.6. $\tilde{R}_{ij} = \{\tilde{R}_{ij}(t) : t \in [0, T]\}$ is an $\{\mathcal{F}_t\}$-adapted, non-decreasing, continuous process which is null at the origin. Since $\tilde{R}_{ij}$ is continuous, it is previsible.

Finally, we define the set of processes $\{\mathcal{Q}_{ij}\}$ which turn out to be the canonical martingales of the Markov chain α.

Definition B.3.7. For each $i, j \in I$, define a process $\mathcal{Q}_{ij} : \Omega \times [0, T] \to [0, \infty)$ by

$$\mathcal{Q}_{ij}(\omega, t) = R_{ij}(\omega, t) - \tilde{R}_{ij}(\omega, t). \tag{B.3.5}$$

Remark B.3.8. As R_{ij} and $\tilde{R}_{ij}$ are $\{\mathcal{F}_t\}$-adapted, càdlàg processes which are null at the origin, then $\mathcal{Q}_{ij}$ is an $\{\mathcal{F}_t\}$-adapted, càdlàg process which is null at the origin.

Remark B.3.9. Upon expanding (B.3.5) by using the definitions of R_{ij} and $\tilde{R}_{ij}$ given in (B.3.1) and (B.3.3), we obtain for $i \neq j$,

$$\mathcal{Q}_{ij}(t) = \sum_{0 < s \leq t} \chi[\alpha(s_-) = i] \, \chi[\alpha(s) = j] - q_{ij} \int_0^t \chi[\alpha(s) = i] \, \mathrm{d}s. \tag{B.3.6}$$

Clearly, $\mathcal{Q}_{ii}(t) = 0$ a.s. for $i = 1, \dots, D$.

Lemma B.3.10. *For all $i, j = 1, \dots, D$,*

$$\mathcal{Q}_{ij} \in \mathcal{M}_{0,loc}((\Omega, \mathcal{F}, \mathbb{P}), \{\mathcal{F}_t\}). \tag{B.3.7}$$

Proof. This argument is taken from Rogers and Williams [46]. The result holds trivially for $\mathcal{Q}_{ii}$ since it is the constant zero martingale. So fix $i \neq j$ and consider the solution $((\Omega, \mathcal{F}, \mathbb{P}), \{\mathcal{F}_t\}, \alpha)$ to the martingale problem for (A^Q, i_0). Set $f := \chi[j]$ in (B.2.2) and write $M^{(j)} = M^{\mathbf{f}}(\alpha)$, to obtain for

$$\begin{aligned} M^{(j)}(t) &:= \chi[\alpha(t) = j] - \chi[\alpha(0) = j] - \int_0^t (Q\chi[j])(\alpha(s)) \, \mathrm{d}s \\ &= \chi[\alpha(t) = j] - \chi[\alpha(0) = j] - \int_0^t q_{\alpha(s), j} \, \mathrm{d}s. \end{aligned} \tag{B.3.8}$$

that $M^{(j)} \in \mathcal{M}_0((\Omega, \mathcal{F}, \mathbb{P}), \{\mathcal{F}_t\})$. For $i \neq j$, define the function

$$H^{(i)}(t) := \chi[\alpha(t_-) = i], \qquad \forall t \in [0, T]. \tag{B.3.9}$$

Then $H^{(i)}(t)$ is left-continuous, which means it is $\{\mathcal{F}_t\}$-previsible, and obviously uniformly bounded. We can immediately conclude that the (Lebesgue-Stieltjes) integral $(H^{(i)} \bullet M^{(j)})$ is a local martingale, that is

$$(H^{(i)} \bullet M^{(j)}) \in \mathcal{M}_{0,\mathrm{loc}}((\Omega, \mathcal{F}, \mathbb{P}), \{\mathcal{F}_t\}). \tag{B.3.10}$$

Recalling Remark B.1.1, for all $t \in [0, T]$ we have

$$
\begin{aligned}
(H^{(i)} \bullet M^{(j)})(t) &:= \int_0^t H^{(i)}(s)\, \mathrm{d}M^{(j)}(s) \\
&\overset{(B.3.8)}{=} \int_0^t \chi[\alpha(s_-) = i]\, \mathrm{d}\chi[\alpha(s) = j] - \int_0^t \chi[\alpha(s) = i]\, q_{\alpha(s),j}\, \mathrm{d}s \\
&= \sum_{0<s\leq t} \chi[\alpha(s_-) = i]\chi[\alpha(s) = j] - q_{ij} \int_0^t \chi[\alpha(s) = i]\, \mathrm{d}s.
\end{aligned}
\tag{B.3.11}
$$

Comparing the representation of $\mathcal{Q}_{ij}$ given by (B.3.6) to the representation of the local martingale $H^{(i)} \bullet M^{(j)}$ given by (B.3.11), we see that they are identical, in other words

$$
\mathcal{Q}_{ij} = H^{(i)} \bullet M^{(j)}. \tag{B.3.12}
$$

Thus we conclude from (B.3.10) that $\mathcal{Q}_{ij} \in \mathcal{M}_{0,\mathrm{loc}}((\Omega, \mathcal{F}, \mathbb{P}), \{\mathcal{F}_t\})$ for $i \neq j$. □

The next step is to examine the integrability properties of the processes R_{ij} and $\tilde{R}_{ij}$. We will see that these are finite-variation processes which have very strong integrability properties. However, we only need to show that they are square-integrable in order to show that $\mathcal{Q}_{ij}$ is a square-integrable martingale.

Lemma B.3.11. *For all $i, j \in I$, $E\left(\tilde{R}_{ij}(t)\right)^n < \infty$ for all $t \in [0, T]$ and for all $n \in \mathbb{N}$.*

Proof. For $i = j$, $R_{ij} = 0$ and the result is trivial. So assume that $i \neq j$. Then

$$
\tilde{R}_{ij}(t) = q_{ij} \int_0^t \chi[\alpha(s) = i](\omega)\, \mathrm{d}s \leq t\, q_{ij} \leq T\, q_{ij} \quad \text{a.s.} \tag{B.3.13}
$$

Thus we obtain $\mathrm{E}\left(\tilde{R}_{ij}(t)\right)^n \leq T^n\, q_{ij}^n < \infty$. □

Remark B.3.12. It is immediate from Lemma B.3.11 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

$$
V_{\tilde{R}_{ij}}(t) = \int_0^t |\, \mathrm{d}\tilde{R}_{ij}(s)| = \tilde{R}_{ij}(t) < \infty \quad \text{a.s.} \tag{B.3.14}
$$

Lemma B.3.13. *For all $i, j \in I$, $E(R_{ij}(t))^n < \infty$ for all $t \in [0, T]$ and for all $n \in \mathbb{N}$.*

Proof. The proof is taken directly from Rogers and Williams [46], Section IV.21. However, we repeat it here as we wish to emphasize the fact that the proof relies only on $\mathcal{Q}_{ij}$ being a *local* martingale with respect to the measure $\mathbb{P}$. For $i = j$, $R_{ij} = 0$ and the result is trivial. So assume that $i \neq j$. Fix $\epsilon > 0$ and introduce the

Doléan's-Dade exponential $Y := \mathcal{E}(\epsilon \mathcal{Q}_{ij})$ (see Theorem C.15.1 and Remark C.15.2) which, after some calculation, reduces to

$$Y(t) = \exp\{-\epsilon \tilde{R}_{ij}(t)\} \; (1+\epsilon)^{R_{ij}(t)}, \tag{B.3.15}$$

Then

$$\mathrm{d}Y(t) = \epsilon \, Y(t_-) \; \mathrm{d}\mathcal{Q}_{ij}(t). \tag{B.3.16}$$

As the process $Y(t_-)$ is an $\{\mathcal{F}_t\}$-adapted, càglàd process, it is locally bounded (see Rogers and Williams [46], Lemma IV.10.2 for this result) and previsible. Furthermore, from (B.3.16), since $\mathcal{Q}_{ij}$ is a local martingale, Y is a local martingale. However, Y is positive, implying that Y is a positive supermartingale (see Rogers and Williams [46], Lemma IV.14.3 for this result). Hence for all $t \in [0,T]$, as $\tilde{R}_{ij}(t) \leq tq_{ij}$ a.s., we have

$$1 = \mathrm{E}Y(0) \geq \mathrm{E}Y(t) \geq \exp\{-\epsilon q_{ij} t\} \, \mathrm{E}\,(1+\epsilon)^{R_{ij}(t)} \tag{B.3.17}$$

and so

$$\mathrm{E}\,(1+\epsilon)^{R_{ij}(t)} \leq \exp\{\epsilon t q_{ij}\}. \tag{B.3.18}$$

For all $n \in \mathbb{N}$ and $x \geq 0$, we have $x^n \leq \exp\{x\}$. Then as R_{ij} is a non-decreasing process which is null at the origin,

$$R_{ij}^n(t) \leq \exp\{R_{ij}(t)\} \leq \exp\{R_{ij}(T)\}. \tag{B.3.19}$$

Taking expectations in (B.3.19) and using (B.3.18) with $\epsilon = e - 1$,

$$\mathrm{E}\,(R_{ij}(t))^n \leq \mathrm{E}\,(\exp\{R_{ij}(T)\}) \leq \exp\{(e-1)Tq_{ij}\} < \infty. \tag{B.3.20}$$

□

Remark B.3.14. It is immediate from Lemma B.3.13 and the fact that R_{ij} is a non-decreasing process that R_{ij} has paths of finite variation over compact intervals, namely

$$V_{R_{ij}}(t) = \int_0^t |\, \mathrm{d}R_{ij}(s)| = R_{ij}(t) < \infty \quad \text{a.s.} \tag{B.3.21}$$

Remark B.3.15. As R_{ij} and $\tilde{R}_{ij}$ have paths of finite variation over compact intervals, then so does $\mathcal{Q}_{ij} = R_{ij} - \tilde{R}_{ij}$.

Lemma B.3.16. *For all $i, j \in I$, $\mathcal{Q}_{ij} \in \mathcal{M}_0^2((\Omega, \mathcal{F}, \mathbb{P}), \{\mathcal{F}_t\})$.*

Proof. From Lemma B.3.10, $\mathcal{Q}_{ij}$ is a local martingale which is null at the origin. We show that $\mathcal{Q}_{ij}$ is L^2-bounded. For all $t \in [0,T]$,

$$\begin{aligned} \mathrm{E}|\mathcal{Q}_{ij}(t)|^2 &\overset{(B.3.7)}{=} \mathrm{E}|R_{ij}(t) - \tilde{R}_{ij}(t)|^2 \\ &\leq 2\mathrm{E}\,(R_{ij}(t))^2 + 2\mathrm{E}\left(\tilde{R}_{ij}(t)\right)^2 \\ &\leq 2\mathrm{E}\,(R_{ij}(T))^2 + 2\mathrm{E}\left(\tilde{R}_{ij}(T)\right)^2. \end{aligned} \tag{B.3.22}$$

Thus

$$\sup_{t\in[0,T]} \mathrm{E}\,|\mathcal{Q}_{ij}(t)|^2 \leq 2\mathrm{E}\,(R_{ij}(T))^2 + 2\mathrm{E}\left(\tilde{R}_{ij}(T)\right)^2 < \infty, \tag{B.3.23}$$

the finiteness of $\mathrm{E}\,(R_{ij}(T))^2$ and $\mathrm{E}\left(\tilde{R}_{ij}(T)\right)^2$ coming from Lemmas B.3.13 and B.3.11. Applying Corollary C.16.2, we have that $\mathcal{Q}_{ij}$ is an L^2-bounded martingale and the lemma is shown. □

Remark B.3.17. Applying Theorem C.13.4 to $\mathcal{Q}_{ij}$, which is $\{\mathcal{F}_t\}$-adapted, càdlàg and has paths of finite variation on compact intervals, we have that $\mathcal{Q}_{ij}$ is a purely discontinuous martingale. Then, from Theorem C.13.6,

$$[\mathcal{Q}_{ij}, \mathcal{Q}_{ij}](t) = \sum_{0<s\leq t} (\triangle \mathcal{Q}_{ij}(s))^2 \quad \text{a.s.} \tag{B.3.24}$$

Lemma B.3.18. *For all* $t \in [0,T]$*,*

$$[\mathcal{Q}_{ij}, W_n](t) = \langle \mathcal{Q}_{ij}, W_n\rangle(t) = 0 \quad a.s. \tag{B.3.25}$$

for $i, j = 1, \ldots, D$ *and* $n = 1, \ldots, N$*.*

Proof. For $i = j$, $\mathcal{Q}_{ii} = 0$ and trivially $[\mathcal{Q}_{ii}, W_n](t) = \langle \mathcal{Q}_{ii}, W_n\rangle(t) = 0$ for $i = 1, \ldots, D$ and $n = 1, \ldots, N$, for all $t \in [0,T]$. So assume $i \neq j$. Applying Theorem C.13.6 to the martingale W_n and the purely discontinuous martingale $\mathcal{Q}_{ij}$, we have

$$[W_n, \mathcal{Q}_{ij}](t) = \sum_{0<s\leq t} \triangle W_n(s)\, \triangle \mathcal{Q}_{ij}(s). \tag{B.3.26}$$

As W_n is continuous, $\triangle W_n = 0$ and, substituting this into (B.3.26), we obtain the result. □

Lemma B.3.19. *For all* $t \in [0,T]$ *and* $i, j, a, b \in I$*, the following hold:*

1. $[\mathcal{Q}_{ij}, \mathcal{Q}_{ij}](t) = R_{ij}(t)$ *a.s.;*
2. $[\mathcal{Q}_{ij}, \mathcal{Q}_{ab}](t) = 0$ *a.s. if* $\{(i,j)\} \neq \{(a,b)\}$*.*

Proof. Suppose first that $i \neq j$. Recalling from Remark B.3.6 that $\tilde{R}_{ij}$ is a continuous process, we have

$$\triangle \mathcal{Q}_{ij}(t) \overset{(B.3.5)}{=} \triangle R_{ij}(t) - \triangle \tilde{R}_{ij}(t) = \triangle R_{ij}(t) \overset{(B.3.1)}{=} \chi[\alpha(t_-) = i]\, \chi[\alpha(t) = j]. \tag{B.3.27}$$

Note from (B.3.27) that $(\triangle \mathcal{Q}_{ij}(t))^2 = \triangle \mathcal{Q}_{ij}(t)$. Substituting (B.3.27) into (B.3.24), we obtain

$$[\mathcal{Q}_{ij}, \mathcal{Q}_{ij}](t) = \sum_{0<s\leq t} (\triangle \mathcal{Q}_{ij}(t))^2 \overset{(B.3.27)}{=} \sum_{0<s\leq t} \chi[\alpha(s_-) = i]\, \chi[\alpha(s) = j] \overset{(B.3.1)}{=} R_{ij}(t) \quad \text{a.s} \tag{B.3.28}$$

Since $\mathcal{Q}_{ii}(t) = 0$ a.s., trivially $[\mathcal{Q}_{ii}, \mathcal{Q}_{ii}](t) = 0$ a.s. for all $t \in [0, T]$.

Now consider a square-bracket quadratic co-variation process of $\mathcal{Q}_{ij}$ and $\mathcal{Q}_{ab}$, for $\{(i,j)\} \neq \{(a,b)\}$.

$$\begin{aligned}[\mathcal{Q}_{ij}, \mathcal{Q}_{ab}](t) &\overset{(B.3.26)}{=} \sum_{0<s\leq t} \triangle \mathcal{Q}_{ij}(s) \triangle \mathcal{Q}_{ab}(s) \\ &\overset{(B.3.27)}{=} \sum_{0<s\leq t} \chi[\alpha(s_-) = i]\, \chi[\alpha(s) = j] \chi[\alpha(s_-) = a]\, \chi[\alpha(s) = b] \\ &= 0 \quad \text{a.s.}\end{aligned} \tag{B.3.29}$$

□

Remark B.3.20. From Lemma B.3.19 and Theorem C.9.4, $\mathcal{Q}_{ij}^2 - R_{ij}$ is a uniformly integrable martingale.

Having shown that R_{ij} is the square-bracket quadratic variation process of $\mathcal{Q}_{ij}$, we show next that $\tilde{R}_{ij}$ is the angle-bracket quadratic variation process of $\mathcal{Q}_{ij}$.

Lemma B.3.21. *For all $t \in [0, T]$ and $i, j, a, b \in I$, the following hold*

1. $\langle \mathcal{Q}_{ij}, \mathcal{Q}_{ij} \rangle(t) = \tilde{R}_{ij}(t)$ *a.s.;*
2. $\langle \mathcal{Q}_{ij}, \mathcal{Q}_{ab} \rangle(t) = 0$ *a.s. for* $\{(i,j)\} \neq \{(a,b)\}$.

Proof. Suppose $i \neq j$ and note that the angle-bracket quadratic variation process $\langle \mathcal{Q}_{ij}, \mathcal{Q}_{ij} \rangle$ of $\mathcal{Q}_{ij}$ exists by the result in Lemma B.3.16 applied to Theorem C.12.4. We show that $\tilde{R}_{ij}$ satisfies all the conditions of being the angle-bracket quadratic variation process of $\mathcal{Q}_{ij}$, as given by Theorem C.9.1. We have the $\tilde{R}_{ij}$ is previsible, continuous, $\{\mathcal{F}_t\}$-adapted, non-decreasing and null at the origin. It remains to show that $\mathcal{Q}_{ij}^2 - \tilde{R}_{ij}$ is a martingale.

From Remark B.3.20, $\mathcal{Q}_{ij}^2 - R_{ij}$ is a martingale. From Lemma B.3.16, $\mathcal{Q}_{ij} = R_{ij} - \tilde{R}_{ij} \in \mathcal{M}_0^2(\{\mathcal{F}_t\}, \mathbb{P})$, so $R_{ij} - \tilde{R}_{ij}$ is certainly a martingale. Then

$$\left(\mathcal{Q}_{ij}^2 - \tilde{R}_{ij}\right) + \left(R_{ij} - \tilde{R}_{ij}\right) = \mathcal{Q}_{ij}^2 - \tilde{R}_{ij} \tag{B.3.30}$$

is also a martingale.

Since $\mathcal{Q}_{ii}(t) = 0$ a.s., trivially $\langle \mathcal{Q}_{ii}, \mathcal{Q}_{ii} \rangle(t) = 0$ a.s.

Similarly, as $[\mathcal{Q}_{ij}, \mathcal{Q}_{ab}](t) = 0$ a.s. then trivially $\langle \mathcal{Q}_{ij}, \mathcal{Q}_{ab} \rangle(t) = 0$ a.s. for $\{(i,j)\} \neq \{(a,b)\}$. □

Remark B.3.22. The processes R_{ij} and $\tilde{R}_{ij}$ were useful in constructing the canonical martingales $\{\mathcal{Q}_{ij}\}$ of the Markov chain α. However, as we see from Lemma B.3.19 and Lemma B.3.21, R_{ij} is the square-bracket quadratic variation process of $\mathcal{Q}_{ij}$

and $\tilde{R}_{ij}$ is the angle-bracket quadratic variation process of $\mathcal{Q}_{ij}$. 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 $[\mathcal{Q}_{ij}]$ to represent the square-bracket quadratic variation process of $\mathcal{Q}_{ij}$ and $\langle \mathcal{Q}_{ij} \rangle$ to represent the angle-bracket quadratic variation process of $\mathcal{Q}_{ij}$.

Remark B.3.23. From Lemma B.3.16, $\mathcal{Q}_{ij} \in \mathcal{M}_0^2((\Omega, \mathcal{F}, \mathbb{P}), \{\mathcal{F}_t\})$. Then from Theorem C.12.4, $\langle \mathcal{Q}_{ij} \rangle$ is the *compensator* of $[\mathcal{Q}_{ij}]$. It then follows from Theorem C.12.3 that for any nonnegative previsible process H we have

$$\mathrm{E} \int_0^T H(t) \; \mathrm{d}[\mathcal{Q}_{ij}](t) = \mathrm{E} \int_0^T H(t) \; \mathrm{d}\langle \mathcal{Q}_{ij} \rangle(t). \tag{B.3.31}$$

Remark B.3.24. The measures that we define next are used to specify the degree of uniqueness of the stochastic integrands of the stochastic integrals which have $\mathcal{Q}_{ij}$ as integrator.

Definition B.3.25. On the measurable space $(\Omega \times [0, T], \mathcal{P}^\star)$ and for each $i, j \in I$, $i \neq j$, define a measure $\nu_{\langle \mathcal{Q}_{ij} \rangle}$ by

$$\nu_{\langle \mathcal{Q}_{ij} \rangle}[A] := \mathrm{E} \int_0^T \chi_A(\omega, t) \; \mathrm{d}\langle \mathcal{Q}_{ij} \rangle(\omega, t), \qquad \forall A \in \mathcal{P}^\star, \tag{B.3.32}$$

and a measure $\nu_{[\mathcal{Q}_{ij}]}$ by

$$\nu_{[\mathcal{Q}_{ij}]}[A] := \mathrm{E} \int_0^T \chi_A(\omega, t) \; \mathrm{d}[\mathcal{Q}_{ij}](\omega, t), \qquad \forall A \in \mathcal{P}^\star. \tag{B.3.33}$$

Lemma B.3.26. *Recalling that Leb represents Lebesgue measure, we have for each $i, j \in I$, $i \neq j$,*

$$\nu_{[\mathcal{Q}_{ij}]} \ll \mathbb{P} \otimes Leb \quad \text{on } \mathcal{P}^\star. \tag{B.3.34}$$

Proof. Fix $i \neq j$. Applying Fubini's theorem to obtain the joint measure, we get from (B.3.32), for all $A \in \mathcal{P}^\star$,

$$\begin{aligned}
\nu_{\langle \mathcal{Q}_{ij} \rangle}[A] &= \mathrm{E} \int_0^T \chi_A(\omega, t) \; \mathrm{d}\langle \mathcal{Q}_{ij} \rangle(\omega, t) \\
&\overset{\text{Lemma B.3.21}}{=} \mathrm{E} \int_0^T \chi_A(\omega, t) \; \mathrm{d}\tilde{R}_{ij}(\omega, t) \\
&\overset{(B.3.3)}{=} \mathrm{E} \int_0^T \chi_A(\omega, t) \, q_{ij} \chi[\alpha(t) = i](\omega) \; \mathrm{d}t \\
&= \int_{\Omega \times [0,T]} \chi_A(\omega, t) \, q_{ij} \chi[\alpha(t) = i](\omega) \; \mathrm{d}(\mathbb{P} \otimes Leb) \\
&= \int_A q_{ij} \chi[\alpha(t) = i](\omega) \; \mathrm{d}(\mathbb{P} \otimes Leb).
\end{aligned} \tag{B.3.35}$$

This shows that the measure $\nu_{\langle \mathcal{Q}_{ij}\rangle}$ is absolutely continuous with respect to the measure $\mathbb{P}\otimes \mathit{Leb}$ on $\mathcal{P}^\star$, in other words

$$\nu_{\langle \mathcal{Q}_{ij}\rangle} \ll \mathbb{P}\otimes \mathit{Leb} \quad \text{on } \mathcal{P}^\star. \tag{B.3.36}$$

For all $A\in\mathcal{P}^\star$, we have

$$\begin{aligned}\nu_{\langle \mathcal{Q}_{ij}\rangle}[A] &\overset{(B.3.32)}{=} \mathrm{E}\int_0^T \chi_A(\omega,t)\;\mathrm{d}\langle \mathcal{Q}_{ij}\rangle(\omega,t)\\ &\overset{(B.3.31)}{=} \mathrm{E}\int_0^T \chi_A(\omega,t)\;\mathrm{d}[\mathcal{Q}_{ij}](\omega,t)\\ &\overset{(B.3.33)}{=} \nu_{[\mathcal{Q}_{ij}]}[A].\end{aligned} \tag{B.3.37}$$

Hence $\nu_{\langle \mathcal{Q}_{ij}\rangle} = \nu_{[\mathcal{Q}_{ij}]}$ on $\mathcal{P}^\star$ and it immediately follows this and (B.3.36) that $\nu_{[\mathcal{Q}_{ij}]} \ll \mathbb{P}\otimes \mathit{Leb}$ on $\mathcal{P}^\star$. □

Remark B.3.27. The next lemma is used in the proof of the martingale representation theorem for square-integrable martingales (Theorem B.4.6). The idea is that if we can show that each $\mathcal{Q}_{ij}$ remains a martingale under a specified change of measure $\tilde{\mathbb{P}}$, then $((\Omega,\mathcal{F},\tilde{\mathbb{P}}),\{\mathcal{F}_t\},\alpha)$ is a solution to the martingale problem for (A^Q, i_0).

Lemma B.3.28. *Suppose we are given a probability measure $\tilde{\mathbb{P}}$ on the filtered measurable space $(\Omega,\mathcal{F},\{\mathcal{F}_t\})$ such that $\mathcal{Q}_{ij}\in\mathcal{M}_0((\Omega,\mathcal{F},\tilde{\mathbb{P}}),\{\mathcal{F}_t\})$ for all $i,j = 1,\dots,D$. Then $((\Omega,\mathcal{F},\tilde{\mathbb{P}}),\{\mathcal{F}_t\},\alpha)$ is a solution to the martingale problem for (A^Q, i_0).*

Proof. Fix a function f on I and define

$$N^f(t) := \sum_{i,j=1}^{D} (f(j)-f(i))\,\mathcal{Q}_{ij}(t). \tag{B.3.38}$$

Then as $\mathcal{Q}_{ij}\in\mathcal{M}_0((\Omega,\mathcal{F},\tilde{\mathbb{P}}),\{\mathcal{F}_t\})$, for all $t\in[0,T]$,

$$\mathrm{E}_{\tilde{\mathbb{P}}}|N^f(t)| \le \sum_{i,j=1}^{D} |f(j)-f(i)|\,\mathrm{E}_{\tilde{\mathbb{P}}}|\mathcal{Q}_{ij}(t)| = 0 \tag{B.3.39}$$

and for all $0\le s\le t\le T$,

$$\begin{aligned}\mathrm{E}_{\tilde{\mathbb{P}}}\left(N^f(t)\,|\,\mathcal{F}_s\right) &\overset{(B.3.38)}{=} \sum_{i,j=1}^{D} (f(j)-f(i))\;\mathrm{E}_{\tilde{\mathbb{P}}}\left(\mathcal{Q}_{ij}(t)\,|\,\mathcal{F}_s\right)\\ &= \sum_{i,j=1}^{D} (f(j)-f(i))\;\mathcal{Q}_{ij}(s)\\ &\overset{(B.3.38)}{=} N^f(s),\end{aligned} \tag{B.3.40}$$

meaning $N^f \in \mathcal{M}_0((\Omega, \mathcal{F}, \tilde{\mathbb{P}}), \{\mathcal{F}_t\})$. We will show that $N^f \equiv M^{\mathbf{f}}(\alpha)$ and thus $M^{\mathbf{f}}(\alpha) \in \mathcal{M}_0((\Omega, \mathcal{F}, \tilde{\mathbb{P}}), \{\mathcal{F}_t\})$.

Note that for all $i \in I$, since $f(i) - f(i) = 0$, we can without loss of generality replace $(f(i) - f(i))\, \mathcal{Q}_{ii}(t)$ with

$$(f(i) - f(i)) \left(\sum_{0<s\leq t} \chi[\alpha(s_-) = i]\chi[\alpha(s) = i] - q_{ii} \int_0^t \chi[\alpha(s_-) = i] \, \mathrm{d}s \right). \tag{B.3.41}$$

Then

$$\begin{aligned}
N^f(t) &\overset{(B.3.38)}{=} \sum_{i,j=1}^{D} (f(j) - f(i))\, \mathcal{Q}_{ij}(t) \\
&\overset{(B.3.7),(B.3.41)}{=} \sum_{i,j=1}^{D} (f(j) - f(i)) \sum_{0<s\leq t} \chi[\alpha(s_-) = i]\chi[\alpha(s) = j] \\
&\quad - \sum_{i,j=1}^{D} (f(j) - f(i))\, q_{ij} \int_0^t \chi[\alpha(s_-) = i] \, \mathrm{d}s \qquad \text{(B.3.42)} \\
&= \sum_{0<s\leq t} (f(\alpha(s)) - f(\alpha(s_-))) \sum_{i,j=1}^{D} \chi[\alpha(s_-) = i]\chi[\alpha(s) = j] \\
&\quad - \int_0^t \sum_{j=1}^{D} (f(j) - f(\alpha(s_-)))\, q_{\alpha(s_-),j} \sum_{i=1}^{D} \chi[\alpha(s_-) = i] \, \mathrm{d}s
\end{aligned}$$

Since the Markov chain α has to be in some state in the state space I at each instant of time, the second summation in the first and second terms of the last line above both equal one, giving

$$\begin{aligned}
N^f(t) &= \sum_{0<s\leq t} (f(\alpha(s)) - f(\alpha(s_-))) - \int_0^t \sum_{j=1}^{D} (f(j) - f(\alpha(s_-)))\, q_{\alpha(s_-),j} \, \mathrm{d}s \\
&= f(\alpha(t)) - f(\alpha(0)) - \int_0^t \sum_{j=1}^{D} f(j)\, q_{\alpha(s_-),j} \, \mathrm{d}s \\
&\quad + \int_0^t f(\alpha(s_-)) \sum_{j=1}^{D} q_{\alpha(s_-),j} \, \mathrm{d}s
\end{aligned} \tag{B.3.43}$$

Recall that the generator Q is conservative, so for all $i \in I$, $\sum_{j=1}^{D} q_{ij} = 0$. Hence the last term in the last line above equals zero. Recalling Remark B.1.1 we then get

$$N^f(t) = f(\alpha(t)) - f(\alpha(0)) - \int_0^t (Qf)(\alpha(s_-))\ \mathrm{d}s$$
$$= f(\alpha(t)) - f(\alpha(0)) - \int_0^t (Qf)(\alpha(s))\ \mathrm{d}s \tag{B.3.44}$$
$$= M^{\mathbf{f}}(\alpha)(t).$$

Hence $M^{\mathbf{f}}(\alpha) \in \mathcal{M}_0((\Omega, \mathcal{F}, \tilde{\mathbb{P}}), \{\mathcal{F}_t\})$ and so $((\Omega, \mathcal{F}, \tilde{\mathbb{P}}), \{\mathcal{F}_t\}, \alpha)$ is a solution to the martingale problem for (A^Q, i_0). □

Remark B.3.29. Suppose that we can show that for some probability measure $\tilde{\mathbb{P}}$ on the filtered measurable space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\})$, we have

$$\tilde{\mathbb{P}}[\alpha(0) = i_0] = 1 \tag{B.3.45}$$

and $\mathcal{Q}_{ij} \in \mathcal{M}_0((\Omega, \mathcal{F}, \tilde{\mathbb{P}}), \{\mathcal{F}_t\})$ for each $i, j = 1, \dots, D$. Then applying Lemma B.3.28, we see that $((\Omega, \mathcal{F}, \tilde{\mathbb{P}}), \{\mathcal{F}_t\}, \alpha)$ is a solution to the martingale problem for (A^Q, i_0).

B.4 A martingale representation theorem

In this section, we prove a martingale representation theorem for processes which are square-integrable martingales and processes which are locally square-integrable martingales with respect to the filtration generated jointly by the N-dimensional standard Brownian motion $\mathbf{W}$ and the finite state space Markov chain α which has generator Q. We will see that we can write these locally square-integrable martingales as the sum of two stochastic integrals; one with Brownian motion as the integrator and the other with the canonical martingales $(\mathcal{Q}_{ij})_{i,j=1}^D$ of the Markov chain, as the integrator.

The highlight of this section is Theorem B.4.6, which is the martingale representation theorem for square-integrable martingales. In Theorem B.4.22, we extend Theorem B.4.6 to locally square-integrable martingales, using a standard argument.

This first lemma is a preliminary result for the martingale representation theorem for square-integrable martingales. The proof follows an argument adapted from Wong and Hajek [51], Proposition 6.7.3.

Lemma B.4.1. *For $Y \in \mathcal{M}_0^2(\{\mathcal{F}_t\}, \mathbb{P})$, we have for each $n = 1, \dots, N$,*

$$\langle Y, W_n \rangle(t) = \int_0^t \Lambda_n(\tau)\ \mathrm{d}\tau \quad a.s., \qquad \forall t \in [0, T], \tag{B.4.1}$$

for Λ_n a real-valued, previsible process satisfying $E\int_0^T |\Lambda_n(t)|\ \mathrm{d}t < \infty$. If $\tilde{\Lambda}_n$ is another real-valued, previsible process satisfying (B.4.1) then

$$E\int_0^T \tilde{\Lambda}_n(t)\ \mathrm{d}t = E\int_0^T \Lambda_n(t)\ \mathrm{d}t. \tag{B.4.2}$$

For each $i, j = 1, \dots, D$, $i \neq j$, we have

$$\langle Y, \mathcal{Q}_{ij} \rangle(t) = \int_0^t \Gamma_{ij}(\tau) \, \mathrm{d}\langle \mathcal{Q}_{ij} \rangle(\tau) \quad a.s., \qquad \forall t \in [0, T], \tag{B.4.3}$$

for Γ_{ij} a real-valued, previsible process satisfying $E \int_0^T |\Gamma_{ij}(t)| \, \mathrm{d}\langle \mathcal{Q}_{ij} \rangle(t) < \infty$. If $\tilde{\Gamma}_{ij}$ is another real-valued, previsible process satisfying (B.4.3) then

$$E \int_0^T \Gamma_{ij}(t) \, \mathrm{d}\langle \mathcal{Q}_{ij} \rangle(t) = E \int_0^T \tilde{\Gamma}_{ij}(t) \, \mathrm{d}\langle \mathcal{Q}_{ij} \rangle(t). \tag{B.4.4}$$

Proof. We will repeatedly use the Kunita-Watanabe inequality in this proof, which can be found in Theorem C.14.3.

For $Y \in \mathcal{M}_0^2(\{\mathcal{F}_t\}, \mathbb{P})$, the angle-bracket quadratic variation process $\langle Y \rangle$ exists and $Y^2 - \langle Y \rangle$ is a uniformly integrable martingale, with

$$\mathrm{E}\langle Y \rangle(T) < \infty. \tag{B.4.5}$$

Fix $n \in \{1, \dots, N\}$. For any bounded, previsible process H we have by the Kunita-Watanabe inequality,

$$\mathrm{E} \int_0^T |H(t)| \, |\, \mathrm{d}\langle Y, W_n \rangle(t)| \leq (\mathrm{E}\langle Y \rangle(T))^{\frac{1}{2}} \left(\mathrm{E} \int_0^T H^2(t) \, \mathrm{d}t \right)^{\frac{1}{2}}. \tag{B.4.6}$$

Motivated by the above inequality, we can define a measure ν_n on $(\Omega \times [0, T], \mathcal{P}^\star)$ by

$$\nu_n(A) := \mathrm{E} \int_0^T \chi_A(\omega, t) \, \mathrm{d}\langle Y, W_n \rangle(\omega, t), \qquad \forall A \in \mathcal{P}^\star, \tag{B.4.7}$$

and another measure μ on $(\Omega \times [0, T], \mathcal{P}^\star)$ by

$$\mu(A) := \mathrm{E} \int_0^T \chi_A(\omega, t) \, \mathrm{d}t, \qquad \forall A \in \mathcal{P}^\star. \tag{B.4.8}$$

From (B.4.6), we see that the measure ν_n is absolutely continuous with respect to the measure μ, in other words $\nu_n \ll \mu$ on $\mathcal{P}^\star$. An application of the Kunita-Watanabe inequality to (B.4.7), shows that ν_n is a finite measure.

$$|\nu_n|(\Omega \times [0, T]) = \mathrm{E} \int_0^T |\, \mathrm{d}\langle Y, W_n \rangle(\omega, t)| \leq (\mathrm{E}\langle Y \rangle(T))^{\frac{1}{2}} \, (\mathrm{E}T)^{\frac{1}{2}} \overset{(B.4.5)}{<} \infty. \tag{B.4.9}$$

From (B.4.8), we easily see that μ is a finite measure, since

$$|\mu|(\Omega \times [0, T]) = \mathrm{E} \int_0^T \mathrm{d}t = \mathrm{E}T < \infty. \tag{B.4.10}$$

Then from Theorem C.1.2, a real-valued, previsible Radon-Nikodým derivative Λ_n exists such that

$$\nu_n(A) = \int_A \Lambda_n \, \mathrm{d}\mu, \qquad \forall A \in \mathcal{P}^\star. \tag{B.4.11}$$

As $|\nu_n|(\Omega \times [0,T]) < \infty$, then Λ_n is integrable, that is

$$\mathrm{E} \int_0^T |\Lambda_n(t)| \; \mathrm{d}t < \infty. \tag{B.4.12}$$

By definition of the Radon-Nikodým derivative Λ_n, the measure ν_n given by (B.4.7) and the measure μ given by (B.4.8), for all bounded, previsible processes H the following equation holds

$$\mathrm{E} \int_0^T H(t) \; \mathrm{d}\langle Y, W_n\rangle(t) = \mathrm{E} \int_0^T H(t)\Lambda_n(t) \; \mathrm{d}t. \tag{B.4.13}$$

Claim B.4.2. For

$$A_n(t) := \langle Y, W_n\rangle(t) - \int_0^t \Lambda_n(\tau) \; \mathrm{d}\tau, \tag{B.4.14}$$

we have $A_n \in \mathcal{M}_0((\Omega, \mathcal{F}, \mathbb{P}), \{\mathcal{F}_t\})$.

First note that we can rewrite (B.4.13) as

$$\mathrm{E} \int_0^T H(t) \; \mathrm{d}A_n(t) = 0. \tag{B.4.15}$$

Taking expectations in (B.4.14) and applying the Kunita-Watanabe inequality, we get for all $t \in [0,T]$,

$$\begin{aligned} \mathrm{E}|A_n(t)| \; &\leq \; \mathrm{E}\,|\langle Y, W_n\rangle(t)| + \mathrm{E}\left|\int_0^t \Lambda_n(\tau) \; \mathrm{d}\tau\right| \\ &\overset{(B.4.13)}{=} 2\,\mathrm{E}|\langle Y, W_n\rangle(t)| \leq 2\; (\mathrm{E}\langle Y\rangle(T))^{\frac{1}{2}} \; (\mathrm{E}T)^{\frac{1}{2}} \overset{(B.4.5)}{<} \infty. \end{aligned} \tag{B.4.16}$$

Thus $A_n(t)$ is integrable. For $0 \leq u < v \leq T$, choose a bounded, previsible process

$$H(\omega, t) := Z(\omega)\,\chi(u,v](t), \tag{B.4.17}$$

where Z is an arbitrary bounded, $\{\mathcal{F}_u\}$–measurable function. Upon substituting H into (B.4.15), we obtain

$$\mathrm{E} \int_0^T Z\,\chi(u,v](t) \; \mathrm{d}A_n(t) = 0 \quad \Rightarrow \quad \mathrm{E}\left(Z \int_u^v \; \mathrm{d}A_n(t)\right) = 0 \tag{B.4.18}$$

Equivalently,

$$\mathrm{E}\left(ZA_n(v)\right) = \mathrm{E}\left(ZA_n(u)\right). \tag{B.4.19}$$

Then the arbitrary choice of Z and the definition of conditional expectation gives

$$\mathrm{E}\left(A_n(v)|\,\mathcal{F}_u\right) = A_n(u) \quad \text{a.s.} \tag{B.4.20}$$

By the arbitrary choice of $0 \leq u < v \leq T$, it follows that A_n is a martingale. From (B.4.14), we have $A_n(0) = 0$ a.s. Thus Claim B.4.2 is proved.

We now prove (B.4.1) by showing that $A_n(t) = 0$ for all $t \in [0,T]$.

Claim B.4.3. We have a.s.

$$\langle Y, W_n \rangle(t) = \int_0^t \Lambda_n(\tau)\ \mathrm{d}\tau, \qquad \forall t \in [0,T]. \tag{B.4.21}$$

From Claim B.4.2, A_n, as defined in (B.4.14), is a martingale. A_n is previsible as the first term on the right-hand side of (B.4.14) is an angle-bracket quadratic co-variation process, and hence previsible, and the second term on the right-hand side is continuous, and therefore also previsible. A_n is of finite variation, since from (B.4.14) we have

$$\int_0^t |\,\mathrm{d}A_n(\tau)| \le \int_0^t |\,\mathrm{d}\langle Y, W_n\rangle(\tau)| + \int_0^t |\Lambda_n(\tau)|\ \mathrm{d}\tau. \tag{B.4.22}$$

The second term on the right-hand side of the above inequality is clearly of finite variation (see (B.4.12)). For the first term on the right-hand side of the above inequality, an application of the Kunita-Watanabe inequality shows that a.s.

$$\int_0^t |\,\mathrm{d}\langle Y, W_n\rangle(\tau)| \le (\langle Y\rangle(t))^{\frac{1}{2}}\, t^{\frac{1}{2}} \le (\langle Y\rangle(T))^{\frac{1}{2}}\, (T)^{\frac{1}{2}} \overset{(B.4.5)}{<} \infty. \tag{B.4.23}$$

Thus A_n is a previsible, finite variation martingale. However, such martingales are equal to zero for all $t \in [0,T]$ with probability one (see, for example, Rogers and Williams [46], Theorem IV.1.7 for this result). Then from (B.4.14),

$$\langle Y, W_n\rangle(t) - \int_0^t \Lambda_n(\tau)\ \mathrm{d}\tau = 0 \quad \text{a.s.}, \qquad \forall t \in [0,T]. \tag{B.4.24}$$

Thus Claim B.4.3 is proved and, as this argument holds for each $n = 1, \ldots, N$, we have shown (B.4.1).

To show uniqueness, we again use Theorem C.1.2. If $\tilde{\Lambda}_n$ is another real-valued, previsible function such that (B.4.11) holds with Λ_n replaced by $\tilde{\Lambda}_n$ then, by Theorem C.1.2, $\tilde{\Lambda}_n = \Lambda_n$ μ-a.e. We see from the definition of μ in (B.4.8) that this gives (B.4.2).

We will now repeat this argument, but with the Brownian motion W_n replaced by the canonical martingales $\mathcal{Q}_{ij}$ of the Markov chain α. From Lemma B.3.16, $\mathcal{Q}_{ij} \in \mathcal{M}_0^2((\Omega, \mathcal{F}, \mathbb{P}), \{\mathcal{F}_t\})$, which means

$$\mathrm{E}\langle \mathcal{Q}_{ij}\rangle(T) < \infty. \tag{B.4.25}$$

Fix $i, j \in \{1, \ldots, D\}$, $i \neq j$. Then for any bounded, previsible process H we have by the Kunita-Watanabe inequality,

$$\mathrm{E}\int_0^T |H(t)|\,|\,\mathrm{d}\langle Y, \mathcal{Q}_{ij}\rangle(t)| \le (\mathrm{E}\langle Y\rangle(T))^{\frac{1}{2}} \left(\mathrm{E}\int_0^T H^2(t)\ \mathrm{d}\langle \mathcal{Q}_{ij}\rangle(t)\right)^{\frac{1}{2}}. \tag{B.4.26}$$

Motivated by the above inequality, we can define a measure η_{ij} on $(\Omega \times [0,T], \mathcal{P}^\star)$ by

$$\eta_{ij}(A) := \mathrm{E}\int_0^T \chi_A(\omega,t)\ \mathrm{d}\langle Y, \mathcal{Q}_{ij}\rangle(\omega,t), \qquad \forall A \in \mathcal{P}^\star, \tag{B.4.27}$$

and another measure ς_{ij} on $(\Omega \times [0,T], \mathcal{P}^\star)$ by

$$\varsigma_{ij}(A) := \mathrm{E}\int_0^T \chi_A(\omega,t)\ \mathrm{d}\langle \mathcal{Q}_{ij}\rangle(\omega,t), \qquad \forall A \in \mathcal{P}^\star. \tag{B.4.28}$$

Then (B.4.26) implies that the measure η_{ij} is absolutely continuous with respect to the measure ς_{ij}, in other words $\eta_{ij} \ll \varsigma_{ij}$ on $\mathcal{P}^\star$. An application of the Kunita-Watanabe inequality to (B.4.27) shows that η_{ij} is a finite measure.

$$\begin{aligned} |\eta_{ij}|(\Omega \times [0,T]) &= \mathrm{E}\int_0^T |\,\mathrm{d}\langle Y, \mathcal{Q}_{ij}\rangle(t)| \\ &\leq (\mathrm{E}\langle Y\rangle(T))^{\frac{1}{2}}\ (\mathrm{E}\langle \mathcal{Q}_{ij}\rangle(T))^{\frac{1}{2}} \overset{(B.4.5),(B.4.25)}{<} \infty. \end{aligned} \tag{B.4.29}$$

From (B.4.28) and the fact that $\langle \mathcal{Q}_{ij}\rangle$ is a non-decreasing process, we easily see that ς_{ij} is a finite measure.

$$|\varsigma_{ij}|(\Omega \times [0,T]) = \mathrm{E}\int_0^T |\,\mathrm{d}\langle \mathcal{Q}_{ij}\rangle(t)| = \mathrm{E}\langle \mathcal{Q}_{ij}\rangle(T) \overset{(B.4.25)}{<} \infty. \tag{B.4.30}$$

Thus, from Theorem C.1.2, a real-valued, previsible Radon-Nikodým derivative Γ_{ij} exists such that

$$\eta_{ij}(A) = \int_A \Gamma_{ij}\ \mathrm{d}\varsigma_{ij}, \qquad \forall A \in \mathcal{P}^\star. \tag{B.4.31}$$

As $|\eta_{ij}|(\Omega \times [0,T]) < \infty$, then Γ_{ij} is integrable, that is

$$\mathrm{E}\int_0^T |\Gamma_{ij}(t)|\ \mathrm{d}\langle \mathcal{Q}_{ij}\rangle(t) < \infty. \tag{B.4.32}$$

By the definition of the Radon-Nikodým derivative Γ_{ij}, the measure η_{ij} given by (B.4.27) and the measure ς_{ij} given by (B.4.28), for all bounded, previsible processes H the following equation holds:

$$\mathrm{E}\int_0^T H(t)\ \mathrm{d}\langle Y, \mathcal{Q}_{ij}\rangle(t) = \mathrm{E}\int_0^T H(t)\Gamma_{ij}(t)\ \mathrm{d}\langle \mathcal{Q}_{ij}\rangle(t). \tag{B.4.33}$$

Claim B.4.4. For

$$B_{ij}(t) := \langle Y, \mathcal{Q}_{ij}\rangle(t) - \int_0^t \Gamma_{ij}(\tau)\ \mathrm{d}\langle \mathcal{Q}_{ij}\rangle(\tau), \tag{B.4.34}$$

we have $B_{ij} \in \mathcal{M}_0((\Omega, \mathcal{F}, \mathbb{P}), \{\mathcal{F}_t\})$.

First note that we can rewrite (B.4.33) as

$$\mathrm{E}\int_0^T H(t)\ \mathrm{d}B_{ij}(t) = 0. \tag{B.4.35}$$

An application of the Kunita-Watanabe inequality shows that for all $t \in [0,T]$,

$$\begin{aligned}\mathrm{E}|B_{ij}(t)| \ &\leq\ \mathrm{E}\,|\langle Y, \mathcal{Q}_{ij}\rangle(t)| + \mathrm{E}\left|\int_0^t \Gamma_{ij}(\tau)\ \mathrm{d}\langle \mathcal{Q}_{ij}\rangle(\tau)\right| \\ &\overset{(B.4.33)}{=} 2\,\mathrm{E}|\langle Y, \mathcal{Q}_{ij}\rangle(t)| \leq 2\ (\mathrm{E}\langle Y\rangle(T))^{\frac{1}{2}}\ (\mathrm{E}\langle \mathcal{Q}_{ij}\rangle(T))^{\frac{1}{2}} \overset{(B.4.5),(B.4.25)}{<} \infty.\end{aligned} \tag{B.4.36}$$

So $B_{ij}(t)$ is integrable. For $0 \leq u < v \leq T$, choose a bounded, previsible process

$$H(\omega, t) := Z(\omega)\,\chi(u,v](t), \tag{B.4.37}$$

where Z is an arbitrary bounded, $\{\mathcal{F}_u\}$–measurable function. Upon substituting H into (B.4.35), we obtain

$$\mathrm{E}\int_0^T (Z\,\chi(u,v](t))\ \mathrm{d}B_{ij}(t) = 0 \quad \Rightarrow \quad \mathrm{E}\left(Z\int_u^v \mathrm{d}B_{ij}(t)\right) = 0 \tag{B.4.38}$$

Equivalently,

$$\mathrm{E}\left(Z\,B_{ij}(v)\right) = \mathrm{E}\left(Z\,B_{ij}(u)\right). \tag{B.4.39}$$

Then the arbitrary choice of Z and the definition of conditional expectation gives

$$\mathrm{E}\left(B_{ij}(v)|\,\mathcal{F}_u\right) = B_{ij}(u) \quad \text{a.s.} \tag{B.4.40}$$

By the arbitrary choice of $0 \leq u < v \leq T$, it follows that B_{ij} is a martingale. From (B.4.34), $B_{ij}(0) = 0$ a.s. Thus Claim B.4.4 is proved.

We will now prove (B.4.3) by showing that $B_{ij} = 0$.

Claim B.4.5. We have a.s.

$$\langle Y, \mathcal{Q}_{ij}\rangle(t) = \int_0^t \Gamma_{ij}(\tau)\ \mathrm{d}\langle \mathcal{Q}_{ij}\rangle(\tau), \qquad \forall t \in [0,T]. \tag{B.4.41}$$

From Claim B.4.4, B_{ij}, as defined by (B.4.34), is a martingale. B_{ij} is previsible as the first term on the right-hand side of (B.4.34) is an angle-bracket quadratic variation process, and hence previsible, and the second term on the right-hand side is continuous, and therefore also previsible. B_{ij} is also of finite variation, since from (B.4.34) we have

$$\int_0^t |\,\mathrm{d}B_{ij}(\tau)| \leq \int_0^t |\,\mathrm{d}\langle Y, \mathcal{Q}_{ij}\rangle(\tau)| + \int_0^t |\Gamma_{ij}(\tau)|\ \mathrm{d}\langle \mathcal{Q}_{ij}\rangle(\tau). \tag{B.4.42}$$

The second term on the right-hand side of the above inequality is clearly of finite variation. For the first term on the right-hand side of the above inequality, an application of the Kunita-Watanabe inequality shows that a.s.

$$\int_0^t |\,\mathrm{d}\langle Y, \mathcal{Q}_{ij}\rangle(\tau)| \leq (\langle Y\rangle(t))^{\frac{1}{2}} \, (\langle \mathcal{Q}_{ij}\rangle(t))^{\frac{1}{2}} \leq (\langle Y\rangle(T))^{\frac{1}{2}} \, (\langle \mathcal{Q}_{ij}\rangle(T))^{\frac{1}{2}} \overset{(B.4.5),(B.4.25)}{<} \infty. \tag{B.4.43}$$

Thus B_{ij} is a previsible, finite variation martingale. However, such martingales are equal to zero for all $t \in [0, T]$ with probability one (see Rogers and Williams [46], Theorem IV.1.7). Then from (B.4.34),

$$\langle Y, \mathcal{Q}_{ij}\rangle(t) - \int_0^t \Gamma_{ij}(\tau) \, \mathrm{d}\langle \mathcal{Q}_{ij}\rangle(\tau) = 0 \quad \text{a.s.,} \qquad \forall t \in [0, T]. \tag{B.4.44}$$

Thus Claim B.4.5 is proved and, as this argument holds for each $i, j = 1, \ldots, D$, $i \neq j$, we have shown (B.4.3).

To show uniqueness, we again use Theorem C.1.2. If $\tilde{\Gamma}_{ij}$ is another real-valued, previsible function such that (B.4.11) holds with Γ_{ij} replaced by $\tilde{\Gamma}_{ij}$ then, by Theorem C.1.2, $\tilde{\Gamma}_{ij} = \Gamma_{ij}$ ς_{ij}-a.e. We see from the definition of ς_{ij} in (B.4.28) that this gives (B.4.4).

□

The next theorem is the martingale representation theorem for square-integrable martingales. The proof follows an argument adapted from Wong and Hajek [51], Proposition 6.7.3. The theorem begins by proving the martingale representation theorem for uniformly bounded square-integrable martingales. It uses Lemma B.4.1 to construct the integrands for bounded square-integrable martingales and also the properties of the martingale problems examined in Section B.4.6. Extending this to any square-integrable martingale is then relatively straightforward.

Theorem B.4.6. *Suppose $Y \in \mathcal{M}_0^2((\Omega, \mathcal{F}, \mathbb{P}), \{\mathcal{F}_t\})$. Then there exists*

$$\boldsymbol{\Lambda} = (\Lambda_1, \ldots, \Lambda_N)^\top \in L^2(\mathbf{W}) \quad \textit{and} \quad \boldsymbol{\Gamma} = (\Gamma_{ij})_{i,j=1}^D \in L^2(\mathcal{Q}) \tag{B.4.45}$$

such that Y has the representation

$$Y(t) = \sum_{n=1}^N \int_0^t \Lambda_n(\tau) \, \mathrm{d}W_n(\tau) + \sum_{i,j=1}^D \int_0^t \Gamma_{ij}(\tau) \, \mathrm{d}\mathcal{Q}_{ij}(\tau) \quad \textit{a.s.,} \qquad \forall t \in [0, T], \tag{B.4.46}$$

with the square-bracket quadratic variation process of Y given by

$$[Y](t) = \sum_{n=1}^N \int_0^t \Lambda_n^2(\tau) \, \mathrm{d}\tau + \sum_{i,j=1}^D \int_0^t \Gamma_{ij}^2(\tau) \, \mathrm{d}[\mathcal{Q}_{ij}](\tau) \quad \textit{a.s.,} \qquad \forall t \in [0, T]. \tag{B.4.47}$$

Moreover, $\mathbf{\Lambda}$ *and* $\mathbf{\Gamma}$ *are unique in the sense that if* $\tilde{\mathbf{\Lambda}} = \left(\tilde{\Lambda}_1, \ldots, \tilde{\Lambda}_N\right)^\top \in L^2(\mathbf{W})$ *and* $\tilde{\mathbf{\Gamma}} = \left(\tilde{\Gamma}_{ij}\right)_{i,j=1}^{D} \in L^2(\mathcal{Q})$ *are such that (B.4.46) holds with* Λ_n *replaced by* $\tilde{\Lambda}_n$ *and* Γ_{ij} *replaced by* $\tilde{\Gamma}_{ij}$*, then* $\mathbf{\Lambda} = \tilde{\mathbf{\Lambda}}$ $(\mathbb{P} \otimes Leb)$*-a.e. and* $\mathbf{\Gamma} = \tilde{\mathbf{\Gamma}}$ $\nu_{[\mathcal{Q}]}$*-a.e.*

Proof. Note that $\langle Y, W_n \rangle$ and $\langle Y, \mathcal{Q}_{ij} \rangle$ exist since $Y, W_n, \mathcal{Q}_{ij} \in \mathcal{M}_0^2((\Omega, \mathcal{F}, \mathbb{P}), \{\mathcal{F}_t\})$ for each $n = 1, \ldots, N$ and $i, j = 1, \ldots, D$.

First we prove the result for uniformly bounded $Y \in \mathcal{M}_0^2((\Omega, \mathcal{F}, \mathbb{P}), \{\mathcal{F}_t\})$. From Lemma B.4.1, we obtain the set of real-valued, previsible processes $\{\Lambda_n\}_{n=1}^N$, which satisfy

$$\langle Y, W_n \rangle(t) = \int_0^t \Lambda_n(\tau) \, \mathrm{d}\tau \quad \text{a.s.}, \qquad \forall t \in [0, T], \tag{B.4.48}$$

and

$$\mathrm{E} \int_0^T |\Lambda_n(t)| \, \mathrm{d}t < \infty. \tag{B.4.49}$$

Setting $\Gamma_{ii} := 0$ $(\mathbb{P} \otimes Leb)$-a.e. for $i = 1, \ldots, D$,, we also obtain the set of real-valued, previsible processes $\{\Gamma_{ij}\}_{i,j=1}^D$, which satisfy for $i \neq j$,

$$\langle Y, \mathcal{Q}_{ij} \rangle(t) = \int_0^t \Gamma_{ij}(\tau) \, \mathrm{d}\langle \mathcal{Q}_{ij} \rangle(\tau), \quad \text{a.s.}, \quad \forall t \in [0, T], \tag{B.4.50}$$

and

$$\mathrm{E} \int_0^T |\Gamma_{ij}(t)| \, \mathrm{d}\langle \mathcal{Q}_{ij} \rangle(t) < \infty. \tag{B.4.51}$$

As $\langle \mathcal{Q}_{ij} \rangle$ is the compensator of $[\mathcal{Q}_{ij}]$, applying the result in Remark B.3.23 to Γ_{ij}, which is previsible, we get

$$\mathrm{E} \int_0^T |\Gamma_{ij}(t)| \, \mathrm{d}[\mathcal{Q}_{ij}](t) = \mathrm{E} \int_0^T |\Gamma_{ij}(t)| \, \mathrm{d}\langle \mathcal{Q}_{ij} \rangle(t) \overset{(B.4.51)}{<} \infty. \tag{B.4.52}$$

For each $k \in \mathbb{N}$, define the indicator function

$$H^{(k)}(t) := \chi \left[\max_{n=1,\ldots,N} |\Lambda_n(t)| \leq k \text{ and } \max_{i,j=1,\ldots,D} |\Gamma_{ij}(t)| \leq k \right]. \tag{B.4.53}$$

For each $t \in [0, T]$, $H^{(k)}(t) \to 1$ a.s. as $k \to \infty$. We also define

$$\tilde{Y}^{(k)}(t) := \int_0^t H^{(k)}(\tau) \, \mathrm{d}Y(\tau) \tag{B.4.54}$$

and

$$\hat{Y}^{(k)}(t) := \sum_{n=1}^N \int_0^t H^{(k)}(\tau) \, \Lambda_n(\tau) \, \mathrm{d}W_n(\tau) + \sum_{i,j=1}^D \int_0^t H^{(k)}(\tau) \, \Gamma_{ij}(\tau) \, \mathrm{d}\mathcal{Q}_{ij}(\tau). \tag{B.4.55}$$

$\tilde{Y}^{(k)}$ and $\hat{Y}^{(k)}$ are in $\mathcal{M}_0^2((\Omega, \mathcal{F}, \mathbb{P}), \{\mathcal{F}_t\})$ as the integrands are uniformly bounded and the integrators are in $\mathcal{M}_0^2((\Omega, \mathcal{F}, \mathbb{P}), \{\mathcal{F}_t\})$.

Remark B.4.7. We aim to show the following for each $t \in [0,T]$:

1. $\lim_{k\to\infty} \mathrm{E}|\tilde{Y}^{(k)}(t) - Y(t)|^2 = 0$;
2. $\tilde{Y}^{(k)}(t) = \hat{Y}^{(k)}(t)$ a.s. for all $k \in \mathbb{N}$; and
3. $\lim_{k\to\infty} \mathrm{E}|\hat{Y}^{(k)}(t) - \sum_{n=1}^{N} \int_0^t \Lambda_n(\tau)\ \mathrm{d}W_n(\tau) - \sum_{i,j=1}^{D} \int_0^t \Gamma_{ij}(\tau)\ \mathrm{d}\mathcal{Q}_{ij}(\tau)|^2 = 0$.

From these three items, we will conclude that for each $t \in [0,T]$,

$$Y(t) = \sum_{n=1}^{N} \int_0^t \Lambda_n(\tau)\ \mathrm{d}W_n(\tau) + \sum_{i,j=1}^{D} \int_0^t \Gamma_{ij}(\tau)\ \mathrm{d}\mathcal{Q}_{ij}(\tau) \quad \text{a.s.} \tag{B.4.56}$$

Claim B.4.8. $\lim_{k\to\infty} \mathrm{E}|\tilde{Y}^{(k)}(t) - Y(t)|^2 = 0$ for all $t \in [0,T]$.

Fix $t \in [0,T]$. Using the Itô isometry,

$$\begin{aligned} &\mathrm{E}\left|\tilde{Y}^{(k)}(t) - Y(t)\right|^2 \\ &\overset{(B.4.54)}{=} \mathrm{E}\left|\int_0^t \left(H^{(k)}(\tau) - 1\right)\ \mathrm{d}Y(\tau)\right|^2 = \mathrm{E}\int_0^t \left(H^{(k)}(\tau) - 1\right)^2\ \mathrm{d}[Y](\tau). \end{aligned} \tag{B.4.57}$$

For the indicator function $H^{(k)}$ given by (B.4.53), we have for each $t \in [0,T]$ that $H^{(k)}(t) \to 1$ a.s. as $k \to \infty$. Then $(H^{(k)} - 1)^2 \to 0$ $(\mathbb{P} \otimes \mathit{Leb})$-a.e. as $k \to \infty$. Applying the Lebesgue Dominated Convergence Theorem to the sequence of functions $\{(H^{(k)} - 1)^2\}_{k\in\mathbb{N}}$ (which is dominated by the number 1), we get

$$\lim_{k\to\infty} \mathrm{E}\int_0^t \left(H^{(k)}(\tau) - 1\right)^2\ \mathrm{d}[Y](\tau) = \mathrm{E}\int_0^t \lim_{k\to\infty} \left(H^{(k)}(\tau) - 1\right)^2\ \mathrm{d}[Y](\tau) = 0. \tag{B.4.58}$$

Letting k go to infinity in (B.4.57) and using (B.4.58), we obtain

$$\lim_{k\to\infty} \mathrm{E}\left|\tilde{Y}^{(k)}(t) - Y(t)\right|^2 = 0, \tag{B.4.59}$$

which proves Claim B.4.8.

For all $k \in \mathbb{N}$ and $t \in [0,T]$, define

$$Z^{(k)}(t) := \tilde{Y}^{(k)}(t) - \hat{Y}^{(k)}(t). \tag{B.4.60}$$

Remark B.4.9. Next we show that $\tilde{Y}^{(k)}(t) = \hat{Y}^{(k)}(t)$ a.s. for all $t \in [0,T]$, which is the second item we want to show in Remark B.4.7. We construct a new measure on $(\Omega, \mathcal{F})$ which is the expectation under the measure $\mathbb{P}$ of the exponential of $Z^{(k)}$ times a specified scalar constant. The scalar constant is to ensure that the exponential is strictly positive. Using a Girsanov theorem and the martingale problems in Section B.2, we show that the new measure is identical to the measure $\mathbb{P}$. It will follow that $Z^{(k)}(t) = 0$ a.s. and hence $\tilde{Y}^{(k)}(t) = \hat{Y}^{(k)}(t)$ a.s.

Remark B.4.10. As $\tilde{Y}^{(k)}, \hat{Y}^{(k)} \in \mathcal{M}_0^2((\Omega, \mathcal{F}, \mathbb{P}), \{\mathcal{F}_t\})$ and $Z^{(k)} = \tilde{Y}^{(k)} - \hat{Y}^{(k)}$ then $Z^{(k)} \in \mathcal{M}_0^2((\Omega, \mathcal{F}, \mathbb{P}), \{\mathcal{F}_t\})$. Thus the angle-bracket quadratic variation process of $Z^{(k)}$ exists.

Fix $k \in \mathbb{N}$.

Claim B.4.11. For each $n = 1, \dots, N$,

$$\langle Z^{(k)}, W_n \rangle(t) = 0 \text{ a.s.}, \qquad \forall t \in [0, T]. \tag{B.4.61}$$

Fix $t \in [0, T]$. From (B.4.48), we have

$$\langle Y, W_n \rangle(t) = \int_0^t \Lambda_n(\tau) \, \mathrm{d}\tau \text{ a.s.}, \qquad \forall t \in [0, T]. \tag{B.4.62}$$

Consider

$$\begin{aligned}
\langle Z^{(k)}, W_n \rangle(t) &\overset{(B.4.60)}{=} \langle \tilde{Y}^{(k)}, W_n \rangle(t) - \langle \hat{Y}^{(k)}, W_n \rangle(t) \\
&\overset{(B.4.54),(B.4.55)}{=} \langle \int_0^{\cdot} H^{(k)}(\tau) \, \mathrm{d}Y(\tau), W_n \rangle(t) \\
&\quad - \langle \sum_{m=1}^{N} \int_0^{\cdot} H^{(k)}(\tau) \, \Lambda_m(\tau) \, \mathrm{d}W_m(\tau) + \sum_{i,j=1}^{D} \int_0^{\cdot} H^{(k)}(\tau) \, \Gamma_{ij}(\tau) \, \mathrm{d}\mathcal{Q}_{ij}(\tau), \; W_n \rangle(t) \\
&= \int_0^t H^{(k)}(\tau) \, \mathrm{d}\langle Y, W_n \rangle(\tau) - \sum_{m=1}^{N} \int_0^t H^{(k)}(\tau) \, \Lambda_m(\tau) \, \mathrm{d}\langle W_m, W_n \rangle(\tau) \\
&\quad + \sum_{i,j=1}^{D} \int_0^t H^{(k)}(\tau) \, \Gamma_{ij}(\tau) \, \mathrm{d}\langle \mathcal{Q}_{ij}, W_n \rangle(\tau).
\end{aligned} \tag{B.4.63}$$

By Lemma B.3.18, $\langle \mathcal{Q}_{ij}, W_n \rangle(t) = 0$ a.s. for $i, j = 1, \dots, D$. Then using (B.4.62), we get

$$\langle Z^{(k)}, W_n \rangle(t) = \int_0^t H^{(k)}(\tau) \Lambda_n(\tau) \, \mathrm{d}\tau - \int_0^t H^{(k)}(\tau) \, \Lambda_n(\tau) \, \mathrm{d}\tau = 0, \tag{B.4.64}$$

and Claim B.4.11 is shown.

Claim B.4.12. For each $i, j = 1, \dots, D$,

$$\langle Z^{(k)}, \mathcal{Q}_{ij} \rangle(t) = 0 \text{ a.s.}, \quad \forall t \in [0, T]. \tag{B.4.65}$$

Fix $t \in [0, T]$. From (B.4.50), we have

$$\langle Y, \mathcal{Q}_{ij} \rangle(t) = \int_0^t \Gamma_{ij}(\tau) \, \mathrm{d}\langle \mathcal{Q}_{ij} \rangle(\tau) \text{ a.s.}, \qquad \forall t \in [0, T]. \tag{B.4.66}$$

Consider

$$
\begin{aligned}
&\langle Z^{(k)}, \mathcal{Q}_{ij}\rangle(t) \overset{(B.4.60)}{=} \langle \tilde{Y}^{(k)}, \mathcal{Q}_{ij}\rangle(t) - \langle \hat{Y}^{(k)}, \mathcal{Q}_{ij}\rangle(t) \\
&\overset{(B.4.54),(B.4.55)}{=} \langle \int_0^{\cdot} H^{(k)}(\tau)\,\mathrm{d}Y(\tau), \mathcal{Q}_{ij}\rangle(t) \\
&\qquad - \langle \sum_{n=1}^{N} \int_0^{\cdot} H^{(k)}(\tau)\,\Lambda_n(\tau)\ \mathrm{d}W_n(\tau) + \sum_{a,b=1}^{D} \int_0^{\cdot} H^{(k)}(\tau)\,\Gamma_{ab}(\tau)\ \mathrm{d}\mathcal{Q}_{ab}(\tau),\ \mathcal{Q}_{ij}\rangle(t) \\
&= \int_0^t H^{(k)}(\tau)\,\mathrm{d}\langle Y, \mathcal{Q}_{ij}\rangle(\tau) - \sum_{m=1}^{N} \int_0^t H^{(k)}(\tau)\,\Lambda_n(\tau)\ \mathrm{d}\langle W_n, \mathcal{Q}_{ij}\rangle(\tau) \\
&\quad - \sum_{a,b=1}^{D} \int_0^t H^{(k)}(\tau)\,\Gamma_{ab}(\tau)\ \mathrm{d}\langle \mathcal{Q}_{ab}, \mathcal{Q}_{ij}\rangle(\tau).
\end{aligned}
\tag{B.4.67}
$$

By Lemma B.3.18, $\langle \mathcal{Q}_{ij}, W_n\rangle(t) = 0$ a.s. for $n = 1, \ldots, N$. Then using (B.4.66) and the results in Lemma B.3.21, we get

$$
\langle Z^{(k)}, \mathcal{Q}_{ij}\rangle(t) = \int_0^t H^{(k)}(\tau)\Gamma_{ij}(\tau)\,\mathrm{d}\langle \mathcal{Q}_{ij}\rangle(\tau) - \int_0^t H^{(k)}(\tau)\,\Gamma_{ij}(\tau)\ \mathrm{d}\langle \mathcal{Q}_{ij}\rangle(\tau) = 0,
\tag{B.4.68}
$$

and Claim B.4.12 is shown.

Claim B.4.13. The jumps of $Z^{(k)}$ are uniformly bounded.

The jump of $Z^{(k)}$ at time t is $\triangle Z^{(k)}(t) = \triangle \tilde{Y}^{(k)}(t) - \triangle \hat{Y}^{(k)}(t)$. Consider $\triangle \tilde{Y}^{(k)}(t)$. By the boundedness of $H^{(k)}$, which takes values 0 or 1, and applying Theorem C.14.5,

$$
\triangle \tilde{Y}^{(k)}(t) \overset{(B.4.54)}{=} \triangle \left(\int_0^t H^{(k)}(\tau)\,\mathrm{d}Y(\tau) \right) = H^{(k)}(t) \triangle Y(t) \le \triangle Y(t) \le C,
\tag{B.4.69}
$$

where $C > 0$ is a constant which uniformly bounds the jumps of Y (recall the assumption that Y is bounded). Now consider $\triangle \hat{Y}^{(k)}(t)$. Upon applying Theorem C.14.5 and using the pathwise continuity of $\mathbf{W}$, we get

$$
\begin{aligned}
&\triangle \hat{Y}^{(k)}(t) \\
&\overset{(B.4.55)}{=} \triangle \left(\sum_{n=1}^{N} \int_0^t H^{(k)}(\tau)\,\Lambda_n(\tau)\ \mathrm{d}W_n(\tau) + \sum_{i,j=1}^{D} \int_0^t H^{(k)}(\tau)\,\Gamma_{ij}(\tau)\ \mathrm{d}\mathcal{Q}_{ij}(\tau) \right) \\
&= \sum_{n=1}^{N} H^{(k)}(t)\,\Lambda_n(t) \triangle W_n(t) + \sum_{i,j=1}^{D} H^{(k)}(t)\,\Gamma_{ij}(t) \triangle \mathcal{Q}_{ij}(t) \\
&= \sum_{i,j=1}^{D} H^{(k)}(t)\,\Gamma_{ij}(t) \triangle \mathcal{Q}_{ij}(t).
\end{aligned}
\tag{B.4.70}
$$

From (B.3.27), we see that the jumps of $\mathcal{Q}_{ij}(t)$ are at most of magnitude 1. Then, as $H^{(k)}\Gamma_{ij}$ is uniformly bounded by the integer k, we get

$$\triangle \hat{Y}^{(k)}(t) \leq kD^2. \tag{B.4.71}$$

Hence from (B.4.69) and (B.4.71) we get

$$\triangle Z^{(k)}(t) \leq C + kD^2, \tag{B.4.72}$$

Thus the jumps of $Z^{(k)}$ are uniformly bounded, which proves Claim B.4.13.

As the jumps of $Z^{(k)}$ are uniformly bounded, we can choose a real number $\lambda > 0$ such that

$$\lambda \triangle Z^{(k)}(t) > -1 \quad \text{a.s.}, \forall t \in [0, T]. \tag{B.4.73}$$

Next define $\mathcal{E}(\lambda Z^{(k)})$, the Doléans-Dade exponential of $\lambda Z^{(k)}$ (see Theorem C.15.1 and Remark C.15.2).

By Remark B.4.10, $Z^{(k)} \in \mathcal{M}_0^2((\Omega, \mathcal{F}, \mathbb{P}), \{\mathcal{F}_t\})$. Therefore, by Remark C.15.2, we have

$$\mathcal{E}(\lambda Z^{(k)}) \in \mathcal{M}_{\text{loc}}((\Omega, \mathcal{F}, \mathbb{P}), \{\mathcal{F}_t\}). \tag{B.4.74}$$

From (B.4.73) and Remark C.15.2, we also have that $\mathcal{E}(\lambda Z^{(k)})$ is strictly positive.

As $\{\big(\mathcal{E}(\lambda Z^{(k)})(t), \mathcal{F}_t\big) : t \in [0, T]\}$ is a local martingale, there exists a sequence $(R^m)_{m \in \mathbb{N}}$ of $\{\mathcal{F}_t\}$–stopping times with $R^m \Uparrow T$ a.s. such that

$$\{\big(\mathcal{E}(\lambda Z^{(k)})(t \wedge R^m), \mathcal{F}_t\big) : t \in [0, T]\} \tag{B.4.75}$$

is a uniformly integrable martingale for every $m \in \mathbb{N}$. In particular, we see that $\mathcal{E}(\lambda Z^{(k)})(\cdot \wedge R^m)$ satisfies the assumptions of Girsanov's theorem (Theorem C.15.5).

Fix $m \in \mathbb{N}$ and define a probability measure $\mathbb{Q}_m$ on the measurable space $(\Omega, \mathcal{F})$ by

$$\mathbb{Q}_m[A] := \mathrm{E}_{\mathbb{P}}\left(\mathcal{E}(\lambda Z^{(k)})(T \wedge R^m); A\right), \quad \forall A \in \mathcal{F}, \tag{B.4.76}$$

with $\mathrm{E}_{\mathbb{P}}$ being used to emphasize that the expectation is with respect to the probability measure $\mathbb{P}$. As $W_n \in \mathcal{M}_0^2((\Omega, \mathcal{F}, \mathbb{P}), \{\mathcal{F}_t\})$ and the process $\langle W_n, \lambda Z^{(k)} \rangle$ exists under $\mathbb{P}$, we can apply Girsanov's theorem (Theorem C.15.5) to W_n to get

$$W_n - \langle W_n, \lambda Z^{(k)} \rangle \in \mathcal{M}_{\text{loc}}((\Omega, \mathcal{F}, \mathbb{Q}_m), \{\mathcal{F}_t\}). \tag{B.4.77}$$

From Claim B.4.11, $\langle W_n, Z^{(k)} \rangle(t) = 0$, so trivially $\langle W_n, \lambda Z^{(k)} \rangle(t) = 0$ for all $t \in [0, T]$. Then from (B.4.77) we get

$$W_n \in \mathcal{M}_{0,\text{loc}}((\Omega, \mathcal{F}, \mathbb{Q}_m), \{\mathcal{F}_t\}). \tag{B.4.78}$$

As the square-bracket quadratic variation process is invariant under the change of measure (see Jacod and Shiryaev [25], Theorem III.3.13), from Lévy's Theorem (Theorem C.14.4) we have that $\mathbf{W}$ is a standard N-dimensional Brownian motion on the filtered probability space $((\Omega, \mathcal{F}, \mathbb{Q}_m), \{\mathcal{F}_t\})$.

Claim B.4.14. $((\Omega, \mathcal{F}, \mathbb{Q}_m), \{\mathcal{F}_t\}, \mathbf{W})$ is a solution to the martingale problem for $(A^\nabla, \mathbf{0})$.

We have already proved that $\mathbf{W}$ is an N-dimensional standard Brownian motion on the filtered probability space $((\Omega, \mathcal{F}, \mathbb{Q}_m), \{\mathcal{F}_t\})$. Then, by Theorem B.2.12, $((\Omega, \mathcal{F}, \mathbb{Q}_m), \{\mathcal{F}_t\}, \mathbf{W})$ is a solution to the martingale problem for $(A^\nabla, \mathbf{0})$ and Claim B.4.14 is shown.

By Theorem B.2.14, the martingale problem for $(A^\nabla, \mathbf{0})$ is well-posed and therefore has the property of uniqueness, as described in Definition B.2.4. From Corollary B.2.13, $((\Omega, \mathcal{F}, \mathbb{P}), \{\mathcal{F}_t\}, \mathbf{W})$ also solves the martingale problem for $(A^\nabla, \mathbf{0})$. As the two solutions are on the same measurable space and the Brownian motion $\mathbf{W}$ has the same finite-dimensional distributions under both measures, then the measures must agree on the filtration generated by the Brownian motion $\mathbf{W}$, that is

$$\mathbb{P}[B] = \mathbb{Q}_m[B], \quad \forall B \in \mathcal{F}_T^{\mathbf{W}}. \tag{B.4.79}$$

Now consider the canonical martingales of the Markov chain α. As $\mathcal{Q}_{ij} \in \mathcal{M}_0^2((\Omega, \mathcal{F}, \mathbb{P}), \{\mathcal{F}_t\})$ and the process $\langle \mathcal{Q}_{ij}, \lambda Z^{(k)} \rangle$ exists under $\mathbb{P}$, then applying Girsanov's theorem (Theorem C.15.5) to $\mathcal{Q}_{ij}$, we get

$$\mathcal{Q}_{ij} - \langle \mathcal{Q}_{ij}, \lambda Z^{(k)} \rangle \in \mathcal{M}_{\text{loc}}((\Omega, \mathcal{F}, \mathbb{Q}_m), \{\mathcal{F}_t\}). \tag{B.4.80}$$

From Claim B.4.12, $\langle \mathcal{Q}_{ij}, Z^{(k)} \rangle(t) = 0$, so trivially $\langle \mathcal{Q}_{ij}, \lambda Z^{(k)} \rangle(t) = 0$ for all $t \in [0, T]$. Then from (B.4.80) we get

$$\mathcal{Q}_{ij} \in \mathcal{M}_{0,\text{loc}}((\Omega, \mathcal{F}, \mathbb{Q}_m), \{\mathcal{F}_t\}). \tag{B.4.81}$$

Claim B.4.15. $\mathcal{Q}_{ij} \in \mathcal{M}_0^2((\Omega, \mathcal{F}, \mathbb{Q}_m), \{\mathcal{F}_t\})$.

Following the proof of Lemma B.3.13, but using the new measure $\mathbb{Q}_m$, we get for all $i, j = 1, \ldots, D$,

$$\mathrm{E}_{\mathbb{Q}_m} \left([\mathcal{Q}_{ij}](t) \right)^n < \infty, \qquad \forall t \in [0, T] \qquad \forall n \in \mathbb{N}, \tag{B.4.82}$$

since the proof relies only on $\mathcal{Q}_{ij}$ being a local martingale with respect to the measure $\mathbb{Q}_m$ (we are using the symbol $\mathrm{E}_{\mathbb{Q}_m}$ to denote expectation with respect to the measure $\mathbb{Q}_m$). Then following the proof of Lemma B.3.16 we get

$$\sup_{t \in [0,T]} \mathrm{E}_{\mathbb{Q}_m} |\mathcal{Q}_{ij}(t)|^2 \leq 2\mathrm{E}_{\mathbb{Q}_m} \left([\mathcal{Q}_{ij}](T) \right)^2 + 2\mathrm{E}_{\mathbb{Q}_m} \left(\langle \mathcal{Q}_{ij} \rangle (T) \right)^2 \overset{(B.3.13),(B.4.82)}{<} \infty. \tag{B.4.83}$$

Then by Corollary C.16.2, $\mathcal{Q}_{ij} \in \mathcal{M}_0^2((\Omega, \mathcal{F}, \mathbb{Q}_m), \{\mathcal{F}_t\})$ and Claim B.4.15 is shown.

Claim B.4.16. $((\Omega, \mathcal{F}, \mathbb{Q}_m), \{\mathcal{F}_t\}, \alpha)$ is a solution to the martingale problem for (A^Q, i_0).

From (B.4.76), the measure $\mathbb{Q}_m$ is absolutely continuous with respect to the measure $\mathbb{P}$. As by assumption, $\mathbb{P}[\alpha(0) = i_0] = 1$, it follows from the absolute continuity that

$$\mathbb{Q}_m[\alpha(0) = i_0] = 1. \tag{B.4.84}$$

We have also proved that $\mathcal{Q}_{ij} \in \mathcal{M}_0^2((\Omega, \mathcal{F}, \mathbb{Q}_m), \{\mathcal{F}_t\})$. Then by Lemma B.3.28, $((\Omega, \mathcal{F}, \mathbb{Q}_m), \{\mathcal{F}_t\}, \alpha)$ is a solution to the martingale problem for (A^Q, i_0) and Claim B.4.16 is shown.

By Theorem B.2.8, the martingale problem for (A^Q, i_0) is well-posed and therefore has the property of uniqueness, as described in Definition B.2.4. From Lemma B.2.6, $((\Omega, \mathcal{F}, \mathbb{P}), \{\mathcal{F}_t\}, \alpha)$ also solves the martingale problem for (A^Q, i_0). As the two solutions are on the same measurable space and the Markov chain α has the same finite-dimensional distributions under both measures, then the measures must agree on the filtration generated by the Markov chain α, that is

$$\mathbb{P}[A] = \mathbb{Q}_m[A], \quad \forall A \in \mathcal{F}_T^\alpha. \tag{B.4.85}$$

Claim B.4.17. $\mathbb{Q}_m = \mathbb{P}$ on $\mathcal{F}$.

From Claim B.4.14, $((\Omega, \mathcal{F}, \mathbb{Q}_m), \{\mathcal{F}_t\}, \mathbf{W})$ is a solution to the martingale problem for $(A^\nabla, \mathbf{0})$. From Claim B.4.16, $((\Omega, \mathcal{F}, \mathbb{Q}_m), \{\mathcal{F}_t\}, \alpha)$ is a solution to the martingale problem for (A^Q, i_0). Then applying Proposition B.2.15, $((\Omega, \mathcal{F}, \mathbb{Q}_m), \{\mathcal{F}_t\}, (\alpha, \mathbf{W}))$ is a solution to the martingale problem for $(A^{Q,\nabla}, (i_0, \mathbf{0}))$. Furthermore, by Proposition B.2.21, α and $\mathbf{W}$ are independent under the measure $\mathbb{Q}_m$. Recall also the assumption that α and $\mathbf{W}$ are independent under the measure $\mathbb{P}$. Then for all $A \in \mathcal{F}_T^\alpha$ and all $B \in \mathcal{F}_T^\mathbf{W}$,

$$\mathbb{P}[A \cap B] = \mathbb{P}[A]\mathbb{P}[B] \overset{(B.4.79),(B.4.85)}{=} \mathbb{Q}_m[A]\mathbb{Q}_m[B] = \mathbb{Q}_m[A \cap B]. \tag{B.4.86}$$

As the sets $\{A \cap B \,:\, A \in \mathcal{F}_T^\alpha,\ B \in \mathcal{F}_T^\mathbf{W}\}$ generate the σ-algebra $\mathcal{F}$ (recall (3.2.4) - (3.2.6)), then $\mathbb{Q}_m = \mathbb{P}$ on $\mathcal{F}$ and Claim B.4.17 is shown.

Claim B.4.18. For all $k \in \mathbb{N}$, $\tilde{Y}^{(k)}(t) = \hat{Y}^{(k)}(t)$ a.s. for all $t \in [0, T]$.

From the definition of the measure $\mathbb{Q}_m$ in (B.4.76) and the equivalence of measures shown in Claim B.4.17, we have

$$\mathrm{E}_\mathbb{P}\left(\mathcal{E}(\lambda Z^{(k)})(T \wedge R^m); A\right) = \mathbb{Q}_m[A] = \mathbb{P}[A] = \mathrm{E}_\mathbb{P}[1; A], \quad \forall A \in \mathcal{F}. \tag{B.4.87}$$

Hence

$$\mathcal{E}(\lambda Z^{(k)})(T \wedge R^m) = 1 \quad \text{a.s.} \tag{B.4.88}$$

As $\{\left(\mathcal{E}(\lambda Z^{(k)})(t \wedge R^m), \mathcal{F}_t\right) \,:\, t \in [0, T]\}$ is a uniformly integrable martingale, we have for all $t \in [0, T]$,

$$\mathcal{E}(\lambda Z^{(k)})(t \wedge R^m) = \mathrm{E}_\mathbb{P}\left(\mathcal{E}(\lambda Z^{(k)})(T \wedge R^m) \,|\, \mathcal{F}_t\right) \quad \mathbb{P}\text{-a.s.} \tag{B.4.89}$$

From (B.4.88), the right-hand side of the above equation equals 1. Thus we obtain $\mathcal{E}(\lambda Z^{(k)})(t \wedge R^m) = 1$ a.s for all $t \in [0, T]$. Since, from Theorem C.15.1, $\mathcal{E}(\lambda Z^{(k)})$ satisfies

$$\mathcal{E}(\lambda Z^{(k)})(t) = 1 + \int_0^t \mathcal{E}(\lambda Z^{(k)})(\tau_-)\ \mathrm{d}(\lambda Z^{(k)})(\tau), \tag{B.4.90}$$

then we must have

$$Z^{(k)}(t) = 0 \qquad \text{a.s.,} \quad \forall t \in [0, T \wedge R^m]. \tag{B.4.91}$$

From (B.4.60), $Z^{(k)}(t) = \tilde{Y}^{(k)}(t) - \hat{Y}^{(k)}(t)$ and hence

$$\tilde{Y}^{(k)}(t) = \hat{Y}^{(k)}(t) \quad \text{a.s.,} \qquad \forall t \in [0, T \wedge R^m]. \tag{B.4.92}$$

As the stopping times $R^m \Uparrow T$ a.s., we have

$$\tilde{Y}^{(k)}(t) = \hat{Y}^{(k)}(t) \quad \text{a.s.,} \qquad \forall t \in [0, T]. \tag{B.4.93}$$

As k was arbitrarily fixed, this proves Claim B.4.18.

Remark B.4.19. We have now shown the second item in Remark B.4.7. For the third item, we will rely heavily upon the properties of square-integrable martingales.

Claim B.4.20. For all $t \in [0, T]$,

$$\lim_{k\to\infty} \mathrm{E}\left|\hat{Y}^{(k)}(t) - \sum_{n=1}^{N} \int_0^t \Lambda_n(\tau)\ \mathrm{d}W_n(\tau) - \sum_{i,j=1}^{D} \int_0^t \Gamma_{ij}(\tau)\ \mathrm{d}\mathcal{Q}_{ij}(\tau)\right|^2 = 0. \tag{B.4.94}$$

We begin by showing that for the process $\bar{Y}$ defined as

$$\bar{Y}(t) := \sum_{n=1}^{N} \int_0^t \Lambda_n(\tau)\ \mathrm{d}W_n(\tau) + \sum_{i,j=1}^{D} \int_0^t \Gamma_{ij}\ \mathrm{d}\mathcal{Q}_{ij}(\tau), \quad \forall t \in [0, T], \tag{B.4.95}$$

we have $\bar{Y} \in \mathcal{M}_0^2((\Omega, \mathcal{F}, \mathbb{P}), \{\mathcal{F}_t\})$. To do this, we show that the integrands Λ_n and Γ_{ij} satisfy

$$\mathrm{E} \sum_{n=1}^{N} \int_0^T \Lambda_n^2(\tau)\ \mathrm{d}\tau + \mathrm{E} \sum_{i,j=1}^{D} \int_0^T \Gamma_{ij}^2(\tau)\, \mathrm{d}[\mathcal{Q}_{ij}](\tau) < \infty. \tag{B.4.96}$$

We then show that the limit in (B.4.94) is attained.

From Claim B.4.18, $\tilde{Y}^{(k)}(T) = \hat{Y}^{(k)}(T)$ a.s. for all $k \in \mathbb{N}$. Upon squaring each side and taking expectations, we can apply the Itô isometry to get

$$\begin{aligned} \mathrm{E}|\hat{Y}^{(k)}(T)|^2 = \mathrm{E}|\tilde{Y}^{(k)}(T)|^2 &\overset{(B.4.54)}{=} \mathrm{E}\left| \int_0^T H^{(k)}(\tau)\ \mathrm{d}Y(\tau)\right|^2 \\ &= \mathrm{E} \int_0^T |H^{(k)}(\tau)|^2\ \mathrm{d}[Y](\tau) \le \mathrm{E}[Y](T) < \infty. \end{aligned} \tag{B.4.97}$$

Letting $k \to \infty$,

$$\lim_{k\to\infty} \mathrm{E}|\hat{Y}^{(k)}(T)|^2 \le \mathrm{E}[Y](T) < \infty. \tag{B.4.98}$$

Consider a square-bracket quadratic variation process of $\hat{Y}^{(k)}$.

$$\begin{aligned}
&[\hat{Y}^{(k)}](T) \\
&\overset{(B.4.55)}{=} \left[\sum_{n=1}^{N} \int_0^{\cdot} H^{(k)}(\tau)\,\Lambda_n(\tau)\ \mathrm{d}W_n(\tau) + \sum_{i,j=1}^{D} \int_0^{\cdot} H^{(k)}(\tau)\,\Gamma_{ij}(\tau)\ \mathrm{d}\mathcal{Q}_{ij}(\tau) \right](T) \\
&= \left[\sum_{n=1}^{N} \int_0^{\cdot} H^{(k)}(\tau)\,\Lambda_n(\tau)\ \mathrm{d}W_n(\tau), \sum_{n=1}^{N} \int_0^{\cdot} H^{(k)}(\tau)\,\Lambda_n(\tau)\ \mathrm{d}W_n(\tau) \right](T) \\
&\quad + 2 \left[\sum_{n=1}^{N} \int_0^{\cdot} H^{(k)}(\tau)\,\Lambda_n(\tau)\ \mathrm{d}W_n(\tau), \sum_{i,j=1}^{D} \int_0^{\cdot} H^{(k)}(\tau)\,\Gamma_{ij}(\tau)\ \mathrm{d}\mathcal{Q}_{ij}(\tau) \right](T) \\
&\quad + \left[\sum_{i,j=1}^{D} \int_0^{\cdot} H^{(k)}(\tau)\,\Gamma_{ij}(\tau)\ \mathrm{d}\mathcal{Q}_{ij}(\tau), \sum_{i,j=1}^{D} \int_0^{\cdot} H^{(k)}(\tau)\,\Gamma_{ij}(\tau)\ \mathrm{d}\mathcal{Q}_{ij}(\tau) \right](T) \\
&= \sum_{n=1}^{N} \sum_{m=1}^{N} \int_0^{T} |H^{(k)}(\tau)|^2\,\Lambda_n(\tau)\Lambda_m(\tau)\ \mathrm{d}[W_n, W_m](\tau) \\
&\quad + 2 \sum_{n=1}^{N} \sum_{i,j=1}^{D} \int_0^{T} |H^{(k)}(\tau)|^2\,\Lambda_n(\tau)\Gamma_{ij}(\tau)\ \mathrm{d}[W_n, \mathcal{Q}_{ij}](\tau) \\
&\quad + \sum_{i,j=1}^{D} \sum_{a,b=1}^{D} \int_0^{T} |H^{(k)}(\tau)|^2\,\Gamma_{ij}(\tau)\Gamma_{ab}(\tau)\,\mathrm{d}[\mathcal{Q}_{ij}, \mathcal{Q}_{ab}](\tau).
\end{aligned} \tag{B.4.99}$$

Using Lemma B.3.18 and Lemma B.3.19, we get

$$[\hat{Y}^{(k)}](T) = \sum_{n=1}^{N} \int_0^{T} |H^{(k)}(\tau)|^2\,\Lambda_n^2(\tau)\ \mathrm{d}\tau + \sum_{i,j=1}^{D} \int_0^{T} |H^{(k)}(\tau)|^2\,\Gamma_{ij}^2(\tau)\,\mathrm{d}[\mathcal{Q}_{ij}](\tau). \tag{B.4.100}$$

As $\hat{Y}^{(k)} \in \mathcal{M}_0^2((\Omega, \mathcal{F}, \mathbb{P}), \{\mathcal{F}_t\})$, we have by the Itô isometry,

$$\begin{aligned}
&\mathrm{E}|\hat{Y}^{(k)}(T)|^2 \\
&\overset{(B.4.100)}{=} \sum_{n=1}^{N} \mathrm{E} \int_0^{T} |H^{(k)}(\tau)|^2\,\Lambda_n^2(\tau)\ \mathrm{d}\tau + \sum_{i,j=1}^{D} \mathrm{E} \int_0^{T} |H^{(k)}(\tau)|^2\,\Gamma_{ij}^2(\tau)\,\mathrm{d}[\mathcal{Q}_{ij}](\tau).
\end{aligned} \tag{B.4.101}$$

For the indicator function $H^{(k)}$ given by (B.4.53), we have that $H^{(k)} \to 1$ $(\mathbb{P} \otimes \mathit{Leb})$-a.e. as $k \to \infty$. So

$$|H^{(k)}|^2\,\Lambda_n^2 \to \Lambda_n^2 \quad (\mathbb{P} \otimes \mathit{Leb})\text{-a.e. as } k \to \infty. \tag{B.4.102}$$

for each $n = 1, \dots, N$. As $H^{(k)}$ is a non-decreasing function, we have

$$0 \leq |H^{(k)}(\omega, t)|^2 \, \Lambda_n^2(\omega, t) \leq |H^{(k+1)}(\omega, t)|^2 \, \Lambda_n^2(\omega, t), \tag{B.4.103}$$

for all $(\omega, t) \in \Omega \times [0, T]$. Thus we can apply the Monotone Convergence Theorem to conclude that for each $n = 1, \dots, N$,

$$\lim_{k\to\infty} \mathrm{E} \int_0^T |H^{(k)}(\tau)|^2 \, \Lambda_n^2(\tau) \, \mathrm{d}\tau = \mathrm{E} \int_0^T \Lambda_n^2(\tau) \, \mathrm{d}\tau. \tag{B.4.104}$$

We also have for each $i, j = 1, \dots, D$,

$$|H^{(k)}|^2 \, \Gamma_{ij}^2 \to \Gamma_{ij}^2 \quad (\mathbb{P} \otimes \mathit{Leb})\text{-a.e. as } k \to \infty. \tag{B.4.105}$$

For each $i \neq j$, setting

$$N := \{(\omega, t) \in \Omega \times [0, T] \, : \, |H^{(k)}(\omega, t)|^2 \, \Gamma_{ij}^2(\omega, t) \nrightarrow \Gamma_{ij}^2(\omega, t)\}, \tag{B.4.106}$$

then by (B.4.105), we have $(\mathbb{P} \otimes \mathit{Leb})(N) = 0$. By Lemma B.3.26 and for $i \neq j$, the measure $\nu_{[\mathcal{Q}_{ij}]}$ is absolutely continuous with respect to the measure $(\mathbb{P} \otimes \mathit{Leb})$ on $\mathcal{P}^\star$. Then $\nu_{[\mathcal{Q}_{ij}]}(N) = 0$ and hence

$$|H^{(k)}|^2 \, \Gamma_{ij}^2 \to \Gamma_{ij}^2 \quad \nu_{[\mathcal{Q}_{ij}]}\text{-a.e. as } k \to \infty. \tag{B.4.107}$$

Also, for each $i, j = 1, \dots, D$,

$$0 \leq |H^{(k)}(\omega, t)|^2 \, \Gamma_{ij}^2(\omega, t) \leq |H^{(k+1)}(\omega, t)|^2 \, \Gamma_{ij}^2(\omega, t), \tag{B.4.108}$$

for all $(\omega, t) \in \Omega \times [0, T]$. Then we can apply the Monotone Convergence Theorem to conclude

$$\lim_{k\to\infty} \mathrm{E} \int_0^T |H^{(k)}(\tau)|^2 \, \Gamma_{ij}^2(\tau) \, \mathrm{d}[\mathcal{Q}_{ij}](\tau) = \mathrm{E} \int_0^T \Gamma_{ij}^2(\tau) \, \mathrm{d}[\mathcal{Q}_{ij}](\tau), \tag{B.4.109}$$

for each $i, j = 1, \dots, D$. Then

$$\begin{aligned}
&\lim_{k\to\infty} \mathrm{E}|\hat{Y}^{(k)}(T)|^2 \\
&\overset{(B.4.101)}{=} \lim_{k\to\infty} \Bigg\{ \sum_{n=1}^{N} \mathrm{E} \int_0^T |H^{(k)}(\tau)|^2 \Lambda_n^2(\tau) \, \mathrm{d}\tau \\
&\qquad\qquad + \sum_{i,j=1}^{D} \mathrm{E} \int_0^T |H^{(k)}(\tau)|^2 \Gamma_{ij}^2(\tau) \, \mathrm{d}[\mathcal{Q}_{ij}](\tau) \Bigg\} \\
&\overset{(B.4.104),(B.4.109)}{=} \sum_{n=1}^{N} \mathrm{E} \int_0^T \Lambda_n^2(\tau) \, \mathrm{d}\tau + \sum_{i,j=1}^{D} \mathrm{E} \int_0^T \Gamma_{ij}^2(\tau) \, \mathrm{d}[\mathcal{Q}_{ij}](\tau).
\end{aligned} \tag{B.4.110}$$

From (B.4.98), $\lim_{k\to\infty} \mathrm{E}|\hat{Y}^{(k)}(T)|^2 < \infty$. Applying this bound to (B.4.110), we get

$$\mathrm{E} \sum_{n=1}^{N} \int_0^T \Lambda_n^2(\tau) \, \mathrm{d}\tau + \mathrm{E} \sum_{i,j=1}^{D} \int_0^T \Gamma_{ij}^2(\tau) \, \mathrm{d}[\mathcal{Q}_{ij}](\tau) < \infty. \tag{B.4.111}$$

Together with the previsibility of Λ_n and Γ_{ij}, which comes from Lemma B.4.1, this shows that

$$\mathbf{\Lambda} = (\Lambda_1, \ldots, \Lambda_N)^\top \in L^2(\mathbf{W}) \text{ and } \mathbf{\Gamma} \equiv (\Gamma_{ij})_{i,j=1}^D \in L^2(\mathcal{Q}). \tag{B.4.112}$$

Hence we can assert the existence of $\bar{Y}$, as defined by (B.4.95), in the space of martingales $\mathcal{M}_0^2((\Omega, \mathcal{F}, \mathbb{P}), \{\mathcal{F}_t\})$.

Next we show that (B.4.94) holds. Using (B.4.95), we can rewrite (B.4.94) in terms of $\bar{Y}$ as

$$\lim_{k\to\infty} \mathrm{E}|\hat{Y}^{(k)}(t) - \bar{Y}(t)|^2 = 0, \quad \forall t \in [0, T]. \tag{B.4.113}$$

As $\hat{Y}^{(k)}, \bar{Y} \in \mathcal{M}_0^2((\Omega, \mathcal{F}, \mathbb{P}), \{\mathcal{F}_t\})$, we can use the Itô isometry and a calculation similar to the one in (B.4.99) to find

$$\begin{aligned}
\mathrm{E}|\hat{Y}^{(k)}(T) - \bar{Y}(T)|^2 &= \mathrm{E}\left[\hat{Y}^{(k)}(\cdot) - \bar{Y}(\cdot)\right](T) \\
&\overset{(B.4.55),(B.4.95)}{=} \mathrm{E}\Bigg[\sum_{n=1}^N \int_0^\cdot \left(H^{(k)}(\tau) - 1\right) \Lambda_n(\tau)\ \mathrm{d}W_n(\tau) \\
&\qquad + \sum_{i,j=1}^D \int_0^\cdot \left(H^{(k)}(\tau) - 1\right) \Gamma_{ij}(\tau)\ \mathrm{d}\mathcal{Q}_{ij}(\tau)\Bigg](T) \\
&= \sum_{n=1}^N \mathrm{E}\int_0^T |H^{(k)}(\tau) - 1|^2\, \Lambda_n^2(\tau)\ \mathrm{d}\tau + \sum_{i,j=1}^D \mathrm{E}\int_0^T |H^{(k)}(\tau) - 1|^2\, \Gamma_{ij}^2(\tau)\, \mathrm{d}[\mathcal{Q}_{ij}](\tau).
\end{aligned} \tag{B.4.114}$$

The sequence of functions $\{|H^{(k)} - 1|^2\, \Lambda_n^2\}_{k\in\mathbb{N}}$ converges $(\mathbb{P} \otimes \mathit{Leb})$-a.e. to 0 and is dominated by the random variable Λ_n^2, which is integrable by (B.4.111). Thus we can apply the Lebesgue Dominated Convergence Theorem to the sequence of functions to get

$$\lim_{k\to\infty} \mathrm{E}\int_0^T |H^{(k)}(\tau) - 1|^2\, \Lambda_n^2(\tau)\ \mathrm{d}\tau = 0, \tag{B.4.115}$$

for each $n = 1, \ldots, N$.

The sequence of functions $\{|H^{(k)} - 1|^2\, \Gamma_{ij}^2\}_{k\in\mathbb{N}}$ also converges $(\mathbb{P} \otimes \mathit{Leb})$-a.e. to 0, for each $i, j = 1, \ldots, D$. For $i \neq j$, setting

$$N := \{(\omega, t) \in \Omega \times [0, T] \,:\, |H^{(k)}(\omega, t) - 1|^2\, \Gamma_{ij}^2(\omega, t) \nrightarrow \Gamma_{ij}^2(\omega, t)\}, \tag{B.4.116}$$

we then have $(\mathbb{P} \otimes \mathit{Leb})(N) = 0$. By Lemma B.3.26 and for $i \neq j$, the measure $\nu_{[\mathcal{Q}_{ij}]}$ is absolutely continuous with respect to the measure $(\mathbb{P} \otimes \mathit{Leb})$ on $\mathcal{P}^\star$. So $\nu_{[\mathcal{Q}_{ij}]}(N) = 0$ and hence

$$|H^{(k)} - 1|^2\, \Gamma_{ij}^2 \to 0 \quad \nu_{[\mathcal{Q}_{ij}]}\text{-a.e. as } k \to \infty, \tag{B.4.117}$$

for each $i, j = 1, \ldots, D$, $i \neq j$. As $|H^{(k)} - 1|^2 \, \Gamma_{ij}^2$ is dominated by the random variable Γ_{ij}^2, which is integrable by (B.4.111), we can apply the Lebesgue Dominated Convergence Theorem to the sequence of functions to get

$$\lim_{k\to\infty} \mathrm{E} \int_0^T |H^{(k)}(\tau) - 1|^2 \, \Gamma_{ij}^2(\tau) \ \mathrm{d}[\mathcal{Q}_{ij}](\tau) = 0, \tag{B.4.118}$$

for each $i, j = 1, \ldots, D$.

Letting $k \to \infty$ in (B.4.114), we obtain

$$\lim_{k\to\infty} \mathrm{E}|\hat{Y}^{(k)}(t) - \bar{Y}(t)|^2 \overset{(B.4.115),(B.4.118)}{=} 0. \tag{B.4.119}$$

This proves Claim B.4.20.

Now we have shown that all three items in Remark B.4.7 hold, we will use these items to prove the martingale representation theorem for a bounded $Y \in \mathcal{M}_0^2((\Omega, \mathcal{F}, \mathbb{P}), \{\mathcal{F}_t\})$.

Using Claim B.4.18 to substitute $\hat{Y}^{(k)}$ for $\tilde{Y}^{(k)}$ in Claim B.4.8, we obtain

$$\lim_{k\to\infty} \mathrm{E}|\hat{Y}^{(k)}(t) - Y(t)|^2 = 0, \quad \forall t \in [0, T]. \tag{B.4.120}$$

Comparing this to Claim B.4.20 and using the uniqueness of limits, we must have

$$Y(t) = \sum_{n=1}^{N} \int_0^t \Lambda_n(\tau) \ \mathrm{d}W_n(\tau) + \sum_{i,j=1}^{D} \int_0^t \Gamma_{ij}(\tau) \ \mathrm{d}[\mathcal{Q}_{ij}](\tau) \quad \text{a.s.}, \qquad \forall t \in [0, T]. \tag{B.4.121}$$

Using (B.4.121) and a calculation similar to the one in (B.4.99), we can calculate the square-bracket quadratic variation process of the bounded martingale Y as

$$[Y](t) = \sum_{n=1}^{N} \int_0^t \Lambda_n^2(\tau) \ \mathrm{d}\tau + \sum_{i,j=1}^{D} \int_0^t \Gamma_{ij}^2(\tau) \, \mathrm{d}[\mathcal{Q}_{ij}](\tau) \quad \text{a.s.} \tag{B.4.122}$$

Upon taking expectations and setting $t = T$,

$$\mathrm{E}[Y](T) = \sum_{n=1}^{N} \mathrm{E} \int_0^T \Lambda_n^2(\tau) \ \mathrm{d}\tau + \sum_{i,j=1}^{D} \mathrm{E} \int_0^T \Gamma_{ij}^2(\tau) \, \mathrm{d}[\mathcal{Q}_{ij}](\tau). \tag{B.4.123}$$

From (B.4.112), we have $\mathbf{\Lambda} \in L^2(\mathbf{W})$ and $\mathbf{\Gamma} \in L^2(\mathcal{Q})$.

Thus we have shown that (B.4.45), (B.4.46) and (B.4.47) hold when Y is a bounded member of $\mathcal{M}_0^2((\Omega, \mathcal{F}, \mathbb{P}), \{\mathcal{F}_t\})$. We show next that these equations hold when Y is any member of $\mathcal{M}_0^2((\Omega, \mathcal{F}, \mathbb{P}), \{\mathcal{F}_t\})$.

Suppose $Y \in \mathcal{M}_0^2((\Omega, \mathcal{F}, \mathbb{P}), \{\mathcal{F}_t\})$. Then $Y(T) \in L^2(\Omega, \mathcal{F}_T, \mathbb{P})$ and there exists a sequence of bounded, $\{\mathcal{F}_T\}$–measurable random variables $\{X_T^{(k)}\}_{k\in\mathbb{N}}$ in the space $L^2(\Omega, \mathcal{F}_T, \mathbb{P})$ which converge in the L^2-norm to $Y(T)$, that is

$$\lim_{k\to\infty} \mathrm{E}|X_T^{(k)} - Y(T)|^2 = 0. \tag{B.4.124}$$

Define for every $k \in \mathbb{N}$,

$$X^{(k)}(t) := \mathrm{E}[X_T^{(k)} \,|\, \mathcal{F}_t], \qquad \forall t \in [0,T]. \tag{B.4.125}$$

For each $k \in \mathbb{N}$, $X^{(k)} \in \mathcal{M}_0^2((\Omega,\mathcal{F},\mathbb{P}),\{\mathcal{F}_t\})$ is bounded, since $X_T^{(k)}$ is bounded. Thus we can apply Theorem B.4.6, which we have proved holds for bounded martingales, to $X^{(k)}$ to obtain

$$\mathbf{\Lambda}^{(k)} = \left(\Lambda_1^{(k)},\ldots,\Lambda_N^{(k)}\right) \in L^2(\mathbf{W}) \text{ and } \mathbf{\Gamma}^{(k)} = \left(\Gamma_{ij}^{(k)}\right)_{i,j=1}^{D} \in L^2(\mathcal{Q}) \tag{B.4.126}$$

such that for all $t \in [0,T]$,

$$X^{(k)}(t) = \sum_{n=1}^{N}\int_0^t \Lambda_n^{(k)}(\tau)\ \mathrm{d}W_n(\tau) + \sum_{i,j=1}^{D}\int_0^t \Gamma_{ij}^{(k)}(\tau)\ \mathrm{d}\mathcal{Q}_{ij}(\tau) \quad \text{a.s.} \tag{B.4.127}$$

Claim B.4.21. $\{\mathbf{\Lambda}^{(k)}\}_{k\in\mathbb{N}}$ is a Cauchy sequence in $L^2(\mathbf{W})$ and $\{\mathbf{\Gamma}^{(k)}\}_{k\in\mathbb{N}}$ is a Cauchy sequence in $L^2(\mathcal{Q})$.

Consider $X^{(k_1)} - X^{(k_2)}$, for some $k_1,k_2 \in \mathbb{N}$, which is a bounded martingale as it is the difference of two bounded martingales. Taking the expectation of the square-bracket quadratic variation process of $X^{(k_1)} - X^{(k_2)}$, we can do a calculation similar to the one in (B.4.99) to find

$$\begin{aligned}
&\mathrm{E}\left[X^{(k_1)} - X^{(k_2)}\right](T)\\
&\overset{(B.4.127)}{=} \mathrm{E}\Bigg[\sum_{n=1}^{N}\int_0^{\cdot}\left(\Lambda_n^{(k_1)} - \Lambda_n^{(k_2)}\right)(t)\ \mathrm{d}W_n(t)\\
&\qquad\qquad + \sum_{i,j=1}^{D}\int_0^{\cdot}\left(\Gamma_{ij}^{(k_1)} - \Gamma_{ij}^{(k_2)}\right)(t)\ \mathrm{d}\mathcal{Q}_{ij}(t)\Bigg](T)\\
&= \sum_{n=1}^{N}\mathrm{E}\int_0^T\left(\Lambda_n^{(k_1)} - \Lambda_n^{(k_2)}\right)^2(t)\ \mathrm{d}t + \sum_{i,j=1}^{D}\mathrm{E}\int_0^T\left(\Gamma_{ij}^{(k_1)} - \Gamma_{ij}^{(k_2)}\right)^2(t)\ \mathrm{d}[\mathcal{Q}_{ij}](t).
\end{aligned} \tag{B.4.128}$$

From (B.4.124), $\{X_T^{(k)}\}_{k\in\mathbb{N}}$ is a Cauchy sequence in $(\Omega,\mathcal{F}_T,\mathbb{P})$. Then as $X^{(k)} \in \mathcal{M}_0^2((\Omega,\mathcal{F},\mathbb{P}),\{\mathcal{F}_t\})$, we can use the Itô isometry to obtain

$$\begin{aligned}
\mathrm{E}\left[X^{(k_1)} - X^{(k_2)}\right](T) = \mathrm{E}|\,X^{(k_1)}(T) - X^{(k_2)}(T)|^2 &\overset{(B.4.125)}{=} \mathrm{E}|\,X_T^{(k_1)} - X_T^{(k_2)}|^2\\
&\to 0 \quad \text{as } k_1,k_2 \to \infty.
\end{aligned} \tag{B.4.129}$$

Thus from (B.4.129) and the nonnegativity of all the terms in (B.4.128), we have for each $n = 1,\ldots,N$,

$$\mathrm{E}\int_0^T\left(\Lambda_n^{(k_1)} - \Lambda_n^{(k_2)}\right)^2(t)\ \mathrm{d}t \to 0 \quad \text{as } k_1,k_2 \to \infty \tag{B.4.130}$$

and for each $i, j = 1, \ldots, D$,

$$\mathrm{E} \int_0^T \left(\Gamma_{ij}^{(k_1)} - \Gamma_{ij}^{(k_2)} \right)^2 (t) \; \mathrm{d}[\mathcal{Q}_{ij}](t) \to 0 \quad \text{as } k_1, k_2 \to \infty. \tag{B.4.131}$$

So $\{\mathbf{\Lambda}^{(k)}\}_{k \in \mathbb{N}}$ is a Cauchy sequence in $L^2(\mathbf{W})$ and $\{\mathbf{\Gamma}^{(k)}\}_{k \in \mathbb{N}}$ is a Cauchy sequence in $L^2(\mathcal{Q})$, and Claim B.4.21 is shown.

Since $L^2(\mathbf{W})$ and $L^2(\mathcal{Q})$ are complete spaces, the limit of any Cauchy sequence in each space exists and is in the space. Denote the limit of the Cauchy sequence $\{\mathbf{\Lambda}^{(k)}\}_{k \in \mathbb{N}}$ by $\mathbf{\Lambda} = (\Lambda_1, \ldots, \Lambda_N)$ and denote the limit of the Cauchy sequence $\{\mathbf{\Gamma}^{(k)}\}_{k \in \mathbb{N}}$ by $\mathbf{\Gamma} = (\Gamma_{ij})_{i,j=1}^D$. As $\mathbf{\Lambda} \in L^2(\mathbf{W})$ and $\mathbf{\Gamma} \in L^2(\mathcal{Q})$, upon defining

$$\bar{Y}(t) := \sum_{n=1}^N \int_0^t \Lambda_n(\tau) \; \mathrm{d}W_n(\tau) + \sum_{i,j=1}^D \int_0^t \Gamma_{ij}(\tau) \; \mathrm{d}\mathcal{Q}_{ij}(\tau), \tag{B.4.132}$$

for all $t \in [0, T]$, we have by a calculation similar to the one in (B.4.99) that

$$\mathrm{E}[\bar{Y}](T) = \sum_{n=1}^N \mathrm{E} \int_0^T \Lambda_n^2(t) \; \mathrm{d}t + \sum_{i,j=1}^D \mathrm{E} \int_0^T \Gamma_{ij}^2(t) \; \mathrm{d}[\mathcal{Q}_{ij}](t) < \infty, \tag{B.4.133}$$

and so $\bar{Y} \in \mathcal{M}_0^2((\Omega, \mathcal{F}, \mathbb{P}), \{\mathcal{F}_t\})$.

As $X^{(k)} \in \mathcal{M}_0^2((\Omega, \mathcal{F}, \mathbb{P}), \{\mathcal{F}_t\})$ then squaring $X^{(k)}(T) - \bar{Y}(T)$ and taking expectations, we can use the Itô isometry and a calculation similar to the one in (B.4.99) to get

$$\begin{aligned} \mathrm{E}|X^{(k)}(T) - \bar{Y}(T)|^2 &= \mathrm{E}[X^{(k)}(\cdot) - \bar{Y}(\cdot)](T) \\ &\overset{(B.4.127),(B.4.132)}{=} \mathrm{E}\Bigg[\sum_{n=1}^N \int_0^{\cdot} \left(\Lambda_n^{(k)} - \Lambda_n \right)(t) \; \mathrm{d}W_n(t) \\ &\qquad + \sum_{i,j=1}^D \int_0^{\cdot} \left(\Gamma_{ij}^{(k)} - \Gamma_{ij} \right)(t) \; \mathrm{d}\mathcal{Q}_{ij}(t) \Bigg](T) \\ &= \sum_{n=1}^N \mathrm{E} \int_0^T \left(\Lambda_n^{(k)} - \Lambda_n \right)^2 (t) \; \mathrm{d}t + \sum_{i,j=1}^D \mathrm{E} \int_0^T \left(\Gamma_{ij}^{(k)} - \Gamma_{ij} \right)^2 (t) \; \mathrm{d}[\mathcal{Q}_{ij}](t). \end{aligned} \tag{B.4.134}$$

As $\mathbf{\Lambda}$ is the limit of $\{\mathbf{\Lambda}^k\}$ in $L^2(\mathbf{W})$ and $\mathbf{\Gamma}$ is the limit of $\{\mathbf{\Gamma}^k\}$ in $L^2(\mathcal{Q})$, then the last line in the above equation converges to zero, that is

$$\lim_{k \to \infty} \mathrm{E}|X^{(k)}(T) - \bar{Y}(T)|^2 = 0. \tag{B.4.135}$$

We also have,

$$\lim_{k \to \infty} \mathrm{E}|X^{(k)}(T) - Y(T)|^2 \overset{(B.4.125)}{=} \lim_{k \to \infty} \mathrm{E}|X_T^{(k)} - Y(T)|^2 \overset{(B.4.124)}{=} 0. \tag{B.4.136}$$

Comparing (B.4.135) and (B.4.136), by the uniqueness of limits we must have that

$$Y(T) = \bar{Y}(T) \overset{(B.4.132)}{=} \sum_{n=1}^{N} \int_0^T \Lambda_n(t) \, \mathrm{d}W_n(t) + \sum_{i,j=1}^{D} \int_0^T \Gamma_{ij}(t) \, \mathrm{d}\mathcal{Q}_{ij}(t) \quad \text{a.s.} \tag{B.4.137}$$

Hence by the martingale property of $Y \in \mathcal{M}_0^2((\Omega, \mathcal{F}, \mathbb{P}), \{\mathcal{F}_t\})$,

$$Y(t) = \sum_{n=1}^{N} \int_0^t \Lambda_n(\tau) \, \mathrm{d}W_n(\tau) + \sum_{i,j=1}^{D} \int_0^t \Gamma_{ij}(\tau) \, \mathrm{d}\mathcal{Q}_{ij}(\tau) \quad \text{a.s.,} \quad \forall t \in [0, T], \tag{B.4.138}$$

for $\mathbf{\Lambda} \in L^2(\mathbf{W})$ and $\mathbf{\Gamma} \in L^2(\mathcal{Q})$. Thus we have shown that (B.4.45) and (B.4.46) hold when $Y \in \mathcal{M}_0^2((\Omega, \mathcal{F}, \mathbb{P}), \{\mathcal{F}_t\})$. Performing a similar calculation to the one in (B.4.99), it is clear that (B.4.47) also holds.

To show the uniqueness of the integrands, let $\tilde{\mathbf{\Lambda}} = \left(\tilde{\Lambda}_1, \ldots, \tilde{\Lambda}_N\right)^\top \in L^2(\mathbf{W})$ and $\tilde{\mathbf{\Gamma}} = \left(\tilde{\Gamma}_{ij}\right)_{i,j=1}^{D} \in L^2(\mathcal{Q})$ be such that

$$Y(T) = \sum_{n=1}^{N} \int_0^T \tilde{\Lambda}_n(t) \, \mathrm{d}W_n(t) + \sum_{i,j=1}^{D} \int_0^T \tilde{\Gamma}_{ij}(t) \, \mathrm{d}\mathcal{Q}_{ij}(t) \quad \text{a.s.} \tag{B.4.139}$$

Subtracting (B.4.139) from (B.4.137), then squaring and taking expectations, we have from the Itô isometry and by a similar calculation to the one in (B.4.99),

$$\begin{aligned} 0 &= \mathrm{E}\left| \sum_{n=1}^{N} \int_0^T \left(\Lambda_n - \tilde{\Lambda}_n\right)(t) \, \mathrm{d}W_n(t) + \sum_{i,j=1}^{D} \int_0^T \left(\Gamma_{ij} - \tilde{\Gamma}_{ij}\right)(t) \, \mathrm{d}\mathcal{Q}_{ij}(t) \right|^2 \\ &= \mathrm{E}\left[\sum_{n=1}^{N} \int_0^{\cdot} \left(\Lambda_n - \tilde{\Lambda}_n\right)(t) \, \mathrm{d}W_n(t) + \sum_{i,j=1}^{D} \int_0^{\cdot} \left(\Gamma_{ij} - \tilde{\Gamma}_{ij}\right)(t) \, \mathrm{d}\mathcal{Q}_{ij}(t) \right](T) \\ &= \sum_{n=1}^{N} \mathrm{E} \int_0^T \left(\Lambda_n - \tilde{\Lambda}_n\right)^2(t) \, \mathrm{d}t + \sum_{i,j=1}^{D} \mathrm{E} \int_0^T \left(\Gamma_{ij} - \tilde{\Gamma}_{ij}\right)^2(t) \, \mathrm{d}[\mathcal{Q}_{ij}](t). \end{aligned} \tag{B.4.140}$$

Since all terms on the right-hand side of the above equation are positive, we have for each $n = 1, \ldots, N$,

$$\mathrm{E} \int_0^T \left(\Lambda_n - \tilde{\Lambda}_n\right)^2(t) \, \mathrm{d}t = 0 \qquad \Leftrightarrow \qquad \Lambda_n = \tilde{\Lambda}_n \quad (\mathbb{P} \otimes \mathit{Leb})\text{-a.e.} \tag{B.4.141}$$

and for each $i, j = 1, \ldots, D$, $i \neq j$,

$$\mathrm{E} \int_0^T \left(\Gamma_{ij} - \tilde{\Gamma}_{ij}\right)^2(t) \, \mathrm{d}[\mathcal{Q}_{ij}](t) = 0 \qquad \Leftrightarrow \qquad \Gamma_{ij} = \tilde{\Gamma}_{ij} \quad \nu_{[\mathcal{Q}_{ij}]}\text{-a.e.} \tag{B.4.142}$$

Thus we obtain the required uniqueness for the integrands. □

Now we extend this result to locally square-integrable martingales. With the convention that for any process $\{H(t) : t \geq 0\}$ and any $\{\mathcal{F}_t\}$-stopping time S, we have

$$H[0,S](\omega,t) := \begin{cases} H(\omega,t) & \text{if } t \in [0,S(\omega)] \\ 0 & \text{if } t \in (S(\omega),\infty], \end{cases} \tag{B.4.143}$$

we define the following spaces of integrands,

$$L^2_{\text{loc}}(\mathbf{W}) := \Big\{\boldsymbol{\lambda} : \Omega \times [0,T] \to \mathbb{R}^N \Big| \text{there exists a sequence of } \{\mathcal{F}_t\}\text{-stopping times } \{S^k\}_{k\in\mathbb{N}} \text{ such that } S^k \Uparrow T \text{ a.s. and } \boldsymbol{\lambda}[0,S^k] \in L^2(\mathbf{W}) \text{ for all } k \in \mathbb{N}\Big\} \tag{B.4.144}$$

and

$$L^2_{\text{loc}}(\mathcal{Q}) := \Big\{\boldsymbol{\gamma} = \{\gamma_{ij}\}_{i,j=1}^D : \Omega \times [0,T] \to \mathbb{R}^{D\times D} \Big| \text{there exists a sequence of } \{\mathcal{F}_t\}\text{-stopping times } \{S^k\}_{k\in\mathbb{N}} \text{ such that } S^k \Uparrow T \text{ a.s. and } \boldsymbol{\gamma}[0,S^k] \in L^2(\mathcal{Q}) \text{ for all } k \in \mathbb{N}\Big\}. \tag{B.4.145}$$

Theorem B.4.22. *Suppose $Y \in \mathcal{M}^2_{0,loc}((\Omega,\mathcal{F},\mathbb{P}),\{\mathcal{F}_t\})$. Then there exists*

$$\boldsymbol{\Lambda} = (\Lambda_1,\ldots,\Lambda_N)^\top \in L^2_{loc}(\mathbf{W}) \quad \textit{and} \quad \boldsymbol{\Gamma} = (\Gamma_{ij})_{i,j=1}^D \in L^2_{loc}(\mathcal{Q}) \tag{B.4.146}$$

such that Y has the representation for all $t \in [0,T]$

$$Y(t) = \sum_{n=1}^N \int_0^t \Lambda_n(\tau)\, \mathrm{d}W_n(\tau) + \sum_{i,j=1}^D \int_0^t \Gamma_{ij}(\tau)\, \mathrm{d}\mathcal{Q}_{ij}(\tau) \quad a.s. \tag{B.4.147}$$

Moreover, $\boldsymbol{\Lambda}$ and $\boldsymbol{\Gamma}$ are unique in the sense that if $\tilde{\boldsymbol{\Lambda}} = \left(\tilde{\Lambda}_1,\ldots,\tilde{\Lambda}_N\right)^\top \in L^2_{loc}(\mathbf{W})$ and $\tilde{\boldsymbol{\Gamma}} = \left(\tilde{\Gamma}_{ij}\right)_{i,j=1}^D \in L^2_{loc}(\mathcal{Q})$ are such that (B.4.147) holds with Λ_n replaced by $\tilde{\Lambda}_n$ and Γ_{ij} replaced by $\tilde{\Gamma}_{ij}$, then $\boldsymbol{\Lambda} = \tilde{\boldsymbol{\Lambda}}$ ($\mathbb{P} \otimes Leb$)-a.e. and $\boldsymbol{\Gamma} = \tilde{\boldsymbol{\Gamma}}$ $\nu_{[\mathcal{Q}]}$-a.e.

Proof. As $Y \in \mathcal{M}^2_{0,\text{loc}}((\Omega,\mathcal{F},\mathbb{P}),\{\mathcal{F}_t\})$, there exists a localizing sequence $\{T^m\}_{m\in\mathbb{N}}$ of $\{\mathcal{F}_t\}$–stopping times such that $Y(\cdot \wedge T^m) \in \mathcal{M}^2_0((\Omega,\mathcal{F},\mathbb{P}),\{\mathcal{F}_t\})$ for each $m \in \mathbb{N}$. We can apply Theorem B.4.6 to $Y(\cdot \wedge T^m)$ for each $m \in \mathbb{N}$ to obtain processes

$$\boldsymbol{\Lambda}^{(m)} = \left(\Lambda_1^{(m)},\ldots,\Lambda_N^{(m)}\right)^\top \in L^2(\mathbf{W}) \quad \text{and} \quad \boldsymbol{\Gamma}^{(m)} = \left(\Gamma_{ij}^{(m)}\right)_{i,j=1}^D \in L^2(\mathcal{Q}) \tag{B.4.148}$$

such that for all $t \in [0,T]$,

$$Y(t \wedge T^m) = \sum_{n=1}^{N} \int_0^t \Lambda_n^{(m)}(\tau)\ \mathrm{d}W_n(\tau) + \sum_{i,j=1}^{D} \int_0^t \Gamma_{ij}^{(m)}(\tau)\ \mathrm{d}\mathcal{Q}_{ij}(\tau) \quad \text{a.s.} \tag{B.4.149}$$

Recall the notation introduced by (B.4.143). Then for $\boldsymbol{\lambda} = (\lambda_1, \ldots, \lambda_N)^\top \in L^2(\mathbf{W})$, we have by Rogers and Williams [46], Theorem IV.27.6i,

$$\sum_{n=1}^{N} \int_0^{t\wedge S} \lambda_n(\tau)\ \mathrm{d}W_n(\tau) = \sum_{n=1}^{N} \int_0^t \lambda_n[0,S](\tau)\ \mathrm{d}W_n(\tau) \quad \text{a.s.} \tag{B.4.150}$$

and for $\boldsymbol{\gamma} = \{\gamma_{ij}\}_{i,j=1}^{D} \in L^2(\mathcal{Q})$, we have

$$\sum_{i,j=1}^{D} \int_0^{t\wedge S} \gamma_{ij}(\tau)\ \mathrm{d}\mathcal{Q}_{ij}(\tau) = \sum_{i,j=1}^{D} \int_0^t \gamma_{ij}[0,S](\tau)\ \mathrm{d}\mathcal{Q}_{ij}(\tau) \quad \text{a.s.} \tag{B.4.151}$$

We check the consistency of the integrands $\boldsymbol{\Lambda}^m$ and $\boldsymbol{\Gamma}^m$, in the sense of showing that $\Lambda_n^{(m)}$ is $(\mathbb{P} \otimes Leb)$-a.e equal to $\Lambda_n^{(m+1)}[0,T^m]$ and $\Gamma_{ij}^{(m)}$ is $\nu_{[\mathcal{Q}_{ij}]}$-a.e equal to $\Gamma_{ij}^{(m+1)}[0,T^m]$.

For all $t \in [0,T]$ we have a.s.

$$\begin{aligned} Y(t \wedge T^m) &= Y((t \wedge T^m) \wedge T^m) \\ &\overset{(B.4.149)}{=} \sum_{n=1}^{N} \int_0^{t\wedge T^m} \Lambda_n^{(m)}(\tau)\ \mathrm{d}W_n(\tau) + \sum_{i,j=1}^{D} \int_0^{t\wedge T^m} \Gamma_{ij}^{(m)}(\tau)\ \mathrm{d}\mathcal{Q}_{ij}(\tau) \\ &\overset{(B.4.150),(B.4.151)}{=} \sum_{n=1}^{N} \int_0^t \Lambda_n^{(m)}[0,T^m](\tau)\ \mathrm{d}W_n(\tau) + \sum_{i,j=1}^{D} \int_0^t \Gamma_{ij}^{(m)}[0,T^m](\tau)\ \mathrm{d}\mathcal{Q}_{ij}(\tau). \end{aligned} \tag{B.4.152}$$

For all $t \in [0,T]$ we also have a.s.

$$\begin{aligned} Y(t \wedge T^m) &= Y((t \wedge T^m) \wedge T^{m+1}) \\ &\overset{(B.4.149)}{=} \sum_{n=1}^{N} \int_0^{t\wedge T^m} \Lambda_n^{(m+1)}(\tau)\ \mathrm{d}W_n(\tau) + \sum_{i,j=1}^{D} \int_0^{t\wedge T^m} \Gamma_{ij}^{(m+1)}(\tau)\ \mathrm{d}\mathcal{Q}_{ij}(\tau) \\ &\overset{(B.4.150),(B.4.151)}{=} \sum_{n=1}^{N} \int_0^t \Lambda_n^{(m+1)}[0,T^m](\tau)\ \mathrm{d}W_n(\tau) \\ &\qquad + \sum_{i,j=1}^{D} \int_0^t \Gamma_{ij}^{(m+1)}[0,T^m](\tau)\ \mathrm{d}\mathcal{Q}_{ij}(\tau). \end{aligned} \tag{B.4.153}$$

Subtracting (B.4.153) from (B.4.152), setting $t = T$ then squaring and taking expectations, we have from the Itô isometry and by a similar calculation to the one in (B.4.99),

$$\begin{aligned}0 &= \mathrm{E}\Bigg|\sum_{n=1}^{N}\int_0^T \left(\Lambda_n^{(m)}[0,T^m]-\Lambda_n^{(m+1)}[0,T^m]\right)(t)\ \mathrm{d}W_n(t)\\ &\qquad + \sum_{i,j=1}^{D}\int_0^T \Big(\Gamma_{ij}^{(m)}[0,T^m]-\Gamma_{ij}^{(m+1)}[0,T^m]\Big)(t)\ \mathrm{d}\mathcal{Q}_{ij}(t)\Bigg|^2\\ &= \mathrm{E}\Bigg[\sum_{n=1}^{N}\int_0^{\cdot} \left(\Lambda_n^{(m)}[0,T^m]-\Lambda_n^{(m+1)}[0,T^m]\right)(t)\ \mathrm{d}W_n(t)\\ &\qquad + \sum_{i,j=1}^{D}\int_0^{\cdot} \Big(\Gamma_{ij}^{(m)}[0,T^m]-\Gamma_{ij}^{(m+1)}[0,T^m]\Big)(t)\ \mathrm{d}\mathcal{Q}_{ij}(t)\Bigg](T)\\ &= \sum_{n=1}^{N}\mathrm{E}\int_0^T \left(\Lambda_n^{(m)}[0,T^m]-\Lambda_n^{(m+1)}[0,T^m]\right)^2(t)\ \mathrm{d}t\\ &\qquad + \sum_{i,j=1}^{D}\mathrm{E}\int_0^T \Big(\Gamma_{ij}^{(m)}[0,T^m]-\Gamma_{ij}^{(m+1)}[0,T^m]\Big)^2(t)\ \mathrm{d}[\mathcal{Q}_{ij}](t).\end{aligned} \tag{B.4.154}$$

Since all the terms on the last line of the above equation are positive, we get

$$\mathrm{E}\int_0^T \left(\Lambda_n^{(m)}[0,T^m]-\Lambda_n^{(m+1)}[0,T^m]\right)^2(t)\ \mathrm{d}t = 0 \quad \text{for } n=1,\ldots,N \tag{B.4.155}$$

and

$$\mathrm{E}\int_0^T \Big(\Gamma_{ij}^{(m)}[0,T^m]-\Gamma_{ij}^{(m+1)}[0,T^m]\Big)^2(t)\ \mathrm{d}[\mathcal{Q}_{ij}](t) = 0 \quad \text{for } i,j=1,\ldots,D. \tag{B.4.156}$$

Hence for all $m\in\mathbb{N}$,

$$\Lambda_n^{(m)}[0,T^m] = \Lambda_n^{(m+1)}[0,T^m] \quad (\mathbb{P}\otimes \mathit{Leb})\text{-a.e. for } n=1,\ldots,N \tag{B.4.157}$$

and

$$\Gamma_{ij}^{(m)}[0,T^m] = \Gamma_{ij}^{(m+1)}[0,T^m] \quad \nu_{[\mathcal{Q}_{ij}]}\text{-a.e. for } i,j=1,\ldots,D,\ i\neq j. \tag{B.4.158}$$

This shows that the integrands are consistent.

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

$$\Lambda_n(\omega,t) := \limsup_{m\to\infty}\Lambda_n^{(m)}(\omega,t) \quad \text{for } n=1,\ldots,N \tag{B.4.159}$$

and

$$\Gamma_{ij}(\omega,t) := \limsup_{m\to\infty}\Gamma_{ij}^{(m)}(\omega,t) \quad \text{for } i,j=1,\ldots,D. \tag{B.4.160}$$

Then $\boldsymbol{\Lambda} = (\Lambda_1, \ldots, \Lambda_N)^\top$ and $\boldsymbol{\Gamma} = \{\Gamma_{ij}\}_{i,j=1}^D$ are previsible processes. As $T^m \Uparrow T$ a.s., then from the consistency of the integrands shown by (B.4.157) and (B.4.158), we get for each $n = 1, \ldots, N$,

$$(\mathbb{P} \otimes \mathit{Leb})\left[(\omega, t) \in \Omega \times [0, T] \,:\, |\Lambda_n(\omega, t)| = +\infty\right] = 0 \tag{B.4.161}$$

and for each $i, j = 1, \ldots, D$, $i \neq j$,

$$\nu_{[\mathcal{Q}_{ij}]}\left[(\omega, t) \in \Omega \times [0, T] \,:\, |\Gamma_{ij}(\omega, t)| = +\infty\right] = 0. \tag{B.4.162}$$

So $\boldsymbol{\Lambda}$ and $\boldsymbol{\Gamma}$ are a.e. real-valued. By the consistency of the integrands $\boldsymbol{\Lambda}^{(m)}$ and $\boldsymbol{\Gamma}^{(m)}$, they satisfy for each $m \in \mathbb{N}$,

$$\Lambda_n[0, T^m] = \Lambda_n^{(m)}[0, T^m] \quad (\mathbb{P} \otimes \mathit{Leb})\text{-a.e. for } n = 1, \ldots, N \tag{B.4.163}$$

and

$$\Gamma_{ij}[0, T^m] = \Gamma_{ij}^{(m)}[0, T^m] \quad \nu_{[\mathcal{Q}_{ij}]}\text{-a.e. for } i, j = 1, \ldots, D,\ i \neq j. \tag{B.4.164}$$

Then $\boldsymbol{\Lambda}[0, T^m] \in L^2(\mathbf{W})$ and $\boldsymbol{\Gamma}[0, T^m] \in L^2(\mathcal{Q})$ for all $m \in \mathbb{N}$, implying that $\boldsymbol{\Lambda} \in L^2_{\text{loc}}(\mathbf{W})$ and $\boldsymbol{\Gamma} \in L^2_{\text{loc}}(\mathcal{Q})$. From (B.4.152),

$$\begin{aligned}
Y(t \wedge T^m) &= \sum_{n=1}^N \int_0^t \Lambda_n^{(m)}[0, T^m](\tau)\ \mathrm{d}W_n(\tau) + \sum_{i,j=1}^D \int_0^t \Gamma_{ij}^{(m)}[0, T^m](\tau)\ \mathrm{d}\mathcal{Q}_{ij}(\tau) \\
&\overset{(B.4.163),(B.4.164)}{=} \sum_{n=1}^N \int_0^t \Lambda_n[0, T^m](\tau)\ \mathrm{d}W_n(\tau) + \sum_{i,j=1}^D \int_0^t \Gamma_{ij}[0, T^m](\tau)\ \mathrm{d}\mathcal{Q}_{ij}(\tau) \\
&\overset{(B.4.150),(B.4.151)}{=} \sum_{n=1}^N \int_0^{t \wedge T^m} \Lambda_n(\tau)\ \mathrm{d}W_n(\tau) + \sum_{i,j=1}^D \int_0^{t \wedge T^m} \Gamma_{ij}(\tau)\ \mathrm{d}\mathcal{Q}_{ij}(\tau).
\end{aligned} \tag{B.4.165}$$

As $\lim_{m \to \infty} Y(t \wedge T^m) = Y(t)$ a.s., upon letting $m \to \infty$ in (B.4.165) we obtain,

$$\begin{aligned}
Y(t) &= \lim_{m \to \infty} \left\{ \sum_{n=1}^N \int_0^{t \wedge T^m} \Lambda_n(\tau)\ \mathrm{d}W_n(\tau) + \sum_{i,j=1}^D \int_0^{t \wedge T^m} \Gamma_{ij}(\tau)\ \mathrm{d}\mathcal{Q}_{ij}(\tau) \right\} \\
&= \sum_{n=1}^N \int_0^t \Lambda_n(\tau)\ \mathrm{d}W_n(\tau) + \sum_{i,j=1}^D \int_0^t \Gamma_{ij}(\tau)\ \mathrm{d}\mathcal{Q}_{ij}(\tau) \quad \text{a.s.}
\end{aligned} \tag{B.4.166}$$

Thus we have shown (B.4.146) and (B.4.147).

To show the uniqueness of the integrands, let $\tilde{\boldsymbol{\Lambda}} = \left(\tilde{\Lambda}_1, \ldots, \tilde{\Lambda}_N\right)^\top \in L^2_{\text{loc}}(\mathbf{W})$ and $\tilde{\boldsymbol{\Gamma}} = \left(\tilde{\Gamma}_{ij}\right)_{i,j=1}^D \in L^2_{\text{loc}}(\mathcal{Q})$ be such that for all $t \in [0, T]$,

$$Y(t) = \sum_{n=1}^N \int_0^t \tilde{\Lambda}_n(\tau)\ \mathrm{d}W_n(\tau) + \sum_{i,j=1}^D \int_0^t \tilde{\Gamma}_{ij}(\tau)\ \mathrm{d}\mathcal{Q}_{ij}(\tau) \quad \text{a.s.} \tag{B.4.167}$$

For each $m \in \mathbb{N}$, we have

$$Y(T \wedge T^m) \overset{(B.4.167)}{=} \sum_{n=1}^{N} \int_0^{T \wedge T^m} \tilde{\Lambda}_n(t) \, \mathrm{d}W_n(t) + \sum_{i,j=1}^{D} \int_0^{T \wedge T^m} \tilde{\Gamma}_{ij}(t) \, \mathrm{d}\mathcal{Q}_{ij}(t)$$
$$= \sum_{n=1}^{N} \int_0^{T} \tilde{\Lambda}_n[0,T^m](t) \, \mathrm{d}W_n(t) + \sum_{i,j=1}^{D} \int_0^{T} \tilde{\Gamma}_{ij}[0,T^m](t) \, \mathrm{d}\mathcal{Q}_{ij}(t) \quad \text{a.s.} \tag{B.4.168}$$

and as $(T^m)_{m\in\mathbb{N}}$ is a localising sequence for Y then $Y(\cdot \wedge T^m) \in \mathcal{M}_0^2((\Omega, \mathcal{F}, \mathbb{P}), \{\mathcal{F}_t\})$. Similarly, from (B.4.165), we obtain the square-integrable martingale

$$Y(T \wedge T^m) = \sum_{n=1}^{N} \int_0^{T} \Lambda_n[0,T^m](t) \, \mathrm{d}W_n(t) + \sum_{i,j=1}^{D} \int_0^{T} \Gamma_{ij}[0,T^m](t) \, \mathrm{d}\mathcal{Q}_{ij}(t). \tag{B.4.169}$$

Subtracting (B.4.168) from (B.4.169), then squaring and taking expectations, we have from the Itô isometry and by a similar calculation to the one in (B.4.99),

$$\begin{aligned}
0 &= \mathrm{E}\Bigg| \sum_{n=1}^{N} \int_0^{T} \left(\Lambda_n[0,T^m] - \tilde{\Lambda}_n[0,T^m]\right)(t) \, \mathrm{d}W_n(t) \\
&\qquad + \sum_{i,j=1}^{D} \int_0^{T} \left(\Gamma_{ij}[0,T^m] - \tilde{\Gamma}_{ij}[0,T^m]\right)(t) \, \mathrm{d}\mathcal{Q}_{ij}(t) \Bigg|^2 \\
&= \mathrm{E}\Bigg[\sum_{n=1}^{N} \int_0^{\cdot} \left(\Lambda_n[0,T^m] - \tilde{\Lambda}_n[0,T^m]\right)(t) \, \mathrm{d}W_n(t) \\
&\qquad + \sum_{i,j=1}^{D} \int_0^{\cdot} \left(\Gamma_{ij}[0,T^m] - \tilde{\Gamma}_{ij}[0,T^m]\right)(t) \, \mathrm{d}\mathcal{Q}_{ij}(t) \Bigg](T) \\
&= \sum_{n=1}^{N} \mathrm{E} \int_0^{T} \left(\Lambda_n[0,T^m] - \tilde{\Lambda}_n[0,T^m]\right)^2(t) \, \mathrm{d}t \\
&\qquad + \sum_{i,j=1}^{D} \mathrm{E} \int_0^{T} \left(\Gamma_{ij}[0,T^m] - \tilde{\Gamma}_{ij}[0,T^m]\right)^2(t) \, \mathrm{d}[\mathcal{Q}_{ij}](t).
\end{aligned} \tag{B.4.170}$$

Since all terms on the last line of the above equation are positive, we have for each $n = 1, \ldots, N$, we have

$$\mathrm{E} \int_0^{T} \left(\Lambda_n[0,T^m] - \tilde{\Lambda}_n[0,T^m]\right)^2(t) \, \mathrm{d}t = 0 \tag{B.4.171}$$

and for each $i, j = 1, \ldots, D$, we have

$$\mathrm{E} \int_0^{T} \left(\Gamma_{ij}[0,T^m] - \tilde{\Gamma}_{ij}[0,T^m]\right)^2(t) \, \mathrm{d}[\mathcal{Q}_{ij}](t) = 0. \tag{B.4.172}$$

Hence for each $n = 1, \ldots, N$, we have

$$\Lambda_n[0, T^m] = \tilde{\Lambda}_n[0, T^m] \quad (\mathbb{P} \otimes Leb)\text{-a.e.} \tag{B.4.173}$$

and for each $i, j = 1, \ldots, D$, $i \neq j$, we have

$$\Gamma_{ij}[0, T^m] = \tilde{\Gamma}_{ij}[0, T^m] \quad \nu_{[\mathcal{Q}_{ij}]}\text{-a.e.} \tag{B.4.174}$$

As this holds for all $m \in \mathbb{N}$ and $T^m \Uparrow T$ a.s., we get for each $n = 1, \ldots, N$,

$$\Lambda_n = \tilde{\Lambda}_n \quad (\mathbb{P} \otimes Leb)\text{-a.e.} \tag{B.4.175}$$

and for each $i, j = 1, \ldots, D$, $i \neq j$,

$$\Gamma_{ij} = \tilde{\Gamma}_{ij} \quad \nu_{[\mathcal{Q}_{ij}]}\text{-a.e.} \tag{B.4.176}$$

which shows the required uniqueness. □

Remark B.4.23. The remaining part of this appendix is not needed for the thesis. However, it shows quite nicely how the martingale representation theorem for square-integrable martingales (Theorem B.4.6) ties in with a decomposition theorem found in Protter [42], Chapter IV, Section 3.

From Protter [42], Chapter IV, Section 3, page 181 we have the following definition.

Definition B.4.24. Two martingales $N, M \in \mathcal{M}_0^2(\{\mathcal{F}_t\}, \mathbb{P})$ are said to be *strongly orthogonal* if their product $L = NM$ is a uniformly integrable martingale.

Remark B.4.25. From Lemma B.3.19 and Theorem C.9.4, $\mathcal{Q}_{ij}\mathcal{Q}_{ab}$ is a uniformly integrable martingale, which implies that $\mathcal{Q}_{ij}$ and $\mathcal{Q}_{ab}$ are strongly orthogonal.

Definition B.4.26. Let $\mathcal{N} \subset \mathcal{M}_0^2(\{\mathcal{F}_t\}, \mathbb{P})$. The *strong orthogonal complement* $\mathcal{N}^\times$ of $\mathcal{N}$ in $\mathcal{M}_0^2(\{\mathcal{F}_t\}, \mathbb{P})$ is defined by

$$\mathcal{N}^\times := \{A \in \mathcal{M}_0^2(\{\mathcal{F}_t\}, \mathbb{P}) \,|\, A \text{ is strongly orthogonal to every } N \in \mathcal{N}\}. \tag{B.4.177}$$

From Jacod and Shiryaev [25], Proposition I.4.15 we have the following proposition. We use it in Lemma B.4.28 to show that any square-integrable martingale which can be expressed as a stochastic integral with the canonical martingales $\mathcal{Q}$ of the Markov chain as integrator, is strongly orthogonal to any square-integrable martingale which can be expressed as a stochastic integral with the N-dimensional Brownian motion $\mathbf{W}$ as integrator.

Proposition B.4.27. *Let $N, M \in \mathcal{M}^2(\{\tilde{\mathcal{F}}_t\}, \tilde{\mathbb{P}})$. There is an equivalence between*

- *N and M are strongly orthogonal; and*
- *$\langle N, M \rangle = 0$.*

Lemma B.4.28. *Defining*

$$S(\mathcal{Q}) := \left\{ A = \left\{ \sum_{i,j=1}^{D} \int_0^t \Gamma_{ij}(\tau)\ \mathrm{d}\mathcal{Q}_{ij}(\tau) : t \in [0,T] \right\} \middle| \boldsymbol{\Gamma} = (\Gamma_{ij})_{i,j=1}^{D} \in L^2(\mathcal{Q}) \right\}, \tag{B.4.178}$$

and

$$S(\mathbf{W}) := \left\{ B = \left\{ \int_0^t \boldsymbol{\Lambda}^\top(\tau)\ \mathrm{d}\mathbf{W}(\tau) : t \in [0,T] \right\} \middle| \boldsymbol{\Lambda} \in L^2(\mathbf{W}) \right\} \tag{B.4.179}$$

we have that

$$S(\mathcal{Q}) = (S(\mathbf{W}))^\times \tag{B.4.180}$$

and

$$S(\mathbf{W}) = (S(\mathcal{Q}))^\times. \tag{B.4.181}$$

Proof. It is clear from their definitions that $S(\mathcal{Q}) \subset \mathcal{M}_0^2(\{\mathcal{F}_t\}, \mathbb{P})$ and $S(\mathbf{W}) \subset \mathcal{M}_0^2(\{\mathcal{F}_t\}, \mathbb{P})$. We show that $S(\mathcal{Q}) = (S(\mathbf{W}))^\times$.

We first show that $S(\mathcal{Q}) \subset (S(\mathbf{W}))^\times$. Fix $A \in S(\mathcal{Q})$, so there exists some $\boldsymbol{\Gamma} = (\Gamma_{ij})_{i,j=1}^{D} \in L^2(\mathcal{Q})$ such that $A(t) = \sum_{i,j=1}^{D} \int_0^t \Gamma_{ij}(\tau)\ \mathrm{d}\mathcal{Q}_{ij}(\tau)$. Now fix $B \in S(\mathbf{W})$, so there exists some $\boldsymbol{\Lambda} \in L^2(\mathbf{W})$ such that $B(t) = \int_0^t \boldsymbol{\Lambda}^\top(\tau)\ \mathrm{d}\mathbf{W}(\tau)$. Then for all $t \in [0,T]$,

$$\begin{aligned} \langle A, B\rangle(t) &= \langle \sum_{i,j=1}^{D} \int_0^{\cdot} \Gamma_{ij}(\tau)\ \mathrm{d}\mathcal{Q}_{ij}(\tau), \int_0^{\cdot} \boldsymbol{\Lambda}^\top(\tau)\ \mathrm{d}\mathbf{W}(\tau)\rangle(t) \\ &= \sum_{n=1}^{N} \sum_{i,j=1}^{D} \int_0^t \Gamma_{ij}(\tau)\Lambda_n(\tau)\ \mathrm{d}\langle W_n, \mathcal{Q}_{ij}\rangle(\tau) \\ &\overset{\text{Lemma B.3.18}}{=} 0 \text{ a.s.} \end{aligned} \tag{B.4.182}$$

Hence, by Proposition B.4.27, AB is a uniformly integrable martingale, that is, A and B are strongly orthogonal. By the arbitrary choice of $B \in S(\mathbf{W})$, A is strongly orthogonal to all $B \in S(\mathbf{W})$, so $A \in (S(\mathbf{W}))^\times$. As $A \in S(\mathcal{Q})$ was also arbitrarily chosen, then $S(\mathcal{Q}) \subset (S(\mathbf{W}))^\times$.

Next we show that $(S(\mathbf{W}))^\times \subset S(\mathcal{Q})$.

Fix $A \in (S(\mathbf{W}))^\times$. From Theorem B.4.6, there exists $\boldsymbol{\Lambda}^A \in L^2(\mathbf{W})$ and $\boldsymbol{\Gamma}^A = (\Gamma_{ij}^A)_{i,j=1}^{D} \in L^2(\mathcal{Q})$ such that

$$A(t) = \int_0^t \left(\boldsymbol{\Lambda}^A\right)^\top(\tau)\ \mathrm{d}\mathbf{W}(\tau) + \sum_{i,j=1}^{D} \int_0^t \Gamma_{ij}^A(\tau)\ \mathrm{d}\mathcal{Q}_{ij}(\tau). \tag{B.4.183}$$

In particular, we have

$$\int_0^t \left(\boldsymbol{\Lambda}^A\right)^\top(\tau)\ \mathrm{d}\mathbf{W}(\tau) \in S(\mathbf{W}) \quad \text{and} \quad \sum_{i,j=1}^{D} \int_0^t \Gamma_{ij}^A(\tau)\ \mathrm{d}\mathcal{Q}_{ij}(\tau) \in S(\mathcal{Q}). \tag{B.4.184}$$

Fix $B \in S(\mathbf{W})$, so there exists some $\mathbf{\Lambda}^B \in L^2(\mathbf{W})$ such that

$$B(t) = \int_0^t \left(\mathbf{\Lambda}^B\right)^\top(\tau)\ \mathrm{d}\mathbf{W}(\tau). \tag{B.4.185}$$

Since $A \in (S(\mathbf{W}))^\times$, the product AB is a uniformly integrable martingale. Then for all $t \in [0,T]$, $[A,B](t) = 0$ a.s., that is

$$\begin{aligned}
0 &= [A,B](t) \\
&= \left[\int_0^\cdot \left(\mathbf{\Lambda}^A\right)^\top(\tau)\ \mathrm{d}\mathbf{W}(\tau) + \sum_{i,j=1}^{D}\int_0^\cdot \Gamma_{ij}^A(\tau)\ \mathrm{d}\mathcal{Q}_{ij}(\tau), \int_0^\cdot \left(\mathbf{\Lambda}^B\right)^\top(\tau)\ \mathrm{d}\mathbf{W}(\tau)\right](t) \\
&= \sum_{n=1}^{N}\sum_{m=1}^{N}\int_0^t \Lambda_n^A(\tau)\Lambda_m^B(\tau)\ \mathrm{d}[W_n, W_m](\tau) + \sum_{n=1}^{N}\sum_{i,j=1}^{D}\int_0^t \Gamma_{ij}^A(\tau)\Lambda_n^B(\tau)\ \mathrm{d}[W_n, \mathcal{Q}_{ij}](\tau) \\
&\overset{\text{Lemma B.3.18}}{=} \sum_{n=1}^{N}\int_0^t \Lambda_m^A(\tau)\Lambda_n^B(\tau)\ \mathrm{d}\tau \text{ a.s.}
\end{aligned} \tag{B.4.186}$$

By the arbitrary choice of B, this means $\mathbf{\Lambda}^A = 0$ $(\mathbb{P} \otimes Leb)$-a.e. Hence

$$A(\cdot) = \sum_{i,j=1}^{D}\int_0^\cdot \Gamma_{ij}^A\ \mathrm{d}\mathcal{Q}_{ij} \in S(\mathcal{Q}) \tag{B.4.187}$$

and by the arbitrary choice of $A \in (S(\mathbf{W}))^\times$, it follows that $(S(\mathbf{W}))^\times \subset S(\mathcal{Q})$.

Hence as we have also shown that $S(\mathcal{Q}) \subset (S(\mathbf{W}))^\times$, we must have equality, that is $S(\mathcal{Q}) = (S(\mathbf{W}))^\times$.

A similar argument shows that $S(\mathbf{W}) = (S(\mathcal{Q}))^\times$. □

From Protter [42], Chapter IV, Section 3, pages 182 we have the following definition and lemma.

Definition B.4.29. A closed subspace F of $\mathcal{M}_0^2(\{\tilde{\mathcal{F}}_t\}, \tilde{\mathbb{P}})$ is called a *stable subspace* if it is stable under stopping, in other words if $M \in F$ and T is a stopping time then $M(\cdot \wedge T) \in F$.

Lemma B.4.30. *If F is any subset of $\mathcal{M}_0^2(\{\tilde{\mathcal{F}}_t\}, \tilde{\mathbb{P}})$ then $F^\times$ is closed and stable.*

Remark B.4.31. $S(\mathbf{W}) \subset \mathcal{M}_0^2(\{\tilde{\mathcal{F}}_t\}, \tilde{\mathbb{P}})$. Hence by Lemma B.4.30, $(S(\mathbf{W}))^\times$ is closed and stable. By Lemma B.4.28, this implies that $S(\mathcal{Q})$ is closed and stable.

From Protter [42], Chapter IV, Section 3, page 183 we have the following theorem.

Theorem B.4.32. *Let F be a stable subset of $\mathcal{M}_0^2(\{\tilde{\mathcal{F}}_t\}, \tilde{\mathbb{P}})$. Then each $M \in \mathcal{M}_0^2(\{\tilde{\mathcal{F}}_t\}, \tilde{\mathbb{P}})$ has a unique decomposition $M = A + B$ with $A \in F$ and $B \in F^\times$.*

Remark B.4.33. From Remark B.4.31 and Theorem B.4.32, every $M \in \mathcal{M}_0^2(\{\tilde{\mathcal{F}}_t\}, \tilde{\mathbb{P}})$ has a unique decomposition $M = A + B$ with $A \in S(\mathbf{W})$ and $B \in S(\mathcal{Q})$. This is the same decomposition that Theorem B.4.6 gives us, except that we have the additional information that $S(\mathbf{W})$ is a closed, stable subset of $\mathcal{M}_0^2(\{\tilde{\mathcal{F}}_t\}, \tilde{\mathbb{P}})$ and its strong orthogonal complement is $S(\mathcal{Q})$. The set $S(\mathbf{W})$ can be considered as the set of continuous martingales in $\mathcal{M}_0^2(\{\tilde{\mathcal{F}}_t\}, \tilde{\mathbb{P}})$. Similarly, $S(\mathcal{Q})$ is a closed, stable subset of $\mathcal{M}_0^2(\{\tilde{\mathcal{F}}_t\}, \tilde{\mathbb{P}})$ and its strong orthogonal complement is $S(\mathbf{W})$. The set $S(\mathcal{Q})$ can be considered as the set of purely discontinuous martingales in $\mathcal{M}_0^2(\{\tilde{\mathcal{F}}_t\}, \tilde{\mathbb{P}})$.

Appendix C

Standard definitions and results

C.1 Measure theory results

The following lemma is a slight adaptation of Karatzas and Shreve [31], Lemma 5.4.2.

Lemma C.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*

$$\delta(\mathbf{z}) := \sup_{\mathbf{p}\in K}\{-\mathbf{p}^\top \mathbf{z}\} \quad \forall \mathbf{z} \in \mathbb{R}^N. \tag{C.1.1}$$

For any given previsible process $\mathbf{p} : \Omega \times [0,T] \to \mathbb{R}^N$, there exists an $\mathbb{R}^N$-valued, previsible $\boldsymbol{\nu}(\cdot)$ such that a.s.

$$\|\boldsymbol{\nu}(t)\| \leq 1, \quad |\delta(\boldsymbol{\nu}(t))| \leq 1 \quad \forall t \in [0,T] \tag{C.1.2}$$

and, setting

$$B := \{(\omega,t) \in \Omega \times [0,T] : p(\omega,t) \in K\}, \tag{C.1.3}$$

we have for all $t \in [0,T]$,

$$\begin{cases} \delta(\boldsymbol{\nu}(t)) + \mathbf{p}^\top(t)\boldsymbol{\nu}(t) = 0 & (\mathbb{P}\otimes Leb)\text{-a.e. on } B \\ \delta(\boldsymbol{\nu}(t)) + \mathbf{p}^\top(t)\boldsymbol{\nu}(t) < 0 & (\mathbb{P}\otimes Leb)\text{-a.e. on } \Omega\times[0,T]\setminus B. \end{cases} \tag{C.1.4}$$

Proof. The proof follows that of Karatzas and Shreve [31], Lemma 5.4.2. The main difference is that Karatzas and Shreve [31], Lemma 5.4.2 is for an $\{\mathcal{F}_t^{\mathbf{W}}\}$-progressively measurable process $\mathbf{p}$, in place of our ($\{\mathcal{F}_t\}$-)previsible $\mathbf{p}$. 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 previsible.

Karatzas and Shreve [31], Lemma 5.4.2 then gives us (C.1.2) and that for all $t \in [0,T]$, we have

$$\begin{cases} \mathbf{p} \in K & \Leftrightarrow \boldsymbol{\nu}(t) = 0 \\ \mathbf{p} \notin K & \Leftrightarrow \delta(\boldsymbol{\nu}(t)) + \mathbf{p}^\top(t)\boldsymbol{\nu}(t) < 0. \end{cases} \tag{C.1.5}$$

However, since $\boldsymbol{\nu}(t) = 0$ implies both that $\mathbf{p}^\top(t)\boldsymbol{\nu}(t) = 0$ and, from (C.1.1), that $\delta(\boldsymbol{\nu}(t)) = \delta(\mathbf{0}) = 0$, then

$$\mathbf{p} \in K \quad \Rightarrow \delta(\boldsymbol{\nu}(t)) + \mathbf{p}^\top(t)\boldsymbol{\nu}(t) = 0. \tag{C.1.6}$$

Finally, noting that $\mathbf{p} \in K$ $(\mathbb{P} \otimes Leb)$-a.e. on B, we obtain (C.1.4). □

From Friedman [16], Theorem 2.12.4, we have the *Radon-Nikodým Theorem.*

Theorem C.1.2. Radon-Nikodým Theorem *Let $(E, \mathcal{E}, \mu)$ be a σ-finite measure space with μ a measure, and let ν be a σ-finite signed measure on $\mathcal{E}$, absolutely continuous with respect to μ. There there exists a real-valued measurable function f on E such that*

$$\nu(A) = \int_A f \,\mathrm{d}\mu \tag{C.1.7}$$

for every measurable set A for which $|\nu|(A) < \infty$. If g is another function such that $\nu(A) = \int_A g \,\mathrm{d}\mu$ for any measurable set A for which $|\nu|(A) < \infty$, then $f = g$ a.e. with respect to μ.

Remark C.1.3. The function f appearing in Theorem C.1.2 is called the *Radon-Nikodým derivative* and is often denoted $\frac{\mathrm{d}\nu}{\mathrm{d}\mu}$. For all bounded, measurable functions H, it satisfies

$$\int_E H \,\mathrm{d}\nu = \int_E H \frac{\mathrm{d}\nu}{\mathrm{d}\mu} \,\mathrm{d}\mu. \tag{C.1.8}$$

From Friedman [16], Theorem 2.10.5, we have the well-known *Fatou's Lemma.*

Theorem C.1.4. Fatou's Lemma *Let $(E, \mathcal{E}, \mu)$ be a σ-finite measure space with μ a measure. Let $\{f^{(m)}\}_{m \in \mathbb{N}}$ be a sequence of nonnegative integrable functions on E. Then*

$$\int_E \liminf_{m\to\infty} f^{(m)} \,\mathrm{d}\mu \leq \liminf_{m\to\infty} \int_E f^{(m)} \,\mathrm{d}\mu. \tag{C.1.9}$$

C.2 Lp-spaces results

From Rogers and Williams [45], Theorem II.70.2, we have *Doob's L^p-inequality.* Although their theorem also deals with submartingales, we state only the martingale part here.

Theorem C.2.1. Doob's L^p-inequality *Let $p > 1$. 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 bounded in $L^p(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}})$. Then $M_\infty := \lim_{t\to\infty} M(t)$ exists a.s. and in $L^p(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}})$, and*

$$E\left(M_\infty \,|\, \tilde{\mathcal{F}}_t\right) = M(t) \quad a.s. \tag{C.2.1}$$

We also have

$$E\left(\sup_{t\geq 0}|M(t)|^p\right) \leq \left(\frac{p}{p-1}\right)^p E|M_\infty|^p. \tag{C.2.2}$$

From Friedman [16], Theorem 3.2.1, we have *Hölder's inequality.*

Theorem C.2.2. Hölder's inequality *Let* $(E, \mathcal{E}, \mu)$ *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* $f \in L^p(E, \mathcal{E}, \mu)$*,* $g \in L^q(E, \mathcal{E}, \mu)$*, then* $fg \in L^1(E, \mathcal{E}, \mu)$ *and*

$$\int_E |fg| \, \mathrm{d}\mu \leq \left(\int_E |f|^p \, \mathrm{d}\mu \right)^{\frac{1}{p}} \left(\int_E |g|^q \, \mathrm{d}\mu \right)^{\frac{1}{q}}. \tag{C.2.3}$$

C.3 Conditional expectation results

From Chow and Teicher [7], Theorem 7.2.4, we have *Hölder's inequality for conditional expectations.*

Theorem C.3.1. 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*

$$E\left(|XY| \,|\mathcal{G}\right) \leq \left(E\left(|X|^p \,|\mathcal{G}\right)\right)^{\frac{1}{p}} \cdot \left(E\left(|Y|^q \,|\mathcal{G}\right)\right)^{\frac{1}{q}}. \tag{C.3.1}$$

From Elliott [14], Lemma 1.9, we have *Jensen's inequality for conditional expectations.*

Lemma C.3.2. 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*

$$f\left(E(X \,|\, \mathcal{G})\right) \leq E(f(X) \,|\, \mathcal{G}). \tag{C.3.2}$$

C.4 General definitions and conventions for stochastic processes

Definition C.4.1. A process $X = \{X(t) : t \in [0, T]\}$ is *càdlàg* (continu à droite avec des limites à gauche) if it is pathwise right-continuous with finite left-hand limits.

If X is càdlàg then we define the process $X_- = \{X(t_-) : t \in [0, T]\}$ as

$$X(0_-) := X(0) \qquad \text{and} \qquad X(t_-) := \lim_{\substack{s \to t \\ s < t}} X(s), \qquad \forall t \in (0, T], \tag{C.4.1}$$

and we also define the process $\triangle X = \{\triangle X(t) : t \in [0, T]\}$ as

$$\triangle X(t) := X(t) - X(t_-), \qquad \forall t \in [0, T]. \tag{C.4.2}$$

Definition C.4.2. A process $X = \{X(t) : t \in [0, T]\}$ is *càglàd* (continu à gauche avec des limites à droite) if the mappings $t \mapsto X(\omega, t)$ are left-continuous with finite right-hand limits on $[0, \infty)$ for all $\omega \in \tilde{\Omega}$.

Definition C.4.3. The *raw filtration* $\{\mathcal{F}_t^X\}$ generated by a stochastic process $X = \{X(t) : t \in [0, T]\}$ is

$$\mathcal{F}_t^X := \sigma\{X(s) : s \in [0, t]\} \quad \forall t \in [0, T]. \tag{C.4.3}$$

Definition C.4.4. 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 $\tilde{\mathcal{F}}$.

Notation C.4.5. 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 $\mathrm{E}_{\mathbb{P}}$. If the expectation is with respect to another measure $\tilde{\mathbb{P}}$, we will write $\mathrm{E}_{\tilde{\mathbb{P}}}$.

Definition C.4.6. 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 $t \mapsto X(\omega, t)$ are non-decreasing on $[0, \infty)$ for all $\omega \in \tilde{\Omega}$.

C.5 Stopping time results

From Jacod and Shiryaev [25], Proposition I.1.28(a), Definition I.1.11(a) and Definition I.1.20a, we have the following proposition. Note that a càd process is a process such that all of its paths are right-continuous.

Proposition C.5.1. *If X is an $\mathbb{R}^n$-valued $\{\mathcal{F}_t\}$-adapted càd process and if B is an open subset of $\mathbb{R}^n$, then $S := \inf\{t : X(t) \in B\}$ is an $\{\mathcal{F}_t\}$-stopping time.*

C.6 Spaces of martingales

Definition C.6.1. $\mathcal{M}((\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}}), \{\tilde{\mathcal{F}}_t\})$ denotes the set of all real-valued, $\{\tilde{\mathcal{F}}_t\}$-adapted processes $M = \{M(t) : t \in [0, T]\}$ on $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}})$ such that $\{(M(t), \tilde{\mathcal{F}}_t) : t \in [0, T]\}$ is a martingale, in other words

- $\mathrm{E}_{\tilde{\mathbb{P}}}|M(t)| < \infty$, for all $t \in [0, T]$; and
- $\mathrm{E}_{\tilde{\mathbb{P}}}\left(M(t) \,\middle|\, \tilde{\mathcal{F}}_s\right) = M(s)$ $\tilde{\mathbb{P}}$-a.s., for all $0 \le s \le t \le T$.

If there is no ambiguity about the measurable space on which the space of martingales $\mathcal{M}((\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}}), \{\tilde{\mathcal{F}}_t\})$ is defined, then we specify only the filtration and probability measure. In that case, we use the notation $\mathcal{M}(\{\tilde{\mathcal{F}}_t\}, \tilde{\mathbb{P}})$ instead of $\mathcal{M}((\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}}), \{\tilde{\mathcal{F}}_t\})$. We continue this convention for all the spaces of processes that we define below.

$\mathcal{M}_0(\{\tilde{\mathcal{F}}_t\}, \tilde{\mathbb{P}})$ denotes the set of $M \in \mathcal{M}(\{\tilde{\mathcal{F}}_t\}, \tilde{\mathbb{P}})$ which are are $\tilde{\mathbb{P}}$-a.s. null at the origin.

$\mathcal{M}^c(\{\tilde{\mathcal{F}}_t\}, \tilde{\mathbb{P}})$ denotes the set of $M \in \mathcal{M}(\{\tilde{\mathcal{F}}_t\}, \tilde{\mathbb{P}})$ which are continuous martingales.

$\mathcal{M}^c_0(\{\tilde{\mathcal{F}}_t\}, \tilde{\mathbb{P}})$ denotes the set of $M \in \mathcal{M}^c(\{\tilde{\mathcal{F}}_t\}, \tilde{\mathbb{P}})$ which are $\tilde{\mathbb{P}}$-a.s. null at the origin.

Definition C.6.2. A martingale M is *square-integrable* if $\mathrm{E}|M(t)|^2 < \infty$, for all $t \in [0, T]$.

Definition C.6.3. A martingale M is *L^2-bounded* if $\sup_{t\in[0,T]} \mathrm{E}|M(t)|^2 < \infty$.

Remark C.6.4. As we are dealing with a finite time interval $[0, T]$, then a martingale M is square-integrable if and only if it is L^2-bounded.

Definition C.6.5. $\mathcal{M}^2(\{\tilde{\mathcal{F}}_t\}, \tilde{\mathbb{P}})$ denotes the set of $M \in \mathcal{M}(\{\tilde{\mathcal{F}}_t\}, \tilde{\mathbb{P}})$ which are square-integrable.

Similarly, $\mathcal{M}^2_0(\{\tilde{\mathcal{F}}_t\}, \tilde{\mathbb{P}})$ denotes the set of $M \in \mathcal{M}_0(\{\tilde{\mathcal{F}}_t\}, \tilde{\mathbb{P}})$ which are square-integrable.

C.7 Spaces of local martingales

Definition C.7.1. For a sequence $\{T^m\}_{m\in\mathbb{N}}$ of $\{\tilde{\mathcal{F}}_t\}$-stopping times, we write

$$T^m \Uparrow T \quad \tilde{\mathbb{P}}\text{-a.s.} \tag{C.7.1}$$

to mean that $\tilde{\mathbb{P}}$-a.s.,

- $0 \le T^m(\omega) \le T^{m+1}(\omega)$ for all $\omega \in \tilde{\Omega}$ and for all $m \in \mathbb{N}$; and
- there exists $M(\omega) \in \mathbb{N}$ such that $T^m(\omega) = T$, for all $m \ge M(\omega)$ and for all $\omega \in \tilde{\Omega}$.

If the measure is $\mathbb{P}$ and there is no ambiguity about the measure, then we will write $T^m \Uparrow T$ a.s.

Definition C.7.2. $\mathcal{M}_{\mathrm{loc}}((\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}}), \{\tilde{\mathcal{F}}_t\})$ denotes the set of all real-valued processes $M = \{M(t) : t \in [0, T]\}$ on $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}})$ which are $\{\tilde{\mathcal{F}}_t\}$-local martingales, that is such that there exists a sequence $\{T^m\}_{m\in\mathbb{N}}$ of $\{\tilde{\mathcal{F}}_t\}$-stopping times such that

1. $T^m \Uparrow T$ $\tilde{\mathbb{P}}$-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 say that the sequence $\{T^m\}_{m\in\mathbb{N}}$ of $\{\tilde{\mathcal{F}}_t\}$-stopping times is a *localizing sequence* for M.

$\mathcal{M}_{0,\mathrm{loc}}(\{\tilde{\mathcal{F}}_t\},\tilde{\mathbb{P}})$ denotes the set of $M \in \mathcal{M}_{\mathrm{loc}}(\{\tilde{\mathcal{F}}_t\},\tilde{\mathbb{P}})$ which are $\tilde{\mathbb{P}}$-a.s. null at the origin.

$\mathcal{M}^2_{\mathrm{loc}}(\{\tilde{\mathcal{F}}_t\},\tilde{\mathbb{P}})$ denotes the set of $M \in \mathcal{M}_{\mathrm{loc}}(\{\tilde{\mathcal{F}}_t\},\tilde{\mathbb{P}})$ which are locally square-integrable martingales.

We say that $M \in \mathcal{M}^2_{\mathrm{loc}}(\{\tilde{\mathcal{F}}_t\},\tilde{\mathbb{P}})$ is a *locally square-integrable martingale.*

C.8 Spaces of finite variation processes

Definition C.8.1. 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 $t \mapsto A(\omega,t)$ is of finite variation, in other words for all $(\omega,t) \in \Omega \times [0,T]$ the variation $V_A(\omega,t)$ of $s \mapsto A(\omega,s)$ over $(0,t]$ is finite:

$$V_A(\omega,t) := \int_{(0,t]} |\,\mathrm{d}A(\omega,s)| = \sup \sum_{i=1}^{n} |A(\omega,s_i) - A(\omega,s_{i-1})| < \infty. \tag{C.8.1}$$

The supremum is taken over all partitions $0 = s_0 < s_1 < \cdots < s_n = t$ of $[0,t]$.

Definition C.8.2. We denote by $\mathcal{FV}((\tilde{\Omega},\tilde{\mathcal{F}},\tilde{\mathbb{P}}),\{\tilde{\mathcal{F}}_t\})$ the set of all real-valued, $\{\tilde{\mathcal{F}}_t\}$-adapted, càdlàg processes on $(\tilde{\Omega},\tilde{\mathcal{F}},\tilde{\mathbb{P}})$ which are of finite variation.

We denote by $\mathcal{FV}_0(\{\tilde{\mathcal{F}}_t\},\tilde{\mathbb{P}})$ the set of $A \in \mathcal{FV}(\{\tilde{\mathcal{F}}_t\},\tilde{\mathbb{P}})$ which are $\tilde{\mathbb{P}}$-a.s. null at the origin.

We denote by $\mathcal{FV}^+((\tilde{\Omega},\tilde{\mathcal{F}},\tilde{\mathbb{P}}),\{\tilde{\mathcal{F}}_t\})$ the set of all real-valued, $\{\tilde{\mathcal{F}}_t\}$-adapted, càdlàg processes on $(\tilde{\Omega},\tilde{\mathcal{F}},\tilde{\mathbb{P}})$ which are non-decreasing.

We denote by $\mathcal{FV}^+_0(\{\tilde{\mathcal{F}}_t\},\tilde{\mathbb{P}})$ the set of $A \in \mathcal{FV}^+(\{\tilde{\mathcal{F}}_t\},\tilde{\mathbb{P}})$ which are $\tilde{\mathbb{P}}$-a.s. null at the origin.

Definition C.8.3. 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

$$\mathrm{E}\left(V_A(\omega,\infty)\right) < \infty, \tag{C.8.2}$$

for V_A given by (C.8.1) and $V_A(\omega,\infty)$ defined as the pointwise limit of $V_A(\omega,\cdot)$.

Definition C.8.4. We denote by $\mathcal{IV}(\{\tilde{\mathcal{F}}_t\},\tilde{\mathbb{P}})$ the set of $A \in \mathcal{FV}(\{\tilde{\mathcal{F}}_t\},\tilde{\mathbb{P}})$ which are of integrable variation.

We denote by $\mathcal{IV}_0(\{\tilde{\mathcal{F}}_t\},\tilde{\mathbb{P}})$ the set of $A \in \mathcal{IV}(\{\tilde{\mathcal{F}}_t\},\tilde{\mathbb{P}})$ which are $\tilde{\mathbb{P}}$-a.s. null at the origin.

We denote by $\mathcal{IV}^+(\{\tilde{\mathcal{F}}_t\},\tilde{\mathbb{P}})$ the set of $A \in \mathcal{FV}^+(\{\tilde{\mathcal{F}}_t\},\tilde{\mathbb{P}})$ which are integrable, that is $\mathrm{E}(A(\infty)) < \infty$.

We denote by $\mathcal{IV}^+_0(\{\tilde{\mathcal{F}}_t\},\tilde{\mathbb{P}})$ the set of $A \in \mathcal{IV}^+(\{\tilde{\mathcal{F}}_t\},\tilde{\mathbb{P}})$ which are $\tilde{\mathbb{P}}$-a.s. null at the origin.

Definition C.8.5. Let $\mathcal{FV}_{\text{loc}}(\{\tilde{\mathcal{F}}_t\}, \tilde{\mathbb{P}})$ denote the set of processes such that for each $A \in \mathcal{FV}_{\text{loc}}$ there exists a sequence of $\{\tilde{\mathcal{F}}_t\}$-stopping times $\{T^m\}_{m\in\mathbb{N}}$ (depending on A) such that $T^m \Uparrow T$ a.s. and each stopped process $A[0, T^m] \in \mathcal{FV}(\{\tilde{\mathcal{F}}_t\}, \tilde{\mathbb{P}})$.

We use similar notation to denote the localized set of processes for any of the set of processes defined above.

C.9 Quadratic co-variation and variation processes

The following theorem is from Jacod and Shiryaev [25], Theorem I.4.2.

Theorem C.9.1. *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 previsible; and*
3. $NM - \langle N, M\rangle \in \mathcal{M}_{loc}(\{\tilde{\mathcal{F}}_t\}, \tilde{\mathbb{P}})$.

Moreover,

$$\langle N, M\rangle = \frac{1}{4}\left(\langle N + M, N + M\rangle - \langle N - M, N - M\rangle\right) \tag{C.9.1}$$

and if $N, M \in \mathcal{M}^2(\{\tilde{\mathcal{F}}_t\}, \tilde{\mathbb{P}})$ then $\langle N, M\rangle \in \mathcal{IV}_0(\{\tilde{\mathcal{F}}_t\}, \tilde{\mathbb{P}})$ and $NM - \langle N, M\rangle$ is a uniformly integrable martingale.

Furthermore, $\langle M, M\rangle$ is non-decreasing.

Remark C.9.2. We call $\langle N, M\rangle$ the *angle-bracket quadratic co-variation process* of N and M.

Remark C.9.3. For any $M \in \mathcal{M}^2_{\text{loc}}(\{\tilde{\mathcal{F}}_t\}, \tilde{\mathbb{P}})$, the process $\langle M, M\rangle$ is called the *angle-bracket quadratic variation process* of M. We will often write $\langle M\rangle$ for $\langle M, M\rangle$.

From Jacod and Shiryaev [25], equation I.4.46 and Proposition I.4.50 and Rogers and Williams [46], Theorem VI.36.6 and Theorem VI.37.8, we have the following theorem.

Theorem C.9.4. *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. $NM - [N, M] \in \mathcal{M}_{loc}(\{\tilde{\mathcal{F}}_t\}, \tilde{\mathbb{P}})$.

Moreover,

$$[N,M] = \frac{1}{4}\left([N+M,N+M] - [N-M,N-M]\right) \tag{C.9.2}$$

and 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, $[M,M]$ *is non-decreasing.*

Remark C.9.5. We call $[N,M]$ the *square-bracket quadratic co-variation process* of N and M.

Remark C.9.6. For any $M \in \mathcal{M}_{\text{loc}}(\{\tilde{\mathcal{F}}_t\}, \tilde{\mathbb{P}})$, the process $[M,M]$ is called the *square-bracket quadratic variation process* of the local martingale M. We will often write $[M]$ for $[M,M]$.

Remark C.9.7. The square-bracket quadratic co-variation process $[N,M]$ exists for all local martingales N, M. This is the main reason for preferring $[N,M]$ to $\langle N,M\rangle$; the angle-bracket quadratic co-variation process $\langle N,M\rangle$ only exists for local martingales which are locally L^2-bounded.

Furthermore, the square-bracket quadratic variation process $[M]$ is invariant under absolutely continuous changes of measure (see Jacod and Shiryaev [25], Theorem III.3.13), unlike the angle-bracket quadratic variation process $\langle M\rangle$.

Remark C.9.8. If $M \in \mathcal{M}^c_{\text{loc}}(\{\tilde{\mathcal{F}}_t\}, \tilde{\mathbb{P}})$, then

$$[M] = \langle M\rangle. \tag{C.9.3}$$

C.10 Purely discontinuous local martingales

From Jacod and Shiryaev [25], Definition I.4.11 we have the following definition.

Definition C.10.1. Two local martingales N and M are called *orthogonal* if their product $L = NM$ is a local martingale.

From Jacod and Shiryaev [25], Definition I.4.11 we have the following definition.

Definition C.10.2. A local martingale M is called a *purely discontinuous local martingale* if $M(0) = 0$ and if it is orthogonal to all continuous local martingales.

C.11 Decompositions of semimartingales

Definition C.11.1. Let $\mathcal{SM}((\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}}), \{\tilde{\mathcal{F}}_t\})$ denote the set of all real-valued processes $\{S(t) : t \in [0,T]\}$ on $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}})$ which are semimartingales, that is the set of all real-valued processes $\{S(t) : t \in [0,T]\}$ on $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}})$ which can be written in the form

$$S = S(0) + M + A, \tag{C.11.1}$$

for some $M \in \mathcal{M}_{0,\text{loc}}((\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}}), \{\tilde{\mathcal{F}}_t\})$ and $A \in \mathcal{FV}_0((\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}}), \{\tilde{\mathcal{F}}_t\})$, which are generally non-unique.

We call M the *local martingale part* of the semimartingale S and we call A the *finite variation part* of the semimartingale S.

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 $\mathcal{SM}((\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}}), \{\tilde{\mathcal{F}}_t\})$.

From Jacod and Shiryaev [25], Theorem I.4.18 we have the following theorem.

Theorem C.11.2. *Any local martingale M admits a unique (up to indistinguishability) decomposition*

$$M = M(0) + M^c + M^d, \tag{C.11.2}$$

where $M^c(0) = M^d(0) = 0$, M^c is a continuous local martingale and M^d is a purely discontinuous local martingale.

Remark C.11.3. 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 Jacod and Shiryaev [25], Proposition I.4.27 we have the following proposition.

Proposition C.11.4. *Let $S \in \mathcal{SM}(\{\tilde{\mathcal{F}}_t\}, \tilde{\mathbb{P}})$. Then for any two arbitrary decompositions*

$$S = S(0) + M + A \tag{C.11.3}$$

and

$$S = S(0) + \tilde{M} + \tilde{A}, \tag{C.11.4}$$

for $M, \tilde{M} \in \mathcal{M}_{0,loc}((\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}}), \{\tilde{\mathcal{F}}_t\})$ and $A, \tilde{A} \in \mathcal{FV}_0((\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}}), \{\tilde{\mathcal{F}}_t\})$, upon applying Theorem C.11.2 to the local martingales $M, \tilde{M} \in \mathcal{M}_{0,loc}((\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}}), \{\tilde{\mathcal{F}}_t\})$ to get the decompositions

$$M = M(0) + M^c + M^d, \tag{C.11.5}$$

and

$$\tilde{M} = \tilde{M}(0) + \tilde{M}^c + \tilde{M}^d, \tag{C.11.6}$$

where $M^c(0) = M^d(0) = \tilde{M}^c(0) = \tilde{M}^d(0) = 0$, $M^c, \tilde{M}^c$ are continuous local martingales and $M^d, \tilde{M}^d$ are purely discontinuous local martingales, we have that M^c and $\tilde{M}^c$ are indistinguishable.

Remark C.11.5. Given any $S \in \mathcal{SM}(\{\tilde{\mathcal{F}}_t\}, \tilde{\mathbb{P}})$, denote by S^c the unique (up to indistinguishability) member of $\mathcal{M}^c_{0,\text{loc}}(\{\tilde{\mathcal{F}}_t\}, \tilde{\mathbb{P}})$ such that for any decomposition

$$S = S(0) + M + A \tag{C.11.7}$$

for $M \in \mathcal{M}_{0,\text{loc}}((\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}}), \{\tilde{\mathcal{F}}_t\})$ and $A \in \mathcal{FV}_0((\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}}), \{\tilde{\mathcal{F}}_t\})$, the continuous part of the local martingale M and S^c are indistinguishable.

C.12 Compensator results

From Jacod and Shiryaev [25], Theorem I.3.17, we have the following result.

Theorem C.12.1. *Let $A \in \mathcal{IV}^+_{0,loc}(\{\tilde{\mathcal{F}}_t\}, \tilde{\mathbb{P}})$. Then there exists a previsible process $\bar{A} \in \mathcal{IV}^+_{0,loc}(\{\tilde{\mathcal{F}}_t\}, \tilde{\mathbb{P}})$ such that $A - \bar{A} \in \mathcal{M}_{0,loc}(\{\tilde{\mathcal{F}}_t\}, \tilde{\mathbb{P}})$.*

Moreover, for any previsible process $\hat{A} \in \mathcal{IV}^+_{0,loc}(\{\tilde{\mathcal{F}}_t\}, \tilde{\mathbb{P}})$ such that $A - \hat{A} \in \mathcal{M}_{0,loc}(\{\tilde{\mathcal{F}}_t\}, \tilde{\mathbb{P}})$, the processes $\bar{A}$ and $\hat{A}$ are indistinguishable.

Remark C.12.2. Given $A \in \mathcal{IV}^+_{0,\mathrm{loc}}(\{\tilde{\mathcal{F}}_t\}, \tilde{\mathbb{P}})$, denote by A^p any $\bar{A} \in \mathcal{IV}^+_{0,\mathrm{loc}}(\{\tilde{\mathcal{F}}_t\}, \tilde{\mathbb{P}})$ given by Theorem C.12.1. The process A^p is called the *compensator* of A and is unique to within indistinguishability.

Theorem C.12.3. *Let $A \in \mathcal{IV}^+_{0,loc}(\{\tilde{\mathcal{F}}_t\}, \tilde{\mathbb{P}})$. For each previsible process $\bar{A} \in \mathcal{IV}^+_{0,loc}(\{\tilde{\mathcal{F}}_t\}, \tilde{\mathbb{P}})$ the following are equivalent:*

1. *$\bar{A}$ is the compensator of A;*
2. *for all nonnegative, previsible processes H,*

$$E\int_0^\infty H(\tau)\,\mathrm{d}A(\tau) = E\int_0^\infty H(\tau)\,\mathrm{d}A^p(\tau). \tag{C.12.1}$$

Rogers and Williams [46], Theorem VI.34.2, we have the following result for the existence of the compensator of a process.

Theorem C.12.4. *Let M be a local martingale null at the origin. Then the following statements are equivalent:*

1. *there exists a previsible increasing process $\langle M\rangle$ null at the origin such that $M^2 - \langle M\rangle$ is a local martingale;*
2. *$M \in \mathcal{M}^2_{0,loc}$;*
3. *the increasing process $[M]$ is locally integrable.*

When these conditions hold, $\langle M\rangle$ is the compensator of $[M]$.

C.13 Purely discontinuous semimartingales

From Jacod and Shiryaev [25], Definition I.4.45, we have the following definition.

Definition C.13.1. The square-bracket quadratic co-variation process of two semimartingales X and Y is

$$[X,Y](t) := X(t)Y(t) - X(0)Y(0) - \int_0^t X(\tau_-)\,\mathrm{d}Y(\tau) - \int_0^t Y(\tau_-)\,\mathrm{d}X(\tau), \tag{C.13.1}$$

which is defined uniquely up to indistinguishability.

From Jacod and Shiryaev [25], Definition I.4.52, we have the following theorem.

Theorem C.13.2. *Let X, Y be semi-martingales and let X^c, Y^c denote their continuous martingale parts, respectively. Then*

$$[X,Y](t) = \langle X^c, Y^c\rangle(t) + \sum_{0\le s\le t} \triangle X(s)\triangle Y(s) \tag{C.13.2}$$

From Protter [42], Chapter II, Section 6, page 71 we have the following definition and theorem.

Definition C.13.3. Let X be a semimartingale and let X^c denote its continuous martingale part. Then X is called a *purely discontinuous semimartingale* if $\langle X^c, X^c\rangle = 0$.

Theorem C.13.4. *If a semimartingale X is adapted, càdlàg, with paths of finite variation then X is a purely discontinuous semimartingale.*

Remark C.13.5. Protter uses the term *quadratic pure jump* for purely discontinuous.

As a special case of Jacod and Shiryaev [25], Theorem I.4.52, we have the following theorem.

Theorem C.13.6. *Let X be a purely discontinuous semimartingale. Then for any semimartingale Y we have*

$$[X,Y](t) = \sum_{0\le s\le t} \triangle X(s)\triangle Y(s). \tag{C.13.3}$$

C.14 Stochastic integration results

From Rogers and Williams [46], Theorem IV.38.3, we have the following *integration-by-parts formula* for semimartingales.

Theorem C.14.1. Integration-by-parts formula *Let X and Y be semimartingales. Then*

$$X(t)Y(t) = X(0)Y(0) + \int_0^t X(\tau_-)\,\mathrm{d}Y(\tau) + \int_0^t Y(\tau_-)\,\mathrm{d}X(\tau) + [X,Y](t). \tag{C.14.1}$$

From Rogers and Williams [46], Theorem VI.39.1, we have the following *Itô's Formula* for semimartingales.

Theorem C.14.2. Itô's Formula *Let $f : \mathbb{R}^N \to \mathbb{R}$ be a function which has continuous derivatives up to order two. Suppose $\mathbf{X} = (X_1, \dots, X_N)$ is a semimartingale* *in* $\mathbb{R}^N$. *Then*

$$f(\mathbf{X}(t)) - f(\mathbf{X}(0)) = \sum_{i=1}^{N} \int_0^t \frac{\partial f}{\partial X_i} \mathbf{X}(\tau_-) \, \mathrm{d}\mathbf{X}(\tau)$$
$$+ \frac{1}{2} \sum_{i=1}^{N} \sum_{j=1}^{N} \int_0^t \frac{\partial^2 f}{\partial X_i X_j} \mathbf{X}(\tau_-) \, \mathrm{d}[(X_i)^c, (X_j)^c](\tau)$$
$$+ \sum_{0 \le \tau \le t} \left(f(\mathbf{X}(\tau)) - f(\mathbf{X}(\tau_-)) - \sum_{i=1}^{N} \frac{\partial f}{\partial X_i} \mathbf{X}(\tau_-) \triangle X_i(\tau) \right), \tag{C.14.2}$$

$(X_i)^c$ *denoting the continuous martingale part of the semimartingale* X_i.

Next are the well-known *Kunita-Watanabe inequalities* (see Elliott [14], Theorem 10.11 and Remark 10.13, Chapter 10).

Theorem C.14.3. *Let* $N, M \in \mathcal{M}_0^2((\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}}), \{\tilde{\mathcal{F}}_t\})$ *and suppose that* H *and* K *are measurable processes. Then*

$$\left(\int_0^\infty |H(s)K(s)| \, |\mathrm{d}\langle N, M\rangle(s)| \right)^2 \le \int_0^\infty H^2(s) \; \mathrm{d}\langle N\rangle(s) \; \int_0^\infty K^2(s) \; \mathrm{d}\langle M\rangle(s) \tag{C.14.3}$$

and

$$\left(\int_0^\infty |H(s)K(s)| \, |\mathrm{d}[N, M](s)| \right)^2 \le \int_0^\infty H^2(s) \; \mathrm{d}[N](s) \; \int_0^\infty K^2(s) \; \mathrm{d}[M](s). \tag{C.14.4}$$

Hence

$$E_{\tilde{\mathbb{P}}} \left(\int_0^\infty |H(s)K(s)| \, |\mathrm{d}\langle N, M\rangle(s)| \right) \le \left(E_{\tilde{\mathbb{P}}} \int_0^\infty H^2(s) \; \mathrm{d}\langle N\rangle(s) \right)^{\frac{1}{2}}$$
$$\cdot \left(E_{\tilde{\mathbb{P}}} \int_0^\infty K^2(s) \; \mathrm{d}\langle M\rangle(s) \right)^{\frac{1}{2}} \tag{C.14.5}$$

and

$$E_{\tilde{\mathbb{P}}} \left(\int_0^\infty |H(s)K(s)| \, |\mathrm{d}[N, M](s)| \right) \le \left(E_{\tilde{\mathbb{P}}} \int_0^\infty H^2(s) \; \mathrm{d}[N](s) \right)^{\frac{1}{2}}$$
$$\cdot \left(E_{\tilde{\mathbb{P}}} \int_0^\infty K^2(s) \; \mathrm{d}[M](s) \right)^{\frac{1}{2}}. \tag{C.14.6}$$

From Protter [42], Chapter II, Section 8, Theorem 40, page 87, we have *Lévy's Theorem*.

Theorem C.14.4. Lévy's Theorem *Let* $\mathbf{X} = (X_1, \ldots, X_m)$ *be an* m*-dimensional continuous local martingale such that*

$$[X_i, X_j](t) = \begin{cases} t & \text{if } i = j \\ 0 & \text{if } i \neq j. \end{cases} \tag{C.14.7}$$

Then $\mathbf{X}$ *is a standard* m*-dimensional Brownian motion.*

From Rogers and Williams [46], Theorem IV.27.6.iv, we have the following result (recall Definition C.4.1).

Theorem C.14.5. *Let* $M \in \mathcal{M}_0^2((\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}}), \{\tilde{\mathcal{F}}_t\})$*. Define the space*

$$L^2(M) := \left\{ H : \tilde{\Omega} \times [0, T] \to \mathbb{R} \,\middle|\, H \text{ is previsible and } E_{\tilde{\mathbb{P}}} \int_0^T H^2(\tau) \, \mathrm{d}[M](\tau) < \infty \right\}. \tag{C.14.8}$$

Then for $H \in L^2(M)$*, we have for all* $t \in [0, T]$*,*

$$\triangle \left(\int_0^t H(\tau) \, \mathrm{d}M(\tau) \right) = H(t) \triangle M(t), \qquad a.s. \tag{C.14.9}$$

C.15 Doléans-Dade exponential results

From Elliott [14], Chapter 13, Theorem 13.5 and Remark 13.6, we have

Theorem C.15.1. *Suppose* $X = \{X(t) : t \geq 0\}$ *is a semimartingale which is null at the origin. Let* X^c *denote its continuous martingale part. Then there is a unique semimartingale* $Z = \{Z(t) : t \geq 0\}$ *such that*

$$Z(t) = 1 + \int_0^t Z(\tau_-) \, \mathrm{d}X(\tau). \tag{C.15.1}$$

Furthermore, $Z(t)$ *is given by the expression*

$$Z(t) = \exp\left\{ X(t) - \frac{1}{2} \langle X^c, X^c \rangle(t) \right\} \prod_{\tau \in [0,t]} (1 + \triangle X(\tau)) \exp\{-\triangle X(\tau)\}, \tag{C.15.2}$$

for $t \geq 0$*, where the infinite product is absolutely convergent almost surely.*

Remark C.15.2. 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 (C.15.1), if X is a local martingale then $\mathcal{E}(X)$ is also a local martingale. Furthermore, from (C.15.2), $\mathcal{E}(X)$ 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 C.9.8, $[X, X] = \langle X, X \rangle$, and hence

$$Z(t) = \exp\left\{ X(t) - \frac{1}{2} [X, X](t) \right\} \quad \text{a.s.} \tag{C.15.3}$$

From Elliott [14], Chapter 13, Corollary 13.58, we also have the following result.

Corollary C.15.3. *If X and Y are semimartingales then*

$$\mathcal{E}(X)\mathcal{E}(Y) = \mathcal{E}(X + Y + [X, Y]). \tag{C.15.4}$$

From Protter [42], 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 C.15.4. Novikov's Criterion *Let M be a continuous local martingale and suppose that*

$$E\left(\exp\left\{\frac{1}{2}[M, M](\infty)\right\}\right) < \infty. \tag{C.15.5}$$

Then $\mathcal{E}(M)$ is a uniformly integrable martingale.

From Elliott [14], Theorem 13.19 we have the following form of *Girsanov's theorem.*

Theorem C.15.5. Girsanov's theorem *Suppose that the Doléans-Dade exponential $Z(t) = \mathcal{E}(X)(t)$ is a uniformly integrable positive martingale and that a new probability measure $\tilde{\mathbb{Q}}$ is defined on $(\tilde{\Omega}, \tilde{\mathcal{F}})$ by putting (the Radon-Nikodým derivative)*

$$\frac{\mathrm{d}\tilde{\mathbb{Q}}}{\mathrm{d}\tilde{\mathbb{P}}} = Z(\infty). \tag{C.15.6}$$

If N is a local martingale under measure $\tilde{\mathbb{P}}$ and the process $\langle N, X\rangle$ exists under $\tilde{\mathbb{P}}$, then

$$\tilde{N}(t) = N(t) - \langle N, X\rangle(t) \tag{C.15.7}$$

is a local martingale under $\tilde{\mathbb{Q}}$.

C.16 Martingale results

From Protter [42], Chapter II, Section 6, Corollary 3, page 73, we have the following corollary which gives conditions for a local martingale to be a martingale.

Corollary C.16.1. *Let M be a local martingale. Then M is a martingale with $E|M(t)|^2 < \infty$ for all $t \geq 0$ if and only if $E[M, M](t) < \infty$ for all $t \geq 0$.*

If $E[M, M](t) < \infty$ then $E|M(t)|^2 = E[M, M](t) < \infty$.

From Protter [42], 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 C.16.2. *If M be a local martingale and $E[M,M](\infty) < \infty$, then M is an L^2-bounded martingale (that is $\sup_{t\geq 0} E|M(t)|^2 = E|M(\infty)|^2 < \infty$). Moreover*

$$E|M(t)|^2 = E[M,M](t) \tag{C.16.1}$$

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

Remark C.16.3. We refer to (C.16.1) as the *Itô isometry*.

C.17 Convex analysis

Definition C.17.1. Let $f : \mathbb{R}^m \to \mathbb{R} \cup \{\pm\infty\}$ be a function on V. The *convex conjugate* of f is a function $f^\star : \mathbb{R}^m \to \mathbb{R} \cup \{\pm\infty\}$ and defined by

$$f^\star(\mathbf{y}) := \sup_{\mathbf{x}\in\mathbb{R}^m} \{\mathbf{x}^\top \mathbf{y} - f(\mathbf{x})\}, \qquad \forall \mathbf{y} \in \mathbb{R}^m. \tag{C.17.1}$$

Definition C.17.2. Let V be a real vector space. Let $f : V \to \mathbb{R} \cup \{\pm\infty\}$ be a function on V. Then f is said to be *convex* if for every x and y in V, we have

$$f(\epsilon x + (1-\epsilon)y) \leq \epsilon f(x) + (1-\epsilon) f(y), \qquad \forall \epsilon \in [0,1]. \tag{C.17.2}$$

Definition C.17.3. Let V be a real normed vector space. Let $f : V \to \mathbb{R} \cup \{\pm\infty\}$ be a function on V. Then f is said to be *lower semi-continuous* (l.s.c.) at $x \in V$ if for every sequence $x^{(m)}$ which converges in the norm to x, we have

$$f(x) \leq \liminf_{m\to\infty} f(x^{(m)}). \tag{C.17.3}$$

The following theorem is a summary of the results we require from Aubin [1].

Theorem C.17.4. *Let V be a vector space. Suppose that for some fixed $d \in \mathbb{R}$, we have the following conditions:*

1. *A is a convex subset of V;*
2. *$f : A \to \mathbb{R}$ is a convex function;*
3. *$g : V \to \mathbb{R}$ is a linear operator; and*
4. *d is in the interior of $\{g(\boldsymbol{\pi}) : \boldsymbol{\pi} \in A\}$.*

Define the Lagrangian

$$L(\mu; \boldsymbol{\pi}) := f(\boldsymbol{\pi}) + \mu(g(\boldsymbol{\pi}) - d), \qquad \forall \mu \in \mathbb{R} \qquad \forall \boldsymbol{\pi} \in A. \tag{C.17.4}$$

Then there exists a Lagrange multiplier $\bar{\mu} \in \mathbb{R}$ *such that*

$$\inf_{\boldsymbol{\pi}\in A} L(\bar{\mu}; \boldsymbol{\pi}) = \sup_{\mu\in\mathbb{R}} \inf_{\boldsymbol{\pi}\in A} L(\mu; \boldsymbol{\pi}) = \inf_{\substack{\boldsymbol{\pi}\in A \\ g(\boldsymbol{\pi})=d}} f(\boldsymbol{\pi}). \tag{C.17.5}$$

Proof. From Conditions 1-3 and Aubin [1], Proposition 2.6.1, we always obtain

$$\sup_{\mu \in \mathbb{R}} L(\mu; \boldsymbol{\pi}) = \begin{cases} f(\boldsymbol{\pi}) & \text{if } \boldsymbol{\pi} \in A \text{ and } g(\boldsymbol{\pi}) = d \\ \infty & \text{otherwise.} \end{cases} \tag{C.17.6}$$

Consequently,

$$\sup_{\mu \in \mathbb{R}} \inf_{\boldsymbol{\pi} \in A} L(\mu; \boldsymbol{\pi}) \leq \inf_{\substack{\boldsymbol{\pi} \in A \\ g(\boldsymbol{\pi}) = d}} f(\boldsymbol{\pi}). \tag{C.17.7}$$

From Conditions 1-4 and Aubin [1], Theorem 2.6.1, there exists $\bar{\mu} \in \mathbb{R}$ such that

$$\inf_{\boldsymbol{\pi} \in A} L(\bar{\mu}; \boldsymbol{\pi}) = \inf_{\substack{\boldsymbol{\pi} \in A \\ g(\boldsymbol{\pi}) = d}} f(\boldsymbol{\pi}). \tag{C.17.8}$$

Since it is also true that

$$\inf_{\boldsymbol{\pi} \in A} L(\bar{\mu}; \boldsymbol{\pi}) \leq \sup_{\mu \in \mathbb{R}} \inf_{\boldsymbol{\pi} \in A} L(\mu; \boldsymbol{\pi}), \tag{C.17.9}$$

combining (C.17.7), (C.17.8) and (C.17.8), we get

$$\sup_{\mu \in \mathbb{R}} \inf_{\boldsymbol{\pi} \in A} L(\mu; \boldsymbol{\pi}) \leq \inf_{\substack{\boldsymbol{\pi} \in A \\ g(\boldsymbol{\pi}) = d}} f(\boldsymbol{\pi}) = \inf_{\boldsymbol{\pi} \in A} L(\bar{\mu}; \boldsymbol{\pi}) \leq \sup_{\mu \in \mathbb{R}} \inf_{\boldsymbol{\pi} \in A} L(\mu; \boldsymbol{\pi}). \tag{C.17.10}$$

Hence we must have equality and the theorem is proved. □

Index and Glossary of Symbols

References

[1] Jean-Pierre Aubin. *Applied Functional Analysis*. Wiley-Interscience, USA, second edition, 2000. 195, 196

[2] Louis Bachelier. Théorie de la spéculation. *Annales scientifique de l'É.N.S.*, 17:21–86, 1900. 5

[3] Jean-Michel Bismut. Conjugate convex functions in optimal stochastic control. *Journal of Mathematical Analysis and Applications*, 44(2):384–404, November 1973. 12, 39, 47, 48, 114, 115

[4] Jean-Michel Bismut. Growth and optimal intertemporal allocations of risk. *Journal of Economic Theory*, 10:239–287, 1975. 12, 14

[5] Fischer Black. Studies of stock price volatility changes. *American Statistical Society, Proceedings of the Business and Economics Statistics Section*, pages 177–181, 1976. 6

[6] René Boel, Pravin Varaiya, and Eugene Wong. Martingales on jump processes. I Representation results. *SIAM Journal of Control*, 13(5):999–1021, 1975. 124, 141

[7] Yuan Shih Chow and Henry Teicher. *Probability theory: independence, interchangeability, martingales*. Springer, USA, third edition, 1997. 183

[8] Paul H. Cootner, editor. *The Random Character of Stock Market Prices*, pages 17–78. MIT Press, 1964. 5

[9] John C. Cox and Chi-fu Huang. Optimal consumption and portfolio policies when assets follow a diffusion process. *Journal of Economic Theory*, 49:33–83, 1989. 13

[10] John C. Cox and Chi-fu Huang. A variational problem arising in financial economics. *Journal of Mathematical Economics*, 20:465–487, 1991. 13

[11] Jakša Cvitanić and Ioannis Karatzas. Convex duality in constrained portfolio optimization. *The Annals of Applied Probability*, 2(4):767–818, 1992. 13, 14

[12] Darrell Duffie and Henry R. Richardson. Mean-variance hedging in continuous time. *The Annals of Applied Probability*, 1(1):1–15, January 1991. 9

[13] Ivar Ekeland and Roger Témam. *Convex Analysis and Variational Problems*. American Elsevier Publishing Company, Inc, New York, USA, 1976. 40, 59, 60

[14] Robert J. Elliott. *Stochastic Calculus and Applications*. Springer-Verlag, New York, USA, 1982. 183, 192, 193, 194

[15] Stewart N. Ethier and Thomas G. Kurtz. *Markov processes*. John Wiley and Sons, USA, 1986. 123, 125, 127, 136, 139

[16] Avner Friedman. *Foundations of Modern Analysis*. Dover, NY, USA, 1982. 182, 183

[17] Stephen F. Gray. Modeling the conditional distribution of interest rates as a regime-switching process. *Journal of Financial Economics*, 42(1):27–62, 1996. 6

[18] J. Michael Harrison and David M. Kreps. Martingales and arbitrage in multiperiod security markets. *Journal of Economic Theory*, 20:381–408, 1979. 6, 12

[19] J. Michael Harrison and Stanley R. Pliska. Martingales and stochastic integrals in the theory of continuous trading. *Stochastic Processes and Applications*, 11:215–260, 1981. 6, 12

[20] J. Michael Harrison and Stanley R. Pliska. A stochastic calculus model of continuous trading: complete markets. *Stochastic Processes and Applications*, 15:313–316, 1983. 6, 12

[21] David G. Hobson and L. C. G. Rogers. Complete models with stochastic volatility. *Mathematical Finance*, 8(1):215–229, January 1998. 6

[22] Ying Hu and Xun Yu Zhou. Constrained stochastic LQ control with random coefficients and application to portfolio selection. *SIAM Journal on Control and Optimization*, 44(2):444–466, 2005. 9, 12, 29

[23] John Hull. *Options, Futures and other Derivative Securities*. Prentice Hall, Englewood Cliffs, NJ, 1993. 6

[24] Philip J. Hunt and Joanne E. Kennedy. *Financial Derivatives in Theory and Practice*. Wiley, 2004. 28

[25] Jean Jacod and Albert N. Shiryaev. *Limit Theorems for Stochastic Processes*. Springer-Verlag, Berlin, 1987. 108, 116, 123, 161, 177, 184, 187, 188, 189, 190, 191

[26] Arnaud Jobert and L. C. G. Rogers. Option pricing with Markov-modulated dynamics. *SIAM Journal on Control and Optimization*, 44(6):2063–2078, 2006. 6

[27] Madhu Kalimipalli and Raul Susmel. Regime-switching stochastic volatility and short-term interest rates. *Journal of Empirical Finance*, 11(3):309–329, June 2004. 6

[28] Ioannis Karatzas, John P. Lehoczky, and Steven E. Shreve. Optimal portfolio and consumption decisions for a small investor on a finite horizon. *SIAM Journal on Control and Optimization*, 25(6):1557–1586, 1987. 13, 30

[29] Ioannis Karatzas, John P. Lehoczky, Steven E. Shreve, and Gan-Lin Xu. Martingale and duality methods for utility maximization in an incomplete market. *SIAM Journal on Control and Optimization*, 29(3):702–730, 1991. 13

[30] Ioannis Karatzas and Steven E. Shreve. *Brownian Motion and Stochastic Calculus*. Springer-Verlag, New York, USA, 1991. 134

[31] Ioannis Karatzas and Steven E. Shreve. *Methods of Mathematical Finance*. Springer-Verlag, New York, USA, 1998. 7, 51, 181

[32] Chantal Labbé. *Contributions to the Theory of Constrained Portfolio Optimization*. PhD thesis, University of Waterloo, Canada, 2004. 117

[33] Chantal Labbé and Andrew J. Heunis. Convex duality in constrained mean-variance portfolio optimization. *Advances in Applied Probability*, 39(1):77–104, March 2007. 14, 43, 75, 98, 99, 107

[34] Xun Li, Xun Yu Zhou, and Andrew E. B. Lim. Dynamic mean-variance portfolio selection with no-shorting constraints. *SIAM Journal on Control and Optimization*, 40(5):1540–1555, 2002. 9, 12

[35] Andrew E. B. Lim and Xun Yu Zhou. Mean-variance portfolio selection with random parameters in a complete market. *Mathematics of Operation Research*, 27(1):101–120, February 2002. 9, 12

[36] Harry Markowitz. Portfolio selection. *Journal of Finance*, 7(1):77–91, March 1952. 1, 9

[37] Robert C. Merton. Lifetime portfolio selection under uncertainty: the continuous-time case. *The Review of Economics and Statistics*, 51(3):247–257, August 1969. 5, 8

[38] Robert C. Merton. Optimum consumption and portfolio rules in a continuous-time model. *Journal of Economic Theory*, 3(4):373–413, December 1971. 6, 12

[39] Paul A. Mossin. Optimal multiperiod portfolio policies. *Journal of Business*, 41:215–229, 1968. 8

[40] James R. Norris. *Markov chains*. Cambridge University Press, Cambridge, England, 1997. 125

[41] Stanley R. Pliska. A stochastic calculus model of continuous trading: optimal portfolios. *Mathematics of Operation Research*, 11(2):371–382, May 1986. 13

[42] Philip E. Protter. *Stochastic Integration and Differential Equations*. Springer-Verlag, New York, USA, second edition, 2005. 177, 179, 191, 192, 194

[43] Daniel Revuz and Marc Yor. *Continuous Martingales and Brownian Motion*. Springer-Verlag, Germany, second edition, 1991. 116

[44] L. C. G. Rogers. Duality in constrained optimal investment and consumption problems: a synthesis. *Paris-Princeton Lectures on Mathematical Finance*, pages 95–131, 2002. 14

[45] L. C. G. Rogers and David Williams. *Diffusions, Markov Processes and Martingales, Volume I Foundations*. Cambridge University Press, 2000. 182

[46] L. C. G. Rogers and David Williams. *Diffusions, Markov Processes and Martingales, Volume II Ito Calculus*. Cambridge University Press, 2000. 109, 124, 125, 129, 132, 141, 142, 143, 144, 153, 156, 173, 187, 190, 191, 193

[47] Paul A. Samuelson. Lifetime portfolio selection by dynamic stochastic programming. *The Review of Economics and Statistics*, 51(3):239–246, August 1969. 8

[48] Martin Schweizer. Mean-variance hedging for general claims. *The Annals of Applied Probability*, 2(1):171–179, February 1992. 9, 10

[49] Steven E. Shreve and Gan-Lin Xu. A duality method for optimal consumption and investment under short-selling prohibition: I, General market coefficients; and II, Constant market coefficients. *The Annals of Applied Probability*, 2:87–112 and 314–328, 1992. 13

[50] Richard H. Stockbridge. Portfolio optimization in markets having stochastic rates. *Stochastic Theory and Control*, 280:447–458, 2002. 7

[51] Eugene Wong and Bruce Hajek. *Stochastic Processes in Engineering Systems*. Springer, Berlin, 1985. 123, 124, 150, 156

[52] Jiongmin Yong and Xun Yu Zhou. *Stochastic Controls: Hamiltonian Systems and HJB Equations*. Springer, 1999. 11

[53] Xun Yu Zhou and Gang George Yin. Markowitz's mean-variance portfolio selection with regime switching: A continuous-time model. *SIAM Journal on Control and Optimization*, 42(4):1466–1482, 2003. 2, 3, 7, 9, 16, 17, 18, 19, 22, 27, 28, 29, 35, 41, 108