OPTIMAL PREDICTION GAMES IN LOCAL ELECTRICITY MARKETS

A thesis submitted to the University of Manchester for the degree of Doctor of Philosophy in the Faculty of Engineering and Physical Sciences

2015

Randall Martyr
School of Mathematics
Contents

Abstract 12

Declaration 13

Copyright Statement 14

The Author 15

Acknowledgements 16

1 Introduction 17

1.1 Research motivation ........................................ 17
   1.1.1 Local electricity markets, stochastic control and games .... 17
   1.1.2 Industrial partnership with Tempus Energy ................. 18
1.2 The British electricity market .................................. 20
   1.2.1 An overview of the market .................................. 20
   1.2.2 Mitigation of market risks for electricity suppliers ... 21
   1.2.3 Mitigation of market risks for renewable generators ... 23
1.3 Summary of topics ............................................ 25
   1.3.1 Chapter 3: Deterministic optimal control ................. 26
   1.3.2 Chapters 4 and 5: Optimal switching ....................... 27
   1.3.3 Chapter 6: Optimal stopping games and optimal switching 28
   1.3.4 Chapter 7: Non-zero sum stochastic games of classic and impulse control ....................................... 29

2 Literature Review 31

2.1 Review for Chapter 3 ........................................ 31
## Optimal control of electric space heating

### 3.1 Introduction

### 3.2 The model

#### 3.2.1 Notation and preliminaries

#### 3.2.2 Indoor temperature evolution

#### 3.2.3 Admissible controls

#### 3.2.4 Performance criterion

### 3.3 Assumptions

### 3.4 Existence of minimising controls

### 3.5 Necessary conditions for optimality

#### 3.5.1 Statement of Pontryagin’s Maximum Principle
3.5.2 Application of the Maximum Principle ......................... 78
3.6 The method of dynamic programming .............................. 79
3.6.1 Classical results ........................................... 79
3.6.2 Viscosity solutions to Hamilton-Jacobi-Bellman PDEs ......... 82
3.6.3 Necessary and sufficient conditions for optimality ............. 84
3.6.4 Stochastic control and viscosity approximation of the HJB equa-

tion ............................................................ 85
3.7 Numerical simulations ............................................ 87
3.7.1 Electricity spot price data .................................... 87
3.7.2 Thermal penalty function parameters .......................... 90
3.7.3 Simulation procedure ....................................... 90
3.7.4 Simulation results .......................................... 92
3.8 Conclusion ...................................................... 99

4 Optimal switching of electric space heating 101
4.1 Introduction ...................................................... 101
4.1.1 Optimal control with switching costs ......................... 101
4.2 Discrete-time optimal switching .................................. 103
4.2.1 Definitions .................................................. 103
4.2.2 The optimal switching problem ............................. 105
4.2.3 Notation, conventions and assumptions ..................... 106
4.3 Optimal stopping and Snell envelopes ............................. 107
4.4 The verification theorem ....................................... 107
4.5 Existence of the optimal processes via backward induction ..... 112
4.6 Numerical example ............................................. 118
4.6.1 Impact of switching costs on the deterministic control policy .... 118
4.6.2 Simulation results with stochastic prices .................... 122
4.7 Conclusion ...................................................... 129

5 Continuous-time optimal switching with signed switching costs 130
5.1 Introduction ...................................................... 130
5.2 Definitions ..................................................... 131
5.2.1 Probabilistic setup ........................................ 131
List of Tables

3.1 Price profile categories according to standard deviation ............... 88
3.2 Default parameter values used to obtain approximate dynamic program-
ing control policies. ..................................................... 91

4.1 Parameter values used to simulate the performance of approximate dy-
namic programming control policies. ........................................ 120
4.2 Default parameter values used to generate the autoregressive error terms
and random spot prices. ......................................................... 124

6.1 Table of parameter values used in numerical solution of the Dynkin game.172

E.1 Occurrence of different categories of price variability from 2005 to 2014
based on the month and type of day. The categories are defined in
Table 3.1. ................................................................. 257
## List of Figures

1.1 A simplified view of Tempus Energy’s Smart Aggregation Manager™ . 19

2.1 Example of a bathtub penalty function for indoor temperature control . 33

3.1 Price variability trend and distribution from 2005 to 2014. . . . . . . . 88
3.2 Median price profile (2005–2014) for each category of price variability. . 89
3.3 The impact of price variability on the percentage financial cost savings. 93
3.4 Controlled trajectories resulting from the dynamic programming strategies. . 94
3.5 The impact of deadband size on financial cost savings. . . . . . . . . . . 95
3.6 Controlled trajectories associated with Profile III for various deadband sizes. . 95
3.7 The impact of the thermal time constant on cost savings. . . . . . . . . 96
3.8 Impact of the thermal gain parameter $T_g$ on the estimated financial cost savings (%). 97
3.9 Controlled trajectories associated with Profile III for different values of the thermal gain $T_g$. 98
3.10 Impact of the ambient temperature parameter $T_a$ on the estimated financial cost savings (%) of approximated optimal control strategies . 98
3.11 Controlled trajectories associated with Profile III for different values of the ambient temperature $T_a$. 99
4.1 Impact of the switching cost value on the performance of the control strategies. 120
4.2 Controlled indoor temperature trajectories under dynamic programming policies with different switching costs. The green step chart shows the MIP in £ per MWh. ............................................ 121
4.3 True and perturbed prices for different price profiles ........................................ 124
4.4 Summary of stochastic optimal control simulation results. ................................. 125
4.5 Indoor temperature trajectories under the deterministic (blue line) and worst performing (red dashed line) stochastic control strategies for different price profiles ........................................ 126
4.6 Impact of increasing uncertainty on the financial cost savings (%). ................. 127
4.7 Impact of increasing uncertainty on the performance of the path-dependent control policies. ........................................ 128
6.1 Comparison of free-boundary and Least-Squares Monte Carlo solutions to the Dynkin game. ........................................ 172
A.1 Dynamic Programming Grid ........................................ 231
Commonly used terms and notation

Abbreviations

BO    battery operator, page 29
BSDE  backward stochastic differential equation, page 40
CCfD  cancellable contract for difference, page 24
CID   Contract for Difference, page 23
DP    dynamic programming, page 26
DPP   dynamic programming principle, page 37
EMR   Electricity Market Reform, page 20
GB    Great Britain, page 20
LCCC  Low Carbon Contracts Company, page 23
MIP   Market Index Price, page 87
PDE   partial differential equation, page 30
PMP   Pontrygain’s Maximum Principle, page 26
SBM   standard Brownian motion, page 49
SDE   stochastic differential equation, page 86
SO    system operator, page 29

Symbols

\( C_b(A; E) \)  the space of bounded continuous functions from \( A \) to \( E \), page 52
Local electricity markets can be defined broadly as “future electricity market designs involving domestic customers, demand-side response and energy storage”. Like current deregulated electricity markets, these localised derivations present specific stochastic optimisation problems in which the dynamic and random nature of the market is intertwined with the physical needs of its participants. Moreover, the types of contracts and constraints in this setting are such that “games” naturally emerge between the agents. Advanced modelling techniques beyond classical mathematical finance are therefore key to their analysis. This thesis aims to study contracts in these local electricity markets using the mathematical theories of stochastic optimal control and games.

Chapter 1 motivates the research, provides an overview of the electricity market in Great Britain, and summarises the content of this thesis. It introduces three problems which are studied later in the thesis: a simple control problem involving demand-side management for domestic customers, and two examples of games within local electricity markets, one of them involving energy storage. Chapter 2 then reviews the literature most relevant to the topics discussed in this work.

Chapter 3 investigates how electric space heating loads can be made responsive to time varying prices in an electricity spot market. The problem is formulated mathematically within the framework of deterministic optimal control, and is analysed using methods such as Pontryagin’s Maximum Principle and Dynamic Programming. Numerical simulations are provided to illustrate how the control strategies perform on real market data.

The problem of Chapter 3 is reformulated in Chapter 4 as one of optimal switching in discrete-time. A martingale approach is used to establish the existence of an optimal strategy in a very general setup, and also provides an algorithm for computing the value function and the optimal strategy. The theory is exemplified by a numerical example for the motivating problem. Chapter 5 then continues the study of finite horizon optimal switching problems, but in continuous time. It also uses martingale methods to prove the existence of an optimal strategy in a fairly general model.

Chapter 6 introduces a mathematical model for a game contingent claim between an electricity supplier and generator described in the introduction. A theory for using optimal switching to solve such games is developed and subsequently evidenced by a numerical example. An optimal switching formulation of the aforementioned game contingent claim is provided for an abstract Markovian model of the electricity market.

The final chapter studies a balancing services contract between an electricity transmission system operator (SO) and the owner of an electric energy storage device (battery operator or BO). The objectives of the SO and BO are combined in a non-zero sum stochastic differential game where one player (BO) uses a classic control with continuous effects, whereas the other player (SO) uses an impulse control (discontinuous effects). A verification theorem proving the existence of Nash equilibria in this game is obtained by recursion on the solutions to Hamilton-Jacobi-Bellman variational PDEs associated with non-zero sum controller-stopper games.
Declaration

No portion of the work referred to in the thesis has been submitted in support of an application for another degree or qualification of this or any other university or other institute of learning.
Copyright Statement

i. The author of this thesis (including any appendices and/or schedules to this thesis) owns certain copyright or related rights in it (the “Copyright”) and s/he has given The University of Manchester certain rights to use such Copyright, including for administrative purposes.

ii. Copies of this thesis, either in full or in extracts and whether in hard or electronic copy, may be made only in accordance with the Copyright, Designs and Patents Act 1988 (as amended) and regulations issued under it or, where appropriate, in accordance with licensing agreements which the University has from time to time. This page must form part of any such copies made.

iii. The ownership of certain Copyright, patents, designs, trade marks and other intellectual property (the “Intellectual Property”) and any reproductions of copyright works in the thesis, for example graphs and tables (“Reproductions”), which may be described in this thesis, may not be owned by the author and may be owned by third parties. Such Intellectual Property and Reproductions cannot and must not be made available for use without the prior written permission of the owner(s) of the relevant Intellectual Property and/or Reproductions.

iv. Further information on the conditions under which disclosure, publication and commercialisation of this thesis, the Copyright and any Intellectual Property and/or Reproductions described in it may take place is available in the University IP Policy (see http://documents.manchester.ac.uk/DocuInfo.aspx?DocID=487), in any relevant Thesis restriction declarations deposited in the University Library, The University Library’s regulations (see http://www.manchester.ac.uk/library/aboutus/regulations) and in The University’s Policy on Presentation of Theses.
The Author

The author of this thesis holds a Bachelor of Science degree in Mathematics and Information Technology (Hons) from the University of the West Indies at the Cave Hill Campus in Barbados, and a Master of Science degree in Mathematical Finance from the University of Manchester.
Acknowledgements

The author would like to thank his supervisor Dr. John Moriarty for his patience and helpful comments. The staff at Tempus Energy Ltd., especially Ms. Sara Bell, must be acknowledged for their support of this research. He acknowledges his family members, especially his parents, for their unwavering love, support and guidance. The author expresses gratitude for all whom he engaged in fruitful discussions at various stages throughout this thesis – particularly Dr. Tiziano De Angelis, Professor Goran Peskir and Dr. Saralees Nadarajah. Above all, he would like to thank Almighty God for His wisdom and strength.
Chapter 1

Introduction

1.1 Research motivation

1.1.1 Local electricity markets, stochastic control and games

This thesis broadly defines local electricity markets as “future electricity market designs involving domestic customers, demand-side response and energy storage”. Like current deregulated electricity markets, these localised derivations present specific stochastic optimisation problems in which the dynamic and random nature of the market is intertwined with the physical needs of its participants. On the one hand, related problems in existing energy markets have already prompted the development of advanced mathematics for their analysis in a burgeoning field known as Quantitative Energy Finance [14]. On the other hand, the types of contracts and constraints in local electricity markets are such that “games” naturally emerge between the agents. Furthermore, the decisions of agents in these markets are inevitably shaped by the present and future states of the underlying electricity market, which is inherently random. This leads to the concept of optimal prediction games in which two or more decision makers with specified objectives can affect the future random payoffs of all those involved.

The concept of “games” appearing in the previous paragraph should be formalised before proceeding. Game theory is the study of the processes which arise when

... an individual pursues an objective(s) in a situation in which other
individuals concurrently pursue other (possibly conflicting, possibly overlapping) objectives. [137, p. 1]

*Non-cooperative* games are those in which conflict can arise due to each player pursuing their own interest, and a game is said to be *zero-sum* if the players’ objectives cancel each other out at a total level. According to [12], *dynamic game theory* merges the fields of classical game theory and *optimal control*, which is the study of certain dynamic optimisation problems. A *stochastic game* is a dynamic game incorporating “chance moves” which influence the state of the game through the random actions of “Nature” or the “Chance Player”, determined by a priori known statistics; informally they are the natural multi-player extensions of stochastic optimal control (or *stochastic control*) problems.

Based on the above descriptions, the optimal prediction games arising in local electricity markets are likely to be at least dynamic and stochastic. Stochastic optimal control is therefore an important tool for analysing them. On the other hand, a basic understanding of the underlying electricity market is also essential for formulating the appropriate stochastic control problems. This thesis helps address these concerns through simple examples based on the electricity market in Great Britain.

### 1.1.2 Industrial partnership with Tempus Energy

This research can also have a strong impact if it supports the creation of practical algorithms that enable householders and small to medium-sized enterprises to participate fully in local electricity markets. The first steps towards this goal have been achieved through an industrial partnership with Tempus Energy, a startup company that partially sponsors the research.

Tempus Energy seeks to remove artificial price barriers from the electricity supply business, bringing transparency and connecting customers with the cheapest available energy. The author assisted the company in their mission by developing predictive algorithms that can shape electricity demand according to the availability of renewable and cheap wholesale energy. Stochastic optimal control played a key role in the development of those algorithms. The partnership between Tempus Energy and the University of Manchester therefore illustrates a niche area where academe and industry
It has also been fruitful in the following ways:

- Tempus Energy was established as an algorithms provider (Alectrona Grid Services Ltd.) in 2012 with a three year research agreement with the University of Manchester (1 PhD student, Randall Martyr).

- The University of Manchester has developed stochastic optimisation techniques underpinning the Smart Aggregation Manager™, the Tempus Energy IT platform enabling a new electricity supply business model based on dynamic demand response (see Figure 1.1 below).

- The company subsequently secured private investment, a £250,000 TSB Smart Grant, and debt funding from the Greater London Authority Growing Places Fund.

- Tempus Energy now employs at least five people and is preparing to launch a UK electricity supply business in 2015.

- More details are available at the website: http://www.tempusenergy.com/

Figure 1.1: A simplified view of Tempus Energy’s Smart Aggregation Manager™.
1.2 The British electricity market

1.2.1 An overview of the market

Great Britain (GB) is home to the world’s first bilateral physical market for electricity, introduced in 2000/1 as the “New Electricity Trading Arrangements” for England and Wales. The market extended its services to Scotland under the name “British Electricity Trading Transmission Arrangements” in 2005, and now has distinct markets for retail, wholesale, imbalance and ancillary services. These markets, which shall be referred to collectively as the British electricity market, are described briefly in the following paragraph.

The retail market concerns the sale and purchase of electricity between consumers and retailers of electricity. The wholesale market is mainly concerned with the sale and purchase of electricity between suppliers and generators. If a wholesale market participant consumes (generates) a different amount of electricity than it is contracted to consume (generate) during a specified time period, it is deemed to have its position corrected during this period by trading in the imbalance market. Finally, participants in the market for ancillary services offer resources which can be procured by the transmission system operator to ensure the security and quality of supply on the power system. The reader may consult the guide [45] for more information on the particular arrangement in Great Britain, and the textbooks [72, 82] for thorough studies on power system economics and the structure of electricity markets.

Ongoing market reforms.

The British electricity market is in the midst of significant change affecting both its structure and the rules for its participants. The most notable among the policies driving this change are the Electricity Balancing Significant Code Review (EBSCR) and Electricity Market Reform (EMR) programme.

The EBSCR, which commenced in August 2011 and was completed in May 2014, proposes adjustments to the pricing structure of the imbalance market in order to increase its efficacy and efficiency. These adjustments will lead to “sharper” imbalance prices, providing an increased incentive for UK electricity suppliers to maintain balance in their demand-supply portfolio [107]. The EMR programme, which was drafted in
2011 and subsequently ratified as the Energy Act 2013, introduces market reforms promoting the delivery of reliable and affordable low carbon energy. The programme, which is comprehensively summarised in [35], has two main components:

- The introduction of a capacity market, which is designed to ensure security of the electricity network by rewarding reliable sources of capacity on both supply and demand sides, especially during times of constrained generation on the system.

- The introduction of Contracts for Difference, which are bilateral financial arrangements between the UK government and certain electricity generators designed to encourage investment in low carbon generation. These are further discussed below in Section 1.2.3.

1.2.2 Mitigation of market risks for electricity suppliers

Consider an electricity supplier that purchases energy according to the spot price in the wholesale market, and sells this energy to its customers at a pre-specified retail price. It faces an imbalance cost if the electricity that it has committed to use (on behalf of its customers) is not the exact amount consumed on the power system. The supplier is therefore faced with price uncertainty in both the wholesale and imbalance markets (in addition to uncertainty in demand).

The typical supplier transfers most of its risk to its customers by choosing a retail price that, on average, is much higher than those in the wholesale market and can compensate for other charges such as imbalance and network (transmission and distribution) use. It can also offer contracts which help hedge its financial risk. For example, it can negotiate with its customers an appropriate time and due compensation for pre-designated reduction or complete curtailment of electricity consumption [72, pp. 90–93] (also see [6]).

Harnessing the power of demand-side management and energy storage.

In a traditional setting, an electricity supplier has minimal control over the consumption patterns of its customers. In recent times, however, power utilities have gained the ability to influence electricity usage at their consumers’ premises through real-time monitoring and control of heating, ventilation and air-conditioning (HVAC) systems,
lighting and other end-use devices. This demand-side management may be achieved either through direct control via a telecommunications system, or indirectly through incentives which can be economic, such as variable pricing tariffs, or environmental such as the level of carbon emissions or supply of renewable energy [119, p. 95]. Electricity consumption is typically shifted away from times of system stress or high energy costs, ideally with a minimum deprivation of core business operations and/or comfort.

In this thesis, an electrical energy storage unit is a buffer used principally or exclusively to counteract the power imbalance between supply and demand [126, p. 179]. A highly topical example is Tesla’s recently announced Powerwall, a rechargeable lithium-ion battery product for home use which stores electricity for domestic consumption, load shifting, and backup power [129]. Together with demand-side management, energy storage offers the potential to improve the British electricity market by

- helping suppliers to reduce their retail prices through avoidance of high wholesale and imbalance charges;
- reducing the number of balancing actions needed by the system operator; and
- increasing competition in the ancillary services market and thereby reducing the cost of these services.

To highlight the significance of the last two points, the National Audit Office [103] reports that National Grid, which is the transmission system operator for the British market, spent approximately £1 billion in procuring balancing services for the year 2013 to 2014. This cost is recovered through a Balancing Services Use of System (BSUoS) levy on generators and suppliers, and is passed on to consumers through their supplier’s retail tariff.

Whilst the potential benefits of demand-side management and energy storage have been recognised by major stakeholders in the British electricity market, the means of achieving them in practice are not as well understood [48]. However, the descriptions provided earlier suggest that they fall within the scope of the local electricity markets described in this thesis. Indeed, Saad et al [119] note that

...the essence of demand-side management revolves around the interactions between various entities with specific objectives that are reminiscent
of the players’ interactions in game theory (p. 95). Demand-side management is perhaps the most natural setting for applying game theory due to the need for combining economic aspects such as pricing with strategic decision making by the various involved entities (p. 99).

Problems involving demand-side management and energy storage may therefore be investigated using the dynamic stochastic analysis proposed for local electricity markets. An example related to demand-side management can be found in Chapter 3 (which is continued in Chapter 4), whilst another involving energy storage is studied in Chapter 7.

1.2.3 Mitigation of market risks for renewable generators

Contracts for Difference.

In the context of Electricity Market Reform, a contract for difference (CfD) is a bilateral arrangement designed to encourage investment in low carbon generation. The contract is struck between a CfD Counterparty, which is the UK government-owned Low Carbon Contracts Company (LCCC), and a generator whose eligibility is determined according to a list of renewable technologies and associated fuels [35]. The generator sells energy into the market as usual and is paid the difference between a pre-agreed strike price and a reference price for the electricity market (the market price). The cost of the contract to the LCCC is funded by consumers through the supplier obligation levy on electricity suppliers [35, p. 57], which essentially makes the LCCC a supplier (the aggregate of all electricity suppliers). The terms and conditions document for CfD contracts [34] highlights other important features of the arrangement:

- A CfD is typically supported for up to 15 years;
- The strike price is adjusted annually for reasons such as inflation (six types of strike price adjustment are listed in [34, p. 49]);
- The LCCC has the right, but not the obligation, to terminate the CfD after the start date (and before the 15 year period) if a “termination event” such as the following occurs:
– the generator becomes insolvent (unable to pay its debts);
– the generator fails to meet a required payment;
– the generator does not deliver the agreed capacity;

• In addition to any cash-flows owed according to the contract, the generator is also liable to a fee (the “Termination Amount”) at the termination date if the LCCC chooses to exercise its termination right.

The strike price is meant to be reflective of the investment cost related to a particular low carbon technology. It will be set administratively by the government during the first few years of the reformed market, a process that will be superseded by one of price discovery through competition by tender or auction [33, pp. 20, 34]. In either case, the price-setting process is very important for the CfD to have the impact that the government intends. If the strike price is set too low then the contract may be unattractive for (new) investment in low carbon generation. On the other hand, if the strike price is set too high the generator benefits while the consumers bare the burden of an unfavourable financial arrangement.

The UK government envisages that the CfD will help achieve the targets of EMR by

i) reducing the generator’s exposure to fluctuating electricity prices, and

ii) protecting consumers from over-payment when the market price exceeds the strike price (since the generator pays back the difference).

However, barring a significant adjustment to the strike price or a carefully drafted termination clause, the long-term maturity of the CfD can create a serious financial burden for either the consumer or the generator. This is the reality in Germany where the renewable energy levy on the consumer’s electricity bill has risen due to a combination of low market prices and high price guarantees for renewable generators. Ironically, one of the main contributors to lower market prices has been the high penetration of renewable generation in the electricity grid, fuelled in part by the price guarantee for those generators [86, 109, 132]. In the author’s opinion, a similar problem in the British electricity market may be circumvented by adopting a cancellable contract for difference (CCfD).
Cancellable contracts for difference.

The proposed CCfD has similar features to the usual CfD but grants both parties the option to terminate the contract early by paying a termination fee. For instance, the generator may choose to exit the contract early if it favours selling its generation on the spot market. Alternatively, the CCfD Counterparty (a supplier) might want to exercise its termination right if the market price is expected to stay below the strike level for a prolonged period. The two-sided exit feature therefore helps alleviate the problem of committing to an unfavourable strike price for a very long period of time.

The CCfD is easy to describe, but its analysis is not a trivial exercise. This is because the timing decisions and random payoffs of both players need to be accounted for. It is therefore an interesting non-trivial example of a stochastic game which can arise naturally in a local electricity market. Chapter 6 below provides a more detailed mathematical description of the contract complete with preliminary results related to its pricing.

1.3 Summary of topics

This thesis is motivated by optimisation problems arising in current and anticipated designs for electricity markets, with particular emphasis on the British market. The main techniques used to study these problems come from the field of optimal control.

Optimal control is a mathematical discipline which emerged in the 1950s through motivating problems in engineering, but has roots extending much further to the classic field of calculus of variations [57, p. 20]. It has seen immense growth since then, stimulated mainly by its numerous applications in other fields such as economics, finance and management sciences. Simultaneously, there has been an explosion in the types of control problems being studied. The references [7, 58, 108] may be consulted for an extensive (yet non-exhaustive) catalogue of the different types of optimal control problems arising in deterministic and stochastic settings.

The optimisation problems in this thesis fall within two main classes of optimal control problems: the classic problem with continuous controls on the one hand, and impulse control problems (a generalisation of optimal stopping) on the other hand. The rest of this section summarises the contents of the thesis according to the chapter
and type of control problem studied. A literature review of the research is provided in Chapter 2.

1.3.1 Chapter 3: Deterministic optimal control

Chapter 3 looks at a particular optimisation problem for a consumer who is directly exposed to electricity spot market prices. This consumer would like to minimise the cost of his electric heater’s (or air conditioner’s) consumption, while simultaneously accounting for (and avoiding) significant deviations of the household’s indoor temperature from an ideal one. A closely related problem in discrete time was studied previously in [27] and was solved using non-linear programming methods. It is formulated alternatively in Chapter 3 as one of deterministic optimal control in continuous time, then solved using Pontryagin’s Maximum Principle (PMP) and Dynamic Programming (DP).

The model of Chapter 3 uses a performance criterion and evolutionary equation that are linear in the control variable, and a bang-bang optimal strategy is consequently derived. This means the electric heater operates optimally in this particular model if it take values at the boundaries of the control set (“on” and “off”) according to a particular mathematical rule. Numerical simulations performed with British electricity market data showed that, relative to a given benchmark, significant cost savings can be achieved while simultaneously keeping the indoor temperature within reasonable limits of the ideal one. Section 3.7.4 below has more details and also highlights the non-trivial nature of the controlled indoor temperature trajectories. Apart from the constraints due to customer preferences, these trajectories appear to be influenced by (at least) the following:

- the location of both local and global price maxima;
- the relative size of price maxima – their absolute values are less relevant due to price normalisation;
- the immediate trend in price movements;
- the rate at which the indoor temperature increases and decreases.
Analogous results can (in principle) be obtained for refrigeration equipment and other thermostatically controlled devices.

### 1.3.2 Chapters 4 and 5: Optimal switching

In a single agent optimal switching problem, a decision maker controls a dynamical system over time by applying a sequence of control modes. These modes are chosen from a discrete set so as to maximise a given objective function. Optimal switching models often arise in the analysis of flexible industrial projects (Real Options Analysis), with examples related to investment in electricity generation [2, 20, 69] and valuation of energy storage [21]. The fourth and fifth chapters study optimal switching problems with multiple modes (two or more) in both discrete and continuous time respectively.

**Discrete-time optimal switching.**

Chapter 4 presents results on discrete-time optimal switching that the author has also submitted for peer review (a preprint version is available at [96]). The solution to a discrete-time optimal switching problem appears frequently in the literature as an approximate solution to a continuous-time optimal switching problem. A backward dynamic programming algorithm is then typically associated with the value function of the discretised optimal switching problem (see [2, 20, 61] for instance). There are, however, very few theoretical results on the solution to the discrete time optimal switching problem *without* an associated continuous-time problem – particularly the derivation of the dynamic programming algorithm. Chapter 4 fills this gap by providing the necessary theory.

The optimisation problem of Chapter 3.1, referred to as \( P_1 \) for the remainder of this section, is similar to an optimal switching problem with zero switching costs due to its bang-bang optimal control solution. While this type of strategy may be optimal, frequent consecutive switches of the control can directly damage electrical equipment and, on a larger scale,

- “induce unstable oscillations in electricity demand” [87, p. 4];
- cause stability problems on the power grid.
This frequent switching behaviour was noticed in the numerical examples of P1. In order to alleviate this problem, P1 was reformulated in Chapter 4 as one of two-mode optimal switching in discrete time.

The numerical results in Chapter 4 demonstrate that the switching behaviour of the algorithm’s control strategies is decreased in the presence of switching costs. Moreover, this reduction can be achieved using small constant values for the switching cost, thereby only slightly impacting the cost reduction benefits realised in Chapter 3. Even further, there was only a slight reduction in the expected achievable cost savings when the spot prices experienced particular “random” perturbations. These perturbations were such that the location and relative size of minima and maxima in the price profile were not significantly distorted. Remember it is only these relative values and not their absolute ones which are important (cf. Section 1.3.1). Thus, reasonable results may be produced by using day ahead market prices (particularly APX, [4]) or more reliable forecasts for the spot price if available.

Continuous-time optimal switching.

Chapter 5 studies the optimal switching problem in a continuous-time setting, presenting results that the author also recently submitted for peer review (a preprint version is available at [97]). In addition to being a natural extension of the discrete-time problem in the previous chapter, the author’s research on continuous-time optimal switching was partially motivated by a link to the stochastic games of timing which are explored subsequently in Chapter 6. The research outcome is an extension of previous results on optimal switching problems to accommodate signed (that is, both positive and negative) and discontinuous switching costs.

1.3.3 Chapter 6: Optimal stopping games and optimal switching

Chapter 6 of this thesis looks at a particular type of two-player zero-sum stochastic game. It is a game of timing where each player tries to optimise her payoff by deciding whether to stop and pay a (possibly random) cost or continue the game. This type of optimal stopping game is called a Dynkin game after Eugene Dynkin who introduced
CHAPTER 1. INTRODUCTION

them [81, p. 1]. The CCfD discussed in Section 1.2.3 is a particular example of this Dynkin game.

It was proved recently in [65] that a perpetual (infinite horizon) continuous-time Dynkin game can be solved by a two-mode optimal switching problem constructed appropriately from the game’s parameters (a relationship between the two in discrete time had already been discovered in [140]). The paper also mentions that a similar proof can be used for a game of finite duration [65, p. 435]. Chapter 6 provides the theory underlying the connection between optimal switching problems and Dynkin games in continuous-time and on a finite time horizon. Besides showing that the game has a value as in [65], the theory has been extended to prove the existence of a Nash equilibrium (saddle point) for the game (as done in [140]). This is accomplished by using the Snell envelope formulation of the optimal switching problem from Chapter 5. This author has also submitted these results for peer review and a pre-print version is available at [95].

The chapter also provides numerical evidence in support of the optimal switching formulation for the Dynkin game by comparing its solution to another one obtained by analytical means. The accuracy of the results and the potential for the optimal switching numerical method to treat high dimensional problems both strengthen the argument for using this approach to value the CCfD. An abstract Markovian model for the CCfD which satisfies the assumptions required to apply the optimal switching theory is then given.

1.3.4 Chapter 7: Non-zero sum stochastic games of classic and impulse control

Chapter 7 proposes the use of dynamic game theory to analyse balancing services contracts in electricity markets. A prototypical example of such a contract is given in which there are two parties: an electricity transmission system operator (SO) and the owner of an electric energy storage device (battery operator or BO). The system operator can choose up to a finite number of times for the BO to deliver electricity (up to a maximum amount). At each request time, the energy delivered by the BO is assumed to instantaneously affect the imbalance between demand and supply on
the power system. Between request times the BO charges the store by purchasing electricity from the spot market. It is assumed that the BO uses the store for no other purpose than for the contract, and charging the store has negligible effect on the system imbalance. The problem for the SO is to choose the appropriate delivery times and energy amounts that minimise the total system imbalance and balancing services cost over a prescribed finite time interval. Simultaneously, the battery operator would like to minimise the cost of its energy purchases and any additional penalties for partial delivery at the request times.

Chapter 7 shows how to combine the optimisation problems for the SO and BO into a non-zero sum stochastic differential game where one player (the BO) uses a classic (continuous) control and the other player (the SO) uses an impulse (discrete) control. A Nash equilibrium solution to this game is defined as a “fair” value for the balancing services contract. This is not unreasonable since game theory is already a popular tool for analysing interactions in electricity markets and power system balancing [54, 119], and the Nash equilibrium is the dominant solution concept employed. It should also be noted, however, that there can be multiple Nash equilibria and it is typically difficult to choose an appropriate one [119, p. 88].

The existence of Nash equilibria in the game is analysed via the Hamilton-Jacobi-Bellman variational partial differential equations (PDEs) associated with stopping and control problems for diffusions. It is shown that the solution to the game can be derived from the solutions to intermediary non-zero sum games of (classic) control and stopping (or classic controller-stopper games). The approach thus taken suggests that, in principle, the results may be extended to models which allow for jumps in the appropriate state variables. However, it is still conceptual since the existence and uniqueness of solutions to the PDEs derived for the diffusion model have not yet been studied. Nevertheless, the discussion points to a possible resolution for this issue.
Chapter 2

Literature Review

This chapter provides a chapter-by-chapter review of the literature that is most relevant to the research conducted in this thesis.

2.1 Review for Chapter 3

Chapter 3 looks at a particular continuous-time optimal control problem for a household with a time-varying pricing structure for electricity use. The consumer would like to minimise the cost of using an electric heater while keeping the household’s indoor temperature relatively close to an ideal one. Inspiration for the performance criterion and model used in Chapter 3 came from a paper by Constantopoulos, Schewepe and Larson [27], where an analogous discrete-time problem was solved using non-linear programming methods. There are papers such as [115, 136] that solved a closely related problem using techniques similar to those in this thesis. However, their models and criteria are different from the one in Chapter 3. Furthermore, as some of the conclusions in [115, 136] are consistent with those in [27], only the review of [27] is given below.
2.1.1 Connection to Constantopoulos, Schweppe and Larson (1991)

The model and performance criterion.

The paper by Constantopoulos, Schweppe and Larson [27] studied the optimal control of electric space conditioning equipment in order to minimise consumption costs in the presence of spot market electricity prices. The authors used a discrete-time model with time parameter set \( T = \{0, 1, \ldots, T\} \). For ease of exposition, the duration of a single control period is assumed here to be one unit of time. It is further assumed that the space conditioning equipment is an electric heater. The model [27, pp. 756] used for the evolution of the building’s indoor temperature, \((x_t)_{t \in \mathbb{T}}\), is a discrete-time analogue of the one used in Chapter 3:

\[
x_{t+1} = e^{-\frac{1}{\tau}} x_t + (1 - e^{-\frac{1}{\tau}}) (T^a_t + u_t T^g) , \quad t = 0, \ldots, T - 1
\]

where \( \tau \) is the thermal time constant (which is related to the insulation), \( T^g \) is the maximum thermal gain of the space conditioning equipment, \((T^a_t)\) is the ambient (external) temperature, and the control variable \((u_t)\) is the amount of energy consumed with (normalised) values in \([0, 1]\). See equations 3.2 and (A.16) below for a comparison. Note that there are other normalisations done in [27, pp. 756], but these are omitted here because they add nothing to the present discussion.

**Remark 2.1.1.** The ambient temperature \((T^a_t)_{t \in \mathbb{T}}\) is assumed constant in Chapter 3 for simplicity. However, given suitable assumptions on this variable the same techniques in the chapter can be used.

The financial cost for period \( t \in \{0, 1, \ldots, T - 1\} \) is the amount of energy consumed \( u_t \) multiplied by the spot price \( S_t \). The quality of service for this period is measured by a penalty \( B(x_{t+1} - \mathcal{T}) \) for the (end of period) deviation of the indoor temperature from an ideal one \( \mathcal{T} \). Additional utility weights \( \alpha \) and \( \beta_t \) are given for the financial and comfort costs respectively. The sum of these weighted costs over the control horizon \( \{0, 1, \ldots, T - 1\} \) gives the criterion to be minimised [27, p. 756]:

\[
J(x, u) = \sum_{t=0}^{T-1} \left[ \alpha u_t S_t + \beta_t B(x_{t+1} - \mathcal{T}) \right].
\]
An analogous performance measure is used in Chapter 3 and the numerical example of Chapter 4.

The penalty function for temperature deviations should effectively describe the changing attitude of the building’s occupants towards temperature deviations from the given setpoint $T^*$. For this purpose, it was suggested in [27] that bathtub-shaped penalty functions are suitable. Figure 2.1 below depicts such a function. The flat base or “deadband” describes a range of most comfortable temperatures. The saturation points indicate temperatures which are deemed uncomfortable for the building’s occupants, and are therefore penalised the heaviest. The penalty remains constant as deviations increase past these saturation points. Note that the bathtub function need not be symmetric.

![Figure 2.1: Example of a bathtub penalty function for indoor temperature control](image)

**Summary of the numerical results.**

This thesis confirmed through numerical experiments that significant financial savings can be achieved while keeping the (controlled) indoor temperature temperatures within the prescribed deadband. These results are consistent with [27], noting that their experiments used the singleton set \{T^*\} for the deadband. This consistency is not surprising for a number of reasons. First, the models of [27] and Chapter 3 are the same when the time parameter in the latter is discretised for numerical implementation. Second, the algorithm implemented in [27] uses optimal control to solve the non-linear program they constructed. However, their solutions are no longer bang-bang (and therefore sub-optimal) since they are approximations obtained from a sequence of solutions to optimal control problems with quadratic cost criteria.

**Remark 2.1.2.** The reader may also be interested in the theoretical connections
between some optimal control problems and non-linear programs, which are discussed in [60, 94] for example.

Two other similarities in the numerical results are summarised below. In each case an observation of [27] is paraphrased then the relevant conclusion drawn from the numerical results of Chapter 3 is presented.

**Observation 1.**

> Regardless of the weather, a constant indoor temperature, which can differ from the setpoint, must be maintained when the price of electricity is constant. [27, p. 752]

Whenever prices are almost constant or on a decreasing trend, the indoor temperature is kept constant by switching the control unit on and off in rapid succession (recall from Section 1.3.1 above that the optimal strategy is bang-bang).

**Observation 2.**

> There are different “optimal” temperatures which depend on the price and ambient temperature, with intervals of zero control occurring in neighbourhoods of local price maxima. [27, pp. 752, 760]

There are periods of *preheating* where non-zero control is used immediately before a local price maximum, in order to momentarily raise the indoor temperature to a higher level within the deadband. Afterwards, there is zero control during the period corresponding to the local price maximum, and the indoor temperature gradually falls to a lower level within the deadband. The rate of decrease (and, to an extent, increase) is determined by the external temperature and the building’s thermal time constant.

### 2.2 Review for Chapter 4

To the best of the author’s knowledge, the references [89, 139, 140] are the only papers that develop a theory of discrete-time optimal switching problems that is closely related to the one in Chapter 4 (the work of [127] looks at a very different two-parameter discrete-time switching problem). The paper [20] shall be used for the review below.
CHAPTER 2. LITERATURE REVIEW

instead of [89]. This is because [89] studied a discrete-time optimal switching game and the theory which is pertinent to the review refers to the previous paper [20]. A comparison to [140] is also presented below, where the paper [139] is omitted since the conclusions are the same as for [140].

2.2.1 Connection to Carmona and Ludkovski (2008)

The discrete-time model for optimal switching.

Most of the notation used below for the comparison to the paper [20] agrees with that used in Chapter 4. The switching problem of interest consists of the following:

• a discrete time parameter set $T = \{0, \ldots, T\}$;
• a probability space $(\Omega, \mathcal{F}, P)$ equipped with a filtration $\mathcal{F} = (\mathcal{F}_t)_{t \in T}$;
• a finite set of (switching) modes $I = \{1, \ldots, m\}$ where $m \geq 2$;
• running rewards modelled by real-valued $\mathcal{F}_t$-adapted stochastic processes $\Psi_j(\cdot)$ where $j = 1, \ldots, m$;
• terminal rewards given by real-valued $\mathcal{F}_T$-measurable random variables $G_j$ where $j = 1, \ldots, m$;
• for every $j, k \in I$, a cost for switching from $j$ to $k$ modelled by a real-valued $\mathcal{F}$-adapted stochastic process $\gamma_{j,k}(\cdot)$. There are two conditions on these costs: for every $i, j, k \in I$ satisfying $i \neq j$ and $j \neq k$, $\gamma_{i,i}(\cdot) = 0$ and $\gamma_{i,k}(t) < \gamma_{i,j}(t) + \gamma_{j,k}(t)$ almost surely for every $t \in T$.

Remark 2.2.1. Switching costs are assumed to satisfying the weaker triangular condition in $\gamma_{i,k}(t) \leq \gamma_{i,j}(t) + \gamma_{j,k}(t)$ in [20, 89], but are assumed strictly positive in [20] and constant in [89].

Let $(t, i) \in T \times I$ be initial parameters for the problem. A switching control is a double sequence $\alpha = (\tau_n, \iota_n)_{n \geq 0}$ of stopping times $\{\tau_n\}_{n \geq 0}$ and $\mathcal{F}_{\tau_n}$-measurable mode indicators $\{\iota_n\}_{n \geq 0}$ satisfying $\tau_0 = t$ and $\iota_0 = i$ amongst other criteria (see Chapter 4 for details). The sequence of times $\{\tau_n\}_{n \geq 0}$ models the decision of when to switch and $\{\iota_n\}_{n \geq 0}$ models the decision of where to switch. There is a class of admissible switching
controls, \( A_{t,i} \), where each control \( \alpha \in A_{t,i} \) should have at most one switch at each time \( s = t, \ldots, T - 1 \). Associated with each control \( \alpha \in A_{t,i} \) is a mode indicator function \((u_s)_{s \geq t}\) which gives the current mode, and is defined by \( u_s = \sum_{n \geq 0} \tau_n 1\{\tau_n \leq s < \tau_{n+1}\} \).

Note that there may be \( \omega \in \Omega \) such that \( \tau_1(\omega) = t \) and the mode indicator jumps immediately from \( \iota_0(\omega) = i \) to \( \iota_1(\omega) \neq i \). For more details see Section 4.2.1 below.

For the switching problem starting at time \( t \in T \) in mode \( i \in I \), the performance of a given strategy \( \alpha \in A_{t,i} \) is the total reward from \( t \) to \( T \) net of the switching costs:

\[
J(\alpha; t, i) = E \left[ \sum_{s=t}^{T-1} \Psi_{u_s}(s) + G_{i_N(\alpha)} - \sum_{n \geq 1} \gamma_{i_{n-1},i_n}(\tau_n) 1\{\tau_n < T\} \mid \mathcal{F}_t \right], \quad \alpha \in A_{t,i}.
\]

The optimal switching problem is to find a strategy \( \alpha^* \in A_{t,i} \) that maximises \( J(\alpha; t, i) \) over all admissible controls. Such a strategy is said to be *optimal*.

**Comparison to existing theoretical results.**

The results of Carmona and Ludkovski [20], in particular their Theorem 3.3, show that the solution to the optimal switching problem can be obtained by solving successive optimal stopping problems. Starting from \((t, i) \in T \times I\), find the first time \( \tau_1^* \) it is optimal to switch from \( i \) and select \( \iota_1^* \) by maximising the time \( \tau_1^* \) reward net of the switching cost. Then start the optimisation problem from \((\tau_1^*, \iota_1^*)\) and solve for \((\tau_2^*, \iota_2^*)\), and so on. The triangular condition on the switching costs ensures that multiple switches at the same time is *sub-optimal*, so it is not difficult to prove that this sequence forms an admissible strategy \( \alpha^* = (\tau_n^*, \iota_n^*)_{n \geq 0} \in A_{t,i} \). In fact, a similar procedure is followed in Chapter 4. The difference is that one directly obtains the dynamic programming equation of [2, 20, 21, 61, 89], in a manner that is distinct from the procedure followed in [20, 89] that is described below.

For a given \( t \in T \), let \( \mathcal{T}_t \) denote the set of stopping times \( \tau \) satisfying \( t \leq \tau \leq T \) almost surely. Define a sequence of processes \( \{Y^k(\cdot, i): k \geq 0\} \) for each \( t = 0, \ldots, T \) and \( i \in I \) by

\[
Y^0(t, i) := E \left[ \sum_{s=t}^{T-1} \Psi_{i}(s) \mid \mathcal{F}_t \right];
\]

\[
Y^k(t, i) := \text{ess sup}_{\tau \in \mathcal{T}_t} E \left[ \sum_{s=t}^{\tau-1} \Psi_{i}(s) + M(Y^{k-1}(\tau, i)) \mid \mathcal{F}_t \right], \quad k \geq 1; \quad (2.1)
\]
where $M(Y^{k-1}(\tau, i))$ is defined for $\tau \in T_t$ by

$$M(Y^{k-1}(\tau, i)) := \max_{j \neq i} \{Y^{k-1}(\tau, j) - \gamma_{i,j}(\tau)\} 1_{\{\tau < T\}}.$$

Theorem 3.3 of [20] (or Proposition 3.1 of [89]) proves that $Y^k(t, i)$ solves the optimal switching problem starting from $(t, i)$ with at most $k \geq 0$ switches remaining. In particular, for $k \geq T$ it holds that $Y^k(0, i) = \text{ess sup}_{\alpha \in \mathcal{A}_0,i} J(\alpha; 0, i)$. Furthermore, an optimal strategy $\alpha^* \in \mathcal{A}_0,i$ is obtained by solving the successive optimal stopping problems in equation (2.1) as described before.

According to Corollary 3.2 of [89], taking $k > T$ gives a coupled dynamic programming principle (DPP): for every $i \in I$ and $t \in T$,

$$Y(t, i) = \text{ess sup}_{\tau \in T_t} \mathbb{E} \left[ \sum_{s=t}^{\tau-1} \Psi_i(s) + \max_{j \neq i} \{Y(\tau, j) - \gamma_{i,j}(\tau)\} 1_{\{\tau < T\}} \bigg| \mathcal{F}_t \right]$$

$$= \text{ess sup}_{\alpha \in \mathcal{A}_t,i} J(\alpha; t, i).$$

(2.2)

This agrees with the results in Chapter 4, and it is noteworthy that this method for proving the existence of $Y(\cdot, i)$ is closer to the method used in Chapter 5 for the analogous continuous-time problem. The next step is to relate $Y(\cdot, i)$ to the Snell envelope of an implicitly defined obstacle process. This is in fact equivalent to defining $Y(\cdot, i)$ recursively as follows (see Chapter 4 for details):

$$Y(T, i) = 0, \quad \text{and for } t = T - 1, \ldots, 0:$$

$$Y(t, i) = \max_{j \neq i} \left\{ -\gamma_{i,j}(t) + \Psi_j(t) + Y(t, j) \right\} \vee \left\{ \Psi_i(t) + \mathbb{E}[Y(t + 1, i) | \mathcal{F}_t] \right\}.$$

(2.3)

Using an argument based on the triangular condition on the switching costs, and recalling $\gamma_{i,i}(\cdot) \equiv 0$, it is stated in [89, p. 502] (and explained further in [61, p. 2037]) that the implicit recursive equation for $Y(t, i)$ can be replaced by an explicit one:

$$Y(t, i) = \max_{j \neq i} \left\{ -\gamma_{i,j}(t) + \Psi_j(t) + \mathbb{E}[Y(t + 1, j) | \mathcal{F}_t] \right\}, \quad t = T - 1, \ldots, 0.$$

(2.4)

This is essentially the argument used in Chapter 4 of this thesis. However, upon following the above procedure one is faced with this question:

Are the processes $Y(\cdot, 1), \ldots, Y(\cdot, m)$, obtained as the limit of the sequences of process $\{Y^k(\cdot, i): k = 0, 1, \ldots\}$, $i = 1, \ldots, m$, actually the (coupled) Snell envelope-like processes associated with the implicitly defined obstacles appearing in (2.2)?
To answer this question, one might argue similarly to the continuous-time case: show that \( \{(Y^k(t,i) + \sum_{s=0}^{t-1} \Psi_i(s)) \}_{t \in T} : k = 0, 1, \ldots \} \) is a non-decreasing sequence of Snell envelopes that converges pointwise on \( T \) to \( (Y(t,i) + \sum_{s=0}^{t-1} \Psi_i(s))_{t \in T} \), and this limit process is the Snell envelope of a suitably defined implicit obstacle process. However, this thesis used an alternative approach which starts with the backward induction formula (2.4) to construct processes \( \tilde{Y}(\cdot, 1), \ldots, \tilde{Y}(\cdot, m) \). The existence and properties of the corresponding Snell envelope, which may be simpler to verify in this case, are used to show that \( \tilde{Y}(\cdot, 1), \ldots, \tilde{Y}(\cdot, m) \) satisfy equation (2.2).

### 2.2.2 Connection to Yushkevish and Gordienko (2002)

**The discrete-time optimal switching model.**

The comparison below reuses much of the notation in [140]. The switching model studied in that paper depends on the following:

- a positive recurrent Markov chain \( \{x_t : t = 0, 1, \ldots, \infty \} \) with state space \( (X, \mathcal{B}(X)) \), where \( \mathcal{B}(X) \) is the Borel \( \sigma \)-algebra on \( X \);

- two real-valued measurable price functions, \( f \) and \( g \), which are defined on \( X \).

Let \( \mathcal{S} \) denote the set of non-decreasing sequences of stopping times \( \mathcal{T} = \{\tau_k : k = 1, 2, \ldots ; 0 \leq \tau_k \leq \infty \} \) (stopping times are defined with respect to the natural filtration of \( x \)). A strategy \( \sigma = (\mathcal{T}_f, \mathcal{T}_g) \) in the switching model is a pair of sequences \( \mathcal{T}_f, \mathcal{T}_g \in \mathcal{S} \), the subscript is used to denote whether the price function \( f \) or \( g \) is at the initial position. Letting \( 1_A \) denote the indicator function of an event \( A \), the reward functional \( \mathcal{J}_f \) (resp. \( \mathcal{J}_g \)) on a finite time horizon \( 0, \ldots, n \) starting with price function \( f \) (resp. \( g \)) is denoted by

\[
\mathcal{J}_f(\mathcal{T}, n) = \sum_{k=1}^{\infty} \left[ f(x_{\tau_{2k-1}})1_{\{\tau_{2k-1} < n\}} - g(x_{\tau_{2k}})1_{\{\tau_{2k} < n\}} \right],
\]

\[
\mathcal{J}_g(\mathcal{T}, n) = \sum_{k=1}^{\infty} \left[ -g(x_{\tau_{2k-1}})1_{\{\tau_{2k-1} < n\}} + f(x_{\tau_{2k}})1_{\{\tau_{2k} < n\}} \right].
\]

Associated with each reward functional is a value function defined by

\[
V_f(x, n) = \sup_{\mathcal{T} \in \mathcal{S}} \mathbb{E}_x[\mathcal{J}_f(\mathcal{T}, n)], \quad V_g(x, n) = \sup_{\mathcal{T} \in \mathcal{S}} \mathbb{E}_x[\mathcal{J}_g(\mathcal{T}, n)]
\]
where \( x \in X \) is the initial state of the Markov chain, and \( E_x \) is the expectation operator under the associated probability distribution \( P_x \). Certain assumptions are made on \( f \) and \( g \) so that the value function is well-defined and finite for every \( n \).

**Comparison of the theoretical results.**

Some of the main differences between the problem studied in [140] and the one in Chapter 4 of this thesis are outlined below:

- The paper [140] is mainly concerned with optimal switching on an infinite time horizon, where a criterion of *average reward* is to be maximised. This criterion is defined in terms of the limiting behaviour as \( n \to \infty \) of the average \( n \)-step reward functional and associated value function. A corresponding concept of *average optimality* for strategies \( \sigma \) is introduced [140, p. 146]. On the other hand, the problem in Chapter 4 concerns the maximisation of a reward functional on a finite time horizon which depends on \( m \geq 2 \) switching costs and running rewards;

- The finite horizon problem in [140] is solved by reduction to a Markov Decision Process ([140, p. 147]). The approach taken in Chapter 4 is to reduce the switching problem to one of iterative optimal stopping, then solve using martingale methods as described previously. This approach does not require Markovian assumptions.

### 2.3 Review for Chapter 5

#### 2.3.1 The continuous-time optimal switching model

Chapter 5 studies the continuous-time analogue of the optimal switching problem discussed above. The performance index is given by

\[
J(\alpha; t, i) = E\left[\int_t^T \psi_{u_s}(s)ds + G_u_T - \sum_{n \geq 1} \gamma_{n-1, i_n}(\tau_n)1_{[\tau_n < T]} \middle| F_t\right], \quad \alpha \in A_{t,i}
\]

where \( 0 < T < \infty \), \( t \in [0, T] \) is an initial time, \( i \in \mathbb{I} = \{1, \ldots, m\} \) is an initial mode, and \( A_{t,i} \) is an appropriate class of *admissible switching controls* (cf. Section 5.2).
2.3.2 An overview of the existing literature

There are several publications on optimal switching problems in continuous time, and a survey of the literature identifies two main approaches: an analytical approach using partial differential equations (PDEs) and a probabilistic one. Methods based on PDEs and associated variational inequalities appeared as early as the 1970s, under the topic of impulsive control for diffusion processes (see [116] and the references therein). A viscosity solutions approach to this type of PDE appeared in the late 1980s to early 1990s (for instance, [128]) and is still the topic of active research [91].

Probabilistic solution methods were being applied since the 1970s and 1980s in various degrees of generality (see [38, 101, 116] for instance), and most of recent research in this area has been a combination of the martingale approach via Snell envelopes ([37, 42]) and the theory of backward stochastic differential equations (BSDE) ([23, 50]). Notable exceptions include the recent paper [24], which used the theory of (super-)harmonic functions to study impulse control problems within a Markovian framework, and [9] which used the Stochastic Perron’s Method – a novel approach that lies between the probabilistic and analytical approaches – to study a new class of optimal switching problems.

Despite all of this interest, very few references ([24, 42, 91, 112]) allowed signed (both positive and negative) switching costs. This is somewhat expected considering the applications that have motivated the bulk of the research. However, this assumption must be relaxed in models where the controller can (partially) recover its investment, or receive a subsidy/grant for investing in a new technology such as renewable (green) energy production [90, 93]. Moreover, signed switching costs arise in the connection between Dynkin games and two-mode optimal switching problems [65, 140]. Motivated by this connection, this author attempted to extend the Snell envelope approach of Djehiche, Hamadène and Popier [37] (also [20]) to optimal switching problems with signed switching costs. Furthermore, in order to make a contribution to Dynkin games that is comparable to the results already available, the author also relaxed the assumption of continuous switching costs made in [37]. The technical terms used below are further explained in Section 5.3.
2.3.3 Extension to discontinuous switching costs

The extension to discontinuous switching costs shall be discussed first. It should be noted that this generalisation was already hinted at in [37, p. 2753]. The authors referred specifically to the paper [68] where the two-mode optimal switching problem had been studied in a filtration generated by a Brownian motion and an independent Poisson random measure. However, an earlier paper by Morimoto [101] had already studied the optimal switching problem in a more general filtration, assuming it was quasi-left-continuous (free of time discontinuities) in order to prove the existence of an optimal switching control strategy. The filtration used in [68] is a particular case of one that is quasi-left-continuous. However, the assumption of constant, positive switching costs (among others) made in [101] was too restrictive for this author’s purposes. A suitable generalisation of the results in [37] was achieved by combining properties of quasi-left-continuous filtrations and quasi-left-continuous switching costs.

2.3.4 Extension to signed switching costs

Extending the results of [37] to the case of negative switching costs proved to be a more difficult task, since negative switching costs have a positive impact on the performance index and raise an integrability concern for an optimal solution (if it exists). The pre-print [42] extended the results of [37] to negative switching costs under the restrictive assumption that the total occurrences of negative switching costs is bounded.

In order to treat signed switching costs, this author started by including an additional integrability criterion in the definition of an admissible control. More precisely, defining $C_n^\alpha$ as the cumulative switching cost after $n \geq 1$ switches under a given strategy $\alpha$:

$$C_n^\alpha := \sum_{k=1}^{n} \gamma_{\tau_{k-1},\tau_k}(\tau_k)1_{\{\tau_k < T\}}$$

then an admissible strategy should additionally satisfy

$$\mathbb{E}\left[\sup_n |C_n^\alpha| \right] < \infty.$$  

This new condition is not restrictive since it is satisfied under the hypotheses assumed in the existing literature. It is also superfluous in the context of [37] where switching costs are assumed positive, since strategies with a strictly positive probability of an infinite number of switches cannot be optimal.
Using this new admissibility criterion, it is possible to prove the existence of càdlàg processes \(Y^1, \ldots, Y^m\) that solve the optimal switching problem. However, further integrability of these processes is required in order to get the Snell-envelope representation in [37]. The required integrability can be verified in at least two cases:

- the number of times negative switching costs occur is limited in a certain sense as in [42];
- the switching costs are time-independent.

In either case, the processes \(Y^1, \ldots, Y^m\) possess the same integrability properties as [37] but with quasi-left-continuous paths instead of continuous ones. Based on the results of [19, 79], the author also believes that the required integrability might be proved in a Markovian setting. Once this can be accomplished, the admissibility of a candidate optimal strategy formed iteratively from the optimal stopping problems involving the processes \(Y^1, \ldots, Y^m\) can be proved.

### 2.4 Review for Chapter 6

#### 2.4.1 The Dynkin game description

Chapter 6 studies the connection between Dynkin games and optimal switching problems. A mathematical formulation of the former is necessary for the literature review below. Let \(0 < T < \infty\) be given and \((\Omega, \mathcal{F}, (\mathcal{F}_t)_{0 \leq t \leq T}, P)\) be a given filtered probability space. For a given \((\mathcal{F}_t)\)-stopping time \(\nu\), let \(T_\nu\) denote the set of stopping times \(\rho\) satisfying \(\nu \leq \rho \leq T\) almost surely, and set \(\mathcal{T} = \mathcal{T}_0\). Let \(\sigma, \tau \in \mathcal{T}\) denote strategies for players \(\textit{MIN}\) and \(\textit{MAX}\) respectively. The payoff of the Dynkin game is defined in terms of the cost to a player \(\textit{MIN}\):

\[
D(\sigma, \tau) = \int_0^{\sigma \wedge \tau} \psi(s) ds + \gamma_- (\sigma) 1_{\{\sigma \leq \tau\}} 1_{\{\sigma < T\}} - \gamma_+ (\tau) 1_{\{\tau < \sigma\}} + \Gamma 1_{\{\sigma = \tau = T\}}. \tag{2.5}
\]

Certain integrability and measurability conditions are imposed to ensure the payoff is well defined. Player \(\textit{MIN}\) chooses \(\sigma\) to minimise the expected payoff \(E[D(\sigma, \tau)]\) whereas \(\textit{MAX}\) chooses \(\tau\) to maximise it. This leads to upper and lower values for the game, \(W^+\) and \(W^-\) respectively, which are defined as follows:

\[
W^+ := \inf_{\sigma \in \mathcal{T}} \sup_{\tau \in \mathcal{T}} E[D(\sigma, \tau)] \quad W^- := \sup_{\tau \in \mathcal{T}} \inf_{\sigma \in \mathcal{T}} E[D(\sigma, \tau)]
\]
If $W^+ = W^-$ then the game is said to have a value. A pair of stopping times $(\sigma^*, \tau^*)$ is said to constitute a Nash equilibrium or a saddle point if the following property is satisfied:

$$E[D(\sigma^*, \tau)] \leq E[D(\sigma^*, \tau^*)] \leq E[D(\sigma, \tau^*)],$$

for arbitrary $\sigma \in \mathcal{T}$ and $\tau \in \mathcal{T}$. The existence of a saddle point implies the game has a value, which is given by:

$$W^+ = E[D(\sigma^*, \tau^*)] = W^-.$$

### 2.4.2 Connection to Guo and Tomecek (2008)

In an infinite time horizon setting, Guo and Tomecek (2008) showed that the solution to a particular two-mode optimal switching problem leads to the value of the Dynkin game with payoff (2.5) and $\Gamma \equiv 0$. The analogue of this construction for a finite time horizon is given below. For $j \in \{0, 1\}$, define a process $\gamma(\cdot, j)$ from the stopping costs of the Dynkin game as follows:

$$\gamma(t, j) := \gamma_-(t)1_{\{j=1\}} + \gamma_+(t)1_{\{j=0\}}.$$

The process $\gamma(\cdot, j)$ is the cost of switching to mode $j$, so that $\gamma(\cdot, j) \equiv \gamma_{j-1, j}(t)$ in the notation of Section 2.3 above. The performance index associated with the auxiliary optimal switching problem starting from $(t, i) \in [0, T] \times \{0, 1\}$ is given by:

$$J(\alpha; t, i) = E \left[ \int_t^T \psi(s)u_sd\tau + \Gamma u_T - \sum_{k \geq 1} \gamma(\tau_k, \iota_k)1_{\{\tau_k < T\}} \left\lvert \mathcal{F}_t \right\rvert, \quad \alpha \in \mathcal{A}_{t, i} \right],$$

where $\mathcal{A}_{t, i}$ is a set of admissible switching control strategies $\alpha = (\tau_n, \iota_n)_{n \geq 0}$, and $u$ is the mode indicator function associated with $\alpha$. The value function is given by $V(t, i) := \esssup_{\alpha \in \mathcal{A}_{t, i}} J(\alpha; t, i)$ and the notation $V(i) := \sup_{\alpha \in \mathcal{A}_i} J(\alpha; i)$ is used when $t = 0$.

Assuming the cost processes are adapted and satisfy some mild integrability conditions, Chapter 6 adapts the arguments leading to Theorem 3.15 of [65] to prove an analogous finite time horizon result.

**Theorem.** The Dynkin game with payoff (2.5) has a value which is given by:

$$\inf_{\sigma} \sup_{\tau} E[D(\sigma, \tau)] = V(1) - V(0) = \sup_{\tau} \inf_{\sigma} E[D(\sigma, \tau)].$$
Chapter 6 goes on to show that, if one imposes further assumptions on the game’s parameters, it is possible to use the results of Chapter 5 to say more about the game’s solution.

**Theorem.** Given further integrability and regularity of the game parameters:

i) There exists a unique pair of processes \((Y^0, Y^1)\) with paths that are càdlàg (right-continuous with left-limits) and quasi-left-continuous, such that \(Y^0\) and \(Y^1\) satisfy the following coupled equations: for \(t \in [0, T]\),

\[
Y^0_t = \text{ess sup}_{\tau \in \bar{T}_t} E\left[ \left\{ Y^1_{\tau} - \gamma_-(\tau) \right\} 1_{\{\tau < T\}} \bigg| \mathcal{F}_t \right],
\]

\[
Y^1_t = \text{ess sup}_{\tau \in \bar{T}_t} E\left[ \int_{\tau}^{T} \psi(s) ds + \Gamma 1_{\{\tau = T\}} + \left\{ Y^0_{\tau} - \gamma_+(\tau) \right\} 1_{\{\tau < T\}} \bigg| \mathcal{F}_t \right].
\] (2.6)

ii) There exists \(\alpha^* \in A_{t,i}\) such that \(Y^i_t = J(\alpha^*; t, i) = V(t, i)\) P-a.s.

iii) Let \(\sigma^*\) (resp. \(\tau^*\)) be the first switching time for the optimal switching problem starting in mode 0 (resp. 1):

\[
\sigma^* = \inf\{t \geq 0: Y^0_t = Y^1_t - \gamma_-(t)\} \wedge T
\]

\[
\tau^* = \inf\{t \geq 0: Y^1_t = Y^0_t - \gamma_+(t)\} \wedge T
\]

Then the pair \((\sigma^*, \tau^*)\) is a saddle point for the Dynkin game.

This result, which does not appear in [65], has a few interesting connections to the literature on Dynkin games as discussed below.

### 2.4.3 Connection to Morimoto (1984)

In the paper [100], Morimoto uses martingale methods to study a Dynkin game on an infinite time horizon. The discounted expected payoff of the game with discount rate \(\mu > 0\) takes the following form:

\[
\tilde{D}(\sigma, \tau) = \int_{0}^{\sigma \wedge \tau} e^{-\mu s} \psi(s) ds + e^{-\mu (\sigma \wedge \tau)} \left( \gamma_-(\sigma) 1_{\{\sigma \leq \tau\}} - \gamma_+(\tau) 1_{\{\tau < \sigma\}} \right).
\] (2.7)

Under suitable hypotheses, Theorem 1 of [100] showed that there exists a Nash equilibrium for the Dynkin game with expected payoff \(E[\tilde{D}(\sigma, \tau)]\).
Theorem ([100], Theorem 1). Let $\mathcal{W}$ be the Banach space of all right-continuous $(\mathcal{F}_t)_{t \geq 0}$-adapted processes $x$ with the norm $\|x\| = \|\sup_t |x_t|\|_{L^\infty} < \infty$. Suppose the following conditions hold:

1. $\psi, \gamma_+, \gamma_- \in \mathcal{W}$;
2. $-\gamma_+ \leq \gamma_-$;
3. there exists a process $z \in \mathcal{W}$ satisfying:

$$-\gamma_+ \leq z \leq \gamma_-,$$

and such that $(\hat{z}_{t\wedge \sigma^*})_{t \geq 0}$ is a supermartingale and $(\hat{z}_{t\wedge \tau^*})_{t \geq 0}$ is a submartingale, where $\hat{z}$, $\sigma^*$ and $\tau^*$ are defined by:

$$\hat{z}_t = e^{-\mu t} z_t + \int_0^t e^{-\mu r} \psi(r) dr,$$

$$\sigma^* = \inf \{t \geq 0: z_t = \gamma_-(t)\}, \quad \tau^* = \inf \{t \geq 0: z_t = -\gamma_+(t)\}.$$

Then $(\sigma^*, \tau^*)$ is a saddle point for the game. Moreover,

$$z_0 = \inf_{\sigma} \sup_{\tau} \mathbb{E}[\hat{D}(\sigma, \tau)] = \sup_{\tau} \inf_{\sigma} \mathbb{E}[\hat{D}(\sigma, \tau)] = \mathbb{E}[\hat{D}(\sigma^*, \tau^*)].$$

Although the problem setting is slightly different to the one in Chapter 6 (the latter is on a finite time horizon), the arguments of Chapter 6 show that the process $z = (z_t)$ defined by $z_t := Y_t^1 - Y_t^0$ (cf. (2.6)) satisfies similar hypotheses to the above theorem and an analogous conclusion is drawn.

### 2.4.4 Connection to Hamadène and Hassani (2006)

A recent paper by Hamadène and Hassani [67] also studied Dynkin games on a finite time horizon using probabilistic methods. Let $E \subset \mathbb{R}^l \setminus \{0\}$ for an integer $l \geq 1$. A probability space $(\Omega, \mathcal{F}, \mathbb{P})$ is given on which a $d$-dimensional Brownian motion $B = (B_t)_{0 \leq t \leq T}$ is defined, and there is an independent Poisson random measure $\mu (\omega, t, de)$ on $\mathbb{R}^+ \times E$. The probability space is equipped with a filtration $\mathbb{F} = (\mathcal{F}_t)_{0 \leq t \leq T}$ generated by the Brownian motion and Poisson random measure (augmented so as to satisfy the usual conditions). Let $\tilde{\mu}$ denote the compensated measure associated with $\mu$. 
For a fixed stopping time $S \in \mathcal{T}$ and deterministic time $T > 0$, define the following payoff for a Dynkin game on $[S, T]$: for $\sigma, \tau \in \mathcal{T}$,

$$\tilde{D}(\sigma, \tau; S) := \int_{S}^{\sigma \wedge \tau} f(r)dr + L_{\tau}\ind{\tau \leq \sigma < T} + U_{\sigma}\ind{\sigma < \tau} + \xi\ind{\sigma = \tau = T}. \quad (2.8)$$

The parameters in this payoff are related to the ones in equation (2.5) by setting $f(t) = \psi(t)$, $L_t = -\gamma_+(t)$, $U_t = \gamma_-(t)$ and $\xi = \Gamma$. The upper and lower values of the game are given by:

$$W^+(S) = \essinf_{\sigma \in \mathcal{T}_S} \esssup_{\tau \in \mathcal{T}_S} \mathbb{E}\left[\tilde{D}(\sigma, \tau; S)\right]$$

$$W^-(S) = \esssup_{\tau \in \mathcal{T}_S} \essinf_{\sigma \in \mathcal{T}_S} \mathbb{E}\left[\tilde{D}(\sigma, \tau; S)\right]$$

The authors used the concept of doubly reflected backward stochastic differential equations (DRBSDEs) to prove the existence of a solution and saddle point for the Dynkin game with payoff (2.8). The key points that are relevant for the comparison are summarised below. The following (imprecise) definition of a solution to a DRBSDE is based on [67, p. 122].

**Definition.** A (global) solution to the DRBSDE associated with a coefficient (or driver) $f(\omega, t, x, z, v)$, terminal value $\xi$ and respective lower and upper barriers, $L = (L_t)$ and $U = (U_t)$, is a quintuple of $\mathbb{F}$-predictable processes $(X_t, Z_t, V_t(\cdot), K^+_t, K^-_t)$ which satisfies:

1. $-dX_t = f(t, X_t, Z_t, V_t(\cdot)) dt + dK^+_t - dK^-_t - Z_t dB_t - \int_{E} V_t(e)\tilde{\mu}(dt, de), t \leq T$ and $X_T = \xi$;
2. $L_t \leq X_t \leq U_t, \forall t \leq T$;
3. $K^\pm$ are continuous, non-decreasing, $K^\pm_0 = 0$ and

$$(X_t - L_t)dK^+_t = (U_t - X_t)dK^-_t = 0.$$

**Remark 2.4.1.** The processes appearing in this comparison and the next should be understood as having certain integrability or regularity properties. The reader can find more details on the classes of processes used in the relevant references.

In the case that the coefficient $f$ does not depend on $v$, $f = f(\omega, t, x, z)$, a local solution to the DRBSDE associated with $(f, \xi, L, U)$ is defined as [67, p. 126]:
CHAPTER 2. LITERATURE REVIEW

**Definition.** A *local* solution to the DRBSDE associated with \((f, \xi, L, U)\) is the unique \(\mathbb{F}\)-optional process \(X\) which satisfies:

1. \(X_T = \xi\) and \(\mathbb{P}\)-a.s. for any \(t \leq T\), \(L_t \leq X_t \leq U_t\);

2. for any \(\nu \in \mathcal{T}\) there exists a quintuple \((\theta_\nu, Z_\nu, V_\nu(\cdot), K_\nu^{\nu+}, K_\nu^{\nu-})\) with \(\theta_\nu \in \mathcal{T}_\nu\) such that \(\mathbb{P}\)-a.s.:
   \[
   \begin{align*}
   &\forall t \in [\nu, \theta_\nu], X_t = X_{\theta_\nu} + \int_{\nu}^{\theta_\nu} f(s, X_s, Z_s^\nu)ds + \int_{\nu}^{\theta_\nu} d\left(K_s^{\nu+} - dK_s^{\nu-}\right) \\
   &\quad - \int_{\nu}^{\theta_\nu} Z_s^\nu dB_s - \int_{\nu}^{\theta_\nu} \int E V_s^\nu(e) \tilde{\mu}(ds, de),
   \end{align*}
   \]
   \[
   \int_{\nu}^{\theta_\nu} (X_s - L_s)dK_s^{\nu+} = \int_{\nu}^{\theta_\nu} (U_s - X_s)dK_s^{\nu-} = 0
   \]

3. Setting \(\sigma_\nu := \inf\{s \geq \nu: X_s = U_s\} \land T\) and \(\tau_\nu := \inf\{s \geq \nu: X_s = L_s\} \land T\), then
   \(\sigma_\nu \land \tau_\nu \leq \theta_\nu\), \(X_{\sigma_\nu} = U_{\sigma_\nu}\) on \(\{\sigma_\nu < T\}\) and \(X_{\tau_\nu} = L_{\tau_\nu}\) on \(\{\tau_\nu < T\}\).

Under appropriate assumptions on the data \((f, \xi, L, U)\), it is shown in [67] that there exists a solution to the corresponding DRBSDE in either the local or global sense. Among the assumptions are that \(L\) and \(U\) are adapted, càdlàg, have jumps only at inaccessible stopping times and satisfy \(L_T \leq \xi \leq U_T\), \(\mathbb{P}\)-a.s. [67, pp. 124,133].

Let \(S \in \mathcal{T}\) be arbitrary and assume that \(X\) is a local solution to the DRBSDE. Define the stopping times \(\sigma_S^*, \tau_S^*\) as follows:

\[
\sigma_S^* := \inf\{t \geq S: X_t = U_t\}, \quad \tau_S^* := \inf\{t \geq S: X_t = L_t\}.
\]

Theorem 3.1 of [67] shows that the following holds true for the pair \((\sigma_S^*, \tau_S^*)\):

(i) \(X_S = \mathbb{E}[\tilde{D}(\sigma_S^*, \tau_S^*; S) \mid \mathcal{F}_S]\);

(ii) For any \(\sigma, \tau \in \mathcal{T}_S\),

\[
\mathbb{E}[\tilde{D}(\sigma_S^*, \tau; S) \mid \mathcal{F}_S] \leq X_S \leq \mathbb{E}[D(\sigma, \tau_S^*; S) \mid \mathcal{F}_S].
\]

In other words, at each stopping time \(S\) the value of the local solution to the DRBSDE gives the value of the Dynkin game on \([S, T]\), and the pair \((\sigma_S^*, \tau_S^*)\) is a saddle point from \(S\) onwards. This result highlights a relationship between \(X\), which is a local solution to the DRBSDE, and the processes \(Y^0, Y^1\) associated with the optimal switching problem (cf. (2.6)).
CHAPTER 2. LITERATURE REVIEW

It is known that there exists a solution to the Dynkin game (and also the DRBSDE) under a hypothesis known as Mokobodski’s condition. This hypothesis is stated as follows:

**Mokobodski’s Condition.** There exist two non-negative càdlàg supermartingales, $h$ and $h'$ such that:

$$L_t \leq h_t - h'_t \leq U_t, \forall t \in [0,T].$$

Theorem 4.2 of [67] showed that the following weaker assumption, referred to as the weak Mokobodski’s condition, is necessary and sufficient for the existence of a (global) solution to the DRBSDE:

**Weak Mokobodski’s Condition.** There exists a sequence $\{\tau_n\}_{n \geq 0}$ of stopping times such that:

1. $\tau_0 = 0$ and for any $n \geq 0$, $\tau_n \leq \tau_{n+1}$ and $P(\{\tau_n < T, \forall n \geq 0\}) = 0$;

2. for any $n \geq 0$ there exists a pair $(h^n, h'^n)$ of non-negative càdlàg supermartingales satisfying:

$$P - a.s., \forall t \leq \tau_n, \quad L_t \leq h^n_t - h'^n_t \leq U_t.$$

The first condition in the Weak Mokobodski’s hypothesis is related to the finiteness property typically required for the switching times of an admissible switching control. Theorem 4.3 of [67] verifies that the Weak Mokobodski’s condition holds if $L_t < U_t$ for all $t \in [0,T]$, which is again a typical assumption for the switching costs in an optimal switching problem. It can be shown that the pair of processes $(h^n, h'^n)$ for $n \geq 1$ are, up to an integral term, essentially the pair $(Y^0, Y^1)$ of (2.6). Thus an analogous conclusion holds without the non-negativity constraint. See the discussion in [95] for more on this.

**Remark 2.4.2.** A connection between the optimal processes solving the two-mode switching problem and a DRBSDE related to a Dynkin game was used in the subsequently published paper [69]. Although there is no explicit reference to a Dynkin game in [69], it is possible to make the connection by following the references cited in that paper. Also note that the game parameters and optimal processes $Y^0$ and $Y^1$ (cf. (2.6)) associated with the Dynkin game (2.5) have sufficient regularity to prove
the existence of a solution and saddle point for the analogous game starting from an arbitrary stopping time $S \in \mathcal{T}$.

### 2.4.5 Connection to Dumitrescu, Quenez and Sulem (2014)

Dynkin games and DRBSDEs were also studied in the recent paper [39]. The notation from the previous section is used below as much as possible.

A probability space $(\Omega, \mathcal{F}, \mathbb{P})$ is given on which a standard Brownian motion (SBM) $B = (B_t)_{0 \leq t \leq T}$ is defined, and there is an independent Poisson random measure $\mu$ on $\mathbb{R}^+ \times E$ where $E = \mathbb{R} \setminus \{0\}$. Let $\tilde{\mu}$ denote the compensated measure associated with $\tilde{\mu}$. The filtration for the probability space $(\mathcal{F}_t)_{0 \leq t \leq T}$ is the (augmented) one generated by the Brownian motion and Poisson random measure. An imprecise definition of a solution to a DRBSDE, which is adapted from [39, p. 5], now follows.

**Definition.** The terminal time $T > 0$ is fixed and the following data are given: a process $f = f(\omega, t, x, z, v)$ known as the coefficient, and respective lower and upper barriers, $L = (L_t)_{0 \leq t \leq T}$ and $U = (U_t)_{0 \leq t \leq T}$, which satisfy $L \leq U$ and $L_T = U_T$ a.s. A solution to the DRBSDE associated with $(f, L, U)$ is a quintuple $(X, Z, V(\cdot), K^+, K^-)$ satisfying:

$$
\begin{align*}
-K^+ \text{ are non-decreasing, predictable and } K^+_0 &= 0, \\
-dX_t &= f(t, X_t, Z_t, V_t(\cdot)) \, dt + dK^+_t - dK^-_t - Z_t dB_t - \int_E V_t(e) \tilde{\mu}(dt, de); \\
X_T &= L_T.
\end{align*}
$$

(2.9)

with

i) $L_t \leq X_t \leq U_t$, $0 \leq t \leq T$ a.s.;

ii) the measures $dK^+_t$ and $dK^-_t$ are mutually singular (see Definition 2.2 of [39]);

iii) $\int_0^T (X_t - L_t)dK^+_t = \int_0^T (U_t - X_t)dK^-_t = 0$ a.s.;

$$
\Delta_\tau K^{+,d} = \Delta_\tau K^{+,c} 1_{\{X_{\tau-} = L_{\tau-}\}}, \text{ a.s. } \forall \tau \in \mathcal{T} \text{ predictable};
$$

$$
\Delta_\tau K^{-,d} = \Delta_\tau K^{-,c} 1_{\{X_{\tau-} = U_{\tau-}\}}, \text{ a.s. } \forall \tau \in \mathcal{T} \text{ predictable},
$$

where $K^{+,c}$ and $K^{+,d}$ denote the continuous and discontinuous parts of $K^+$ respectively.
CHAPTER 2. LITERATURE REVIEW

Under additional hypotheses on the data \((f, L, U)\), including a version of Mokobodski’s hypothesis, it is proved in [39] that there exists a solution to the DRBSDE. In the special case where the coefficient \(f\) does not depend on \((x, z, v)\), results on the existence of a saddle point for the following (standard) Dynkin game can be obtained. For \(S \in \mathcal{T}\) and stopping times \(\sigma, \tau \in \mathcal{T}_S\), the payoff on \([S, T]\) is given by

\[
\tilde{D}(\sigma, \tau; S) = \int_{\sigma \wedge \tau}^{\sigma} f(r)dr + L_\sigma 1_{\{\tau \leq \sigma\}} + U_\tau 1_{\{\sigma < \tau\}}.
\] (2.10)

Following [39, p. 7], consider two adapted càdlàg processes \(\tilde{L}^f\) and \(\tilde{U}^f\) which are defined by:

\[
\tilde{L}^f_t := L_t - E\left[ L_T + \int_t^T f(s)ds \mid \mathcal{F}_t \right], \quad \tilde{U}^f_t := U_t - E\left[ L_T + \int_t^T f(s)ds \mid \mathcal{F}_t \right].
\] (2.11)

The processes \((\tilde{L}^f, \tilde{U}^f)\) satisfy all of the properties of \((L, U)\) and also \(\tilde{L}^f_T = \tilde{U}^f_T = 0\) a.s. Lemma 3.2 of [39] verifies that there exists a pair of non-negative, càdlàg supermartingales \((J_{+}^f, J_{-}^f)\) which is the minimal solution to the system:

\[
\begin{cases}
J_{+}^f_t = \text{ess sup}_{\tau \in \mathcal{T}_t} E\left[ J_{-}^f_{\tau} + \tilde{L}^f_{\tau} \mid \mathcal{F}_t \right] \\
J_{-}^f_t = \text{ess sup}_{\tau \in \mathcal{T}_t} E\left[ J_{+}^f_{\tau} - \tilde{U}^f_{\tau} \mid \mathcal{F}_t \right]
\end{cases}
\] (2.12)

The system of Snell envelopes in equation (2.12) also appears in [83] where results for Dynkin games analogous to [39] were established using martingale methods. The following theorems proved in [39] are of interest to the comparison.

**Theorem** ([39], Theorem 3.3). Let \(L\) and \(U\) be two adapted càdlàg processes with \(L_T = U_T\) a.s. and \(L_t \leq U_t, \forall t \in [0, T]\) a.s. Let \(\bar{X}\) be the càdlàg process defined by:

\[
\bar{X}_t := J_{+}^f_t - J_{-}^f_t + E\left[ L_T + \int_t^T f(s)ds \mid \mathcal{F}_t \right]; \quad 0 \leq t \leq T.
\] (2.13)

Provided \(L, U, J_{+}^f\) and \(J_{-}^f\) satisfy additional integrability hypotheses, there exists a quadruple of process \((\bar{X}, Z, V(\cdot), K^+, K^-)\) such that \((\bar{X}, Z, V(\cdot), K^+, K^-)\) is a solution to the DRBSDE (2.9). Furthermore, \(\kappa = K^+ - K^-\), where \(K^+\) (resp. \(K^-\)) is the non-decreasing predictable process in the Meyer decomposition of \(J_{+}^f\) (resp. \(J_{-}^f\)).

**Theorem** ([39], Theorem 3.5). Suppose the hypotheses of the previous theorem are satisfied. Then there exists a unique solution \((X, Z, V(\cdot), K^+, K^-)\) to the DRBSDE (2.9) associated with the coefficient \(f = f(t)\). Furthermore,
CHAPTER 2. LITERATURE REVIEW

- For each $S \in \mathcal{T}$, $X_S$ is the common value function for the Dynkin game with the payoff (2.10);

- If the processes $K^+$ and $K^-$ are continuous, then for each $S \in \mathcal{T}$ the pair of stopping times $(\sigma_S^*, \tau_S^*)$ defined by:

\[
\sigma_S^* := \inf \{ t \geq S : X_t = U_t \}; \quad \tau_S^* := \inf \{ t \geq S : X_t = L_t \},
\]

is a saddle point for the Dynkin game with the payoff (2.10).

This shows another relation to the processes $(Y^0, Y^1)$ (cf. (2.6)) solving the auxiliary optimal switching problem associated with the Dynkin game (2.5). Note also that if $K^+$ and $K^-$ are continuous then $J^{+,f}$ and $J^{-,f}$ are regular supermartingales (also of class $[D]$ by the hypotheses), and are therefore quasi-left-continuous due to the assumption on the filtration.

Remark 2.4.3. Also discussed in [39, 83] is the connection between the pair of supermartingales $(J^{+,f}, J^{-,f})$ and Mokobodski’s hypothesis. This thesis does not explore this connection in detail, but one can draw similar conclusions using the results of Chapter 6, as this author explains in the preprint [95].

2.5 Review for Chapter 7

Chapter 7 presents preliminary results for a particular finite horizon non-zero sum stochastic differential game. In this game, one player uses a classic control with continuous effects whilst the other player uses an impulse control with discrete, discontinuous effects. Impulse control is a generalisation of optimal stopping similar to the switching problems of the previous chapters. It has various applications in economics and mathematical finance including the valuation of multiple exercise (swing) options (see [28], [77] and Chapter 3 of [14]).

Optimisation problems which involve classic and impulse controls have previously been studied in the context of a single decision maker (examples include [36, 108, 141]), and for zero-sum games in either stochastic or deterministic settings ([142] and the references therein). The extension to non-zero sum stochastic games of classic and impulse control undertaken in this thesis appears to be new. However, Chapter 7
shows that questions related to the existence and characterisation of Nash equilibria for the non-zero sum game of mixed impulse and classic controls may be answered by analysing solutions to related non-zero sum games of control and stopping (controller–stopper games). The closest references for the problem studied in Chapter 7 include:

1. zero-sum games with impulse controls:
   - the paper by Zhang [142], which studied a zero-sum version of the aforementioned game on an infinite time horizon, and
   - the articles by Cosso [28] and Stettner [123] which studied zero-sum games between two players using impulse controls;

2. non-zero sum controller–stopper games:
   - the paper by Wang et al [134] which studied a non-zero sum controller–stopper game using partial differential equations similar to the ones appearing in Chapter 7.

Other references on controller-stopper games are less relevant because they are either zero-sum (for example, [10]), or are non-zero sum but with payoff very different to the one appearing in Chapter 7 (for example, [80]).

2.5.1 Connection to Zhang (2011)

The recent paper by Zhang [142] is very close to the game studied in Chapter 7. Using viscosity solutions theory, the author analysed the quasi-variational partial differential equation associated with a zero-sum stochastic game of mixed classic and impulse controls. The main contributions and their relation to the results in Chapter 7 are discussed below.

Definitions.

Notation. Let \((\Omega, \mathcal{F}, \mathbb{P})\) be a complete probability space on which a \(d\)-dimensional standard Brownian motion \(B = (B(t))_{t \geq 0}\) is defined. This space is equipped with a filtration \(\mathbb{F} = (\mathcal{F}_t)_{t \geq 0}\) which is the one generated by \(B\) and completed by the \(\mathbb{P}\)-null sets of \(\mathcal{F}\). For suitable spaces \(A\) and \(E\), let \(C(A; E)\) (respectively \(C_b(A; E)\)) denote
the space of continuous functions (respectively bounded and continuous) from $A$ to $E$.

If $E = \mathbb{R}$ then the notation $C(A)$ and $C_b(A)$ is used instead. Let $C^k(\mathbb{R}^n)$ denote the space of functions $\phi$ such that $\phi$ and its derivatives up to order $k \geq 1$ are in $C(\mathbb{R}^n)$. Let $U \subset \mathbb{R}^d$ and $K \subset \mathbb{R}^n$ be compact sets.

The action sets and controlled process. The maximising player $\text{MAX}$ chooses a control $(u(t))_{t \geq 0}$ from the set $\mathcal{U}$ of all $U$-valued $\mathbb{F}$-progressively measurable processes. This is the set of classic controls. The minimising player $\text{MIN}$ chooses controls from the set $\mathcal{V}$ consisting of elements $\xi(\cdot) = \sum_{i \geq 1} \xi_i 1_{(\tau_i, \infty)}(\cdot)$ where:

- $\{\tau_i\}_{i \geq 1}$ is a non-decreasing sequence of $\mathbb{F}$-stopping times satisfying $\tau_i \to \infty$ almost surely;
- $\xi_i : \Omega \to K$ is an $\mathcal{F}_{\tau_i}$-measurable random variable.

The process $\xi(\cdot) \in \mathcal{V}$ is called an impulse control. Starting from some initial state $x \in \mathbb{R}^n$ at time 0, the players influence a state variable $(X^x_t)_{t \geq 0}$ over time through their choice of controls $u \in \mathcal{U}$, $\xi \in \mathcal{V}$:

$$X^x(t) = X^x_{t_0} = x + \int_0^t b(X^x(s), u(s))ds + \int_0^t \sigma(u(s))dB(s) + \xi(t), \quad t \geq 0$$

where $b : \mathbb{R}^n \times U \to \mathbb{R}^n$ and $\sigma : U \to \mathbb{R}^{n \times d}$ are given measurable functions.

Remark 2.5.1. The impulse control has an additive effect on the state variable in [28, 142] whereas it assigns a new random value to the state variable in [123]. The impulse control of Chapter 7 is slightly more general since the new value it assigns to the state variable is a function of the impulse amount and the value of the state prior to the impulse. This type of impulse effect can be found in [108] for example.

Performance functional, lower and upper game values. Let $\lambda > 0$ be a given discount factor, $f \in C_b(\mathbb{R}^n \times U)$ and $L \in C_b(K; (0, \infty))$ be given cost functions. The players' controls also determine the value of the following performance functional $J_x(u, \xi)$ associated with the game:

$$J_x(u, \xi) = \mathbb{E} \left[ \int_0^\infty e^{-\lambda t} f(X^x(t), u(t))dt + \sum_{i \geq 1} e^{-\lambda \tau_i} L(\xi_i) 1_{(\tau_i, \infty)} \right]$$ (2.14)
In order to define the upper and lower values of the game, [142] introduces non-anticipating (or Elliott-Kalton) strategies for the players. These are mappings $\alpha : \mathcal{U} \rightarrow \mathcal{V}$ and $\beta : \mathcal{V} \rightarrow \mathcal{U}$ that select controls for players MIN and MAX respectively in a particular manner based on the other player’s control choice. It suffices to note that $\alpha(u) \in \mathcal{V}$ and $\beta(\xi) \in \mathcal{U}$ for every $u \in \mathcal{U}$ and $\xi \in \mathcal{V}$. Let $\mathcal{A}$ and $\mathcal{B}$ denote the sets of non-anticipating strategies for MIN and MAX respectively. The lower and upper values for the zero-sum game, $V^-$ and $V^+$ respectively, are defined by

$$
V^-(x) = \inf_{\alpha \in \mathcal{A}} \sup_{u \in \mathcal{U}} J_x(u, \alpha(u)), \quad V^+(x) = \sup_{\beta \in \mathcal{B}} \inf_{\xi \in \mathcal{V}} J_x(\beta(\xi), \xi).
$$

Note that this is different from the definition of the upper and lower values used previously for the Dynkin game (cf. Section 2.4.1). The game with functional (2.14) is said to have a value if $V^-(x) = V^+(x)$ for all $x \in \mathbb{R}^n$, and the common value $V$ is also called the game’s solution. For a given $x \in \mathbb{R}^n$, if there exist controls $u^* \in \mathcal{U}$ and $\xi^* \in \mathcal{V}$ such that $V^-(x) = V^+(x) = J_x(u^*, \xi^*)$, then $(u^*, \xi^*)$ is called an optimal strategy for the game.

**Existence of the game value and a verification theorem.**

Zhang shows that the game with performance functional (2.14) has a value and provides a verification theorem for a candidate pair $(u^*, \xi^*)$ to be optimal. These were achieved by studying a particular quasi-variational partial differential inequality (QVI).

**The quasi-variational inequality.** Let $\mathcal{N}$ be the operator acting on functions $\phi \in C(\mathbb{R}^n)$ as follows

$$
\mathcal{N}[\phi](x) = \min_{\xi \in \mathcal{K}} \{\phi(x + \xi) + L(\xi)\}.
$$

Let $\mathcal{S}^n$ denote the set of $n \times n$ symmetric real-valued matrices. The Hamiltonian $H : \mathbb{R}^n \times \mathbb{R} \times \mathbb{R}^n \times \mathcal{S}^n$ is defined by

$$
H(x, p, P, Q) = \sup_{u \in \mathcal{U}} \{f(x, u) - \lambda p + \langle b(x, u), P \rangle + \frac{1}{2} \text{Tr}[(\sigma \sigma^T)(u)Q]\}
$$

where $\langle \cdot, \cdot \rangle$ is the usual inner product, $\mathsf{T}$ and $\text{Tr}$ are the transpose and trace matrix operators respectively. Let $D_x \phi$ and $D_x^2 \phi$ denote the gradient vector and Hessian
matrix of a function $\phi \in C^2(\mathbb{R}^n)$. The QVI associated with the stochastic game of mixed classic and impulse controls is

$$\min \{ H(x, V(x), D_x V(x), D^2_x V(x)), N[V](x) - V(x) \} = 0, \forall x \in \mathbb{R}^n. \quad (2.15)$$

In general, there may not be classic a $C^2(\mathbb{R}^n)$ solution to (2.15) so the author resorted to viscosity theory to analyse this equation (see Chapter 3 for results in the first-order case). He shows that the lower and upper values $V^-$ and $V^+$ are bounded, uniformly continuous functions which solve (2.15) in the viscosity sense. Furthermore, the solution is unique in this case and the game has a value $V$ which is the bounded, uniformly continuous viscosity solution to (2.15).

**Dynamic programming principle.** In order to prove the existence of the game’s value, Zhang established the following DPP for $V^-$ and $V^+$: for any $t > 0$

$$V^-(x) = \inf_{\alpha \in \mathcal{A}} \sup_{u \in \mathcal{U}} \mathbb{E} \left[ \int_0^t e^{-\lambda f(X^x(s), u(s))} ds + \sum_{\tau_i \leq t} e^{-\lambda \tau_i} L(\xi_i) + e^{-\lambda t} V^-(X^x(t)) \right], \quad (2.16)$$

where $X^x(t) = X^{x, u, \alpha(u)}(t)$, $\alpha(u) = (\tau_i, \xi_i)_{i \geq 1}$;

$$V^+(x) = \sup_{\beta \in \mathcal{B}} \inf_{\xi \in \mathcal{V}} \mathbb{E} \left[ \int_0^t e^{-\lambda f(X^x(s), \beta(\xi)(s))} ds + \sum_{\tau_i \leq t} e^{-\lambda \tau_i} L(\xi_i) + e^{-\lambda t} V^+(X^x(t)) \right], \quad (2.17)$$

where $X^x(t) = X^{x; \beta(\xi), \xi}(t)$.

The DPP shows that the lower and upper game values can be calculated from related games on a shorter time horizon $[0, t]$ with their future values provided as terminal data. This may be compared to a result of Chapter 7, Section 7.5, which obtains the solution to the game as a recursion on the solutions to other games with random time horizon and terminal value both depending on a single impulse.

**Verification theorem.** Zhang closes the paper by proving a verification theorem for optimal strategies under the assumption of a classic $C^2(\mathbb{R}_n)$ solution $v$ to (2.15). Let $\mathcal{L}^u$, $u \in \mathcal{U}$, be operators acting on functions $\phi \in C^2(\mathbb{R}_n)$ as follows

$$\mathcal{L}^u[\phi](x) = -\lambda \phi(x) + \langle b(x, u), D_x \phi(x) \rangle + \frac{1}{2} \text{Tr}[(\sigma \sigma^T)(u)D^2_x \phi(x)], \quad x \in \mathbb{R}^n.$$
CHAPTER 2. LITERATURE REVIEW

Let $C$ denote the *continuation region* defined by

$$C = \{ x \in \mathbb{R}^n : v(x) < \mathcal{N}[v](x), \quad H(x, v(x), D_x v(x), D_x^2 v(x)) = 0 \}.$$ 

The mixed classic and impulse control associated with $v$ is defined as the pair $(u^*, \xi^*(\cdot) = \sum_{i \geq 1} \xi^*_i 1_{\{\tau^*_i < \infty\}})$ which satisfies, if it exists,

$$\mathbb{P} \left( \left\{ \forall (t, X^*(t)) \in \mathbb{R}_+ \times C : u^*(t) \in \arg \max_{u \in U} \left\{ \mathcal{L}^u[v](X^*_t) + f(X^*(t), u) \right\} \right\} \right) = 1;$$

$$\tau^*_0 = 0, \quad \xi^*_0 = 0;$$

$$\tau^*_i = \inf \{ t > 0 : v(X^*(t)) = \mathcal{N}[v](X^*(t)) \};$$

$$\xi^*_i = \arg \min_{\xi \in K} \{ v(X^*(\tau^*_i) + \xi) + L(\xi) \};$$

$$\vdots$$

$$\tau^*_n = \inf \{ t > \tau^*_{n-1} : v(X^*(t)) = \mathcal{N}[v](X^*(t)) \};$$

$$\xi^*_n = \arg \min_{\xi \in K} \{ v(X^*(\tau^*_n) + \xi) + L(\xi) \}.$$ 

where $(X^*(t))_{t \geq 0}$ is the state variable corresponding to $(u^*, \xi^*)$. This is analogous to Theorem 7.5.4 in Chapter 7 below which first assumes the existence of smooth solutions to certain partial differential equations, and then uses these solutions to construct a pair of strategies for the game that is optimal in the sense of *Nash equilibrium*.

2.5.2 Connection to Cosso (2013)

There is recent work by Andrea Cosso [28] on a zero-sum game involving impulse control strategies. Unlike Zhang’s paper [142] where one of the player’s used a classic control, both of the players in [28] use impulsive strategies. Moreover, the game in [28] is on a finite time horizon as in Chapter 7.

Definitions.

**Notation.** Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a complete probability space on which a $d$-dimensional standard Brownian motion $B = (B(t))_{t \geq 0}$ is defined. This space is equipped with a filtration $\mathbb{F} = (\mathcal{F}_t)_{0 \leq t \leq T}$ which is the one generated by $B$ and completed by the $\mathbb{P}$-null sets of $\mathcal{F}$. Let $U \subset \mathbb{R}^n$ and $W \subset \mathbb{R}^n$ be convex cones satisfying $W \subset U$. Let $T > 0$ be the terminal time and $t \in [0, T]$ be an initial time.
The action sets and controlled process. For a set $K$, a $K$-impulse control on $[t, T]$ is a process $\kappa(\cdot) = \sum_{i \geq 1} \kappa_i 1_{[\bar{t}_i, T]}(\cdot)$ where:

- $\{\delta_i\}_{i \geq 1}$ is a non-decreasing sequence of $\mathcal{F}$-stopping times taking values in $[t, T] \cup \{+\infty\}$;
- $\kappa_i : \Omega \rightarrow K$ is an $\mathcal{F}_{\delta_i}$-measurable random variable.

The control set $\mathcal{U}_t$ for the maximising player $MAX$ is the set of all $U$-impulse controls $\xi(\cdot) = \sum_{i \geq 1} \xi_i 1_{[r_i, T]}(\cdot)$ on $[t, T]$. The minimising player $MIN$ chooses controls from the set $\mathcal{W}_t$ consisting of $W$-impulse controls $\eta(\cdot) = \sum_{i \geq 1} \eta_i 1_{[\rho_i, T]}(\cdot)$ on $[t, T]$. Starting from some initial state $x \in \mathbb{R}^n$ at time $t$, the players influence a state variable $(X^{t,x}_s)_{s \geq t}$ over time through their choice of controls $\xi \in \mathcal{U}_t, \eta \in \mathcal{W}_t$:

$$X^{t,x}(s) = X^{t,x,\xi,\eta}(s) = x + \int_t^s b(r, X^{t,x}(r)) dr + \int_t^s \sigma(r, X^{t,x}(r)) dB(r)
+ \sum_{i \geq 1} \xi_i 1_{[r_i, T]}(s) \prod_{l \geq 1} 1_{\{r_l \neq \rho_l\}} + \eta(s), \quad s \geq t \quad \mathbb{P} \text{-a.s.}$$

with $X^{t,x}(t^-) = x$ and $b : \mathbb{R}_+ \times \mathbb{R}^n \rightarrow \mathbb{R}^n$, $\sigma : \mathbb{R}_+ \times \mathbb{R}^n \rightarrow \mathbb{R}^{n \times d}$ are given measurable functions. The infinite product in the state equation reflects the decision to only allow the action of $MIN$ to take effect if both players act at the same time.

Performance functional, lower and upper game values. Let $f \in C_b(\mathbb{R}_+ \times \mathbb{R}^n)$, $g \in C_b(\mathbb{R}^n)$, $L \in C(\mathbb{R}_+ \times U; (0, \infty))$, $M \in C(\mathbb{R}_+ \times W; (0, \infty))$ be given functions. The players’ controls also determine the value of the following performance functional $J_x(\eta, \xi)$ associated with the game:

$$J(t, x; \xi, \eta) = E \left[ \int_t^T f(s, X^{t,x}(s)) ds - \sum_{i \geq 1} L(\tau_i, \xi_i) 1_{\{\tau_i \leq T\}} \prod_{l \geq 1} 1_{\{\tau_l \neq \rho_l\}}
+ \sum_{k \geq 1} M(\rho_k, \eta_k) 1_{\{\rho_k \leq T\}} + g(X^{t,x}(T)) \right]$$

(2.18)

As done previously, the infinite product in (2.18) reflects the decision to only allow the action of $MIN$ to take effect if both players act at the same time. Elliott-Kalton / non-anticipating strategies for the players are used to define the upper and lower values of the game. As discussed before, these are mappings $\alpha : \mathcal{U}_t \rightarrow \mathcal{W}_t$ and $\beta : \mathcal{W}_t \rightarrow \mathcal{U}_t$ that select controls for players $MIN$ and $MAX$ respectively in a particular manner based on the other player’s control choice. It suffices to note that $\alpha(\xi) \in \mathcal{W}_t$ and
$\beta(\eta) \in U_t$ for every $\xi \in U_t$ and $\eta \in W_t$. Let $A_t$ and $B_t$ denote the sets of non-anticipating strategies on $[t, T]$ for $\text{MIN}$ and $\text{MAX}$ respectively. The lower and upper values for the zero-sum game, $V^-$ and $V^+$ respectively, are defined by

$$V^-(t, x) = \inf_{\alpha \in A_t} \sup_{\xi \in U_t} J(t, x; \xi, \alpha(\xi)), \quad V^+(t, x) = \sup_{\beta \in B_t} \inf_{\eta \in W_t} J(t, x; \beta(\eta), \eta).$$

If $V^-(t, x) = V^+(t, x)$ for all $(t, x) \in \mathbb{R}_+ \times \mathbb{R}^n$, then the game with functional (2.18) is said to have a solution which is given by the common value $V$.

### Existence of the game value.

Cosso proves the game with performance functional (2.18) has a value by studying a particular QVI. The same QVI works for both value functions since the two players do not act simultaneously on the system [28, p. 2107]. However, a different QVI is obtained if the action of $\text{MAX}$ takes effect when both players act at the same time [28, p. 2108]. Cosso also gives sufficient conditions for value of the game to agree across the two cases.

### The quasi-variational inequality.

Let $L$ be the operator acting on functions $\phi \in C^{1,2}(\mathbb{R}_+ \times \mathbb{R}^n)$ as follows: for each $t \geq [0, T]$,

$$L[\phi](t, x) = \langle b(t, x), D_x \phi(t, x) \rangle + \frac{1}{2} \text{Tr}[(\sigma \sigma^T)(u)D^2_x \phi(t, x)], \quad x \in \mathbb{R}^n$$

where $\langle \cdot, \cdot \rangle$ is the inner product, $T$ and $\text{Tr}$ are the transpose and trace matrix operators respectively, and $D_x \phi$ and $D^2_x \phi$ denote the gradient vector and Hessian matrix of spatial derivatives. Define operators $\mathcal{M}_{\text{sup}}$ and $\mathcal{M}_{\text{inf}}$ acting on functions $\phi \in C(\mathbb{R}_+ \times \mathbb{R}^n)$ by

$$\mathcal{M}_{\text{sup}}[\phi](t, x) = \sup_{\xi \in U} \{\phi(t, x + \xi) - L(t, \xi)\}, \quad \mathcal{M}_{\text{inf}}[\phi](t, x) = \inf_{\eta \in W} \{\phi(t, x + \eta) + M(t, \eta)\}.$$ 

The QVI associated with the non-zero sum stochastic game of impulse controls is

$$\begin{cases} 
\max\{\min\{-\frac{\partial V}{\partial t} - L[V] - f, V - \mathcal{M}_{\text{sup}}[V]\}, V - \mathcal{M}_{\text{inf}}[V]\} = 0, & [0, T) \times \mathbb{R}^n \\
V(T, x) = g(x) & \forall x \in \mathbb{R}^n.
\end{cases} \quad (2.19)
$$

In general, classic $C^{1,2}(\mathbb{R}_+ \times \mathbb{R}^n)$ solutions to (2.19) may not exist, so the author looks for a solution in the viscosity sense. The paper shows that the lower and upper values $V^-$ and $V^+$ are bounded, uniformly continuous functions which solve (2.19).
in this weaker sense. Furthermore, a viscosity solution is shown to be unique in this
class of functions, thereby showing that the game has a value \( V \) which is the bounded,
uniformly continuous viscosity solution to (2.19).

**Dynamic programming principle.** In the course of proving the main result, Cosso
established the following DPP for the stochastic differential game: for any \( 0 \leq t \leq s < T \) and \( x \in \mathbb{R}^n \),

\[
V^- (t, x) = \inf_{\alpha \in \mathcal{A}_t} \sup_{\xi \in \mathcal{U}_t} \mathbb{E} \left[ \int_t^s f(r, X^{t,x}(r)) \, ds - \sum_{i \geq 1} L(\tau_i, \xi_i) 1_{\{\tau_i \leq s\}} \prod_{l \geq 1} 1_{\{\tau_i \neq \rho_l\}} \right. \\
+ \sum_{k \geq 1} M(\rho_k, \eta_k) 1_{\{\rho_k \leq s\}} + V^- (s, X^{t,x}(s)) \]  

(2.20)

where \( X^{t,x} = X^{t,x;\alpha(\xi)} \), \( \alpha(\xi) = (\rho_i, \eta_i)_{i \geq 1} \);

\[
V^+ (t, x) = \sup_{\beta \in \mathcal{B}_t} \inf_{\eta \in \mathcal{W}_t} \mathbb{E} \left[ \int_t^s f(r, X^{t,x}(r)) \, ds - \sum_{i \geq 1} L(\tau_i, \xi_i) 1_{\{\tau_i \leq s\}} \prod_{l \geq 1} 1_{\{\tau_i \neq \rho_l\}} \right. \\
+ \sum_{k \geq 1} M(\rho_k, \eta_k) 1_{\{\rho_k \leq s\}} + V^+ (s, X^{t,x}(s)) \]  

(2.21)

where \( X^{t,x} = X^{t,x;\beta(\eta),\eta} \), \( \beta(\eta) = (\tau_i, \xi_i)_{i \geq 1} \).

The DPP (2.20)–(2.21) for Cosso’s paper [28] is analogous to the one in Zhang’s paper
[142] (cf. (2.16)–(2.17)), and a similar analogy to Section 7.5 of Chapter 7 is therefore
taken.

**Optimal strategies.** Unlike Zhang [142], Cosso does not prove the existence of
optimal strategies for the zero-sum game of impulse control. This was left as an
open problem and a discussion [28, pp. 2129–2130] highlighted the main difficulties
in adapting previous methods to solve it. Reference is made to a paper by Stettner
[123], which is discussed below, in which the existence of optimal strategies is obtained
as a limiting argument on the values of related zero-sum games with at most \( n \geq 0 \)
impulses. However, it was not clear to the author how one constructs such a convergent
sequence in the present case.

### 2.5.3 Connection to Stettner (1982)

The paper by Stettner [123] considers three kinds of infinite-horizon, zero-sum, two-
player optimal stopping games in a Markov framework:
1. a Dynkin game where the players’ strategies are stopping times;

2. a stochastic game where the players’ strategies are impulse controls with a constant time delay;

3. a stochastic game where one player uses an instantaneous impulse control and the other one uses a stopping time.

The second and third types of game are the ones of interest to this review and the main results are described below.

**Notation.**

Let $E$ be a set which, when endowed with a suitable metric and topology, becomes a Polish space. The example used below is Euclidean space $E = \mathbb{R}^k$, $k \geq 1$, with its usual metric and topology. Let $\Omega = D([0, \infty); E)$ denote the space of right-continuous with left-hand limits (càdlàg) $E$-valued functions with parameter set $[0, \infty)$. Let $(x_t)_{t \geq 0}$ denote the canonical process on $\Omega$: for $t \geq 0$ and $\omega \in \Omega$, $x_t(\omega) = \omega(t)$. Let $F = (F_t)_{t \geq 0}$ denote a suitable completion of the filtration generated by $(x_t)_{t \geq 0}$ and $\mathcal{F}$ be a completed $\sigma$-algebra containing $F_t$, $t \geq 0$. For $x \in E$, $P_x$ denotes a probability measure on $(\Omega, \mathcal{F})$ under which the canonical process $X = (x_t)_{t \geq 0}$ is a time-homogeneous strong Markov process with respect to $F$ and $P_x(\{x_0 = x\}) = 1$. The games described below are between two players $MIN$ and $MAX$ which can affect the evolution of the Markov process over time. Each game’s payoff is discounted at a constant rate $\lambda > 0$.

**Game with impulse controls having a constant time delay.**

**Definitions.** Let $U^{(1)}$ and $U^{(2)}$ be subsets of the state space $E$ and $h > 0$ be a constant time delay. Consider a stochastic game in which $MIN$ and $MAX$ use impulse controls, $\alpha = (\sigma_n, \xi_n)_{n \geq 1} \in \mathcal{A}$ and $\beta = (\tau_n, \eta_n)_{n \geq 1} \in \mathcal{B}$ respectively, where $\{\sigma_n\}_{n \geq 1}$ and $\{\tau_n\}_{n \geq 1}$ are increasing sequences of $F$-stopping times, and $\xi_n: \Omega \to U^{(1)}$ and $\eta_n: \Omega \to U^{(2)}$ are $\mathcal{F}_{\sigma_n}$-measurable and $\mathcal{F}_{\tau_n}$-measurable random variables respectively. Under the strategy $\alpha = (\sigma_n, \xi_n)_{n \geq 1}$, Player $MIN$ decides at time $\sigma_n$ to shift $X$ to the new state $\xi_n$, and this is realised at time $\sigma_n + h$ provided that:

(i) there was no decision by either player to shift the process $X$ during $[\sigma_n - h, \sigma_n]$;
(ii) there is no other shift of the process by either player during \([\sigma_n, \sigma_n + h]\).

The shift rules for Player \(\text{MAX}\) are analogous. If \(\text{MIN}\) and \(\text{MAX}\) decide to shift the process at the same time then only the action of \(\text{MAX}\) is allowed. The combination of the strategies \(\alpha\) and \(\beta\) has the effect of a single impulse control \(\gamma = (\rho_n, \zeta_n)_{n \geq 1}\) defined by [123, p. 13]:

\[
\rho_1 := \sigma_{r_1} \land \tau_{s_1}; \quad \zeta_1 := \begin{cases} 
\xi_{r_1}, & \text{if } \sigma_{r_1} < \tau_{s_1} \\
\eta_{s_1}, & \text{if } \tau_{s_1} \leq \sigma_{r_1}
\end{cases}
\]

\[
\vdots
\]

\[
\rho_n := \sigma_{r_n} \land \tau_{s_n}; \quad \zeta_n := \begin{cases} 
\xi_{r_n}, & \text{if } \sigma_{r_n} < \tau_{s_n} \\
\eta_{s_n}, & \text{if } \tau_{s_n} \leq \sigma_{r_n}
\end{cases}
\]

where \(r_1 = s_1 = 1\) and for \(n \geq 2,\)

\[
r_n := \min\{i: \sigma_i \geq \rho_{n-1} + h\}, \quad s_n := \min\{i: \tau_i \geq \rho_{n-1} + h\}
\]

with the convention that the minimum over the empty set is \(+\infty\), and also assuming that \(\sigma_{\infty} = \tau_{\infty} = +\infty\). The impulse control \(\gamma\) induces a probability measure \(P^\gamma_x\) on \((\Omega, \mathcal{F})\). This probability measure is such that for \(n \geq 1,\) the canonical process \(X\) starts from \(x \in E\) at time 0 and behaves like a Markov process on each interval \([\rho_n, \rho_{n+1})\) with initial state \(\zeta_n\) at time \(\rho_n\).

**Performance criterion.** Let \(f \in C(E)\) and \(c, d \in C(E \times E)\) be given functions. The performance functional for the game under the strategies \(\alpha\) and \(\beta\) is given by

\[
J_x(\alpha, \beta) = E^\gamma_x \left[ \int_0^\infty e^{-\lambda s} f(x_s) ds + \sum_{i=1}^\infty e^{-\lambda (\sigma_{r_i} \land \tau_{s_i})} \left[ c(x_{\sigma_{r_i}}, \xi_{r_i}) 1_{\{\sigma_{r_i} < \tau_{s_i}\}} + d(x_{\tau_{s_i}}, \eta_{s_i}) 1_{\{\tau_{s_i} \leq \sigma_{r_i}\}} \right] \right], \quad x \in A, \beta \in B. \tag{2.22}
\]

Player \(\text{MIN}\) tries to minimise (2.22) whereas \(\text{MAX}\) tries to maximise it, and a zero-sum game is thereby formed. There are lower and upper values for this game, \(V^-\) and \(V^+\) respectively, which are defined by

\[
V^-(x) := \sup_{\beta \in B} \inf_{\alpha \in A} J_x(\alpha, \beta), \quad V^+(x) := \inf_{\alpha \in A} \sup_{\beta \in B} J_x(\alpha, \beta).
\]
The game has a solution if $V^-(x) = V^+(x)$ for all $x \in E$ and the common value is denoted by $V(\cdot)$. A Nash equilibrium or saddle point solution is a pair $(\alpha^*, \beta^*) \in A \times B$ satisfying: for every $x \in E$,

$$J_x(\alpha^*, \beta) \leq J_x(\alpha^*, \beta^*) \leq J_x(\alpha, \beta^*) \quad \forall \alpha \in A, \beta \in B.$$ 

In this case, neither player can benefit by deviating unilaterally from the Nash equilibrium. Moreover, the existence of a saddle point implies the game has a solution with value $V(x) = J_x(\alpha^*, \beta^*)$ for $x \in E$.

**Solution via a Dynkin game.** Stettner proved the existence of a saddle point solution to the zero-sum game described above. In order to do this, he studied the connection between the upper value $V^+(x)$ and the upper value for a type of Dynkin game. Define operators $\mathcal{M}^1$ and $\mathcal{M}^2$ acting on functions $w \in C(E)$ by

$$\mathcal{M}^1[w](x) = \inf_{\xi \in U^{(1)}} \left[ c(x, \xi) + e^{-\lambda h}w(\xi) \right], \quad \mathcal{M}^2[w](x) = \sup_{\eta \in U^{(2)}} \left[ d(x, \eta) + e^{-\lambda h}w(\eta) \right]$$

and a function $\Phi : E \to \mathbb{R}$ by

$$\Psi(x) := \mathbb{E}_x \left[ \int_0^h e^{-\lambda s} f(x_s)ds \right].$$

Let $\sigma$ and $\tau$ be $\mathbb{F}$-stopping times and consider the following equation:

$$\bar{V}(x) = \inf_{\sigma} \sup_{\tau} \mathbb{E}_x \left[ \int_0^{\sigma \wedge \tau} e^{-\lambda s} f(x_s)ds + e^{-\lambda(\sigma \wedge \tau)} \times \right.$$

$$\left. \left[ \mathcal{M}^1[\bar{V}](x_\sigma)1_{\{\sigma < \tau\}} + \mathcal{M}^2[\bar{V}](x_\tau)1_{\{\tau \leq \sigma\}} + \Psi(x_{\sigma \wedge \tau}) \right] \right]. \quad (2.23)$$

Stettner showed in Section 2 of [123] that (2.23) admits a unique solution $\bar{V} \in C_b(E)$. Furthermore, this solution is also the upper value of the impulse control game, $\bar{V} = V^+$. He then uses this new representation for $V^+$ to define strategies $\alpha^* \in A$, $\beta^* \in B$ such that $V^+(x) = J_x(\alpha^*, \beta^*)$ for every $x \in E$, and $(\alpha^*, \beta^*)$ is a saddle point solution to the impulse control game.

**Relation to this thesis.** Besides the clear differences in the type of game, the nature of the solution in [123] is quite different from the one in Chapter 7 (or even Cosso’s paper [28]). Stettner’s impulse control game with delay was essentially solved using the solution to a single equation reminiscent to the upper value of a Dynkin
game. On the other hand, the proposed solution to the game of Chapter 7 was given as a recursion on the solutions to intermediate games of control and stopping.

The techniques of [123] and Chapter 7 also require the study of certain non-linear equations involving the generator of the Markov process. In the case of the former, the non-linear equations arise from the penalisation method whereas those in the latter are Hamilton-Jacobi-Bellman partial differential equations and variational inequalities. Even the probabilistic setup used to analyse the game above is different (and slightly simpler) than the one used in Chapter 7. The setup in the latter is similar to the one for the third game studied by Stettner.

Game between an impulse controller and a stopper.

Consider now the game between MIN who uses an impulse control similar to the above and MAX who controls the duration of the game via a stopping time. Unlike the previous example, the Markov process is now shifted instantaneously to its new state under the impulse control. The probabilistic model in this case is more involved but has a clear enough interpretation. A new probability space $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}})$ is used where $\tilde{\Omega} = \Omega^N$ and $\tilde{\mathcal{F}} = \mathcal{F}^\otimes \mathbb{N}$ are the $N$-fold (countable) products of the previous sample space and $\sigma$-algebra. The basic idea is to have an infinite (but countably) number of copies of a Markov process $(\Omega, X, \mathcal{F}_t, \mathbb{P}_x)$, and the impulse controlled process is essentially one of these copies between impulses.

The probabilistic model. Introduce the following projection subspaces of $\tilde{\Omega}$: $\Omega_1 = \Omega$, $\mathcal{F}_t^1 = \mathcal{F}_t$ and for $n \geq 1$, $\Omega_{n+1} = \Omega_n \times \Omega$, $\mathcal{F}_{t}^{n+1} = \mathcal{F}_t^n \otimes \mathcal{F}_t$. Let $\mathbb{F}^n = (\mathcal{F}_t^n)_{t \geq 0}$ for $n \geq 1$. The controller MIN chooses new states for a controlled process $(y_t)_{t \geq 0}$ from a set $U \subseteq E$ according to an impulse control strategy $\alpha = (\sigma_n, \xi_n)_{n \geq 1} \in A$. In this case, each $\sigma_n$ is an $\mathbb{F}^n$-stopping time and $\xi_n : \Omega_n \rightarrow U$ is an $\mathcal{F}^n_{\sigma_n}$-measurable random variable. Strategies for MAX belong to the set $\mathcal{T}$ of $[0, \infty]$-valued random variables $\tau$ on $(\tilde{\Omega}, \tilde{\mathcal{F}})$ such that, for $n \geq 0$, $\tau$ is an $\mathbb{F}^{n+1}$-stopping time on $[\sigma_n, \sigma_{n+1})$.

The controlled process $(y_t)_{t \geq 0}$ on $(\tilde{\Omega}, \tilde{\mathcal{F}})$ is of the form: $\omega = (\omega_1, \omega_2, \ldots) \in \tilde{\Omega}$,

$$y_t(\omega) = x_t^{k+1}(\omega_{k+1}), \quad t \in [\sigma_k, \sigma_{k+1}), \quad k \geq 0$$

where $\sigma_0 = 0$ and $x^{k+1}$ is the canonical process on the $(k+1)^{st}$ copy of $(\tilde{\Omega}, \tilde{\mathcal{F}})$. The impulse strategy $\alpha \in A$ induces a probability measure $\mathbb{P}_x^\alpha \equiv \tilde{\mathbb{P}}$ on $(\tilde{\Omega}, \tilde{\mathcal{F}})$ under which
the controlled process \((y_t)_{t \geq 0}\) starts from \(x \in E\) and behaves like a Markov process on the intervals \([\sigma_k, \sigma_{k+1})\) for \(k \geq 0\). This is realised by the projection \(P^n\) of the measure \(P^n_x\) on the space \(\Omega^n\). This should be compared to the construction in Section 7.3.2 below.

**Performance criterion.** The projections \(P^n\) of the measure \(P^n_x\) are used in the definition of the following performance functional for the game:

\[
J^x(\alpha, \tau) = \lim_{N \to \infty} \mathbb{E}^N_x\left[ \int_0^{\sigma_N \wedge \tau} e^{-\lambda s} f(y_s) ds + \sum_{n=1}^N \left[ e^{-\lambda \sigma_n} c(x_{\sigma_n}, \xi_n) 1_{\{\sigma_n < \tau\}} + e^{-\lambda \tau} \Psi(y_\tau) 1_{\{\tau \leq \sigma_n\}} \right] \right], \quad \alpha \in \mathcal{A}, \tau \in \mathcal{T}
\]

where \(f \in C(E; \mathbb{R}_+), \Psi \in C(E)\) and \(c \in C(E \times U; (0, \infty))\) are given functions. Just as in the previous case, there are lower and upper values \(V^-\) and \(V^+\) for this game. The concept of the game’s value and saddle point solution are also defined analogously.

**Solution via a Dynkin game.** Stettner first studies the lower value for the payoff (2.24), \(V^-(x) = \sup_{\tau \in \mathcal{T}} \inf_{\alpha \in \mathcal{A}} J^x(\alpha, \tau)\). He shows that it is continuous and uses it to construct strategies \(\alpha^* \in \mathcal{A}\) and \(\tau^* \in \mathcal{T}\) such that \(V^-(x) = J^x(\alpha^*, \tau^*)\) and \((\alpha^*, \tau^*)\) is a saddle point solution. Continuity of \(V^-\) is established by a limiting argument on the lower values \(\{V^-_n\}_{n \geq 0}\) for a similar game, but with at most \(n \geq 0\) impulses. These lower values are shown to have the following recursive optimal stopping game representation:

\[
V^-_0 = \sup_{\tau} \mathbb{E}\left[ \int_0^\tau e^{-\lambda s} f(x_s) ds + e^{\lambda \tau} \Psi(x_\tau) \right]
\]

\[
\vdots
\]

\[
V^-_n = \sup_{\tau} \inf_{\sigma} \mathbb{E}\left[ \int_0^{\sigma \wedge \tau} e^{-\lambda s} f(x_s) ds + e^{-\lambda \sigma} \left[ M[v_{n-1}] (x_\sigma) \lor \Psi(x_\sigma) \right] 1_{\{\sigma < \tau\}} + e^{-\lambda \tau} \Psi(x_\tau) 1_{\{\tau \leq \sigma\}} \right], \quad n \geq 1
\]

where \(M[v](x) = \inf_{\xi \in U} [c(x, \xi) + v(\xi)]\) and \(\sigma, \tau\) are stopping times with respect to the canonic space \((\Omega, \mathcal{F})\). The lower values \(\{V^-_n\}_{n \geq 0}\) are continuous and Stettner argues that they converge uniformly to the limit \(V^-\).

**Relation to this thesis.** There are a few similarities between these results of [123] and those in Chapter 7. The solution to the non-zero sum game of classic and impulse
control in this thesis was shown to be related to the solutions to non-zero sum games of control and stopping, which are defined in a recursive manner similar to (2.25) (cf. Section 7.5). However, Chapter 7 only considers finitely many impulses. This is mainly because the motivating problem suggests that a finite number \( N \geq 0 \) of impulses is sensible. On a more technical note, monotonicity in \( N \geq 0 \) of the solutions with \( N \) impulses, which is usually the argument used in the references, is not guaranteed for the game in Chapter 7. This also makes the problem with infinitely many impulses more difficult (recall the discussion on Cosso’s paper [28] above).

2.5.4 Connection to Wang et al (2013)

Wang et al [134] studied a non-zero sum stochastic differential game between two players on an infinite time horizon. Player I (the controller) uses a classic control which affects the state variable and payoff in a continuous way. On the other hand, Player II (the stopper) controls the duration of the game through its choice of stopping rule. Each player has its own payoff which it aims to maximise, and the game is non-zero sum since these payoffs may be asymmetric. The situation is different from the one in [80], which is more like a non-zero sum Dynkin game with mixed (control and stopping) strategies for both players. The authors in [134] use a dynamic programming approach to obtain a verification theorem for the existence of a Nash equilibrium to their game, and provide an example in which they can characterise this solution. The connection between their verification theorem and the one in Chapter 7 is discussed below.

Definitions.

Notation. Let \((\Omega, \mathcal{F}, \mathbb{P})\) be a complete probability space on which there is a \(d\)-dimensional standard Brownian motion \( B = (B(t))_{t \geq 0} \), and a Poisson random measure \( \mu \) on \( \mathbb{R}^+ \times E, \ E = \mathbb{R}^d \setminus \{0\} \), which is independent of \( B \). This space is equipped with a filtration \( \mathbb{F} = (\mathcal{F}_t)_{t \geq 0} \) which is the one generated by the Brownian motion and Poisson random measure. Let \( \nu \) and \( \tilde{\mu} \) denote the Lévy measure and compensated measure associated with \( \mu \) respectively.
State equation and strategies. Let $U \subset \mathbb{R}^k$ be given. Let $(u(t))_{t \geq 0}$ be an $\mathbb{F}$-adapted càdlàg $U$-valued process which is used to control the evolution of a state process $(Y_t)_{t \geq 0}$ with the following dynamics,

$$
\begin{cases}
    dY(t) = dY(t)b(Y(t), u(t))dt + \sigma(Y(t), u(t))dB(t) + \int_{E} \theta(Y(t^-), u(t^-), z)\tilde{\mu}(dt, dz), \\
    Y(0) = y \in \mathbb{R}^d
\end{cases}
$$

where $b: \mathbb{R}^d \times U \to \mathbb{R}^d$, $\sigma: \mathbb{R}^d \times U \to \mathbb{R}^{d \times d}$ and $\theta: \mathbb{R}^d \times U \times \mathbb{R}^d \to \mathbb{R}^{d \times d}$ are given functions. The notation $Y^u$ may be used to emphasise the dependence of the state on the control.

The admissible control set $\mathcal{U}$ consists of those controls $(u(t))_{t \geq 0}$ which ensure the existence of a unique strong solution $(Y^u_t)_{t \geq 0}$ to the given state equation. A control $(u(t))_{t \geq 0}$ is said to be a feedback control if $u(t) = u(Y(t)), t \geq 0$, where $u: \mathbb{R}^d \to U$ is a Borel-measurable map. The notation $\mathcal{U}$ is used for the set of all feedback controls in $\mathcal{U}$. The state $(Y^u_t)_{t \geq 0}$ is controlled by $(u(t))_{t \geq 0}$ within a given set $\mathcal{S} \subset \mathbb{R}^d$. Suppose $y \in \mathcal{S}$ and let $\rho_* \in \mathbb{F}$ be the first exit time of $(Y_t)_{t \geq 0}$ from $\mathcal{S}$:

$$
\rho_* := \inf\{t > 0 : Y(t) \notin \mathcal{S}\}
$$

where the infimum of the empty set is $+\infty$. The stopper chooses a strategy $\tau$ from the set $\mathcal{T}$ of $\mathbb{F}$-stopping times satisfying $\tau \leq \rho_*$ almost surely.

Performance functionals. Let $f_1, f_2: \mathbb{R}^d \times U \to \mathbb{R}$ and $g_1, g_2: \mathbb{R}^d \to \mathbb{R}$ be given functions defining running and terminal payoffs respectively. The performance functional for player $I$ ($II$) is denoted by $J_1$ ($J_2$) and is given by: for $i = 1, 2$,

$$
J_i(y; u, \tau) = E_y \left[ \int_0^\tau f_i(Y(t), u(t))dt + g_i(Y(\tau))1_{\tau<\infty} \right], \quad u \in \mathcal{U}, \tau \in \mathcal{T}. \tag{2.26}
$$

Player $I$ ($II$) tries to maximise $J_1$ ($J_2$) by the choice of strategy $u$ ($\tau$). The concept of a Nash equilibrium solution is typical in such cases. A pair $(u^*, \tau^*) \in \mathcal{U} \times \mathcal{T}$ is a Nash equilibrium solution to the game if, for every $y \in \mathcal{S}$,

$$
J_1(y; u^*, \tau^*) \geq J_1(y; u, \tau^*), \quad \forall u \in \mathcal{U}; \\
J_2(y; u^*, \tau^*) \geq J_2(y; u^*, \tau), \quad \forall \tau \in \mathcal{T}.
$$
A verification theorem. For each \( v \in U \), let \( L^v \) be an operator acting on functions \( \phi \in C^2(\mathbb{R}^k) \) as follows:

\[
L^v[\phi](y) = \langle b(y, v), D_x \phi(y) \rangle + \frac{1}{2} \text{Tr}[(\sigma \sigma^T)(y, v)D^2_x \phi(y)]
\]

\[
+ \sum_{j=1}^d \int_{\mathbb{R}} \{ \phi(y + \theta^{(j)}(y, v, z_j)) - \phi(y) - \langle D_x \phi(y), \theta^{(j)}(y, v, z_j) \rangle \} \nu_j(dz_j)
\]

where \( \theta^{(j)} \) is the \( j \)th column of the matrix \( \theta \). This definition extends to operators \( L^u \) associated with feedback control functions \( u : \mathbb{R}^d \to U \) in the following way:

\[
L^u[\phi](y) = L^{u(y)}[\phi](y) = L^v[\phi](y), \quad u(y) = v \in U.
\]

A similar definition holds for controls \( u \in U \). For a subset \( A \) of Euclidean space, let \( A^o, \bar{A} \) and \( \partial A \) denote its interior, closure and boundary respectively.

The verification theorem of Wang et al [134] posits the existence of two functions \( \phi_1, \phi_2 : \mathcal{S} \to \mathbb{R} \) where for \( i = 1, 2 \):

1. \( \phi_i \in C^1(\mathcal{S}^o) \cap C(\bar{\mathcal{S}}) \)
2. \( \phi_i(y) \geq g_i(y), \quad \forall y \in \mathcal{S} \)

Define continuation regions:

\( C_i = \{ y \in \mathcal{S} : \phi_i(y) > g_i(y) \} \)

3. \( E \left[ \int_0^\tau 1_{\partial C_i}(Y^u(t))dt \right] = 0, \quad \forall y \in \mathcal{S}, u \in \mathcal{U} \)
4. \( \partial C_i \) is a Lipschitz surface
5. \( \phi_i \in C^2(\mathcal{S} \setminus \partial C_i) \) and the second order derivatives are locally bounded near \( \partial C_i \)
6. \( C_1 = C_2 =: C \)

Define \( \mathcal{H}[\phi_i] \) by \( \mathcal{H}[\phi_i](y) = \sup_{v \in U} \{ L^v[\phi_i](y) + f_i(y, v) \}, \quad y \in \mathcal{S} \)

7. \[
\begin{cases}
\mathcal{H}[\phi_i](y) = 0, & y \in C \\
\mathcal{H}[\phi_i](y) \leq 0, & y \in \mathcal{S} \setminus C
\end{cases}
\]
8. \( \exists \tilde{u} \in \mathfrak{U} \) such that

\[
L^\tilde{u}[\phi_i](y) + f_i(y, \tilde{u}(y)) = \mathcal{H}[\phi_i](y), \quad y \in \mathcal{S}
\]

9. \( E_y \left[ |\phi_i(Y^u(\tau))| + \int_0^\tau |L^u[\phi_i](Y^u(t))|dt \right] < \infty, \quad \forall u \in \mathcal{U}, \tau \in \mathcal{T} \).

For \( u \in \mathcal{U} \), define \( \tau^u_0 \) by

\[
\tau^u_0 = \inf\{ t > 0 : Y^u(t) \notin C \} \tag{2.27}
\]
and suppose, for all $u \in U$ and $y \in \mathcal{S}$, that $\{\phi_i(Y^u(\tau)) : \tau \in T, \tau \leq \tau^u_\mathcal{C}\}$ is uniformly integrable. Then, $(\hat{u}, \tau^\mathcal{C}_\mathcal{C})$ is a Nash equilibrium for the game with payoff (2.26) and

$\phi_1(y) = \sup_{u \in U} J_1(y; u, \tau^\mathcal{C}_\mathcal{C}) = J_1(y; \hat{u}, \tau^\mathcal{C}_\mathcal{C})$

$\phi_2(y) = \sup_{\tau \in T} J_2(y; \hat{u}, \tau) = J_1(y; \hat{u}, \tau^\mathcal{C}_\mathcal{C})$

**Discussion.** The verification theorem of Wang et al is close to Proposition 7.5.2 of Chapter 7. The latter is also a verification theorem for a non-zero sum game of control and stopping (albeit with an additional impulse term). However, the verification theorem in [134] suggests the stopper simultaneously considers the performance criterion of the controller when optimising its own payoff. Proposition 7.5.2, on the other hand, suggests a weaker coupling only through the influence that the controller has on the generator of the underlying state process.

Continuing the previous discussion, in [134] there are continuation regions $\mathcal{C}_1, \mathcal{C}_2$ for the candidate value functions of both players, whereas in Proposition 7.5.2 below the continuation region is defined using only the candidate value function for the stopper. Outside of the single continuation region in Proposition 7.5.2, the candidate value function of the controller must satisfy a given boundary condition. The verification theorem of [134] seeks an optimal strategy which simultaneously optimises the Hamiltonian $\mathcal{H}$ of both players. The analogous condition in Proposition 7.5.2 below is only required for the controller. The hypotheses in the verification theorem of [134] are therefore stronger than the ones in Chapter 7 and may be harder to validate.
Chapter 3

Optimal control of electric space heating

3.1 Introduction

Let there be a building (small commercial or residential) that uses electricity for heating during specified hours in a single day (for example 7:00 a.m. to 10:00 p.m.). Suppose the cost rate for the building’s electricity use varies throughout the day just as in the British electricity spot market. A natural optimisation problem for the building’s owner (henceforth householder) is to minimise the cost of electricity consumption whilst keeping indoor temperatures within a pre-defined acceptable range. This problem is formulated below as one of deterministic optimal control, and its solution is the topic of this chapter.

The following section presents the mathematical model for the optimal control problem. Section 3.3 outlines the standing assumptions on the problem data. The existence of an optimal control within a particular class of controls is discussed in Section 3.4. Conditions that are necessary for a control to be optimal are obtained in Section 3.5 by applying Pontryagin’s Maximum Principle (PMP). Section 3.6 then utilises the method of dynamic programming (DP) to get necessary and sufficient conditions for a candidate control variable to be optimal. The results of numerical simulations based on the DP method are presented in Section 3.7 and a conclusion is given afterwards. Appendix A collects various results not presented in the main text.
3.2 The model

3.2.1 Notation and preliminaries

The times of interest to the control problem are given by an interval \([0, T]\), where \(0 < T < \infty\). For notational convenience define the set \(Q := [0, T) \times \mathbb{R}\) and its closure \(\bar{Q} = [0, T] \times \mathbb{R}\). The boundary of a set \(\mathcal{O}\) is denoted by \(\partial \mathcal{O}\). For a measurable subset \(D\) of \(\mathbb{R}^n\) and a positive integer \(k\), a function \(\psi\) is said to be of class \(C^k(D)\) (respectively \(C^\infty(D)\)) if \(\psi\) agrees on \(D\) with a function \(\tilde{\psi}\) which, on some open set \(\mathcal{O}\) containing \(D\), has continuous partial derivatives of orders up to and including \(k\) (of all orders respectively). First and second order partial derivatives of this function \(\psi\) with respect to the variable \(x\) are denoted by \(W_x\) and \(W_{xx}\) respectively. All temperature values referred to below are in °C.

3.2.2 Indoor temperature evolution

Let \(x(t)\) denote the value in °C of the building’s indoor temperature at time \(t \in [0, T]\). It is controlled over time via a Lebesgue-measurable function \(u: [0, T] \to U\), where the control set \(U = [0, 1]\) represents the energy demand of the electric heater relative to its maximum value. The indoor temperature is assumed to evolve according to the following first order “lumped capacitance” model, which describes thermal transfer within a medium by treating it as a single space [75]:

\[
\begin{align*}
\begin{cases}
  x(s) = y, & y \in \mathbb{R}; \\
  dx(t) = f(t, x(t), u(t))dt, & t \geq s
\end{cases}
\end{align*}
\]

(3.1)

\[
f(t, x(t), u(t)) = -\frac{1}{\tau} (x(t) - T_a - T_g u(t)),
\]

(3.2)

where \(s \in [0, T]\), \(\tau > 0\), \(T_g, T_a \in \mathbb{R}\). Equation (3.1) is commonly referred to as the equation of motion of the system and \(f: [0, T] \times \mathbb{R} \times U \to \mathbb{R}\) is the system velocity [57, p. 24]. A solution \(x\) to (3.1) is called the trajectory corresponding to the control \(u(\cdot)\) and the initial condition \(x(s) = y\) [57, p. 24]. The pair \((t, x(t))\) on \([s, T]\) may also be referred to as a trajectory.

The thermal time constant, \(\tau\), determines the building’s responsiveness to changes in its surrounding temperature. The constant \(T_g\) is the maximum thermal gain of the
space conditioning equipment in °C. The constant $T_a$ is the ambient temperature that is approached asymptotically when the space conditioning equipment is turned off.

The model (3.2) is simple and allows for quick implementation and execution on a computer. These are desirable qualities given the limited processing power that may be available to an auxiliary control algorithm for a residential appliance. For example, the paper [27] used a discretised version of the model in a space conditioning energy controller that is responsive to electricity spot prices, and the references [75, 102, 105, 121] used (3.2) to model the control of thermostatic devices in aggregated loads.

### 3.2.3 Admissible controls

**Definition 3.2.1.** Let $\mathcal{U}_d(s, y)$ denote the class of Lebesgue-measurable functions $u: [s, T] \to U$ such that there exists a unique solution $x = x(\cdot)$ to (3.1) on $[s, T]$ satisfying $x(s) = y$.

With $f$ as in (3.2) (or more generally $f$ satisfying Assumption 1 below), one verifies that every Lebesgue-measurable function $u: [s, T] \to U$ is admissible. See Chapter II in [26] or the references given by [57, p. 63] for instance. For arbitrary admissible controls $u(\cdot) \in \mathcal{U}_d(s, y)$ and $u'(\cdot) \in \mathcal{U}_d(t, x(t)), t \in [s, T]$, if one defines a function $\tilde{u}$ on $[s, T]$ by,

$$
\tilde{u}(r) = \begin{cases} 
  u(r), & \text{for } r \in [s, t] \\
  u'(r), & \text{for } r \in (t, T)
\end{cases}
$$

then $\tilde{u} \in \mathcal{U}_d(s, y)$. Moreover, if $\tilde{x}$ denotes the solution to (3.1) corresponding to the control $\tilde{u}(\cdot)$, then for any $r \in [s, T]$ the restriction of $\tilde{u}$ to $[r, T]$, $\tilde{u}_r(\cdot)$, satisfies $\tilde{u}_r(\cdot) \in \mathcal{U}_d(r, \tilde{x}(r))$.

### 3.2.4 Performance criterion

Let $s \in [0, T]$ be the initial time and $x(s) = y \in \mathbb{R}$ be the initial value of the indoor temperature. Let $S(t)$ represent the *positive* price of electricity at time $t \in [0, T]$, and $\tilde{S}(t) := \frac{S(t)}{\sup_{r \in [s, T]} S(r)}$ denote the *normalised* spot price of electricity. Note that $\tilde{S}(t) \in [0, 1]$ for all $t \in [0, T]$. 
Remark 3.2.1. It is reasonable to assume positive spot prices for the British spot market, but it should be noted that such prices can become negative elsewhere. This has happened in countries such as Germany where there is a high penetration of renewable energy sources [52].

Let \( B: \mathbb{R} \to [0, 1] \) define a penalty function associated with the value of the indoor temperature, and \( \kappa_1, \kappa_2 \) be positive constants. Define two mappings \( L: \bar{Q} \times U \to \mathbb{R}_+ \) and \( \psi: \mathbb{R} \to \mathbb{R}_+ \), the running cost and terminal cost respectively, as follows:

\[
L(t, x, u) = u\tilde{S}(t) + \kappa_1 B(x), \quad \psi(x) = \kappa_2 B(x). \tag{3.3}
\]

Note that \( L \) (resp. \( \psi \)) is bounded on \( \bar{Q} \times U \) (resp. \( \mathbb{R} \)) since \( U = [0, 1] \), \( \kappa_1 \) and \( \kappa_2 \) are positive constants and \( \tilde{S}() \), \( B() \) are bounded. These functions define the following criterion for evaluating the performance of a control \( u() \) given the initial data:

\[
J(s, y, u) = \int_s^T L(t, x(t), u(t)) \, dt + \psi(x(T)), \quad (s, y) \in Q, \quad u \in \mathcal{U}_d(s, y). \tag{3.4}
\]

The performance index (3.4) consists of a financial cost term \( u()\tilde{S}() \) and a comfort cost term \( B(x()) \) in accordance with the previous description of the problem. These costs have been normalised to make them comparable and, similar to [27, p. 756], the constants \( \kappa_1 \) and \( \kappa_2 \) are used as additional weighting factors.

The goal of the optimal control problem is to minimise (3.4) over controls \( u() \in \mathcal{U}_d(s, y) \). The optimal performance is often expressed using the value function \( V(s, y) \), which is defined using the initial time and state:

\[
V(s, y) = \inf_{u \in \mathcal{U}_d(s, y)} J(s, y, u). \tag{3.5}
\]

A control \( u^* \in \mathcal{U}_d(s, y) \) satisfying \( V(s, y) = J(s, y, u^*) \) is said to be optimal within this particular class of controls.

3.3 Assumptions

Assumption 1. The system velocity \( f: \bar{Q} \times U \to \mathbb{R} \) in (3.1) is continuous. Furthermore, there exist positive constants \( C_1 \) and \( C_2 \) such that for all \( t \in [0, T], x, x' \in \mathbb{R} \) and \( u \in U \):

\[
|f(t, x, u)| \leq C_1 (1 + |x| + |u|)
\]

\[
|f(t, x', u) - f(t, x, u)| \leq C_2 |x' - x|(1 + |u|) \tag{3.6}
\]
Remark 3.3.1. Notice that the function $f$ in equation (3.2) does not depend explicitly on the variable $t$ and $(x, u) \mapsto f(t, x, u)$ is $C^1(\mathbb{R} \times U)$ for every $t \in [0, T]$. Furthermore, the derivatives $f_x = \frac{1}{\tau}$ and $f_u = \frac{T}{\tau}$ are constant and $|f(t, 0, 0)| = \left|\frac{T}{\tau}\right|$ is bounded. The hypotheses of Proposition (A.1.1) in are therefore verified, and Assumption 1 holds with $f$ as in equation (3.2). Moreover, the constants $C_1, C_2$ in (3.6) can be chosen so that $C_1 = C_2$, and the term $|u|$ in can be omitted since $U$ is bounded.

Assumption 2. The running cost $L$ and terminal cost $\psi$ in the performance index (3.4) satisfy:

(i) The mapping $L: \bar{Q} \times U \rightarrow \mathbb{R}_+$ is continuous and bounded. Furthermore, $x \mapsto L(t, x, u)$ is Lipschitz for every $t \in [0, T]$ and $u \in U$ fixed;

(ii) The mapping $\psi: \mathbb{R} \rightarrow \mathbb{R}_+$ is Lipschitz continuous and bounded.

Remark 3.3.2. The properties of the cost functions $L$ and $\psi$ depend on the model for the spot price of electricity $S(\cdot)$ as well as a suitable specification for the thermal penalty $B(\cdot)$. In the British electricity market, spot prices change at half-hour intervals. In the present case, this means there is a finite, increasing sequence of times $s = t_0 < t_1 < \ldots < t_m = T$ such that $S(t) = s_i$ for $t \in [t_i, t_{i+1})$, $s_i > 0$ for $i = 0, \ldots, m - 1$. Letting $1_A$ denote the characteristic / indicator function of a set $A$, the spot price $S(t)$ in this case can be expressed as a step function,

$$S(t) = \sum_{i=0}^{m-1} s_i 1_{[t_i, t_{i+1})}(t). \quad (3.7)$$

Therefore, the functions $L$ and $\psi$ may not be $C^1$ in practice. Nevertheless, a very close $C^1$ (in fact $C^\infty$) approximation may be obtained through a standard mollification technique – see Appendix C.5 of [53] or Appendix C of [58] for more details.

### 3.4 Existence of minimising controls

This section shows that the deterministic optimal control problem is well posed in a certain sense. More specifically, it asserts that there exists at least one Lebesgue-measurable control variable $u(\cdot)$ that minimises the performance index (3.4).

Define $\mathcal{S}$ to be a subset of $\mathbb{R}^4$ determined by the end conditions

$$\mathcal{S} = \{ e = (s, T, y, x(T)) : x(T) \in \mathbb{R} \} . \quad (3.8)$$
This is the constraint set for a free terminal point problem. For the purposes of this section the performance index is written in the following form

\[ J(s, y, u) = \int_s^T L(t, x(t), u(t)) dt + \phi(e) \]  

(3.9)

with \( e \in S \) and \( \phi(e) = \psi(x(T)) \).

**Definition 3.4.1.** Let \( x_s \in \mathbb{R} \) and \( u: [s, T] \to U \) be Lebesgue-measurable. Then \( (x_s, u) \) is said to be a feasible pair if

(i) there exists a solution \( x(\cdot) \) to (3.1) in the following sense,

\[ x(t) = x_s + \int_s^t f(r, x(r), u(r)) dr, \quad s \leq t \leq T. \]  

(3.10)

(ii) the above solution \( x(\cdot) \) satisfies the end conditions \( e \in S \).

Let \( \mathcal{W} \) denote the set of feasible pairs.

**Proposition 3.4.1** (Corollary III.4.1, [57]). Suppose Assumptions 1 and 2 hold for \( f \) and \( L \) respectively. If the following additional hypotheses hold:

(i) There exists at least one feasible pair \( (x_s, u) \);

(ii) The set \( U \) is closed;

(iii) Let \( \mu = \inf_{(x_s, u) \in \mathcal{W}} J(s, x_s, u) \) and \( \mu_1 > \mu \). The function \( \phi(\cdot) \) is continuous on \( S \) and there exists a compact set \( S' \subset S \) such that \( e \in S \) and \( J(s, x_s, u) \leq \mu_1 \) implies \( e \in S' \);

(iv) The set \( U \) is convex, \( f(t, x, u) = \alpha(t, x) + \beta(t, x)u, L(t, x, \cdot) \) is convex on \( U \);

(v) \( L(t, x, u) \geq c_1 |u|^{\beta} - c_2, \) where \( c_1 > 0, \ \beta > 1, c_2 \in \mathbb{R} \).

Then there exists \( (x^*_s, u^*) \in \mathcal{W} \) that minimises \( J(s, x_s, u) \) on \( \mathcal{W} \).

**Theorem 3.4.2.** Suppose the assumptions in Section 3.3 hold. Then there exists a Lebesgue-measurable function \( u^* : [s, T] \to U \) such that \( J(s, y, u^*) = \inf_{u \in U_y(s, y)} J(s, y, u) \).

**Proof.** The proof follows by verification of the hypotheses outlined in Proposition 3.4.1.

(i). Note that \( \mathcal{W} = \{(y, u): u(\cdot) \in U_y(s, y)\} \) for fixed \( y \in \mathbb{R} \) and \( s \in [0, T) \), which is not empty.
(ii). Remember that $U = [0, 1]$ which is closed.

(iii). The set $S$ was defined in (3.8) above. Since the terminal cost function $\psi$ is continuous by Assumption 2, the mapping $\phi: S \to \mathbb{R}$, which is defined by $\phi(e) = \psi(x(T))$ is continuous on $S$.

Note that $\mu = \inf_{(x_s,u) \in \mathcal{W}} J(s,x_s,u) \geq 0$ by non-negativity of the cost functions. Let $u_1: [s,T] \to U$ be an arbitrary Lebesgue-measurable function. Then the pair $(y,u_1)$ is feasible. Define $\mu_1 \in \mathbb{R}_+$ by $\mu_1 := J(s,y,u_1)$. Equation (3.10) and the growth condition on $f$ in (3.6) show that for all $s \leq t \leq T$,

$$|x(t)| \leq |y| + \int_s^t |f(r,x(r),u(r))|dr \leq |y| + \int_s^t C_1(1 + |x(r)| + |u(r)|)dr.$$  

Moreover, since $|u(r)| \leq 1$ for all $r \in [0,T]$, it is true that

$$|y| + \int_s^t C_1(1 + |x(r)| + |u(r)|)dr \leq |y| + \int_s^t C_1(2 + |x(r)|)dr$$

$$= D + C_1 \int_s^t |x(r)|dr$$

with $D = |y| + 2C_1(t - s)$. Application of Gronwall’s inequality [57, p. 198] gives,

$$|x(t)| \leq D \exp \left( C_1(t - s) \right), \quad s \leq t \leq T.$$  

(3.11)

The function $x(\cdot)$ is therefore bounded on $[s,T]$ by $M_1 := D \exp \left( C_1(T - s) \right)$ independently of the control variable. If one then defines a compact set $S' \subset S$ by

$$S' := \{ e = (s,y,T,x_1) \in S : |x_1| \leq M_1 \}$$

for any $(x_s,u) \in \mathcal{W}$, $e \in S$ and $J(s,x_s,u) \leq \mu_1$ imply $e \in S'$ as required.

(iv). It is not difficult to verify that intervals on the real line are convex sets. Next, (3.2) shows that $f(t,x,u) = \alpha(t,x) + \beta(t,x)u$ with $\alpha(t,x) = -\frac{1}{p}(x - T^a)$ and $\beta(t,x) = \frac{T}{p}$. By convexity of $U$ and linearity of $u \mapsto L(t,x,u)$ it follows that $L(t,x,\cdot)$ is convex on $U$.

(v). Let $S_{\text{min}}$ be the minimum spot price on $[s,T]$, $S_{\text{min}} = \inf_{t \in [s,T]} S(t)$, define $c_1 := \frac{S_{\text{min}}}{S_{\text{max}}} > 0$ and take any constant $c_2 \geq 0$. Notice that $|u|^\beta \leq u$ for $u \in [0,1]$ and any $\beta > 1$. Since $\tilde{S}(t) \geq \frac{S_{\text{min}}}{S_{\text{max}}}$ for every $t \in [0,T]$ and $\kappa_1 B(\cdot) \geq 0$, one arrives at

$$L(t,x,u) = u\tilde{S}(t) + \kappa_1 B(x) \geq c_1|u|^\beta - c_2.$$  

$\square$
3.5 Necessary conditions for optimality

Section 3.4 showed that there exists a control \( u^* \in \mathcal{U}_d(s, y) \) that minimises the performance index (3.4). Pontryagin’s Maximum Principle helps characterise such extremal controls by providing necessary conditions that the given control and corresponding trajectory must satisfy.

3.5.1 Statement of Pontryagin’s Maximum Principle

**Proposition 3.5.1** (Theorem I.6.3 of [58]). Suppose Assumptions 1 and 2 hold and that furthermore:

(i) The mappings \( x \mapsto L(t, x, u) \) and \( x \mapsto f(t, x, u) \) are \( C^1(\mathbb{R}) \) for every \( t \in [0, T] \) and \( u \in U \) fixed;

(ii) The mapping \( x \mapsto \psi(x) \) is \( C^1(\mathbb{R}) \).

Let \( u^*(\cdot) \in \mathcal{U}_d(s, y) \) and denote by \( x^*(\cdot) \) the corresponding solution to (3.1) satisfying the initial condition \( x^*(s) = y \). For \( u^*(\cdot) \) to be optimal it is necessary that:

(i) there exists an absolutely continuous function \( P(\cdot) \) solving the linear differential equation

\[
\frac{d}{dt}P(t) = -P(t)f_x(t, x^*(t), u^*(t)) - L_x(t, x^*(t), u^*(t)), \quad \forall t \in [s, T] \tag{3.12}
\]

with the transversality condition

\[
P(T) = \psi_x(x^*(T)). \tag{3.13}
\]

(ii) for almost every \( t \in [s, T] \):

\[
P(t)f(t, x^*(t), u^*(t)) + L(t, x^*(t), u^*(t)) = -H(s, x^*(t), P(t)) \tag{3.14}
\]

where the (maximised) Hamiltonian \( H : [s, T] \times \mathbb{R} \times \mathbb{R} \to \mathbb{R} \) is defined by

\[
H(t, x, p) = \sup_{v \in U} \{-p \cdot f(t, x, v) - L(t, x, v)\}. \tag{3.15}
\]

**Remark 3.5.1.** As noted in [58, p. 21], the full statement of the Maximum Principle includes a Lagrange multiplier term \( \lambda \geq 0 \) on \( L_x \) and \( L \) in equations (3.12) and (3.14).
If one can verify that $\lambda \neq 0$, then it is possible to take $\lambda = 1$ since the equations are homogeneous in $\lambda$ (also see Remark II.5.1 of [57]). This is the case when the Maximum Principle is applied to the free terminal point problem (see Chapter II, Section 11 of [57]), which is why the multiplier is not explicit in its statement.

**Sufficient conditions for optimality.**

Pontryagin’s Maximum Principle also provides sufficient conditions for a control to be optimal in the special case of linear control systems. Suppose the following additional hypotheses are satisfied:

1. The system equation is of the form,

   \[ dx = \left( A(t)x(t) + C(t)u(t) \right) dt \]  

   where $A(t)$ and $C(t)$ are continuous functions.

2. The cost function $L$ (resp. $\psi$) in the performance index is continuous, and also $C^1$ and convex in $(x,u)$ (resp. $x$).

Under these conditions, Theorem II.11.6 of [57] asserts that the Maximum Principle also gives sufficient conditions for a control $u = u(\cdot)$ to be optimal in the free terminal point problem.

Note that it is possible to transform the system equation (3.1)–(3.2) into the special form (3.16) above. Indeed, defining the new state variable $\bar{x}(t) := x(t) - T^a$ shows that it evolves according to,

\[ \frac{d}{dt} \bar{x}(t) = \left( A(t)\bar{x}(t) + C(t)u(t) \right) \]

with $A(t) = -\frac{1}{\tau}$ and $C(t) = \frac{T^a}{\tau}$. The cost functions can then be redefined for the new state variable $\bar{x}$ by translating the arguments with respect to $x$. The cost function is convex in $u$ (recall the discussion in the proof of Theorem 3.4.2), and a sufficiently smooth convex function $B(\cdot)$ can chosen for the thermal deviation penalty. The latter often occurs in the engineering literature (see [115, 136] for instance). Pontryagin’s Maximum Principle can then be used to obtain necessary and sufficient conditions for a control $u^*(\cdot)$ to be optimal.
3.5.2 Application of the Maximum Principle

Suppose the model of Section 3.2 satisfies the hypotheses of Proposition 3.5.1. Let \( u^*(\cdot) \in \mathcal{U}_d(s, y) \) be a candidate optimal control and \( x^*(\cdot) \) the corresponding trajectory. Well known results on linear differential equations confirm the existence of a unique solution \( P(\cdot) \) to the following differential equation:

\[
\begin{cases}
\frac{d}{dt} P(t) = \frac{1}{\tau} P(t) - \kappa_1 B_x(x^*(t)), & \forall t \in [s, T] \\
P(T) = \kappa_2 B_x(x^*(T)).
\end{cases}
\]

If \( u^*(\cdot) \) is indeed optimal, Proposition 3.5.1 states that it necessarily maximises the Hamiltonian (3.14):

\[
\begin{align*}
- P(t) \left( \frac{1}{\tau} \left( x^*(t) - T^a - T^g u^*(t) \right) \right) + u^*(t) \tilde{S}(t) + \kappa_1 B(x^*(t)) \\
= - \sup_{v \in U} \left\{ P(t) \left( \frac{1}{\tau} \left( x^*(t) - T^a - T^g v \right) \right) - (v \tilde{S}(t) + \kappa_1 B(x^*(t))) \right\}
\end{align*}
\]

for almost every \( t \in [s, T] \). Grouping together the terms multiplied by \( v \) in the supremum term of (3.18) gives the following result.

**Theorem 3.5.2.** Suppose the hypotheses of Proposition 3.5.1 have been verified for the model of Section 3.2. Then an optimal control \( u^*(\cdot) \in \mathcal{U}_d(s, y) \), if it exists, must satisfy:

\[
u^*(t) = \begin{cases} 
0, & \tilde{S}(t) > - P(t) \frac{T^g}{\tau}; \\
1, & \tilde{S}(t) < - P(t) \frac{T^g}{\tau},
\end{cases}
\]

for almost every \( t \in [s, T] \), where \( P(\cdot) \) is the unique solution to the adjoint equation (3.17).

**Remark 3.5.2.** The value of the control \( u^*(t) \) in (3.19) is undetermined whenever \( \tilde{S}(t) = - P(t) \frac{T^g}{\tau} \). If this happens for an extended period of time the control problem become singular and the corresponding trajectories are called singular arcs [57, p. 37], [78].

Theorem 3.5.2 shows that an optimal strategy must be of threshold (bang-bang) type, where the action/inaction region depends on the (normalised) price of electricity relative to a time-dependent level. However, the hypothesis that the thermal penalty function \( B(\cdot) \) is continuously differentiable is not satisfied for the numerical example of Section 3.7 below, and Theorem 3.5.2 is just a conceptual result in this case.
Nevertheless, there are various extended versions of the Maximum Principle in which the hypotheses on the data may be relaxed. For example, the Maximum Principle still holds if $L$ is measurable in $t$ ([138], Theorem III.2.1). Furthermore, it holds under even weaker hypotheses outlined in [25, 133], but these require concepts of non-smooth analysis that are beyond the scope of this chapter.

### 3.6 The method of dynamic programming

Given initial data $(s, y) \in Q$, let $U_d(s, y)$ be the set of admissible controls as in Definition 3.2.1. The dynamic programming (DP) method makes the connection between the value function $V(s, y)$ in (3.5) and a particular PDE (3.22) referred to as the Hamilton-Jacobi-Bellman (HJB) equation. The theory presented below follows Chapter IV of [57] and Chapters I and II of [58].

#### 3.6.1 Classical results

This section presents some classic (pre-viscosity theory) results on dynamic programming.

**Theorem 3.6.1.** Suppose Assumptions 1 and 2 are satisfied. For any $h > 0$ such that $s + h \leq T$, the value function satisfies Bellman’s “Principle of Optimality”: for any $u \in U_d(s, y)$,

$$V(s, y) \leq \int_s^{s+h} L(t, x(t), u(t)) dt + V(s + h, x(s + h)).$$

(3.20)

Moreover, if $(s, y)$ is any point of $Q$ at which $V$ is differentiable, then $V(s, y)$ satisfies the partial differential inequality

$$V_s(s, y) + V_y(s, y)f(s, y, v) + L(s, y, v) \geq 0$$

(3.21)

for every $v \in U$. If there exists an optimal control $u^\ast \in U_d(s, y)$, then the Hamilton-Jacobi-Bellman equation is satisfied with $v^\ast = \lim_{h \to 0} u^\ast(s + h)$:

$$0 = V_s + V_y f(s, y, v^\ast) + L(s, y, v^\ast) = V_s + \inf_{v \in U} \{ V_y f(s, y, v) + L(s, y, v) \}.$$  

(3.22)

**Proof.** Let $u \in U_d(s, y)$ be admissible. Then a solution $x$ to the differential equation (3.1) with $x(s) = y$ exists for all $s \leq t \leq T$. Note that the set $U_d(s + h, x(s + h))$
is not empty since a control which is admissible on \([s + h, T]\) can be obtained by restricting \(u\) to the interval \([s + h, T]\). Let \(\tilde{u} \in \mathcal{U}_d(s + h, x(s + h))\) be another admissible control and \(\tilde{x}\) denote the solution to (3.1) corresponding to \(\tilde{u}\) and the initial condition \((s + h, x(s + h))\). Then an admissible control \(u_h \in \mathcal{U}_d(s, y)\) can be defined by

\[
u_h(r) = \begin{cases} u(r), & \text{for } r \in [s, s + h] \\ \tilde{u}(r), & \text{for } r \in (s + h, T] \end{cases}
\]  

(3.23)

Let \(x_h\) denote the solution to (3.1) corresponding to the control \(u_h\). From its definition (3.5), the value function \(V(s, y)\) satisfies,

\[
V(s, y) \leq \int_s^T L(t, x_h(t), u_h(t))dt + \psi(x_h(T))
= \int_s^{s+h} L(t, x(t), u(t))dt + \int_s^T L(t, \tilde{x}(t), \tilde{u}(t))dt + \psi(\tilde{x}(T)).
\]

(3.24)

Since (3.24) holds for any \(\tilde{u} \in \mathcal{U}_d(s + h, x(s + h))\), the value function also satisfies

\[
V(s, y) \leq \int_s^{s+h} L(t, x(t), u(t))dt + \inf_{\tilde{u} \in \mathcal{U}_d(s+h, x(s+h))} \left\{ \int_{s+h}^T L(t, \tilde{x}(t), \tilde{u}(t))dt + \psi(\tilde{x}(T)) \right\} 
= \int_s^{s+h} L(t, x(t), u(t))dt + V(s + h, x(s + h))
\]

(3.25)

which is (3.20).

Next, suppose \(V\) is differentiable at \((s, y)\). Subtract \(V(s, y)\) from both sides of (3.20) and divide by \(h\) to get

\[
0 \leq \frac{1}{h} \int_s^{s+h} L(t, x(t), u(t))dt + \frac{V(s + h, x(s + h)) - V(s, y)}{h}
\]

(3.26)

Let \(\lim_{h \to 0} u(s + h) = v\) and recall \(x(s) = y\) and \(\frac{d}{ds} x(s) = f(s, y, v)\). Letting \(h \to 0\) in (3.26) and applying the chain rule for differentiation gives (3.21).

Now suppose \(u^*\) in \(\mathcal{U}_d(s, y)\) is optimal. Similar arguments to those establishing (3.25) can be used to verify that

\[
V(s, y) = \int_s^{s+h} L(t, x^*(t), u^*(t))dt + V(s + h, x^*(s + h)).
\]

(3.27)

A revision of the earlier argument upon substitution of equation (3.20) by equation (3.27) and using equalities instead of inequalities, \(u^*\) achieves the equality in (3.21) and the HJB equation (3.22) holds. \(\square\)
CHAPTER 3. OPTIMAL CONTROL OF ELECTRIC SPACE HEATING

Generalised solutions to the HJB equation.

Definition 3.6.1. Let \([0, T]\) with \(T > 0\) be a fixed interval and \(Q = [0, T) \times \mathbb{R}\). A function \(W\) is called a generalised solution to the HJB equation in \(Q\) if \(W\) is locally Lipschitz and satisfies (3.22) for almost every \((s, y) \in Q\).

Definition 3.6.1 is motivated by Theorem 3.6.1 and, by Rademacher’s Theorem [53, p. 296], the fact that locally Lipschitz functions are differentiable almost everywhere. Theorem 3.6.2 below gives sufficient conditions under which the value function \(V(s, y)\) is such a generalised solution.

Theorem 3.6.2. Suppose Assumptions 1 and 2 are satisfied. Then the value function \(V(s, y) = \inf_{u \in U_d(s, y)} J(s, y, u)\) is locally Lipschitz on \([0, T] \times \mathbb{R}\).

Proof. The proof is the same as Theorem IV.4.2 of [57] and is provided here for completeness. The goal is to show that \((s, y) \mapsto J(s, y, u)\) is locally Lipschitz independently of \(u \in U_d(s, y)\). Since \(V(s, y)\) would then be the infimum of a family of functions with the same Lipschitz constant, the result follows.

Take any \(s_1, s_2 \in [0, T]\) and suppose without loss of generality that \(s_2 > s_1\). Let \(O_a\) (resp. \(C_a\)) be the open (resp. closed) “ball” of radius \(a > 0\) centred at 0:

\[
O_a = \{y \in \mathbb{R}: |y| < a\}, \quad C_a = \{y \in \mathbb{R}: |y| \leq a\}. \tag{3.28}
\]

Take any Lebesgue-measurable function \(u: [0, T] \to U\) and, for \(i \in \{1, 2\}\), let \(x^i = x(\cdot ; s_i, y_i)\) be the solution to (A.4) corresponding to the restriction of \(u\) to \([s_i, T]\) and such that \(x^i(s_i) = y_i\) for \(y_i \in O_a\). The difference in performance values satisfies

\[
\begin{align*}
|J(s_2, y_2, u) - J(s_1, y_1, u)| &\leq \int_{s_2}^{T} |L(t, x^2(t), u(t)) - L(t, x^1(t), u(t))| \, dt \\
&\quad + \int_{s_1}^{s_2} |L(t, x^1(t), u(t))| \, dt \\
&\quad + |\psi(x^2(T)) - \psi(x^1(T))|. \tag{3.29}
\end{align*}
\]

Using the Lipschitz condition in \(x\) for \(L\) and \(\psi\) together with the Lipschitz condition (A.8) of Lemma A.1.2 in the appendix (also recall Remark 3.3.1), the following holds on \(C_a\):

\[
\begin{align*}
|L(t, x^2(t), u(t)) - L(t, x^1(t), u(t))| &\leq K_L K \left[|y_2 - y_1|^2 + |s_2 - s_1|^2\right]^{1/2} \\
|\psi(x^2(T)) - \psi(x^1(T))| &\leq K_{\psi} K \left[|y_2 - y_1|^2 + |s_2 - s_1|^2\right]^{1/2} \tag{3.30}
\end{align*}
\]
where \( s_2 \leq t \leq T, \ K > 0 \) is a constant and \( K_L \) and \( K_\psi \) are the respective Lipschitz constants for \( L \) and \( \psi \). From (3.30), (3.29) and a bound \( M_L \) on the function \( L \), one gets

\[
|J(s_2, y_2, u) - J(s_1, y_1, u)| \leq (TK_L + K_\psi)K \left[ |y_2 - y_1|^2 + |s_2 - s_1|^2 \right]^{1/2} + M_L|s_2 - s_1|
\]

whence follows,

\[
|J(s_2, y_2, u) - J(s_1, y_1, u)| \leq \tilde{K} \left[ |y_2 - y_1|^2 + |s_2 - s_1|^2 \right]^{1/2} \tag{3.31}
\]

for some constant \( \tilde{K} \) which is independent of the control \( u \). Letting \( a \to \infty \) shows that \( J(s, y, u) \) is locally Lipschitz in \( (s, y) \) independently of \( u \in U_d(s, y) \).

### 3.6.2 Viscosity solutions to Hamilton-Jacobi-Bellman PDEs

Strong assumptions on the smoothness of a solution to the HJB equation (3.22) are sometimes imposed in order to establish sufficient (and necessary) conditions for optimality, see for instance [57], Theorem IV.4.4, or Theorem I.5.1 of [58]. However, there are many control problems of interest where the value function is not smooth, and therefore does not satisfy the HJB equation in the classical \( C^1 \) sense (see Example IV.2.3 of [138] for instance). The concept of a viscosity solution, which appeared in the 1980s, provides a consistent framework for the interpretation and analysis of solutions to (3.22) for a large class of optimal control problems. The definitions and results given here for viscosity solutions in the first-order (deterministic case) are taken from [58].

**Definition 3.6.2** ([58], Definition II.4.2). Let \( O \) be an open set of \( \mathbb{R} \), \( Q = [0, T) \times O \), \( W \in C(\bar{O}) \) and \( F \) be a continuous function satisfying:

\[
F(s, y, p, A + B, V) \leq F(s, y, p, A, V)
\]

for all \( (s, y) \in Q, \ p \in \mathbb{R}, \ A \in \mathbb{R}, \ B \geq 0, \ V \in \mathbb{R} \). Suppose for initial data \( (s, y) \in Q \), the following equation is given

\[
- \frac{\partial}{\partial s}W(s, y) + F(s, y, W_y(s, y), W_{yy}(s, y), W(s, y)) = 0.
\]
(a) One says that $W$ is a viscosity subsolution of equation (3.33) in $Q$ if for each $w \in C^\infty(Q)$,
\[
- \frac{\partial}{\partial s}w(\bar{s}, \bar{y}) + F(\bar{s}, \bar{y}, w_y(\bar{s}, \bar{y}), w_{yy}(\bar{s}, \bar{y}), w(\bar{s}, \bar{y})) \leq 0 \tag{3.34}
\]
at every $(\bar{s}, \bar{y}) \in Q$ which is a local maximum of $W - w$ on $\bar{Q}$, with $W(\bar{s}, \bar{y}) = w(\bar{s}, \bar{y})$.

(b) $W$ is a viscosity supersolution of equation (3.33) in $Q$ if for each $w \in C^\infty(Q)$,
\[
- \frac{\partial}{\partial s}w(\bar{s}, \bar{y}) + F(\bar{s}, \bar{y}, w_y(\bar{s}, \bar{y}), w_{yy}(\bar{s}, \bar{y}), w(\bar{s}, \bar{y})) \geq 0 \tag{3.35}
\]
at every $(\bar{s}, \bar{y}) \in Q$ which is a local minimum of $W - w$ on $\bar{Q}$, with $W(\bar{s}, \bar{y}) = w(\bar{s}, \bar{y})$.

(c) $W$ is a viscosity solution of equation (3.33) in $Q$ if it is both a viscosity subsolution and viscosity supersolution to equation (3.33) in $Q$.

**Remark 3.6.1.** Definition 3.6.2 for viscosity (sub/super) solutions encompasses both first order and second order PDEs. The function $w$ in the definition belongs to a space $\mathcal{D}$ of test functions [58, p. 65]. By Remark II.6.1 of [58], in the case of first order non-linear PDEs one can use $\mathcal{D} := C^1(Q) \cap \mathcal{B}(\bar{Q})$, where $\mathcal{B}(\bar{Q})$ represents the space of real-valued functions defined on $\bar{Q}$ which are bounded from below.

**Remark 3.6.2.** Boundary conditions were excluded in the above definitions of viscosity super and subsolutions. If the boundary condition for a solution to (3.33) is $W = \psi$, then $W \leq \psi$ (respectively, $W \geq \psi$) is used in the case of subsolutions (respectively, supersolutions).

Recall the Hamiltonian function defined in (3.15). The following equivalent and convenient expression for the HJB equation (3.22) can be used:
\[
\begin{cases}
- \frac{\partial}{\partial s}W(s, y) + H(s, y, W_y(s, y)) = 0, \quad \forall (s, y) \in Q; \\
W(T, y) = \psi(y), \quad y \in \mathbb{R}.
\end{cases} \tag{3.36}
\]

Writing the HJB equation in this form shows its relation to equation (3.33) in Definition 3.6.2 with $H \equiv F$. 
Proposition 3.6.3 ([58], Theorem II.7.1). Let \( O \) and \( Q \) be as in Definition 3.6.2 with \( O = \mathbb{R} \) and \( \bar{Q} = [0, T] \times \mathbb{R} \). Suppose controls \( u = u(\cdot) \) are taken from the class \( \mathcal{U}_d(s, y) \) in Definition 3.2.1. Also assume that \( f \) and \( L \) are continuous functions on \( \bar{Q} \times U \), \( U \) is bounded and \( f \) satisfies the hypotheses of Lemma A.1.2. Then the value function \( V = V(s, y) \) is a viscosity solution to the dynamic programming equation (3.36) in \( Q \) provided it is continuous on \( \bar{Q} \).

Theorem 3.6.4. The value function \( V(s, y) = \inf_{u \in \mathcal{U}_d(s, y)} J(s, y, u) \) with \( J(s, y, u) \) defined in (3.4) is a continuous viscosity solution to (3.22) in \( Q \).

Proof. By Theorem 3.6.2, the value function is locally Lipschitz on \( \bar{Q} \) which is sufficient to conclude it is continuous there. Proposition 3.6.3 can then be applied to establish the result.

Remark 3.6.3. Theorem 3.6.4 does not state that the value function is the unique viscosity solution to the dynamic programming equation. The reference [58] proves uniqueness under the additional assumption that the value function is bounded and uniformly continuous on \( \bar{Q} \) – see their Corollary II.9.1 and Remark II.9.1. However, Theorem IV.2.5 of [138] shows that the value function is the unique continuous viscosity solution to (3.4) under assumptions similar to those in Section 3.3.

3.6.3 Necessary and sufficient conditions for optimality

If there is a \( C^1 \) solution to the HJB equation (3.36), then it is straightforward to obtain a verification theorem establishing necessary and sufficient conditions for a control \( u^* \in \mathcal{U}_d(s, y) \) to be optimal. In particular, this candidate optimal control, if it exists, maximises the Hamiltonian in (3.36) (equivalently, minimises (3.22)) for almost all \( t \in [s, T] \). However, the discussion in Chapter III, Section 3.4 of [7] confirms that viscosity theory also allows for the verification of an optimal control \( u^* \in \mathcal{U}_d(s, y) \), albeit using weaker notions of differentiability.

Let \( u^* \in \mathcal{U}_d(s, y) \) be a candidate optimal control and \( x^* \) be the associated state variable. Suppose for simplicity that the value function \( V \) is differentiable at \( (t, x^*(t)) \).
CHAPTER 3. OPTIMAL CONTROL OF ELECTRIC SPACE HEATING

Then, using the definitions for \( f \) and \( L \), the control \( u^* \) satisfies

\[
\begin{align*}
  u^*(t) &= \arg \min_{v \in U} \left\{ V_y(t, x^*(t)) f(t, x^*(t), v) + L(t, x^*(t), v) \right\} \\
  &= \arg \min_{v \in U} \left\{ V_y(t, x^*(t)) \left[ -\frac{1}{\tau}(x^*(t) - T^u - T^g v) \right] + v\tilde{S}(t) + \kappa_1 B(x^*(t)) \right\} \\
  &= \arg \min_{v \in U} \left\{ v[V_y(t, x^*(t))\frac{T^g}{\tau} + \tilde{S}(t)] \right\}.
\end{align*}
\] (3.37)

Equation (3.37) provides the following characterisation for an optimal control, which is similar to the one derived from Pontryagin’s Maximum Principle (cf. (3.19)),

\[
u^*(t) = \begin{cases} 
0, & \text{if } \tilde{S}(t) > -V_y(t, x^*(t))\frac{T^g}{\tau}; \\
1, & \text{if } \tilde{S}(t) < -V_y(t, x^*(t))\frac{T^g}{\tau}.
\end{cases}
\] (3.38)

Suppose for \( s \leq t < T \) that the optimal \( u^*(\cdot) \) is right-continuous and the value function \( V \) is differentiable at \((t, x^*(t))\) under the optimal trajectory. Theorem I.6.2 of [58] shows that the function \( P(\cdot) \) defined by

\[
P(t) = V_y(t, x^*(t)),
\] (3.39)

satisfies the adjoint equation (3.12) and the transversality condition. Furthermore, the Hamiltonian is maximised for almost all \( t \in [s, T] \) (cf. (3.18)).

3.6.4 Stochastic control and viscosity approximation of the HJB equation

It can be quite difficult to rigorously justify the dynamic programming procedure for discontinuous feedback controls such as (3.38). However, by perturbing the indoor temperature \( x \) with small white-noise (zero-mean Guassian) errors, one can obtain a stochastic version of the optimal control problem in Section 3.2 which is easier to analyse. These results and their relation to the deterministic problem are discussed below. More details can be found in the article [56] and Chapter VI, Section 9 of [57].

Stochastic indoor temperatures.

Let \((s, y) \in [0, T) \times \mathbb{R} \) be given and \(( \Omega, \mathcal{F} ) \) be a given measurable space equipped with a filtration \( \mathbb{F} = (\mathcal{F}_t)_{0 \leq t \leq T} \). Let \( u : \tilde{Q} \to U \) be a Borel-measurable function. For each \( \varepsilon > 0 \), the perturbed indoor temperature is a stochastic process \( x^\varepsilon \) which is assumed
to be an \( \mathbb{F} \)-adapted solution to the following stochastic differential equation (SDE):

\[
\begin{aligned}
\mathbb{P}_s y &\quad \text{almost surely,} \\
\left\{
\begin{array}{l}
x^\varepsilon(s) = y; \\
dx^\varepsilon(t) = f(t, x^\varepsilon(t), u(t, x^\varepsilon(t))) \, dt + \sqrt{\varepsilon} \, dW(t), \\
t \geq s
\end{array}
\right.
\end{aligned}
\tag{3.40}
\]

where \( \mathbb{P}_s y \) is a probability measure on \( (\Omega, \mathcal{F}) \), \( f \) is the system velocity defined in (3.2), and \( W = (W(t))_{0 \leq t \leq T} \) is a Brownian motion. Let \( \mathcal{U} \) denote the set of admissible feedback controls which, among other criteria, ensures a solution \( x^\varepsilon \) to (3.40) exists and is unique in probability law – see Chapter VI, Section 3 of [57] for more details.

**Remark 3.6.4.** Proposition A.1.3 in the appendix shows that for any initial data \((s, y) \in [0, T) \times \mathbb{R}\) and all feedback controls \( u : \tilde{Q} \to U \) there exists a solution \( x^\varepsilon \) to (3.40) which is unique in probability law. Taking the temperature evolution equation (3.2) into account, the process \( x^\varepsilon \) described by the SDE (3.40) is an Ornstein–Uhlenbeck process with a controlled, time-dependent mean-reversion level.

**Performance criterion and control problem.**

The performance criterion for the stochastic control problem is given by:

\[
J^\varepsilon(s, y, u) = \mathbb{E}_s y \left[ \int_s^T L(t, x^\varepsilon(t), u(t)) \, dt + \psi(x^\varepsilon(T)) \right], \quad u \in \mathcal{U}
\tag{3.41}
\]

where \( u(t) := u(t, x^\varepsilon(t)), t \in [s, T] \), and the costs \( L \) and \( \psi \) are as in Section 3.2. The value function \( V^\varepsilon \) for the stochastic control problem is

\[
V^\varepsilon(s, y) = \inf_{u \in \mathcal{U}} J^\varepsilon(s, y, u).
\tag{3.42}
\]

**Summary of the relevant results.**

Under suitable assumptions on \( L \) and \( \psi \), it is possible to show that \( V^\varepsilon \) is the unique classical \((C^{1,2})\) solution to the second order HJB equation

\[
\begin{aligned}
\left\{
\begin{array}{l}
-W_s - \varepsilon W_{yy} + H(s, y, W_y) = 0, \\
W(T, y) = \psi(y),
\end{array}
\right. \\
(s, y) \in Q,
\end{aligned}
\tag{3.43}
\]

where \( H \) is the Hamiltonian function appearing in (3.15) above – see Chapter IV, Theorems 4.2 and 4.3 of [58]. It is then possible to prove that there exists an optimal
admissible feedback control $u^* \in \mathcal{U}$ for the stochastic control problem. This result is actually simpler to prove than in the corresponding deterministic case [57, p. 170].

One can also show that as $\varepsilon \to 0$, $V^\varepsilon$ converges uniformly to the value function $V$ for the deterministic control problem, and $u^{*,\varepsilon}$ converges uniformly to an optimal feedback control $u^*$ on compact subsets within a suitable region of $Q = [0, T) \times \mathbb{R}$. Since $V^\varepsilon$ is a classical (and therefore viscosity) solution to (3.43) for each $\varepsilon > 0$, a stability result such as Lemma II.6.2 of [58] or Proposition 4.5.9 of [138] shows that $V := \lim_{\varepsilon \downarrow 0} V^\varepsilon$ is a viscosity solution to the limit equation of (3.43) – the first-order HJB equation (3.36). This approach is referred to as the “viscosity approximation” of the first-order HJB equation, or “vanishing viscosity” of the second-order HJB equation.

3.7 Numerical simulations

This section presents numerical results for the deterministic optimal control problem presented in Section 3.2. The results were produced using a backward induction dynamic programming procedure which is discussed in Section A.2 of the appendix. The cost associated with a control strategy obtained from the DP algorithm was calculated using a discretised version of the performance criterion. The cost of perfect comfort control in which indoor temperature are kept constant at an ideal value $T$ can then be used as an independent benchmark (see [27, p. 759] for other measures).

Remark 3.7.1. Values for the perfect comfort control $u(\cdot)$ are readily estimated using the system dynamics (3.2). In particular, if the initial temperature is equal to the ideal one $T$, setting $f(t, x(t), u(t)) = 0$ in equation (3.2) leads to

$$u(t) = \frac{x(t) - T^u}{T^s}, \quad \forall t \in [s, T].$$

3.7.1 Electricity spot price data

Market Index Price.

The Market Index Price (MIP) is a representative electricity spot market price used for half-hourly imbalance settlement in the British electricity market [46]. The simulations described below utilised MIP data for 1st January, 2005 up to 31st December, 2014.
obtained from the ELEXON portal [47]. The standard deviation of the MIP for each
day in the sample was calculated and served as a measure of price variability. The
trend of the standard deviation is shown in Figure 3.1 below whereas Table 3.1 shows
the categories used to investigate the relationship between price variability and cost
savings.

Figure 3.1: Price variability trend and distribution from 2005 to 2014.

(a) Price variability trend from 2005 to 2014. (b) Variation of MIP daily standard deviation
from 2005 to 2014.

(c) Percentage occurrence for five categories of
price variability.

Table 3.1: Price profile categories according to standard deviation

<table>
<thead>
<tr>
<th>Category</th>
<th>I</th>
<th>II</th>
<th>III</th>
<th>IV</th>
<th>V</th>
</tr>
</thead>
<tbody>
<tr>
<td>Standard Deviation</td>
<td>&lt; 7.5</td>
<td>[7.5, 15)</td>
<td>[15, 22.5)</td>
<td>[22.5, 30)</td>
<td>≥ 30</td>
</tr>
</tbody>
</table>

Figure 3.1(a) reveals that market prices can be quite volatile, although their volatil-
ity has been comparatively low in the past three to four years. It also shows that price
volatility is seasonal, with low price variability more likely in spring and summer than
in winter, and on weekends rather than on weekdays. These well known characteris-
tics of electricity prices (see [135]) are highlighted again in the appendix, Table E.1.
Figure 3.1(b) highlights the variation of the standard deviation within the sample. It reveals that the bulk of the sample consists of days with daily price standard deviation between 6 and 12 (£ per MWh). Figure 3.1(c) above highlights the distribution of daily price profiles according to the categories in Table 3.1.

**Representative price profiles.**

The daily MIP profiles were grouped according to the five categories in Table 3.1 and the median value of the price per settlement period was calculated. The representative profiles thus formed are depicted in Figure 3.2 below. The phrase “Profile x”, where “x” refers to one of I–V, is used instead of “Category x representative profile”.

![Median price profiles according to variability](image)

Figure 3.2: Median price profile (2005–2014) for each category of price variability.

The median profiles are used in the majority of the simulations below. They retain the characteristic afternoon and evening local maxima that occur around 12:00 and 17:00 respectively (the so-called *on-peak* periods). These peaks are most pronounced for Profile V and become less noticeable as the price variability decreases. The impact of this feature on the controlled temperature trajectories is highlighted below in Section 3.7.4.

**Remark 3.7.2.** The standard deviation values for Profiles I to V were, to the nearest integer, 4, 8, 14, 20 and 33 respectively. The standard deviations for Profiles III and IV were therefore below the ranges given in Table 3.1. This does not, however, greatly affect subsequent comparison of their results.
### 3.7.2 Thermal penalty function parameters

Section 1.3.1 in the introductory chapter suggested that bathtub-shaped functions provide an appropriate model for the thermal discomfort penalty $B(\cdot)$. The simulations below used the following definition for $B(\cdot)$: for $x \in \mathbb{R}$,

$$
B(x) = B(x; \mathcal{T}, \delta, \Gamma) = \begin{cases} 
0, & \text{if } |x - \mathcal{T}| \leq \delta \\
\frac{(x-\mathcal{T})^2-\delta^2}{\Gamma^2-\delta^2}, & \text{if } \delta < |x - \mathcal{T}| < \Gamma \\
1, & \text{if } |x - \mathcal{T}| \geq \Gamma 
\end{cases} \quad (3.44)
$$

where $0 \leq \delta < \Gamma$ and $\mathcal{T} \in \mathbb{R}$ are parameters. Figure 2.1 depicting the bathtub shape in the thesis introduction was generated using this specification. The parameter $\delta$ in (3.44) defines the deadband symmetrically about the temperature setpoint $\mathcal{T} = 22^\circ C$. The maximum service loss $\Gamma > \delta$ determines a (symmetric) tolerance level for temperature deviations according to pre-defined saturation points.

The references [5, 74] provide general advice on the legal and typical temperature ranges in working and classroom environments. The legal minimum for workplaces is $16^\circ C$ if the type of work is not physical, but a legal maximum temperature is not defined. For environments where physical activity is low, such as sick rooms in schools, the generally accepted temperature is about $21^\circ C$. Consequently, the following ranges for the deadband were deemed suitable for the simulations: $21^\circ C$ to $23^\circ C$, $20^\circ C$ to $24^\circ C$ and $19^\circ C$ to $25^\circ C$.

### 3.7.3 Simulation procedure

Each simulation below relates to an optimisation problem for one day between the hours of 07:00 and 22:00. These are the times that the householder’s demand for electric space heating is assumed to be flexible and can be optimised. Outside of this period the indoor temperature is kept as close as possible to the ideal setting $\mathcal{T}$ using other means not modelled below.

**Parameter settings.**

Parameter values which were common to the simulations are discussed in the following lines. Table 3.2 summarises the discussion. Let the interval $[0, T]$ represent the control
Table 3.2: Default parameter values used to obtain approximate dynamic programming control policies.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Temperature set point ($T$)</td>
<td>22 °C</td>
</tr>
<tr>
<td>Deadband ($\delta$)</td>
<td>2 °C</td>
</tr>
<tr>
<td>Saturation ($\Gamma$)</td>
<td>$2\delta$</td>
</tr>
<tr>
<td>Min and max temperatures</td>
<td>$x_{\text{min}} = T - (\Gamma + 0.1)$, $x_{\text{max}} = T - (\Gamma + 0.1)$</td>
</tr>
<tr>
<td>Initial indoor temperature ($y$)</td>
<td>22 °C</td>
</tr>
<tr>
<td>Ambient temperature ($T^a$)</td>
<td>3.8 °C</td>
</tr>
<tr>
<td>Thermal gain ($T^g$)</td>
<td>30 °C</td>
</tr>
<tr>
<td>Thermal time constant ($\tau$)</td>
<td>480 minutes</td>
</tr>
<tr>
<td>Time step ($h$)</td>
<td>$\frac{1}{30}$</td>
</tr>
<tr>
<td>Spatial discretisation ($\triangle x$)</td>
<td>0.001</td>
</tr>
<tr>
<td>Weighting constants</td>
<td>$\kappa_1 = 1$, $\kappa_2 = 5$</td>
</tr>
</tbody>
</table>

horizon with 0 and $T > 0$ representing the times 07:00 and 22:00 respectively. The interval $[0, T]$ can be partitioned into $M$ non-overlapping subintervals $[t_{i-1}, t_i)$, $i = 1, \ldots, M$ of equal length $h > 0$, where $M$ is a multiple of 30 – the number of half-hour periods between 07:00 and 22:00 – and $0 = t_0 < t_1 < \ldots < t_M = T$. A single half-hour period between 07:00 and 22:00 therefore coincides with intervals $[t_i \frac{M}{30}, t_{(i+1)} \frac{M}{30})$ for $i = 0, \ldots, 29$. The time step $h$ for the simulations was set to $\frac{1}{30}$, which is a resolution of 1-minute in real time.

The deadband ($\delta$) and saturation ($\Gamma$) parameters for the running cost satisfied $1 \leq \delta \leq 3$ and $\Gamma = 2\delta$. The value of the indoor temperature was represented using discrete values ranging from $x_{\text{min}} = T - (\Gamma + 0.1)$ to $x_{\text{max}} = T + (\Gamma + 0.1)$ according to a uniform spacing $\triangle x = 0.001$. This state space was chosen to take advantage of the bathtub penalty function (3.44) being constant outside of the interval $(T - \Gamma, T + \Gamma)$.

A separate parametrisation of the terminal cost was used in order to heavily penalise trajectories which terminate away the setpoint $T$. In order to achieve the desired effect, the corresponding deadband and saturation parameters, which are denoted by $\delta'$ and $\Gamma'$ respectively, had to be sufficiently small whereas the weighting factor $\kappa_2$ had to be sufficiently large. The values $\delta' = 0$, $\Gamma' = 0.1$ and $\kappa_2 = 5$ worked reasonably well in numerical experiments. The other weighting factor $\kappa_1$ in the running cost was set
The initial value of the indoor temperature and the set point $T$ were both set to 22°C. The ambient temperature $T_a$ was assumed to be constant, and a default value of 3.8°C was used. This particular value is the 1981-2010 average outdoor temperature for winter [98]. The thermal gain $T_g$ of the heater was given a default value of 30°C, which is the same value used by [102, p. 2] in their study of heating and cooling loads.

Remember that the time constant $\tau$ determines building’s thermal responsiveness to changes in the surrounding temperature. The highest value of $\tau$ used in the simulations was 600 minutes. This maximum value is higher than the value of 64 minutes used in [75, p. 4145] and 180 minutes in [102, p. 2]. It is, however, much lower than the value of 25 hours (1500 minutes) used in [27, p. 761] and some of the values for domestic buildings quoted in [3, 114]. Note that the algorithm uses the value $\tau_{30}$ so that the thermal time constant has same unit of time as the spot price.

### 3.7.4 Simulation Results

#### Impact of Price Variability

Figure 3.3 below summarises the performance of the optimal control strategies for the default simulation parameters. Figure 3.3(a) shows the financial cost savings achieved by the dynamic programming algorithm for the representative profiles. The box plots in Figure 3.3(b) show the interquartile range of financial cost savings per variability category for all of the MIP price profiles. A horizontal line within each box identifies the median whilst minimum and maximum values are recorded on the whiskers.

Figure 3.3(a) suggests that the achievable savings increases with respect to the price variability. This trend is confirmed for the entire sample by Figure 3.3(b), which shows that the median values of the savings increased by at least 1% in tandem with the price variability category. Figure 3.3(b) also shows that the savings can be quite modest (5%) for low variability profiles, whereas they can be quite significant (45%) for high variability profiles. It reveals furthermore that, in the majority of the sample, the dynamic programming algorithm was able to achieve at least 10% savings relative to the financial cost of perfect comfort. Despite differences in data and simulation procedures, these savings are not too far off from those reported in [27, p. 762].
CHAPTER 3. OPTIMAL CONTROL OF ELECTRIC SPACE HEATING

(a) Savings achieved by dynamic programming control strategies for representative profiles. (b) Savings achieved by dynamic programming control strategies for all profiles from 2005 to 2014.

Figure 3.3: The impact of price variability on the percentage financial cost savings.

Cost optimal indoor temperature trajectories.

The controlled indoor temperature trajectories (henceforth “controlled trajectories”) for four representative price profiles are illustrated in Figure 3.4 below. The controlled trajectory associated with Profile II was omitted since it did not differ much from the one for Profile I.

The results show that the controlled trajectories adhered to the implicit state constraints on the temperature imposed by the bathtub penalty function. This was not surprising since optimal control problems can account for state constraints by simply including a very large penalty for violating them [73]. However, the results also suggest a fundamental preheating control strategy:

1. the indoor temperature (and therefore consumption) is increased to a point within the deadband prior to the occurrence of local maxima in the spot price;

2. there is little to no consumption during the periods of locally high prices so that the indoor temperature decreases naturally to another point in the deadband at near zero cost.

This strategy is seen clearly in the results for Profiles III to V. The timing, frequency and duration of these preheating periods also appear to vary according to the price variability.

Figures 3.4(c) and 3.4(d) show that the controlled trajectories can cover most of the range of comfortable temperatures (deadband) when price variability is high. Figure 3.4(a), however, shows that the indoor temperature may be kept close to the
minimum value of the deadband for most of the control horizon. In both cases the algorithm seems to keep the indoor temperature at the edge of the deadband whenever prices are either relatively flat or on a downward trend. This might be explained by a lack of incentive to change the indoor temperature at those times since the prices are not changing unfavourably.

**Impact of the deadband size.**

The previous section highlighted the location, size and frequency of price maxima as possible determinants of the achievable savings. There are other factors, many of them physical, that also influence the potential for reduced energy costs. The temperature deadband, which is an indicator of the customer’s flexibility, is one of these key physical elements determining the potential cost savings.

Figure 3.5 displays the financial cost savings attained for each representative profile
as the deadband parameter $\delta$ was increased from 1 to 3. The other parameters were kept fixed according to the values in Table 3.2. The results in this case show that the savings increased linearly with $\delta$ over the range investigated, thereby confirming a trade-off between thermal comfort and cost savings: the householder saves more on consumption costs if he is willing to bear temperatures further away from the setpoint.

Numerical experiments verified that the controlled indoor temperatures still adhered to the constraints imposed by the bathtub penalty function. These results are shown in Figure 3.6 below for Profile III. These illustrations also confirm that the “strategies”, preheating and otherwise, identified in Section 3.7.4 were mostly unperturbed by the change in deadband.

![Figure 3.5: The impact of deadband size on financial cost savings.](image)

![Figure 3.6: Controlled trajectories associated with Profile III for various deadband sizes.](image)
Impact of the thermal time constant.

The thermal time constant $\tau$, which is closely related to the building’s insulation, is expected to have a significant impact on the achievable cost savings. This is because it determines how fast (or slow) the building responds to changes in the surrounding temperature. Its impact was investigated by comparing the simulation results for values of $\tau$ between 60 and 600 in increments of 10.

![Cost savings versus thermal time constant](image)

Figure 3.7: The impact of the thermal time constant on cost savings.

As pointed out in [114], there are potential gains in the savings if the thermal time constant is increased. However, Figure 3.7 shows that these gains are only realised if prices are sufficiently variable. For all profiles the cost savings increased with $\tau$ for values of $\tau$ less than 120. As $\tau$ was increased beyond 120, the cost savings exhibited different behaviour (on average) according to the price profile:

- for Profiles I and II the cost savings quickly plateaued and then went on a decreasing trend;

- for Profile III the cost savings increased moderately and the rate of increase was noticeably slower at high values of $\tau$;

- for Profiles IV and V the cost savings increased rapidly, but there is an eventual slow down in this increase for Profile IV.

These seemingly counter-intuitive observations might be explained as follows. Recalling Figure 3.4 concerning the controlled trajectories, when price variability is low the indoor temperature is kept close to the lower end of the deadband for most of the control horizon. The preheating strategies in the default case ($\tau = 180$) had the
capability to then quickly increase the internal temperature. This strategy is more difficult to achieve for high values of $\tau$ since the temperature evolution is much slower, which can be understood intuitively by the insulation analogy and mathematically by the evolutionary equation (3.2). This explains why profiles with low variability were adversely affected by the increase in $\tau$.

Alternatively, when price variability is high the indoor temperature assumes more values within the deadband due to the substantial differences between price minima and maxima. Slow temperature evolution is favourable in this case as long as the temperature can still be raised quickly enough for preheating strategies. Therefore, one can surmise that the cost savings will start to decrease for Profiles IV and V when the value of $\tau$ hits a critical level beyond 600.

**Impact of the thermal gain.**

Figure 3.8: Impact of the thermal gain parameter $T^g$ on the estimated financial cost savings (%).

The dynamic programming algorithm was run with the default parameters in Table 3.2 and values of the thermal gain parameter $T^g$ between 20°C and 60°C in steps of 5°C. The impact of the change in $T^g$ on the % cost savings was investigated and Figure 3.8 records the outcome. The numerical experiment reveals that the cost savings increased in tandem with $T^g$ for values up to about 35°C, and then decreased slightly for higher values of $T^g$.

Remember that $T^g$ affects the temperature gain of the heater. The increased cost savings may then be attributed to the ability to avoid high prices by performing preheating strategies more effectively. Figure 3.9 below supports this claim by illustrating the trajectories under Profile III with $T^g \in \{20, 40, 60\}$. The slopes of the trajectory
plot on ascents are noticeably steeper for $T^g = 60$ as opposed to $T^g = 20$, and the case $T^g = 40$ is somewhat in the middle. The preheating strategy was most effective in the case $T^g = 60$. Also notice for higher values of $T^g$ that a shorter time is needed to control the temperature for it to terminate near the setpoint. This highlights a relation between the parameter $T^g$ and the terminal penalty, as well as the combined effect they have on the % cost savings.

Figure 3.9: Controlled trajectories associated with Profile III for different values of the thermal gain $T^g$.

**Impact of the ambient temperature.**

Figure 3.10: Impact of the ambient temperature parameter $T^a$ on the estimated financial cost savings (%) of approximated optimal control strategies

Figure 3.10 shows the impact that the value of the average ambient temperature $T^a$ can have on the reported cost savings. The results correspond to values of $T^a$ within the range $-5$°C to $15$°C in steps of $1$°C. The experiment showed that it the savings on consumption costs are higher when $T^a$ is closer to the set-point. Figure 3.11 confirms that, similar to the thermal gain parameter $T^g$, the increase in savings with
respect to $T_a$ is related to the efficacy of preheating strategies. However, less energy is required to keep the indoor temperature within the deadband range in such cases. The reader should therefore be mindful that the increase in the financial cost savings (%) does not necessarily imply a higher cost reduction in nominal terms (£).

![Controlled trajectories associated with Profile III for different values of the ambient temperature $T_a$.](image)

(a) Ambient temperature $-5\,^{\circ}\text{C}$  (b) Ambient temperature $5\,^{\circ}\text{C}$  (c) Ambient temperature $15\,^{\circ}\text{C}$

Figure 3.11: Controlled trajectories associated with Profile III for different values of the ambient temperature $T_a$.

### 3.8 Conclusion

This chapter showed how optimal control can be used to solve an optimisation problem which involves both financial and physical aspects of electricity use. It utilised simple models to describe the cost of an electric heater’s consumption, the evolution of the internal temperature over time, and the householder’s perception of thermal comfort. Within the optimal control framework, Section 3.4 verified that the problem was well posed by confirming the existence of an optimal control that minimises the performance criterion. Sections 3.5 and 3.6 further characterised this control by providing necessary and sufficient conditions for its optimality via Pontryagin’s Maximum Principle and dynamic programming respectively. Both approaches confirmed for the particular model that the heater should be turned on, and at full power, only if the electricity spot price is below a certain time-dependent threshold.

Section 3.7 provided numerical illustrations of the deterministic control strategies based on the dynamic programming approach. The results showed that optimal control strategies can achieve significant savings over the naive strategy of keeping the temperature constant at an ideal value. Moreover, price variability plays a key role in
determining the characteristics of the control strategies and corresponding temperature trajectories in the following ways:

- On days of low price variability, the control strategies kept the temperature close to a value which is comfortable but cheaper to maintain than the ideal one;

- When price variability is high, the control strategies manipulate the temperature within a larger subset of comfortable temperatures;

- In both cases, the strategies pre-empt consumption during periods of local price maxima by consuming electricity to raise the temperature during the preceding cheaper periods.

Finally, the effect of the simulation’s parameters on the financial cost savings was investigated and explained using the behaviour of the corresponding control strategies.
Chapter 4

Optimal switching of electric space heating

4.1 Introduction

The previous chapter demonstrated how optimal control theory could benefit a householder who manages an electric heater’s energy consumption under a variable electricity price. The theory showed that, for the proposed model, the heater is operated optimally by alternating between the full power (1) and off (0) modes according to a variable threshold for the electricity price. Unfortunately, this solution may have adverse physical consequences due to frequent switches between the two operational modes. A simple way to reduce frequent switching is to introduce a penalty for changing the operational mode.

4.1.1 Optimal control with switching costs

Discrete-time notation.

Let $[0, T]$ be the control horizon from the previous chapter and $\{t_0, t_1, \ldots, t_M\}$ represent the times used to partition $[0, T]$ in the numerical examples. Therefore, $0 = t_0 < t_1 < \ldots < t_M = T$ with a constant increment $0 < h = t_k - t_{k-1}$, $k \in \{1, \ldots, M\}$. The time interval $[0, T]$ is replaced below by an enumeration $\mathbb{T} = \{0, 1, \ldots, M\}$ of the partitioning times $\{t_0, t_1, \ldots, t_M\}$, and the customary notation $T \equiv M \geq 1$ for an integer-valued terminal time is used instead. The set of values for the optimal control
CHAPTER 4. OPTIMAL SWITCHING OF ELECTRIC SPACE HEATING

in the previous chapter is denoted by \( \mathbb{I} = \{0, 1\} \).

Temperature evolution model.

Let \( X = \{X_t : t = 0, \ldots, T\} \) represent the sequence of indoor temperature values. This variable is assumed to satisfy the following discrete-time system:

\[
\begin{cases}
X_0 = x \in \mathbb{R}; \\
X_{t+1} = e^{-\frac{h}{\tau}} X_t + (T_f + T^a u_t)(1 - e^{-\frac{h}{\tau}}), & t = 0, \ldots, T - 1
\end{cases}
\]  

(4.1)

where \( u : T \to \mathbb{I} \) is control variable that is constant on each interval \([t_{k-1}, t_k)\), \( k = 1, \ldots, T \). This type of control is described more precisely in Section 4.2 below. The system (4.1) is the discrete-time analogue to the differential equation (3.2).

Costs and performance criterion.

Let \( S = \{S_t : t = 0, 1, \ldots, T\} \) denote the spot price of electricity. Recall the thermal penalty function \( B : \mathbb{R} \to [0, 1] \) and weighting factors \( \kappa_1 \) and \( \kappa_2 \) from (3.3) in the previous chapter. A running cost \( \Psi_i \) and terminal cost \( \Gamma_i \) are defined for \( i \in \{0, 1\} \) by

\[
\Psi_i(S_t, X_t) = h \times \left[ \frac{i \cdot S_t}{\max_{r \in T} S_r} + \kappa_1 B(X_t) \right], \quad t = 0, \ldots, T - 1;
\]

\[
\Gamma_i(X_T) = \kappa_2 B(X_T)
\]

Define costs for switching from \( i \in \{0, 1\} \) to \( j \in \{0, 1\} \) by \( \gamma_{i,j}(\cdot) = 0 \) and

\[
\gamma_{i,j}(\cdot) = \gamma_{j,i}(\cdot) = \gamma; \quad j = 1 - i, \ \gamma > 0
\]

where \( \gamma \) is chosen so that it does not dominate the other costs.

The performance index in this case of switching costs is given informally by

\[
\hat{J}(u; 0, i) = h \sum_{t=0}^{T-1} \left[ u_t \frac{S_t}{\max_{r \in T} S_r} + \kappa_1 B(X_t) \right] + \kappa_2 B(X_T) + \gamma N(u)
\]

\[\approx J^{std}(0, x, u) + \gamma N(u)\]

where \( N(u) \) is the number of switches under the control \( u \), and \( J^{std}(0, x, u) \) is the performance index (3.4) of Chapter 3 evaluated under \( u \) – see (4.43) below for the formal definition. The new objective is to minimise \( \hat{J}(u; 0, i) \) over all \( \mathbb{I} \)-valued \( T \)-indexed controls \( u \).
Layout of the chapter.

This chapter investigates the switching problem above using optimal stopping theory. Section 4.2 formalises the optimisation problem using a more general model with discrete-parameter stochastic processes. Section 4.3 recalls key results on the martingale approach to optimal stopping in discrete time. Section 4.4 constructs an optimal switching control strategy using a hypothetical system of stochastic processes. Existence of the latter is verified afterwards in Section 4.5, and leads to a dynamic programming algorithm for computing the value function and optimal strategy. Section 4.6 presents some numerics for the motivating problem above including an example with stochastic electricity spot prices. Section 4.7 concludes the chapter and Appendix B fills in the details that are omitted from some of the results in this chapter’s text. The author has also uploaded the theoretical results of this chapter to the arXiv preprint server [96].

Remark 4.1.1. The continuous-time, deterministic version of the optimal switching problem can also be studied using partial differential equations. Under suitable assumptions, it is possible to derive a dynamic programming principle and an HJB equation for the value function in the form of a coupled system of quasi-variational inequalities [7, pp. 211–212]. The HJB equation for the optimal control problem in the previous chapter is recovered when the switching costs tend to zero [7, pp. 411–414].

4.2 Discrete-time optimal switching

4.2.1 Definitions

Probabilistic setup.

Let \((\Omega, \mathcal{F}, \mathbb{P})\) be a given complete probability space. The expectation operator with respect to \(\mathbb{P}\) is denoted by \(\mathbb{E}\), and the indicator function of a set or event \(A\) is written as \(1_A\). Let \(\mathbb{T} = \{0, 1, \ldots, T\}\) represent a sequence of integer-valued times with \(0 < T < \infty\). The probability space is equipped with a filtration \(\mathbb{F} = (\mathcal{F}_t)_{t \in \mathbb{T}}\), and it is assumed that \(\mathcal{F}_0\) is the trivial \(\sigma\)-algebra, \(\mathcal{F}_0 = \{\emptyset, \Omega\}\), and \(\mathcal{F} = \mathcal{F}_T\). The notation a.s. stands for “almost-surely”. For a given \(\mathbb{F}\)-stopping time \(\nu\), the notation \(\mathcal{T}_\nu\) is used for the set of \(\mathbb{F}\)-stopping times \(\tau\) such that \(\nu \leq \tau \leq T\) \(\mathbb{P}\)-a.s. (define \(\mathcal{T} \equiv \mathcal{T}_0\)). Martingales,
stopping times and other relevant concepts are understood to be defined with respect to the filtered probability space \((\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})\). The usual convention to suppress the dependence on \(\omega \in \Omega\) is used below.

**Optimal switching definitions.**

The following data for the optimal switching problem are given:

1. A set of operational modes \(\mathbb{I} = \{1, 2, \ldots, m\}\), where \(2 \leq m < \infty\);

2. A reward received at time \(T\) for being in mode \(i \in \mathbb{I}\), which is modelled by an \(\mathcal{F}_T\)-measurable real-valued random variable \(\Gamma_i\);

3. A running reward received while in mode \(i \in \mathbb{I}\), which is represented by a real-valued adapted process \(\Psi_i = (\Psi_i(t))_{t \in \mathbb{T}}\);

4. A cost for switching from mode \(i \in \mathbb{I}\) to \(j \in \mathbb{I}\), which is modelled by a real-valued adapted process \(\gamma_{i,j} = (\gamma_{i,j}(t))_{t \in \mathbb{T}}\).

Define a class of *admissible switching controls* as follows:

**Definition 4.2.1.** Let \(t \in \mathbb{T}\) and \(i \in \mathbb{I}\) be given. An admissible switching control starting from time \(t\) in mode \(i\) is a sequence \(\alpha = (\tau_n, \iota_n)_{n \geq 0}\) with the following properties:

1. For \(n \geq 0\), \(\tau_n \in \mathbb{T}_t\) and satisfies \(t = \tau_0 \leq \tau_1 \leq \tau_2 \leq \ldots \leq T\); if \(n \geq 1\) then \(\mathbb{P}(\{\tau_n < T\} \cap \{\tau_n = \tau_{n+1}\}) = 0\).

2. For \(n \geq 0\), each \(\iota_n : \Omega \rightarrow \mathbb{I}\) is \(\mathcal{F}_{\tau_n}\)-measurable; \(\iota_0 = i\) and \(\iota_n \neq \iota_{n+1}\) \(\mathbb{P}\)-a.s. for \(n \geq 0\).

Let \(\mathcal{A}_{t,i}\) denote the class of admissible switching controls (also called strategies) for the initial condition \((t, i) \in \mathbb{T} \times \mathbb{I}\).

The switching control \(\alpha = (\tau_n, \iota_n)_{n \geq 0}\) models the controller’s decision to switch at time \(\tau_n\), \(n \geq 1\), from the active mode \(\iota_{n-1}\) to another one \(\iota_n\). The condition \(\mathbb{P}(\{\tau_n < T\} \cap \{\tau_n = \tau_{n+1}\}) = 0\) for \(n \geq 1\) means there is at most one switch at each time. The above definition is similar to the one given in [140, p. 145].
Definition 4.2.2. Associated with each \( \alpha = (\tau_n, t_n)_{n \geq 0} \in A_{t,i} \) are the following objects:

- a mode indicator function \( u \): \( \Omega \times \mathbb{T} \to \mathbb{I} \) defined by,

\[
u_s = \sum_{n \geq 0} t_n 1_{\{\tau_n \leq s < \tau_{n+1}\}}, \quad s \geq t. \tag{4.3}
\]

- the (random) total number of switches before \( T \)

\[
N(\alpha) = \sum_{n \geq 1} 1_{\{\tau_n < T\}} \tag{4.4}
\]

Note that the mode indicator function \( u \) may jump immediately from \( i_0(\omega) = i \) to \( i_1(\omega) \neq i \) if \( \tau_1(\omega) = t \).

4.2.2 The optimal switching problem

Define the following performance index for switching controls with initial mode \( i \in \mathbb{I} \) at time \( t \in \mathbb{T} \):

\[
J(\alpha; t, i) = \mathbb{E} \left[ \sum_{s=t}^{T-1} \Psi u_s(s) + \Gamma_{t, N(\alpha)} - \sum_{n \geq 1} \gamma_{t_{n-1}, t_n}(\tau_n) 1_{\{\tau_n < T\}} \right] \quad F_t, \quad \alpha \in A_{t, i} \tag{4.5}
\]

where \( t_{N(\alpha)} \) is the last mode switched to before \( T \). The optimisation problem is to maximise the objective function \( J(\alpha; t, i) \) over all admissible controls \( \alpha \in A_{t,i} \). The value function \( V \) for the optimal switching problem is defined as a random function of the initial time and mode \((t, i)\):

\[
V(t, i) = \text{ess sup}_{\alpha \in A_{t,i}} J(\alpha; t, i). \tag{4.6}
\]

A switching control \( \alpha^* \in A_{t, i} \) is said to be optimal if it achieves the essential supremum in equation (4.6): \( \mathbb{P}\text{-a.s.,} \)

\[
V(t, i) = J(\alpha^*; t, i) \geq J(\alpha; t, i) \quad \forall \alpha \in A_{t,i}.
\]

Remark 4.2.1. Note that the analogous minimisation problem can be treated by negating the performance index (4.5) and value function (4.6). This follows from the relationship between the essential supremum and essential infimum: if \( \Phi \) is a set of random variables on \((\Omega, \mathcal{F}, \mathbb{P})\), the essential infimum of \( \Phi \) with respect to \( \mathbb{P} \) is defined as [59, p. 496]:

\[
\text{ess inf} \Phi = \text{ess inf}_{\varphi \in \Phi} \varphi := - \text{ess sup}_{\varphi \in \Phi} (-\varphi).
\]
Remark 4.2.2. Processes or functions with super(sub)-scripts in terms of the mode indicators \( \{ \nu_n \} \) are interpreted in the following way:

\[
Y^{\nu_n} = \sum_{j \in \mathbb{I}} 1_{\{ \nu_n = j \}} Y^j
\]

\[
\gamma_{\nu_n-1, \nu_n} (\cdot) = \sum_{j \in \mathbb{I}} \sum_{k \in \mathbb{I}} 1_{\{ \nu_{n-1} = j \}} 1_{\{ \nu_n = k \}} \gamma_{j,k} (\cdot)
\]

Note that the summations are finite.

Remark 4.2.3. The costs appearing in equation (4.5) may depend on a state variable \((X_t)_{t \in T}\), and the state variable may also depend on the switching control. The performance index and value function in this case depend on the initial value of the state \(X_0\). Subtleties associated with the existence of the controlled stochastic process and measurability of the value function in this case are not discussed here. However, the book [16] provides an in-depth discussion relating to these issues.

4.2.3 Notation, conventions and assumptions

The convention that \( \sum_{t=v}^{t} (\cdot) = 0 \) for any integers \( t \) and \( v \) with \( t < v \) is used. The following terminology is referred to in later developments:

- For a constant \( p \geq 1 \), let \( L^p \) denote the class of random variables \( Z \) satisfying

\[
\mathbb{E} [|Z|^p] < \infty.
\]

- Similarly, let \( S^p \) denote the class of adapted processes \( X \) satisfying

\[
\mathbb{E} \left[ \max_{t \in T} |X_t|^p \right] < \infty.
\]

Assumptions.

Assumption 3. For each \( i \in \mathbb{I} \), \( \Gamma_i \in L^2 \) and is \( \mathcal{F}_T \)-measurable, \( \Psi_i \in S^2 \), and for every \( j \in \mathbb{I} \), \( \gamma_{i,j} \in S^2 \).

Assumption 3 ensures that the performance index (4.5) is well defined and allows application of optimal stopping theory later. The following standard assumption on the switching costs [65, 140] is also imposed:

Assumption 4. For every \( i, j, k \in \mathbb{I} \) and \( \forall t \in T \): \( \mathbb{P} \)-a.s.,

\[
i. \quad \gamma_{i,i} (t) = 0 \tag{4.7}
\]

\[
ii. \quad \gamma_{i,k} (t) \leq \gamma_{i,j} (t) + \gamma_{j,k} (t), \quad \text{if } i \neq j \text{ and } j \neq k. \tag{4.8}
\]
Condition (4.7) says there is no additional cost for staying in the same mode. The second condition ensures that when going from one mode $i$ to another mode $k$, it is never profitable to immediately visit an intermediate mode $j$. By taking $k = i$ and using (4.7), condition (4.8) also shows that it is unprofitable to switch immediately back and forth between modes.

### 4.3 Optimal stopping and Snell envelopes

This section collects some results on discrete-time optimal stopping problems which are important below.

**Proposition 4.3.1.** Let $U = (U_t)_{t \in T}$ be an adapted, $\mathbb{R}$-valued process such that $U \in S^1$. Then there exists an adapted, integrable $\mathbb{R}$-valued process $Z = (Z_t)_{t \in T}$ such that $Z$ is the smallest supermartingale which dominates $U$. The process $Z$ is called the Snell envelope of $U$ and it enjoys the following properties.

1. For any $t \in T$, $Z_t$ is defined by:

   $$Z_t = \underset{\tau \in T}{\text{ess sup}} \mathbb{E}[U_\tau | \mathcal{F}_t].$$  

   (4.9)

   Moreover, $Z$ can also be defined recursively as follows: $Z_T := U_T$ and $Z_t := U_t \lor \mathbb{E}[Z_{t+1} | \mathcal{F}_t]$ for $t = T - 1, \ldots, 0$.

2. For any $\theta \in T$, the stopping time $\tau^*_\theta = \inf\{t \geq \theta : Z_t = U_t\}$ is optimal after $\theta$ in the sense that:

   $$Z_\theta = \mathbb{E}[U_{\tau^*_\theta} | \mathcal{F}_\theta] = \underset{\tau \in T}{\text{ess sup}} \mathbb{E}[U_\tau | \mathcal{F}_\theta], \quad \text{P - a.s.}$$  

   (4.10)

3. For any $t \in T$ given and fixed, the stopped process $(Z_{r \wedge \tau^*_\theta})_{1 \leq r \leq T}$ is a martingale.

These results are standard and can be found in the references [59, 104, 111].

### 4.4 The verification theorem

This section proposes a probabilistic solution to the optimal switching problem. The approach follows that of [37] in continuous time, which postulates the existence of
a particular system of $m$ stochastic processes and verifies (Theorem 4.4.1) that the components of this system solve the optimal switching problem with given initial conditions. The existence of these candidate optimal processes is proved in the following section (Theorem 4.5.2).

Suppose there exist $m$ real-valued, adapted processes $Y^i = (Y^i_t)_{t \in T}$, $i \in \mathbb{I}$, such that $Y^i \in \mathcal{S}^2$ and

$$Y^i_t = \operatorname{ess sup}_{\tau \in T_i} \mathbb{E} \left[ \sum_{s=t}^{\tau-1} \Psi_i(s) + \Gamma_i \mathbf{1}_{\tau=T} + \max_{j \neq i} \left\{ Y^j_{\tau} - \gamma_{ij}(\tau) \right\} \mathbf{1}_{\tau<T} \middle| F_t \right]. \quad (4.11)$$

For $i \in \mathbb{I}$, define the implicit gain process $(U^i_t)_{t \in T}$ by,

$$U^i_t = \max_{j \neq i} \left\{ Y^j_{\tau} - \gamma_{ij}(t) \right\} \mathbf{1}_{t<T} + \Gamma_i \mathbf{1}_{t=T}. \quad (4.12)$$

Then equation (4.11) becomes,

$$Y^i_t = \operatorname{ess sup}_{\tau \in T_i} \mathbb{E} \left[ \sum_{s=t}^{\tau-1} \Psi_i(s) + U^i_\tau \middle| F_t \right]. \quad (4.13)$$

Note that the assumptions on $Y^i$ and the costs guarantee that $U^i \in \mathcal{S}^2$ for every $i \in \mathbb{I}$. Recalling Proposition 4.3.1, $(Y^1_t + \sum_{s=0}^{t-1} \Psi_i(s))_{t \in T}$ can be identified as the Snell envelope of the process $(U^1_t + \sum_{s=0}^{t-1} \Psi_i(s))_{t \in T}$ (also see Lemma B.1.1).

**Theorem 4.4.1** (Verification Theorem). Let $i \in \mathbb{I}$ be the active mode at some fixed initial time $t \in T$ and suppose $Y^1, \ldots, Y^m$ as defined in equation (4.11) are in $\mathcal{S}^2$. Define sequences of random times $\{\tau^*_n\}_{n \geq 0}$ and mode indicators $\{t^*_n\}_{n \geq 0}$ as follows:

$$\begin{align*}
\tau^*_0 &= t, \quad t^*_0 = i \quad \text{and for } n \geq 1: \\
\tau^*_n &= \inf \left\{ \tau^*_{n-1} \leq s \leq T : Y^*_s = U^*_s \right\}, \quad t^*_n = \sum_{j=1}^{N_{t^*_n}} j \mathbf{1}_{A^*_j} \\
\end{align*} \quad (4.14)$$

where $A^*_j = A^*_j(\omega)$ is the event:

$$A^*_j := \left\{ Y^j_{\tau^*_n} - \gamma_{i_n,j}(\tau^*_n) = \max_{k \neq i_n,j} \left\{ Y^k_{\tau^*_n} - \gamma_{i_n,k}(\tau^*_n) \right\} \right\}.$$

Then, $\alpha^* = (\tau^*_n, t^*_n)_{n \geq 0} \in \mathcal{A}_{\alpha,t,i}$ and satisfies

$$Y^i_t = J(\alpha^*, t, i) = \operatorname{ess sup}_{\alpha \in \mathcal{A}_{t,i}} J(\alpha; t, i) \quad \text{a.s.} \quad (4.15)$$

**Proof.** The proof is essentially the same as Theorem 1 of [37]. Recall the definition of $U^i$ in equation (4.12). At time $t$, $Y^i_t$ is given by
CHAPTER 4. OPTIMAL SWITCHING OF ELECTRIC SPACE HEATING

\[ Y_t^i = \text{ess sup}_{\tau \in \mathcal{T}_i} \mathbb{E} \left[ \sum_{s=t}^{\tau-1} \Psi_i(s) + U^i_t \left| \mathcal{F}_t \right. \right]. \]

If \( t = T \), then \( \tau_0^* = T \), \( t_0^* = i \) which leads to \( Y_T^i = \Gamma_i \), and the claim follows trivially since \( \Gamma_i = J(\alpha; T, i) = V(T, i) \) almost surely for any switching control \( \alpha \in \mathcal{A}_{T,i} \).

Suppose now that \( t < T \). Lemma B.1.2 in the appendix verifies that \( \alpha^* \in \mathcal{A}_{t,i} \).

Note that the infimum in equation (4.14) is always attained since \( Y^i_T = U^i_T \) a.s. for every \( i \in I \). The stopping time \( \tau_1^* \) in (4.14) is optimal after \( t \) by Proposition 4.3.1.

Using this together with the definition of \( \iota_1^* \) gives, almost surely,

\[
Y_t^i = \text{ess sup}_{\tau \in \mathcal{T}_i} \mathbb{E} \left[ \sum_{s=t}^{\tau-1} \Psi_i(s) + U^i_t \left| \mathcal{F}_t \right. \right] \\
= \mathbb{E} \left[ \sum_{s=t}^{\tau-1} \Psi_i(s) + U^i_t \left| \mathcal{F}_t \right. \right] \\
= \mathbb{E} \left[ \sum_{s=t}^{\tau-1} \Psi_i(s) + \Gamma_i \mathbf{1}_{\{\tau_1^* = t\}} + \max_{j \neq i} \left\{ \frac{Y^j_{\tau_1^*}}{\tau_1^*} - \gamma_{i,j}(\tau_1^*) \right\} \mathbf{1}_{\{\tau_1^* < T\}} \left| \mathcal{F}_t \right. \right] \\
= \mathbb{E} \left[ \sum_{s=t}^{\tau-1} \Psi_i(s) + \Gamma_i \mathbf{1}_{\{\tau_1^* = t\}} + \left\{ \frac{Y^j_{\tau_1^*}}{\tau_1^*} - \gamma_{i,j}(\tau_1^*) \right\} \mathbf{1}_{\{\tau_1^* < T\}} \left| \mathcal{F}_t \right. \right] \quad (4.16)
\]

Lemma B.1.1 in the appendix confirms that \( Y^i_\tau \) satisfies:

\[
Y^i_\tau = \text{ess sup}_{\tau \in \mathcal{T}_i} \mathbb{E} \left[ \sum_{r=\tau}^{\tau-1} \Psi_i(r) + U^i_r \left| \mathcal{F}_r \right. \right] \quad \text{on} \quad [\tau^*, T]. \quad (4.17)
\]

Use (4.17) together with the definition and optimality of \( \tau_2^* \) and \( \iota_2^* \) to get, almost surely,

\[
1_{\{\tau_1^* < T\}} Y^i_{\tau_1^*} = \text{ess sup}_{\tau \in \mathcal{T}_{\tau_1^*}} \mathbb{E} \left[ \sum_{r=\tau}^{\tau-1} \Psi_i(r) + U^i_r \left| \mathcal{F}_{\tau_1^*} \right. \right] 1_{\{\tau_1^* < T\}} \\
= \mathbb{E} \left[ \sum_{r=\tau_1^*}^{\tau_2^* - 1} \Psi_i(r) + \Gamma_i \mathbf{1}_{\{\tau_2^* = T\}} \\
+ \left\{ \frac{Y^j_{\tau_2^*}}{\tau_2^*} - \gamma_{i,j,\iota_2^*}(\tau_2^*) \right\} \mathbf{1}_{\{\tau_2^* < T\}} \left| \mathcal{F}_{\tau_1^*} \right. \right] 1_{\{\tau_1^* < T\}} \quad (4.18)
\]
Combining (4.16) and (4.18) gives the following expression for $Y^i_t$: almost surely,

$$Y^i_t = \mathbb{E} \left[ \sum_{s=t}^{\tau^i_1-1} \Psi_i(s) + \Gamma_{i_1} \mathbf{1}_{(\tau^i_1 = T)} + \left\{ Y^i_{\tau^i_1} - \gamma_{i,1}^* (\tau^i_1) \right\} \mathbf{1}_{(\tau^i_1 < T)} \mid \mathcal{F}_t \right]$$

$$= \mathbb{E} \left[ \sum_{s=t}^{\tau^i_1-1} \Psi_i(s) + \sum_{r=\tau^i_1}^{\tau^i_2-1} \Psi_i(r) \mid \mathcal{F}_{\tau^i_1} \right] \mathbf{1}_{(\tau^i_1 < T)} \mid \mathcal{F}_t$$

$$+ \mathbb{E} \left[ \Gamma_{i_1} \mathbf{1}_{(\tau^i_1 = T)} + \mathbb{E} \left[ \Gamma_{i_2} \mathbf{1}_{(\tau^i_2 = T)} \mid \mathcal{F}_{\tau^i_1} \right] \mathbf{1}_{(\tau^i_1 < T)} \mid \mathcal{F}_t \right]$$

$$- \mathbb{E} \left[ \gamma_{i,1}^* (\tau^i_1) \mathbf{1}_{(\tau^i_1 < T)} + \gamma_{i,2}^* (\tau^i_2) \mathbf{1}_{(\tau^i_2 < T)} \mathbf{1}_{(\tau^i_1 < T)} \mid \mathcal{F}_{\tau^i_1} \mid \mathcal{F}_t \right]$$

$$+ \mathbb{E} \left[ \mathbb{E} \left[ Y^i_{\tau^i_2} \mathbf{1}_{(\tau^i_2 < T)} \mid \mathcal{F}_{\tau^i_1} \right] \mathbf{1}_{(\tau^i_1 < T)} \mid \mathcal{F}_t \right]$$

$$= \mathbb{E} \left[ \sum_{s=t}^{\tau^i_1-1} \Psi_i(s) + \sum_{r=\tau^i_1}^{\tau^i_2-1} \Psi_i(r) \right] \mathbf{1}_{(\tau^i_1 < T)} \mid \mathcal{F}_t$$

$$+ \mathbb{E} \left[ \mathbb{E} \left[ \Gamma_{i_2} \mathbf{1}_{(\tau^i_2 = T)} \mid \mathcal{F}_{\tau^i_1} \right] \mathbf{1}_{(\tau^i_1 < T)} \mid \mathcal{F}_t \right]$$

$$- \mathbb{E} \left[ \gamma_{i,2}^* (\tau^i_2) \mathbf{1}_{(\tau^i_2 < T)} \mathbf{1}_{(\tau^i_1 < T)} \mid \mathcal{F}_{\tau^i_1} \mid \mathcal{F}_t \right]$$

$$+ \mathbb{E} \left[ \mathbb{E} \left[ Y^i_{\tau^i_2} \mathbf{1}_{(\tau^i_2 < T)} \mid \mathcal{F}_{\tau^i_1} \right] \mathbf{1}_{(\tau^i_1 < T)} \mid \mathcal{F}_t \right]$$

(4.19)

Since $\mathbf{1}_{(\tau^i_1 = T)}$, $\mathbf{1}_{(\tau^i_1 < T)}$ and $\gamma_{i,1}^* (\tau^i_1)$ are all $\mathcal{F}_{\tau^i_1}$-measurable, they can be brought inside the conditional expectation with respect to $\mathcal{F}_{\tau^i_1}$ in equation (4.19): almost surely,

$$Y^i_t = \mathbb{E} \left[ \sum_{s=t}^{\tau^i_1-1} \Psi_i(s) + \sum_{r=\tau^i_1}^{\tau^i_2-1} \Psi_i(r) \right] \mathbf{1}_{(\tau^i_1 < T)} \mid \mathcal{F}_t$$

$$+ \mathbb{E} \left[ \mathbb{E} \left[ \Gamma_{i_2} \mathbf{1}_{(\tau^i_2 = T)} \mid \mathcal{F}_{\tau^i_1} \right] \mathbf{1}_{(\tau^i_1 < T)} \mid \mathcal{F}_t \right]$$

$$- \mathbb{E} \left[ \gamma_{i,2}^* (\tau^i_2) \mathbf{1}_{(\tau^i_2 < T)} \mathbf{1}_{(\tau^i_1 < T)} \mid \mathcal{F}_{\tau^i_1} \mid \mathcal{F}_t \right]$$

$$+ \mathbb{E} \left[ \mathbb{E} \left[ Y^i_{\tau^i_2} \mathbf{1}_{(\tau^i_2 < T)} \mid \mathcal{F}_{\tau^i_1} \right] \mathbf{1}_{(\tau^i_1 < T)} \mid \mathcal{F}_t \right]$$

Using $i^*_0 = i$ and the definition of the mode indicator $u^*$ (cf. (4.3)) gives,

$$\sum_{s=t}^{\tau^i_1-1} \Psi_i(s) = \sum_{r=\tau^i_1}^{\tau^i_2-1} \Psi_i(r) + \sum_{r=\tau^i_1}^{\tau^i_2-1} \Psi_i(r) \mathbf{1}_{(\tau^i_1 < T)}.$$
\[
\begin{align*}
Y_t^i &= E \left[ \sum_{s=t}^{\tau_2} \Psi_{u_1^i}(s) + \sum_{k=0}^{1} \Gamma_{\tau_k^i} 1_{\{\tau_k^i < T\}} 1_{\{\tau_{k+1}^i = T\}} - \sum_{n=1}^{2} \gamma_{n-1,1}^i (\tau_n^i) 1_{\{\tau_n^i < T\}} \right] F_t \\
&\quad + E \left[ Y_{\tau_2}^i 1_{\{\tau_2^i < T\}} \right] F_t. 
\end{align*}
\]  

(4.20)

Let \( N(\alpha^*) \) be the total number of switches under \( \alpha^* \) (cf. (4.4)). Since \( \alpha^* \) is admissible by Lemma B.1.2 in the appendix, for \( n \geq 1 \) the switching times satisfy \( \tau_n^* < \tau_{n+1}^* \) on \( \{ n \leq N(\alpha^*) \} \) and \( \tau_n^* = T \) on \( \{ n > N(\alpha^*) \} \). Repeating the procedure of substituting for \( Y_{\tau_n^i}^i \) with \( n = 2, 3, \ldots \) yields

\[
Y_t^i = E \left[ \sum_{s=t}^{T-1} \Psi_{u_1^i}(s) + \sum_{k=0}^{N(\alpha^*)} \Gamma_{\tau_k^i} 1_{\{\tau_k^i < T\}} 1_{\{\tau_{k+1}^i = T\}} - \sum_{n=1}^{N(\alpha^*)} \gamma_{n-1,1}^i (\tau_n^i) 1_{\{\tau_n^i < T\}} \right] F_t. 
\]  

(4.21)

The sum of the terminal rewards collapses to a single term,

\[
\sum_{k=0}^{N(\alpha^*)} \Gamma_{\tau_k^i} 1_{\{\tau_k^i < T\}} 1_{\{\tau_{k+1}^i = T\}} = \Gamma_{N(\alpha^*)} P - a.s. 
\]  

(4.22)

From equations (4.21) and (4.22), one arrives at the following representation for \( Y_t^i \):

\[
Y_t^i = E \left[ \sum_{s=t}^{T-1} \Psi_{u_1^i}(s) + \Gamma_{N(\alpha^*)} - \sum_{n \geq 1} \gamma_{n-1,1}^i (\tau_n^i) 1_{\{\tau_n^i < T\}} \right] F_t = J(\alpha^*; t, i) 
\]  

(4.23)

Now let \( \alpha = (\tau_n, \iota_n)_{n \geq 0} \in \mathcal{A}_{t,i} \) be any admissible control. The verification theorem can be completed by showing that \( J(\alpha^*; t, i) \geq J(\alpha; t, i) \) a.s. First, note that \( J(\alpha^*; T, i) = J(\alpha; T, i) \) when \( t = T \), so assume henceforth that \( t < T \). Then, due to possible sub-optimality of the pair \((\tau_1, \iota_1)\), it is true that

\[
Y_t^i = \text{ess sup}_{\tau \in \mathcal{T}_i} E \left[ \sum_{s=t}^{\tau-1} \Psi_{\iota}(s) + U_t^i \right] F_t \\
\geq E \left[ \sum_{s=t}^{\tau_1-1} \Psi_{\iota}(s) + U_t^i \right] F_t \\
= E \left[ \sum_{s=t}^{\tau_1-1} \Psi_{\iota}(s) + \Gamma_t 1_{\{\tau_1 = T\}} + \max_{j \neq i} \{ Y_j^t - \gamma_{i,j} (\tau_1) \} 1_{\{\tau_1 < T\}} \right] F_t \\
\geq E \left[ \sum_{s=t}^{\tau_1-1} \Psi_{\iota}(s) + \Gamma_t 1_{\{\tau_1 = T\}} + \{ Y_{\tau_1}^t - \gamma_{i,\iota_1} (\tau_1) \} 1_{\{\tau_1 < T\}} \right] F_t. 
\]
Repeating the arguments leading to (4.23), replacing the equalities (=) with inequalities (≥) due to possible sub-optimality of \((\tau_n, \iota_n)\) for \(n \geq 2\), eventually leads to

\[
Y_i^t \geq E \left[ \sum_{s=t}^{T-1} \Psi_{u_s}(s) + \sum_{k=0}^{N(\alpha)} \Gamma_{t_k} 1_{\tau_k < T} 1_{\tau_{k+1} = T} - \sum_{n=1}^{N(\alpha)+1} \gamma_{\iota_{n-1}, \iota_n} (\tau_n) 1_{\tau_n < T} \middle| \mathcal{F}_t \right] \\
+ E \left[ Y_{t N(\alpha)+1} 1_{\tau_{N(\alpha)+1} < T} \middle| \mathcal{F}_t \right] \\
= E \left[ \sum_{s=t}^{T-1} \Psi_{u_s}(s) + \Gamma_{t N(\alpha)} - \sum_{n \geq 1} \gamma_{\iota_{n-1}, \iota_n} (\tau_n) 1_{\tau_n < T} \middle| \mathcal{F}_t \right] \\
= J(\alpha; t, i)
\]

and proves that the strategy \(\alpha^*\) is optimal.

\[\square\]

4.5 Existence of the optimal processes via backward induction

This section addresses the existence of the processes \(Y^1, \ldots, Y^m\) used in Theorem 4.4.1. The proof is a constructive one and verifies that the explicit dynamic programming scheme cited in [20, 61] and elsewhere indeed solves the optimal switching problem.

**Lemma 4.5.1** (Backward Induction). For each \(i \in I\), define the process \(\tilde{Y}_i^t = (\tilde{Y}_i^t)_{t \in T}\) recursively as follows:

\[
\tilde{Y}_i^T = \Gamma_i, \quad \text{and for } t = T - 1, \ldots, 0:\n\tilde{Y}_i^t = \max_{j \neq i} \left\{ -\gamma_{i,j}(t) + \Psi_j(t) + E \left[ \tilde{Y}_j^{t+1} \middle| \mathcal{F}_t \right] \right\} \lor \left\{ \Psi_i(t) + E \left[ \tilde{Y}_i^{t+1} \middle| \mathcal{F}_t \right] \right\}. \quad (4.24)
\]

Then \(\tilde{Y}_i^t\) is \(\mathbb{F}\)-adapted and in \(S^2\).

**Proof.** One verifies that \(\tilde{Y}_i^t\) is a well-defined \(\mathbb{F}\)-adapted process by proceeding recursively for \(t = T, \ldots, 0\) using (4.24), noting that the conditional expectations are well-defined by the integrability conditions on the rewards and switching costs. Note that in this discrete-time setting, showing \(\tilde{Y}_i^t \in S^2\) is equivalent to showing \(\tilde{Y}_i^t \in L^2\) for every \(t \in T\). Since \(\tilde{Y}_i^T = \Gamma_i \in L^2\) for all \(i \in I\), the claim is true for \(t = T\). Suppose by induction on \(t = T - 1, \ldots, 0\) that \(\tilde{Y}_j^{t+1} \in L^2\) for all \(j \in I\). The backward induction
Upon defining a new process, formula (4.24) gives:

\[
|\tilde{Y}_t^i| = \left| \max_{j \neq i} \left\{ -\gamma_{i,j}(t) + \Psi_j(t) + E \left[ \tilde{Y}_{t+1}^j | F_t \right] \right\} \right| \\
\leq \left| \max_{j \neq i} \left\{ -\gamma_{i,j}(t) + \Psi_j(t) + E \left[ \tilde{Y}_{t+1}^j | F_t \right] \right\} \right| + \left| \left\{ \Psi_i(t) + E \left[ \tilde{Y}_{t+1}^i | F_t \right] \right\} \right| \\
\leq 2 \max_{j \in I} E \left[ |\tilde{Y}_{t+1}^j| | F_t \right] + \max_{j,k \in I} \max_{r \in \mathbb{T}} |\gamma_{j,k}(r)| + 2 \max_{j \in I} \max_{r \in \mathbb{T}} |\Psi_j(r)| 
\]

(4.25)

Note that by Jensen’s inequality ([117, p. 139]), the conditional expectation satisfies,

\[
E \left[ \left| E \left[ |\tilde{Y}_{t+1}^j| | F_t \right] \right|^2 \right] \leq E \left[ |\tilde{Y}_{t+1}^j|^2 \right] < +\infty,
\]

and is therefore also in \( L^2 \). In addition to the observation that \( I \) is finite and \( \Psi_j, \gamma_{j,k} \in \mathcal{S}^2 \) for every \( j, k \in I \), the random variable on the right-hand side of (4.25) is in \( L^2 \). Therefore \( \tilde{Y}_t^i \in L^2 \), which holds for every \( i \in I \) since \( i \) was arbitrary. The case \( t = T - 1 \) has already been verified so the proof by induction is complete.

\[ \square \]

**An explicit Snell envelope system.**

A connection between \( \tilde{Y}_t^i \) in Lemma 4.5.1 and the Snell envelope becomes apparent upon defining a new process \( \left( \hat{Y}_t^i \right)_{t \leq T} \) for every \( i \in I \) by,

\[
\hat{Y}_t^i := \hat{Y}_t^i + \sum_{s=0}^{t-1} \Psi_i(s), \quad (4.26)
\]

A backward induction formula for \( \hat{Y}_t^i \) is then obtained by adding the \( F_t \)-measurable term \( \sum_{s=0}^{t-1} \Psi_i(s) \) to both sides of (4.24):

\[
\hat{Y}_t^i = \hat{U}_t^i, \quad \text{and for } t = T - 1, \ldots, 0 :
\]

\[
\hat{Y}_t^i = \hat{U}_t^i \lor E \left[ \hat{Y}_{t+1}^i | F_t \right]. \quad (4.27)
\]

where \( \left( \hat{U}_t^i \right)_{t \leq T} \) is the following **explicit gain process**:

\[
\hat{U}_t^i := \max_{j \neq i} \left\{ -\gamma_{i,j}(t) + \sum_{s=0}^{t-1} (\Psi_i(s) - \Psi_j(s)) + E \left[ \hat{Y}_{t+1}^j | F_t \right] \right\} L_{1_{t=T}} \\
+ \left\{ \sum_{s=0}^{T-1} \Psi_i(s) + \Gamma_i \right\} L_{1_{t=T}} \\
= \sum_{s=0}^{t-1} \Psi_i(s) + \max_{j \neq i} \left\{ -\gamma_{i,j}(t) - \sum_{s=0}^{t-1} \Psi_j(s) + E \left[ \hat{Y}_{t+1}^j | F_t \right] \right\} L_{1_{t=T}} \\
+ \Gamma_i L_{1_{t=T}}. \quad (4.28)
\]
CHAPTER 4. OPTIMAL SWITCHING OF ELECTRIC SPACE HEATING 114

The processes \((\hat{Y}_t^i)_{t \in T}\) and \((\hat{U}_t^i)_{t \in T}\) belong to \(S^2\) by integrability properties of the rewards, switching costs and as \(\hat{Y}^i \in S^2\). Proposition 4.3.1 verifies that the backward induction formula uniquely defines \((\hat{Y}_t^i)_{t \in T}\) as the Snell envelope of \((\hat{U}_t^i)_{t \in T}\).

**Theorem 4.5.2 (Existence).** Let \((\tilde{Y}_t^i)_{t \in T}, i \in \mathbb{I}\), be the processes defined by backward induction (4.24). Then, \(P\)-a.s. for every \(t \in T\):

\[
\tilde{Y}_t^i = \text{ess sup}_{\tau \in T_t} \left[ \sum_{s=t}^{\tau-1} \Psi_i(s) + \Gamma_i 1_{\{\tau=T\}} + \max_{j \neq i} \left\{ \tilde{Y}_\tau^j - \gamma_{i,j}(\tau) \right\} 1_{\{\tau<T\}} \right] \quad (4.29)
\]

Therefore, \(\tilde{Y}^1, \ldots, \tilde{Y}^m\) satisfy the verification theorem.

**Proof.** For notational convenience, introduce a new process \((\hat{W}_t^i)_{t \in T}\) which is defined for \(t \in T\) by

\[
\hat{W}_t^i := \sum_{s=0}^{t-1} \Psi_i(s) + \Gamma_i 1_{\{t=T\}} + \max_{j \neq i} \left\{ -\gamma_{i,j}(t) - \sum_{s=0}^{t-1} \Psi_j(s) + \tilde{Y}_s^j \right\} 1_{\{t<T\}}. \quad (4.30)
\]

Note that \(\hat{W}^i \in S^2\) by the properties of \(\Gamma, \Psi_i, \gamma_{i,j}\) and \(\tilde{Y}^j\) for \(i, j \in \mathbb{I}\). Equation (4.29) can be proved if

\[
\forall t \in T : \hat{Y}_t^i = \text{ess sup}_{\tau \in T_t} \mathbb{E} \left[ \hat{W}_\tau^i \bigg| F_t \right], \quad P - a.s. \quad (4.31)
\]

Indeed, if equation (4.31) is true then by (4.26): \(P - a.s.,\)

\[
\tilde{Y}_t^i = \text{ess sup}_{\tau \in T_t} \mathbb{E} \left[ \hat{W}_\tau^i \bigg| F_t \right] - \sum_{s=0}^{t-1} \Psi_i(s)
\]

\[
= \text{ess sup}_{\tau \in T_t} \left[ \sum_{s=0}^{t-1} \Psi_i(s) + \Gamma_i 1_{\{\tau=T\}} + \max_{j \neq i} \left\{ \tilde{Y}_s^j - \gamma_{i,j}(\tau) \right\} 1_{\{\tau<T\}} \right] F_t - \sum_{s=0}^{t-1} \Psi_i(s)
\]

\[
= \text{ess sup}_{\tau \in T_t} \left[ \sum_{s=0}^{t-1} \Psi_i(s) + \Gamma_i 1_{\{\tau=T\}} + \max_{j \neq i} \left\{ \tilde{Y}_s^j - \gamma_{i,j}(\tau) \right\} 1_{\{\tau<T\}} \right] F_t.
\]

In order to prove (4.31), first note by Proposition 4.3.1 the Snell envelope of \(\hat{W}^i\), exists and, denoting it by \(Z^i = (Z_t^i)_{t \in T}\), satisfies

\[
Z_t^i = \text{ess sup}_{\tau \in T_t} \mathbb{E} \left[ \hat{W}_\tau^i \bigg| F_t \right]
\]

as well as the backward induction formula

\[
Z_T^i = \hat{W}_T^i, \quad \text{and for } t = T-1, \ldots, 0 : \quad Z_t^i = \hat{W}_t^i \vee \mathbb{E} \left[ Z_{t+1}^i \bigg| F_t \right]. \quad (4.32)
\]
Thus establishing (4.31) is equivalent to showing that \( \hat{Y}^i \) is a modification of \( Z^i \) defined in (4.32). Note that by Proposition II.36.5 of [117] this would also mean that \( \hat{Y}^i \) and \( Z^i \) are indistinguishable. This shall be proved inductively.

Note that \( Z^i_t = \hat{W}^i_t = \hat{U}^i_t = \hat{Y}^i_t \) almost surely for every \( i \in I \). Suppose inductively for \( t = T - 1, \ldots, 0 \) that for all \( i \in I \), \( Z^i_{t+1} = \hat{Y}^i_{t+1} \) \( P \)-a.s. For every \( i \in I \) define the stopping time \( \theta^i_t \) as follows

\[
\theta^i_t = \inf \left\{ t \leq s \leq T : \hat{Y}^i_s = \hat{U}^i_s \right\} \tag{4.33}
\]
noting that \( t \leq \theta^i_t \leq T \) almost surely. The following lines will establish \( Z^i_t = \hat{Y}^i_t \) separately on the events \( \{ \theta^i_t = t \} \) and \( \{ \theta^i_t > t \} \equiv \{ \theta^i_t \geq t + 1 \} \). Since \( P(\{ \theta^i_t \geq t \}) = 1 \), the previous claim would lead to \( Z^i_t = \hat{Y}^i_t \) almost surely and the induction argument proves that \( \hat{Y}^i \) is a modification of \( Z^i \).

**Case 1:** \( Z^i_t = \hat{Y}^i_t \) on \( \{ \theta^i_t = t \} \).

Since \( \hat{Y}^j \) is the Snell envelope of \( \hat{U}^j \) for \( j \in I \), it is a supermartingale and by definition of \( \hat{U}^j \) and \( \hat{W}^j \) this leads to \( \hat{U}^j \leq \hat{W}^j \leq Z^j \) for \( i \in I \). Then, using (4.33), the backward induction formula and the induction hypothesis one gets

\[
\hat{W}^i_t \geq \hat{U}^i_t = \hat{Y}^i_t \geq E[\hat{Y}^i_{t+1} | \mathcal{F}_t] = E[Z^i_{t+1} | \mathcal{F}_t] \quad \text{on} \quad \{ \theta^i_t = t \}. \tag{4.34}
\]

Using the backward induction formula (4.32) for \( Z^i \) and (4.34) above also shows that

\[
\hat{W}^i_t = Z^i_t \quad \text{on} \quad \{ \theta^i_t = t \} \tag{4.35}
\]
Using (4.35) and finiteness of \( I \) shows that there exists an \( \mathcal{F}_{\theta^i_t} \)-measurable mode \( j_\ast \) (that is, \( j_\ast \) is an \( \mathcal{F}_{\theta^i_t} \)-measurable \( I \)-valued random variable) such that

\[
Z^i_t = \sum_{s=0}^{t-1} \Psi_i(s) - \gamma_{i,j_\ast}(t) - \sum_{s=0}^{t-1} \Psi_{j_\ast}(s) + \hat{Y}^{j_\ast}_t, \tag{4.36}
\]

\[
\begin{align*}
\gamma^i_j(t) = \arg \max_{j \neq i} \left\{ -\gamma_{i,j}(t) - \sum_{s=0}^{t-1} \Psi_j(s) + \hat{Y}^j_s \right\} \quad \text{on} \quad \{ \theta^i_t = t \}.
\end{align*}
\]

Next, it will be shown that

\[
1_{\{ \theta^i_t = t \}} \hat{Y}^{j_\ast}_t = 1_{\{ \theta^i_t = t \}} E[\hat{Y}^{j_\ast}_{t+1} | \mathcal{F}_t] \quad P \text{ - a.s.} \tag{4.37}
\]

Since \( j_\ast \) is \( \mathcal{F}_{\theta^i_t} \)-measurable, one has for \( t \leq r \leq T \)

\[
E \left[ \sum_{j \in I} 1_{\{ j_* = j \}} \hat{Y}^j_r \big| \mathcal{F}_t \right] = \sum_{j \in I} 1_{\{ j_* = j \}} E[\hat{Y}^j_r | \mathcal{F}_t] \leq \sum_{j \in I} 1_{\{ j_* = j \}} \hat{Y}^j_t \quad \text{on} \quad \{ \theta^i_t = t \}
\]
so that \( \hat{Y}^{j^*} := \sum_{j \in I} \hat{Y}^j \mathbf{1}_{\{j=x\}} \) is a supermartingale on \([\theta^i_T, T]\). Now if (4.37) is not true, by the supermartingale property of \( \hat{Y}^{j^*} \) the event \( A^i_t \) defined by

\[
A^i_t := \{ \hat{Y}^{j^*}_t > \mathbb{E}[\hat{Y}^{j^*}_{t+1} | \mathcal{F}_t] \} \cap \{ \theta^i_t = t \}
\]

has positive probability. If \( \mathbb{P}(A^i_t) > 0 \), there exists an \( \mathcal{F}_{\theta^i_t} \)-measurable mode \( k_* \) such that

\[
\hat{Y}^{j^*}_t = \sum_{s=0}^{t-1} \Psi{j_i_s}(s) - \gamma_{j^*,k_*}(t) - \sum_{s=0}^{t-1} \Psi{k_r_s}(s) + \mathbb{E}[\hat{Y}^{k_*}_{t+1} | \mathcal{F}_t] \quad \text{on} \quad \{ \theta^i_t = t \}
\]

This leads to

\[
Z^i_t = \sum_{s=0}^{t-1} \Psi{i_s}(s) - \gamma_{i,j_*}(t) - \sum_{s=0}^{t-1} \Psi{j_*}(s) + \hat{Y}^{j^*}_t
\]

\[
= \sum_{s=0}^{t-1} \Psi{i_s}(s) - \gamma_{i,j_*}(t) - \sum_{s=0}^{t-1} \Psi{j_*}(s) + \sum_{s=0}^{t-1} \Psi{k_r_s}(s)\]

\[
- \gamma_{j^*,k_*}(t) - \sum_{s=0}^{t-1} \Psi{k_r_s}(s) + \mathbb{E}[\hat{Y}^{k_*}_{t+1} | \mathcal{F}_t]\]

\[
< \sum_{s=0}^{t-1} \Psi{i_s}(s) - \gamma_{i,k_*}(t) - \sum_{s=0}^{t-1} \Psi{k_r_s}(s) + \hat{Y}^{k_*}_t \quad \text{on} \quad A^i_t
\]

where the inequality comes from the no-arbitrage condition (4.8) and the supermartingale property of \( \hat{Y}^{k_*} \) on \([\theta^i_t, T]\). However, this contradicts the optimality of \( j_* \) and therefore shows that (4.37) holds.

Using (4.36) together with (4.37) and the definition of \( \hat{Y}^i \) yields:

\[
Z^i_t = \sum_{s=0}^{t-1} \Psi{i_s}(s) - \gamma_{i,j_*}(t) - \sum_{s=0}^{t-1} \Psi{j_*}(s) + \mathbb{E}[\hat{Y}^{j^*}_{t+1} | \mathcal{F}_t] \leq \hat{Y}^i_t \quad \text{on} \quad \{ \theta^i_t = t \}
\]

and using this with (4.34) and (4.35) gives:

\[
\hat{W}^i_t = \hat{U}^i_t = \hat{Y}^i_t = Z^i_t \quad \text{on} \quad \{ \theta^i_t = t \}.
\]

**Case 2:** \( Z^i_t = \hat{Y}^i_t \quad \text{on} \quad \{ \theta^i_t \geq t + 1 \} \).

Note that \( \{ \theta^i_t \geq t + 1 \} = \{ \theta^i_t > t \} \) and is therefore \( \mathcal{F}_t \)-measurable. The properties of the Snell envelopes \( \hat{Y}^i \) and \( Z^i \) together with the induction hypothesis then give

\[
\hat{Y}^i_t = \mathbb{E}[\hat{Y}^i_{t+1} | \mathcal{F}_t] = \mathbb{E}[Z^i_{t+1} | \mathcal{F}_t] \leq Z^i_t \quad \text{on} \quad \{ \theta^i_t \geq t + 1 \}.
\]

Suppose that \( Z^i_t \) is a strict supermartingale on \( \{ \theta^i_t \geq t + 1 \} \) with positive probability. Then the \( \mathcal{F}_t \)-measurable event \( B^i_t \) defined by

\[
B^i_t := \{ \mathbb{E}[Z^i_{t+1} | \mathcal{F}_t] < Z^i_t \} \cap \{ \theta^i_t \geq t + 1 \}
\]
has positive probability, and implies the existence of an $\mathcal{F}_t$-measurable mode $j_*$ such that

$$
\hat{Y}_i^t < Z_i^t = \sum_{s=0}^{t-1} \Psi_i(s) - \gamma_{i,j_*}(t) - \sum_{s=0}^{t-1} \Psi_{j_*}(s) + \hat{Y}_{i,j_*}^t,
$$

$$
\tag{4.41}
$$

$$
\begin{align*}
\hat{Y}_i^t &= \arg \max_{j \neq i} \left\{-\gamma_{i,j}(t) - \sum_{s=0}^{t-1} \Psi_{j}(s) + \hat{Y}_{j}^t\right\} \text{ on } B_i^t.
\end{align*}
$$

But by definition of $\hat{Y}_i^t$, it is true that

$$
\hat{Y}_i^t \geq \sum_{s=0}^{t-1} \Psi_i(s) - \gamma_{i,j_*}(t) - \sum_{s=0}^{t-1} \Psi_{j_*}(s) + \hat{Y}_{i,j_*}^t
$$

and using this in (4.41) shows

$$
\mathbb{E}[\hat{Y}_{i+1}^t | \mathcal{F}_t] < \hat{Y}_{i}^t \text{ on } B_i^t.
$$

Using the same arguments leading up to (4.38), one can show that there exists an $\mathcal{F}_t$-measurable mode $k_*$ such that

$$
Z_i^t = \sum_{s=0}^{t-1} \Psi_i(s) - \gamma_{i,j_*}(t) - \sum_{s=0}^{t-1} \Psi_{j_*}(s) + \hat{Y}_{i,j_*}^t
$$

$$
= \sum_{s=0}^{t-1} \Psi_i(s) - \gamma_{i,j_*}(t) - \sum_{s=0}^{t-1} \Psi_{j_*}(s) + \sum_{s=0}^{t-1} \Psi_{k_*}(s)
$$

$$
- \gamma_{j_*,k_*}(t) - \sum_{s=0}^{t-1} \Psi_{k_*}(s) + \mathbb{E}[\hat{Y}_{i+1}^{k_*} | \mathcal{F}_t]
$$

$$
< \sum_{s=0}^{t-1} \Psi_i(s) - \gamma_{i,k_*}(t) - \sum_{s=0}^{t-1} \Psi_{k_*}(s) + \hat{Y}_{i,k_*}^t \text{ on } B_i^t
$$

which contradicts the optimality of $j_*$. Thus $P(B_i^t) = 0$ and this shows

$$
1_{\{\theta_i \geq t+1\}} \mathbb{E}[Z_{i+1}^t | \mathcal{F}_t] = Z_i^t 1_{\{\theta_i \geq t+1\}}, \ P - \text{a.s.}
$$

$$
\tag{4.42}
$$

Finally, putting (4.42), (4.40) and (4.39) together gives

$$
\hat{Y}_i^t = Z_i^t, \ P - \text{a.s.}
$$

Since the case $t = T - 1$ is true and $i \in I$ was arbitrary, the proof by induction is complete. Therefore, for every $i \in I$,

$$
\forall t \in T : \hat{Y}_i^t = Z_i^t, \ P - \text{a.s.}
$$

which means $\hat{Y}^t$ is a modification of (and therefore indistinguishable from) $Z^t$, whence (4.31) follows.
4.6 Numerical example

This section presents numerical results for the optimal switching problem described in this chapter’s introduction. Since this study is inspired by the classic optimal control problem of Chapter 3, the reader should refer to Section 3.7 of the previous chapter for an explanation of the terminology. Another numerical example which implicitly uses the theoretical results above can be found in Chapter 6.

Let $\hat{S} = \{\hat{S}_t: t = 0, 1, \ldots, T\}$ denote the (possibly random) spot price of electricity and $X = \{X_t: t = 0, 1, \ldots, T\}$ represent the indoor temperature which evolves according to (4.1). The performance index and value function in this case are given by:

$$\hat{J}(\alpha; 0, i) = E \left[ \sum_{t=0}^{T-1} \Psi_{u_t}(\hat{S}_t, X_t) + \Gamma_{iN(\alpha)}(\hat{S}_T, X_T) + \gamma N(\alpha) \right]$$

$$\hat{V}(0, i) = \mathop{\text{ess inf}}_{\alpha \in A_i} \hat{J}(\alpha; 0, i)$$

In equation (4.43), $N(\alpha) = \sum_{n \geq 1} 1_{\{\tau_n < T\}}$ is the number of switches under a given control $\alpha \in A_i$. Let $X_{t+1}^i$ denote the one-step solution of (4.1) with $u_t = i$ for $t = 0, \ldots, T - 1$ and $i \in \{0, 1\}$. The value function $\hat{V}$ can be computed path-wise for every $i \in I$ using a backward induction formula that is analogous to (4.24) (cf. Theorem 4.5.2):

$$\hat{V}(T, i) = \Gamma_i(\hat{S}_T, X_T^i), \quad \text{and for } t = T - 1, \ldots, 0 :$$

$$\hat{V}(t, i) = \min_{j \in I} \left\{ \gamma_{i,j} + E[\Psi_j(\hat{S}_t, X_t^j) + \hat{V}(t+1, j) | F_t] \right\}$$

4.6.1 Impact of switching costs on the deterministic control policy

The first set of simulations investigates the impact of the switching cost parameter $\gamma$ on the performance index and controlled indoor temperature trajectories. For these simulations the spot price $\hat{S}$ is deterministic and is denoted by $S$ to avoid confusion. Recalling the costs $\Psi$ and $G$ defined in (4.2) above, the performance of a strategy $\alpha \in A_i$ is given by

$$\hat{J}(\alpha; 0, i) := h \cdot \left[ \sum_{t=0}^{T-1} \frac{u_t \cdot S_t}{\max_{r \in T} S_r} + \kappa_1 B(X_t) \right] + \kappa_2 B(X_T) + \gamma N(\alpha).$$
CHAPTER 4. OPTIMAL SWITCHING OF ELECTRIC SPACE HEATING

Description of the numerical procedure.

The algorithm used below to compute the value function for the deterministic optimal switching problem is similar to the one used for the classic optimal control problem in the previous chapter. In fact, by setting $\gamma = 0$ in equation (4.44) one recovers the backward dynamic programming algorithm in Appendix A.2. The value of the indoor temperature $X$ is represented using a discrete state space similar to the numerical example of Chapter 3. This spatial discretisation is used to construct a “grid” of values $(t, y, i)$ where $t = 0, \ldots, T$ is the current time, $y = X_t$ is the current value of the indoor temperature, and $i \in \{0, 1\}$ is the current operating mode of the electric heater. The algorithm computes an approximation to the value function $\hat{V}$ and optimal strategy on this grid using the dynamic programming equation (4.44) and interpolation / extrapolation where appropriate.

The controlled trajectory of the indoor temperature is simulated using a feedback control map $\mathcal{U}$ obtained as output from the algorithm. More specifically, starting from initial value $X_0 = x$, the indoor temperature $X_{t+1}$ for $t = 0, \ldots, T - 1$ is calculated one step at a time using (4.1) with $u_t = \mathcal{U}(t, X_t, u_{t-1})$, $u_{t-1} \in \{0, 1\}$ being the previous consumption state. The notation $u_{-1}$ in this case represents the initial mode at time 0, reflecting the possibility of an immediate change at the start. The performance under the simulated control strategy and controlled trajectory are then calculated using (4.45).

Evaluation of control policies.

If one subtracts the contribution of the switching costs from the total in equation (4.45), the result is

$$C(\alpha; i, \gamma) := \hat{J}(\alpha; 0, i) - \gamma N(\alpha)$$

$$= h \sum_{t=0}^{T-1} \left[ u_t \frac{S_t}{\max_{r \in T} S_r} + \kappa_1 B(X_t) \right] + \kappa_2 B(X_T)$$

$$\approx J_{std}(0, x, u)$$

where $J_{std}(0, x, u)$ is the performance index (3.4) of Chapter 3 evaluated under the piecewise constant control $u$. The cost $C$ in (4.46) can therefore be used to compare the impact of the switching cost parameter in relation to the classic control problem.
studied previously. This is done in the following section. Table 4.1 below summarises the parameter values for the simulations which were carried out.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Temperature set point $T$</td>
<td>$22^\circ$C</td>
</tr>
<tr>
<td>Deadband parameter $\delta$</td>
<td>$2^\circ$C</td>
</tr>
<tr>
<td>Initial indoor temperature $x$</td>
<td>$22^\circ$C</td>
</tr>
<tr>
<td>Ambient temperature $T^a$</td>
<td>$3.8^\circ$C</td>
</tr>
<tr>
<td>Thermal gain $T^g$</td>
<td>$30^\circ$C</td>
</tr>
<tr>
<td>Thermal time constant $\tau$</td>
<td>480 minutes</td>
</tr>
<tr>
<td>Time step $h$</td>
<td>$\frac{1}{30}$</td>
</tr>
<tr>
<td>Spatial discretisation $\Delta x$</td>
<td>0.001</td>
</tr>
<tr>
<td>Weighting constants $\kappa_1, \kappa_2$</td>
<td>1, 5</td>
</tr>
<tr>
<td>Initial mode $i$</td>
<td>0</td>
</tr>
<tr>
<td>Switching cost values $\gamma$</td>
<td>$0.0001 \times k$ for $k = 0, 1, \ldots, 100$.</td>
</tr>
<tr>
<td>Deterministic spot prices $S$</td>
<td>MIP values for representative Profile III, settlement periods 15 to 44.</td>
</tr>
</tbody>
</table>

Table 4.1: Parameter values used to simulate the performance of approximate dynamic programming control policies.

The tradeoff between cost minimisation and switching effects.

Figure 4.1: Impact of the switching cost value on the performance of the control strategies.

(a) Total cost $C$ (blue with markers) versus the switching cost value.

(b) Financial cost savings (green with markers) versus the switching cost value.

Figure 4.1 shows the cost $C$ (cf. (4.46)), percentage savings and the number of switches under the approximate dynamic programming control policy for different values of the switching cost $\gamma$. The cost $C$ generally increased with the value of $\gamma$, with
deviations to this trend likely due to approximation errors (supported by numerical tests). On the other hand, the number of switches decreased as the value of $\gamma$ increased. Both of these trends can be explained intuitively using the backward induction formula (4.44): the controller may be forced to make a decision that minimises the value function for a high value of $\gamma$ that would otherwise have been sub-optimal had the switching cost been lower.

Remember that switching costs were added to the deterministic classic control problem of Chapter 3 so as to alleviate the frequent switching behaviour of the optimal bang-bang control strategy. Figure 4.1 suggests that this is achieved for relatively small values of the switching cost $\gamma$. Furthermore, the control strategies for such values of $\gamma$ can achieve financial cost savings that are close to the results of Chapter 3. For instance, when $\gamma$ was increased from 0 to 0.001 the number of switches dropped from 329 to 47 whilst the financial cost savings (%) decreased by only 0.2%. Figure 4.2 below affirms the reduction in switching with plots of the controlled trajectories for $\gamma \in \{0, 0.0001, 0.001, 0.01\}$.

Figure 4.2: Controlled indoor temperature trajectories under dynamic programming policies with different switching costs. The green step chart shows the MIP in £ per MWh.

The controlled temperature trajectory when $\gamma = 0$ is the same as the one obtained for the classic optimal control problem of Chapter 3. The control policies corresponding to $\gamma \in \{0.0001, 0.001, 0.01\}$ all managed to successfully avoid peak prices. However, instead of keeping the temperature nearly constant at times (caused by frequent switching), policies corresponding to higher switching costs produced larger temperature swings. These swings were caused by the longer intervals between the electric heater’s switching times. The plots also suggest that higher switching costs can cause
the indoor temperature to assume values slightly further away from the deadband region. These occurrences might be reduced by increasing the penalty for such deviations.

### 4.6.2 Simulation results with stochastic prices

**A simple model for random spot prices.**

Suppose that the spot price \( \hat{S} \) is now random. For simplicity, \( \hat{S} \) was modelled by adding a relatively small “noise” parameter to the true prices \( S = \{S_t: t = 0, 1, \ldots, T\} \). This ensures the structure of the price profile is similar to the true one and, by altering the size of the noise, helps illustrate the effect of price uncertainty on the control strategies. Let \( r_0 = 0, r_{n+1} = T \) and \( \{r_0, r_1, \ldots, r_n\} \subset \mathbb{T} \) be an increasing sequence of times defining the (half-hour) intervals \( [r_k, r_{k+1}) \) for the spot price (see Remark 3.3.2). Let \( \mathcal{N}(\mu, \nu^2) \) denote the class of normally distributed random variables with mean \( \mu \in \mathbb{R} \) and variance \( \nu^2 \in \mathbb{R}_+ \).

The model for the random price \( \hat{S} \) is given by the equation

\[
\begin{align*}
\hat{S}_t &= \sum_{k=0}^{n} \left[ S_{r_k} + \frac{S_{r_k}}{\max_{j} S_{r_j}} \cdot \varepsilon_k \right] \mathbf{1}_{[r_k, r_{k+1})}(t), \quad t = 0, 1, \ldots, T - 1 \\
\hat{S}_T &= \hat{S}_{r_n} 
\end{align*}
\]

(4.47)

In equation (4.47), \( \varepsilon_k \) is a random error which is multiplied by the factor \( \frac{S_{r_k}}{\max_{j} S_{r_j}} \) in order to prevent large errors dominating low prices within off-peak hours. The model for the random error terms \( \{\varepsilon_k\}_{k=0}^{n} \) is a first-order autoregressive process:

\[
\begin{align*}
\varepsilon_0 &\sim \mathcal{N}(0, \nu^2); \\
\varepsilon_k &= \beta \varepsilon_{k-1} + \zeta_k, \quad k = 1, \ldots, n. 
\end{align*}
\]

(4.48)

In equation (4.48), \( \beta \in (0, 1) \) is the correlation coefficient and \( \zeta_k \sim \mathcal{N}(0, \nu^2) \) is a white noise error term. Notice that for \( k \geq 1 \), \( \varepsilon_k \) satisfies

\[
\mathbb{E}[\varepsilon_k] = \mathbb{E}[\beta \varepsilon_{k-1} + \zeta_k] = \mathbb{E}[\beta \varepsilon_{k-1}] = \ldots = \beta^k \mathbb{E}[\varepsilon_0] = 0.
\]

Hence the random price \( \hat{S}_t \) and deterministic one \( S_t \) are equal in expectation. The white noise errors were generated in *antithetic pairs*, which is a simple variance reduction technique for Monte Carlo based algorithms [62, pp. 205–207]. Notwithstanding
its simplicity, this model is reasonable since it captures the autocorrelation in the prices. Furthermore, it satisfies the Markovian hypothesis which underlies the Least-Squares Monte Carlo regression (LSMC) algorithm described below.

Description of numerical procedure.

The performance of the switching control strategy $\alpha \in \mathcal{A}_i$ under random spot prices is defined by

$$\hat{J}(\alpha; 0, i) := \mathbb{E} \left[ h \cdot \left( \sum_{t=0}^{T-1} \frac{u_t \cdot \hat{S}_t}{\mathbb{E}[\max_{r \in T} \hat{S}_r]} + \kappa_1 B(X_t) \right) + \kappa_2 B(X_T) \right]. \quad (4.49)$$

Under Markovian assumptions on $\hat{S}$, the conditional expectations appearing in equation (4.44) can be approximated by the LSMC algorithm outlined in Appendix B.2. This numerical scheme takes as input $N_S$ sample paths of the spot price $\hat{S}$, where $N_S$ is a positive integer. Letting $\hat{S}^{(l)}$ denote sample path number $l = 1, 2, \ldots, N_S$, the expectation $\mathbb{E} \left[ \max_{r \in T} \hat{S}_r \right]$ appearing in (4.49) was approximated by the empirical average

$$\mathbb{E} \left[ \max_{r \in T} \hat{S}_r \right] \approx \frac{1}{N_S} \sum_{l=1}^{N_S} \max_{r \in T} \hat{S}_r^{(l)}.$$

The algorithm verifies that $\frac{1}{N_S} \sum_{l=1}^{N_S} \max_{r \in T} \hat{S}_r^{(l)} \neq 0$ to ensure the expectation is well defined.

Coupled with the spatial discretisation in the temperature variable, the LSMC procedure used in the simulations is the same as the “mixed interpolation Tsitsiklis-van Roy scheme” described in [21]. It provides near optimal feedback control strategies corresponding to each sample path of the random spot price. These path-dependent control strategies were used with the deterministic spot price profile, and metrics such as the relative %-savings and performance index values were obtained. These metrics were then contrasted with the results from the algorithm using deterministic prices only. Parameter settings for the simulations were the same as those in Table 4.1 above unless stated otherwise, and the switching cost $\gamma = 0.001$ was fixed throughout.

Table 4.2 records the default values used to generate the random errors in the simulations. The value of the standard deviation $\nu$ of the errors added to each price profile was set to one fourth of the standard deviation of the respective price series. The correlation value $\beta = 0.8$ was selected only for illustration. Figure 4.3 shows four
CHAPTER 4. OPTIMAL SWITCHING OF ELECTRIC SPACE HEATING  124

<table>
<thead>
<tr>
<th>Price Profile</th>
<th>I</th>
<th>III</th>
<th>V</th>
</tr>
</thead>
<tbody>
<tr>
<td>Correlation ($\beta$)</td>
<td>0.8</td>
<td>0.8</td>
<td>0.8</td>
</tr>
<tr>
<td>Standard Deviation ($\nu$)</td>
<td>1.0</td>
<td>4.0</td>
<td>8.0</td>
</tr>
<tr>
<td>Number of sample paths ($N_S$)</td>
<td>100</td>
<td>100</td>
<td>100</td>
</tr>
</tbody>
</table>

Table 4.2: Default parameter values used to generate the autoregressive error terms and random spot prices.

Figure 4.3: True and perturbed prices for different price profiles

Figure 4.3: True and perturbed prices for different price profiles

perturbed price paths that were generated for three different price profiles using the values in Table 4.2 and equation (4.47). The figures show that the sampled random price profiles have the same shape as the corresponding true ones concerning the location of peaks and troughs, provided these were pronounced well enough in the original profile.
Graphical summary of simulations with default parameters.

Figure 4.4 below summarises the outcome of the optimal control simulations with 100 perturbed paths for the three different price profiles. The plots show the financial cost savings (%), performance index values, and number of switches observed when the path-dependent stochastic control policies are used with the true prices.

Figure 4.4(a) shows that the cost savings achieved by the path-dependent control policies were generally lower than the savings attained by the deterministic policy. This is expected since the control strategy obtained in the deterministic setting should be near optimal. However, the worst performing path-dependent control policies (in terms of financial cost) still achieved savings within 0.5% of the corresponding deterministic policy’s savings.

Some of the path-dependent control policies appear to slightly outperform their deterministic counterparts with regards to the cost savings. Figure 4.4(b) also shows that the deterministic policy does not have the lowest performance amongst all of the policies tested. This highlights the sub-optimality of the deterministic policies, which is expected since the algorithm is a numerical approximation to the true solution (also see the discussion in Appendix A.2). Nevertheless, there are only small differences between the deterministic and path-dependent results in such cases.
Controlled trajectories for the worst performing path-dependent policies. It is instructive to compare the indoor temperature trajectories under the path-dependent control policies and deterministic ones, as it may help explain any differences in performance. Out of the 100 randomised paths that were simulated, the path-dependent control policy that had the highest (worst) performance index value for the corresponding deterministic prices was selected. Remember that a preheating strategy, as discussed in Chapter 3, describes a pattern where the indoor temperature undergoes relatively sharp and prolonged increases prior to local maxima in the electricity price.

Figure 4.5: Indoor temperature trajectories under the deterministic (blue line) and worst performing (red dashed line) stochastic control strategies for different price profiles

The controlled temperature trajectories under the deterministic and worst performing path-dependent control policies are illustrated in Figure 4.5 for the three different price profiles. The preheating strategies under the path-dependent control policies tend to agree closely with their deterministic counterparts around the periods when the price increase was significant (such as the peak period). On the other hand, a
The impact of spot price uncertainty on control policy performance.

![Graphs showing financial cost savings under stochastic control policies for different profiles](image)

(a) Financial cost savings (%) for the Category I profile
(b) Financial cost savings (%) for the Category III profile
(c) Financial cost savings (%) for the Category V profile

Figure 4.6: Impact of increasing uncertainty on the financial cost savings (%).

Figure 4.6 above summarises the effect of the standard deviation of the random errors on the financial cost savings. For each of the price profiles, the average savings achieved by the path-dependent control policies tends to decrease as the error standard deviation rises. The range of the cost savings also noticeably increases with the error standard deviation. These results were expected since the sample paths of
the random prices are likely to be significantly different from the true ones when the
error dispersion is high. What is interesting is that the percentage cost saved under
the worst performing path-dependent control policy was still within 3% of the deter-
ministic policy’s savings in all of the simulations. Moreover, the median values of the
savings under the path-dependent control policies were well within 0.5% of the deter-
ministic policy’s savings. The control policies may therefore still perform reasonably
well when spot prices are subject to forecast errors which do not significantly alter the
shape of the normalised price profile.

(a) Performance index values for the Category I profile

(b) Performance index values for the Category III profile

(c) Performance index values for the Category V profile

Figure 4.7: Impact of increasing uncertainty on the performance of the path-dependent
control policies.

Figure 4.7 summarises the effect of the standard deviation of the random errors
on the performance index values. Again, the median value is close to the respective
value of the deterministic policy and is almost invariant under the change of error
standard deviation. The range of results increases with the error dispersion as ex-
pected, and the minimum values highlight possible sub-optimality of the deterministic
policy. Notwithstanding these small errors, which can be expected from a numerical approximation, the potential for the algorithm to perform well under uncertainty is apparent from the results shown.

4.7 Conclusion

This chapter used discrete-parameter martingale and backward induction arguments to develop a theory for solving discrete-time optimal switching problems on a finite time horizon. It provided a rigorous derivation of a dynamic programming procedure that has frequently appeared in the literature on numerical methods for continuous-time optimal switching problems. This algorithm was then applied to an appropriate reformulation of the previous chapter’s electric heating control problem.

Numerical experiments proved that the introduction of sufficiently small switching costs

- reduces the number of times the electric heater;
- achieves financial cost savings which are close to those reported in Chapter 3.

Furthermore, this performance may still be achieved when the controller only has access to predictions of the spot price. For this to occur, the price forecasts should at least be accurate in determining the location and relative size of price minima and maxima in the normalised prices.
Chapter 5

Continuous-time optimal switching with signed switching costs

5.1 Introduction

Chapter 4 studied the problem of optimal switching for a stochastic system in discrete-time, then applied the results to a discretised version of the continuous-time classic optimal control problem of Chapter 3. It is therefore natural to consider the corresponding theory of optimal switching for stochastic systems with continuous time parameter (as mentioned previously, a rigorous treatment of deterministic systems can be found in [7]). Moreover, the theory in continuous time is central to the following chapter’s discussion on the relation between Dynkin games and optimal switching problems. This author recently submitted the present chapter’s content for peer review and a preprint version is available at [97].

Section 5.2 below introduces the probabilistic model and optimal switching problem. Preliminary concepts from the general theory of stochastic processes and optimal stopping are recalled in Section 5.3. The modelling assumptions for the optimal switching problem are given in Section 5.4. A verification theorem establishing the relationship between the optimal switching problem and iterative optimal stopping is given in Section 5.5. Additional sufficient conditions confirming the hypotheses of the verification theorem are given in Section 5.6. Section 5.7 concludes the chapter and Appendix C collects proofs not contained in the main text.
5.2 Definitions

5.2.1 Probabilistic setup

The optimal switching problem below is studied on a time horizon $[0, T]$, where $0 < T < \infty$. It is assumed that a complete filtered probability space, $(\Omega, \mathcal{F}, \mathbb{P})$, has been given and the filtration $\mathbb{F} = (\mathcal{F}_t)_{0 \leq t \leq T}$ satisfies the usual conditions of right-continuity and augmentation by the $\mathbb{P}$-null sets. Let $\mathbb{E}$ denote the corresponding expectation operator and $1_A$ represent the indicator function of a set (event) $A$. The shorthand notation a.s. means “almost surely”. Let $T$ denote the set of $\mathbb{F}$-stopping times $\nu$ which satisfy $0 \leq \nu \leq T$ $\mathbb{P}$-a.s. For a given $S \in T$, write $T_S = \{\nu \in T : \nu \geq S \ \mathbb{P} - a.s.\}$. Unless otherwise stated, a stopping time is assumed to be defined with respect to $\mathbb{F}$. For notational convenience the dependence on $\omega \in \Omega$ is often suppressed.

5.2.2 Problem definition

The optimal switching problem involves choosing from a finite set of modes, $\mathbb{I} = \{1, \ldots, m\}$ with $m \geq 2$, which dynamically affect some measure of economic performance over the horizon $[0, T]$. The instantaneous profit in mode $i \in \mathbb{I}$ is a mapping $\psi_i : \Omega \times [0, T] \rightarrow \mathbb{R}$. There is a cost for switching from mode $i$ to $j$ at time $t$ which is given by $\gamma_{i,j} : \Omega \times [0, T] \rightarrow \mathbb{R}$. There is also a reward for being in mode $i \in \mathbb{I}$ at time $T$ which is a random, real-valued quantity denoted by $\Gamma_i$. The assumptions on these costs / rewards are discussed below in Section 5.4.

Definition 5.2.1 (Admissible Switching Control Strategies). Let $t \in [0, T]$ and $i \in \mathbb{I}$ be given. An admissible switching control strategy starting from $(t, i)$ is a double sequence $\alpha = (\tau_n, \iota_n)_{n \geq 0}$ of stopping times $\tau_n \in \mathcal{T}_i$ and mode indicators $\iota_n$ such that:

1. $\tau_0 = t$ and the sequence $\{\tau_n\}_{n \geq 0}$ is non-decreasing;
2. Each $\iota_n : \Omega \rightarrow \mathbb{I}$ is $\mathcal{F}_{\tau_n}$-measurable; $\iota_0 = i$ and $\iota_n \neq \iota_{n+1}$ for $n \geq 0$;
3. Only a finite number of switching decisions can be made before the terminal time $T$:

$$\mathbb{P} \left( \{\tau_n < T, \forall n \geq 0\} \right) = 0. \quad (5.1)$$
4. The family of random variables $\{C^\alpha_n\}_{n \geq 1}$, where $C^\alpha_n$ is the total cost of the first $n \geq 1$ switches

$$C^\alpha_n := \sum_{k=1}^{n} \gamma_{t_{k-1}, t_k}(\tau_k)1_{\{\tau_k < T\}}$$

satisfies

$$\mathbb{E}\left[\sup_n |C^\alpha_n|\right] < \infty.$$ 

(5.3)

Let $A_{t,i}$ denote the set of admissible switching control strategies (henceforth, just strategies). We write $A_i$ when $t = 0$ and drop the subscript $i$ if it is not important for the discussion.

Let $N(\alpha)$ be an $\mathcal{F}_T$-measurable random variable representing the number of switches before $T$ under a strategy $\alpha \in A$:

$$N(\alpha) = \sum_{n \geq 1} 1_{\{\tau_n < T\}}, \quad \alpha \in \mathcal{A}.$$ 

(5.4)

Associated with each strategy $\alpha \in \mathcal{A}$ is a mode indicator function $u: \Omega \times [0, T] \to \mathbb{I}$ that gives the active mode at each time:

$$u_t := t_0 1_{[\tau_0, \tau_1]}(t) + \sum_{n \geq 1} t_n 1_{(\tau_n, \tau_{n+1})}(t), \quad t \in [0, T].$$ 

(5.5)

This definition for the mode indicator function $u$ follows [37, 65] among others.

**Remark 5.2.1.** For a switching control strategy $\alpha \in \mathcal{A}$, define the family of random variables $\mathcal{C}$ by $\mathcal{C} := \{C^\alpha_n : n \geq 1\}$, where $C^\alpha_n$ was defined in (5.2). Let $C^\alpha$ be the total switching cost under $\alpha$:

$$C^\alpha := \sum_{n \geq 1} \gamma_{t_{n-1}, t_n}(\tau_n)1_{\{\tau_n < T\}}$$

By the finiteness condition (5.1), we have:

$$C^\alpha = \lim_{n \to \infty} C^\alpha_n \quad \mathbb{P} - \text{a.s.}$$ 

(5.6)

If $\mathcal{C}$ is uniformly integrable, then Theorem 1.21 of [51] shows that the limit $C^\alpha$ is integrable and convergence takes place in $L^1$ (notation described in Section 5.3.2 below). However, a counterexample given in [17] shows that one cannot assert

$$\lim_{n \to \infty} \mathbb{E}\left[|C^\alpha_n| \mid \mathcal{B}\right] = \mathbb{E}\left[|C^\alpha| \mid \mathcal{B}\right] \quad \text{a.s. for every } \sigma\text{-algebra } \mathcal{B} \subset \mathcal{F}.$$ 

(5.7)
Condition (5.3) has been introduced in order to assert (5.7) for every \( \alpha \in \mathcal{A} \). In the papers [23, 70], switching costs are assumed to be non-negative and \( C^\alpha \) must be square-integrable for \( \alpha \in \mathcal{A} \) to be admissible. Since these conditions imply \( C^\alpha_n \leq C^\alpha \in L^2 \) for all \( n \geq 1 \), they are sufficient for condition (5.3) to hold.

For a fixed time \( t \in [0, T] \) and given mode \( i \in \mathcal{I} \), the performance index for the optimal switching problem starting at \( t \) in mode \( i \) is given by the total profit and cost of switching over \( [t, T] \):

\[
J(\alpha; t, i) = E\left[ \int_t^T \psi_u(s)ds + \Gamma_{\tau_T} - \sum_{n \geq 1} \gamma_{\tau_{n-1}, \tau_{n}}(\tau_{n}) \mathbf{1}_{\{\tau_{n} < T\}} \Big| \mathcal{F}_t \right], \quad \alpha \in \mathcal{A}_{t,i}. \quad (5.8)
\]

The optimisation problem is to find a strategy \( \alpha^* \in \mathcal{A}_{t,i} \) that maximises the performance index:

\[
J(\alpha^*; t, i) = \underset{\alpha \in \mathcal{A}_{t,i}}{\text{ess sup}} J(\alpha; t, i) =: V(t, i). \quad (5.9)
\]

The random function \( V(t, i) \) is called the value function of the optimal switching problem. Processes or functions with super(sub)-scripts in terms of the random mode indicators \( \iota_n \) are interpreted in the same way as in the previous chapter (cf. Remark 4.2.2).

\section*{5.3 Preliminaries}

\subsection*{5.3.1 Some results from the general theory of stochastic processes}

Certain arguments below rely on some knowledge of the general theory of stochastic processes. The relevant results are merely cited below, and the reader is kindly referred to the references \([51, 76, 118]\) for further details.

**Right-continuous with left-limits processes.**

An adapted process \( X = (X_t)_{0 \leq t \leq T} \) is said to be càdlàg if it is right-continuous and admits left limits:

\[
X_t = \lim_{u \to t, u > t} X_u \quad \text{for every } t \geq 0,
\]

\[
X_{t-} = \lim_{s \to t, s < t} X_s \quad \text{exists finitely for every } t > 0.
\]
The left-limits process associated with a càdlàg process $X$ is denoted by $X_- = (X_{t-})_{0 < t \leq T}$. Define the process $\Delta X$ by $\Delta X := X - X_-$ and let $\Delta_t X := X_t - X_{t-}$ denote the size of the jump in $X$ at $t \in (0, T]$. These definitions follow the convention used in [117, 118] for predictable processes.

**Remark 5.3.1.** Since the filtered probability space satisfies the usual conditions, a càdlàg modification can be obtained for any martingale defined with respect to the given filtration. See Doob’s Regularity Theorem in Section II.67 of [117]. Unless stated otherwise, it is assumed that this càdlàg modification is used.

**Predictable and totally inaccessible stopping times.**

A random time $S$ is an $\mathcal{F}$-measurable mapping $S : \Omega \rightarrow [0, T]$. For two random times $\rho$ and $\tau$, the stochastic interval $[\rho, \tau]$ is defined as:

$$[\rho, \tau] = \{(\omega, t) \in \Omega \times [0, T] : \rho(\omega) \leq t \leq \tau(\omega)\}.$$ 

Stochastic intervals $(\rho, \tau], [\rho, \tau), (\rho, \tau)$ are defined analogously. A random time $S > 0$ is said to be predictable if the stochastic interval $[0, S]$ is measurable with respect to the predictable $\sigma$-algebra (the $\sigma$-algebra on $\Omega \times (0, T]$ generated by the adapted processes with paths that are left-continuous with right-limits on $(0, T]$). Note that every predictable time is a stopping time [76, p. 17]. By Meyer’s previsibility (predictability) theorem ([118], Theorem VI.12.6), a stopping time $S > 0$ is predictable if and only if it is announceable in the following sense: there exists a sequence of stopping times $\{S_n\}_{n \geq 0}$ satisfying $S_n(\omega) \leq S_{n+1}(\omega) < S(\omega)$ for all $n$ and $\lim_n S_n(\omega) = S(\omega)$.

**Quasi-left-continuous processes and filtrations.**

A càdlàg process $X$ is called quasi-left-continuous if $\Delta_S X = 0$ a.s. for every predictable time $S$ (Definition I.2.25 of [76]). The strict pre-$S$ $\sigma$-algebra associated with a random time $S > 0$, $\mathcal{F}_{S^-}$, is defined as [118, p. 345]:

$$\mathcal{F}_{S^-} = \sigma \left( \{A \cap \{S > u\} : 0 \leq u \leq T, A \in \mathcal{F}_u\} \right).$$ 

According to [118, p. 346], a filtration $\mathbb{F} = (\mathcal{F}_t)_{0 \leq t \leq T}$ which satisfies the usual conditions is said to be *quasi-left-continuous* if $\mathcal{F}_S = \mathcal{F}_{S^-}$ for every predictable time $S$. The
following equivalence result for quasi-left-continuous filtrations can be found in [118], Theorem VI.18.1 and [51], Theorem 5.36.

**Proposition 5.3.1** (Characterisation of quasi-left-continuous filtrations). The following statements are equivalent:

1. $\mathcal{F}$ satisfies the usual conditions (right-continuous and $\mathbb{P}$-complete) and is quasi-left-continuous;

2. For every bounded (and then for every uniformly integrable) càdlàg martingale $M$ and every predictable time $S$, $M_S = M_S$ a.s.;

3. If $\{S_n\}$ is an increasing sequence of stopping times with limit $\lim_n S_n = S$, then

$$\mathcal{F}_S = \bigvee_n \mathcal{F}_{S_n}.$$ 

**Remark 5.3.2.** The quasi-left-continuous condition on $\mathcal{F}$ is satisfied in many applications. For example, it holds when the filtration is the (completed) natural one generated by a Lévy or Feller-Dynkin process with right-continuous paths such as a Brownian motion or Poisson process -- see Chapter 3 of [113] or Chapter 3 of [117].

### 5.3.2 Some notation

Here is a list of some notation that is frequently used below:

1. For $1 \leq p < \infty$, let $L^p$ denote the set of random variables $Z$ satisfying $E[\|Z\|^p] < \infty$.

2. Let $\mathcal{Q}$ denote the set of adapted, càdlàg processes which are quasi-left-continuous.

3. Let $\mathcal{M}^2$ denote the set of progressively measurable processes $X = (X_t)_{0 \leq t \leq T}$ satisfying,

$$E\left[\int_0^T |X_t|^2 dt\right] < \infty.$$ 

4. Let $\mathcal{S}^2$ denote the set of adapted, càdlàg processes $X$ satisfying:

$$E\left[\left(\sup_{0 \leq t \leq T} |X_t|\right)^2\right] < \infty.$$
CHAPTER 5. CONTINUOUS-TIME OPTIMAL SWITCHING

5.3.3 Properties of Snell envelopes

The theory of optimal stopping, particularly the method of essential supremum (Snell envelope) is utilised frequently in the results below. Recall that a progressively measurable process $X$ is said to belong to class $[D]$ if the set of random variables $\{X_\tau, \tau \in T\}$ is uniformly integrable.

**Proposition 5.3.2.** Let $U = (U_t)_{0 \leq t \leq T}$ be an adapted, $\mathbb{R}$-valued, càdlàg process that belongs to class $[D]$. Then there exists a unique (up to indistinguishability), adapted $\mathbb{R}$-valued càdlàg process $Z = (Z_t)_{0 \leq t \leq T}$ such that $Z$ is the smallest supermartingale which dominates $U$. The process $Z$ is called the Snell envelope of $U$ and it enjoys the following properties.

1. For any stopping time $\theta$:
   \[ Z_\theta = \text{ess sup}_{\tau \in T_\theta} E[U_\tau | F_\theta], \text{ and therefore } Z_T = U_T. \] (5.10)

2. Meyer decomposition: There exist a uniformly integrable right-continuous martingale $M$ and two non-decreasing, adapted, predictable and integrable processes $A$ and $B$, with $A$ continuous and $B$ purely discontinuous, such that for all $0 \leq t \leq T$,
   \[ Z_t = M_t - A_t - B_t, \quad A_0 = B_0 = 0. \] (5.11)
   Furthermore, the jumps of $B$ satisfy $\{ \Delta B > 0 \} \subset \{ Z_- = U_- \}$.

3. Let a stopping time $\theta$ be given and let $\{\tau_n\}_{n \geq 0}$ be an increasing sequence of stopping times tending to a limit $\tau$ such that each $\tau_n \in T_\theta$ and satisfies $E[U_{-\tau_n}] < \infty$. Suppose the following condition is satisfied for any such sequence,
   \[ \limsup_{n \to \infty} U_{\tau_n} \leq U_\tau \] (5.12)
   Then the stopping time $\tau_\theta^*$ defined by
   \[ \tau_\theta^* = \inf\{ t \geq \theta : Z_t = U_t \} \wedge T \] (5.13)
   is optimal after $\theta$ in the sense that:
   \[ Z_\theta = E[Z_{\tau_\theta^*} | F_\theta] = E[U_{\tau_\theta^*} | F_\theta] = \text{ess sup}_{\tau \in T_\theta} E[U_\tau | F_\theta]. \] (5.14)
4. For every $\theta \in T$, if $\tau^*_\theta$ is the stopping time defined in equation (5.13), then the stopped process $(Z_{t \wedge \tau^*_\theta})_{0 \leq t \leq T}$ is a (uniformly integrable) càdlàg martingale.

5. Let $\{U^n\}_{n \geq 0}$ and $U$ be adapted, càdlàg and of class $[D]$ and let $Z^{U^n}$ and $Z$ denote the Snell envelopes of $U^n$ and $U$ respectively. If the sequence $\{U^n\}_{n \geq 0}$ is increasing and converges pointwise to $U$, then the sequence $\{Z^{U^n}\}_{n \geq 0}$ is also increasing and converges pointwise to $Z$. Furthermore, if $U \in S^2$ then $Z \in S^2$.

References for these properties can be found in the appendix of [66] and other references such as [43, 99, 111]. Proof of the fifth property can be found in Proposition 2 of [37]. The following result concerning integrability of the components in the Doob-Meyer decomposition is also important.

**Proposition 5.3.3.** For $0 \leq t \leq T$, let $Z_t = M_t - A_t$ where

1. the process $Z = (Z_t)_{0 \leq t \leq T}$ is in $S^2$;
2. the process $M = (M_t)_{0 \leq t \leq T}$ is a càdlàg, quasi-left-continuous martingale with respect to $\mathbb{F}$;
3. the process $A = (A_t)_{0 \leq t \leq T}$ is an $\mathbb{F}$-adapted càdlàg increasing process.

Then $A$ (and therefore $M$) is also in $S^2$.

**Proof.** The proof essentially uses an integration by parts formula on $(A_T)^2$ and the decomposition $Z = M - A$ in the hypothesis. See Proposition A.5 of [66] for further details, noting that continuity of $M$ in their hypothesis can be substituted with quasi-left-continuity without changing the proof. Also see the *energy formula* and related inequalities in [32].

5.4 Assumptions

Apart from assuming the usual conditions on the stochastic basis $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$, the following assumption is made on the filtration $\mathbb{F}$.

**Assumption 5.** The filtration $\mathbb{F}$ satisfies the usual conditions of right-continuity and $\mathbb{P}$-completeness and is also quasi-left-continuous.
The standing assumptions on the processes defining the performance index of the optimal switching problem (5.8) shall now be stated. For every \(i, j \in \mathbb{I}\) suppose:

1. the instantaneous profit satisfies \(\psi_i \in \mathcal{M}^2\);

2. the switching cost satisfies \(\gamma_{i,j} \in \mathcal{Q} \cap \mathcal{S}^2\).

3. the terminal data \(\Gamma_i \in \mathcal{L}^2\) and is \(\mathcal{F}_T\)-measurable.

The switching costs are also assumed to satisfy the following assumptions which are common for optimal switching problems (see [65, 70] for instance):

**Assumption 6.** For every \(i, j, k \in \mathbb{I}\) and \(\forall t \in [0, T]\), \(\mathbb{P}\)-a.s.:

\[
\begin{align*}
&i. \quad \gamma_{i,i}(t) = 0 \quad (5.15) \\
&i.i. \quad \gamma_{i,k}(t) < \gamma_{i,j}(t) + \gamma_{j,k}(t), \quad \text{if } i \neq j \text{ and } j \neq k, \quad (5.16) \\
&i.ii. \quad \Gamma_i \geq \max_{j \neq i} \{ \Gamma_j - \gamma_{i,j}(T) \}. \quad (5.17)
\end{align*}
\]

Condition (5.15) shows there is no cost for staying in the same mode whilst conditions (5.16) and (5.17) rule out possible arbitrage opportunities.

## 5.5 A verification theorem

Throughout this section it is supposed that there exist processes \(Y^1, \ldots, Y^m\) in \(\mathcal{Q} \cap \mathcal{S}^2\) defined by

\[
\begin{align*}
Y^i_t &= \text{ess sup}_{\tau \in T_i} \mathbb{E} \left[ \int_t^\tau \psi_i(s) ds + \Gamma_i 1_{\{\tau = T\}} + \max_{j \neq i} \{Y^j_\tau - \gamma_{i,j}(\tau)\} 1_{\{\tau < T\}} \right] \bigg| \mathcal{F}_t, \\
Y^i_T &= \Gamma_i.
\end{align*}
\]

Sufficient conditions ensuring the existence of \(Y^1, \ldots, Y^m\) with these properties are given in Section 5.6 below. Theorem 5.5.2 below verifies that the solution to the optimal switching problem (5.9) can be written in terms of these \(m\) stochastic processes.

In preparation of this verification theorem, a few preliminary observations are useful. For every \(i \in \mathbb{I}\), let \(U^i = (U^i_t)_{0 \leq t \leq T}\) be a càdlàg process defined by

\[
U^i_t := \Gamma_i 1_{\{t = T\}} + \max_{j \neq i} \{Y^j_t - \gamma_{i,j}(t)\} 1_{\{t < T\}}, \quad 0 \leq t \leq T. \quad (5.19)
\]
Recall for every $i, j \in \mathbb{I}$ that $Y^i, \gamma_{i,j} \in \mathcal{Q} \cap \mathcal{S}^2$, $\Gamma_i \in L^2$ by assumption. Hence the process $U^i \in \mathcal{S}^2$ and is therefore of class $[D]$.

By rewriting equation (5.18) for $Y^i_t$ as follows,

$$Y^i_t = \text{ess sup}_{\tau \in \mathcal{T}_t} \mathbb{E} \left[\int_0^\tau \psi_i(s)ds + U^i_t \left| \mathcal{F}_\tau \right.\right]$$

$$= \text{ess sup}_{\tau \in \mathcal{T}_t} \mathbb{E} \left[\int_0^\tau \psi_i(s)ds + \int_0^t \psi_i(s)ds - \int_0^t \psi_i(s)ds + U^i_t \left| \mathcal{F}_\tau \right.\right]$$

$$= \text{ess sup}_{\tau \in \mathcal{T}_t} \mathbb{E} \left[\int_0^\tau \psi_i(s)ds + U^i_t \left| \mathcal{F}_\tau \right.\right] - \int_0^t \psi_i(s)ds, \quad \mathbb{P}-\text{a.s.} \quad (5.20)$$

Proposition 5.3.2 can be used to formally identify $\left(Y^i_t + \int_0^t \psi_i(s)ds\right)_{0 \leq t \leq T}$ as the Snell envelope of $\left(U^i_t + \int_0^t \psi_i(s)ds\right)_{0 \leq t \leq T}$. Note that the last line in equation (5.20) follows by $\mathcal{F}_\tau$-measurability of the integral term. The following lemma shows how the Snell envelope property extends to processes defined in terms of random mode indicators.

**Lemma 5.5.1.** For each $i \in \mathbb{I}$, let $U^i$ and $Y^i$ be defined as in equations (5.19) and (5.20) respectively. Let $\tau_n \in \mathcal{T}$ be given and $\iota_n: \Omega \to \mathbb{I}$ be $\mathcal{F}_{\tau_n}$-measurable. Suppose that $Y^1, \ldots, Y^m$ are in $\mathcal{Q} \cap \mathcal{S}^2$. Then,

$$Y^{i\in n} = \text{ess sup}_{\tau \in \mathcal{T}_t} \mathbb{E} \left[\int_0^\tau \psi_{i\in n}(s)ds + U^i_t \left| \mathcal{F}_\tau \right.\right], \quad \mathbb{P}-\text{a.s.} \quad \forall \tau_n \leq t \leq T. \quad (5.21)$$

Furthermore, there exist a uniformly integrable càdlàg martingale $M^{i\in n} = (M^{i\in n}_t)_{\tau_n \leq t \leq T}$ and a predictable, continuous, increasing process $A^{i\in n} = (A^{i\in n}_t)_{\tau_n \leq t \leq T}$ such that

$$Y^{i\in n}_t + \int_0^t \psi_{i\in n}(s)ds = M^{i\in n}_t - A^{i\in n}_t, \quad \mathbb{P}-\text{a.s.} \quad \forall \tau_n \leq t \leq T. \quad (5.22)$$

**Proof.** The claim (5.21) is established by showing that $Y^{i\in n}_t + \int_0^t \psi_{i\in n}(s)ds$ is the Snell envelope of $U^i_t + \int_0^t \psi_{i\in n}(s)ds$ for $\tau_n \leq t \leq T$. The proof will only be sketched due to its similarity to the first few lines of Theorem 1 in [37].

By equation (5.20) and the discussion proceeding it, $Y^i_t + \int_0^t \psi_i(s)ds$ is the Snell envelope of $U^i_t + \int_0^t \psi_i(s)ds$ on $[0, T]$ for every $i \in \mathbb{I}$. As the indicator function $1_{\{\iota_n = i\}}$ is non-negative and $\mathcal{F}_{\tau_n}$-measurable (and therefore $\mathcal{F}_\tau$-measurable for $t \geq \tau_n$), one can show that $\left(Y^i_t + \int_0^t \psi_i(s)ds\right)1_{\{\iota_n = i\}}$ is the smallest càdlàg supermartingale dominating $\left(U^i_t + \int_0^t \psi_i(s)ds\right)1_{\{\iota_n = i\}}$ on $[\tau_n, T]$. By summing over $i \in \mathbb{I}$ (recall $\mathbb{I}$ is finite), this shows $\left(Y^{i\in n}_t + \int_0^t \psi_{i\in n}(s)ds\right)$ is the smallest càdlàg supermartingale dominating $\left(U^{i\in n}_t + \int_0^t \psi_{i\in n}(s)ds\right)$ for $\tau_n \leq t \leq T$. In particular,

$$Y^{i\in n}_t + \int_0^t \psi_{i\in n}(s)ds = \text{ess sup}_{\tau \in \mathcal{T}_t} \mathbb{E} \left[\int_0^\tau \psi_{i\in n}(s)ds + U^{i\in n}_t \left| \mathcal{F}_\tau \right.\right], \quad \mathbb{P}-\text{a.s.} \quad \forall t \leq \tau_n \leq T,$$
and equation (5.21) follows by $\mathcal{F}_t$-measurability of the integral term for $t \geq \tau_n$.

For the claim (5.22), use the unique Meyer decomposition of the Snell envelope (property 2 of Proposition 5.3.2) to show that for every $i \in I$,

$$Y_t = Y_t^i + \int_0^t \psi_i(s)ds = M_t^i - A_t^i$$

for $t \in [0, T]$, where $M^i = (M^i_t)_{0 \leq t \leq T}$ is a càdlàg uniformly integrable martingale and $A^i = (A^i_t)_{0 \leq t \leq T}$ is a predictable, increasing process. The Snell envelope $(Y_t^i + \int_0^t \psi_i(s)ds)_{0 \leq t \leq T}$ is in $\mathcal{Q} \cap \mathcal{S}^2$ since $Y^i \in \mathcal{Q} \cap \mathcal{S}^2$ and $\psi_i \in \mathcal{M}^2$. This means the Snell envelope is a regular supermartingale of class $[D]$ and Theorem VII.10 of [32] asserts that its compensator, $A^i$, is continuous.

Using the Meyer decomposition shows

$$Y_t = Y_t^n + \int_0^t \psi_n(s)ds := \sum_{i \in I} \left( Y_t^i + \int_0^t \psi_i(s)ds \right) 1_{\{\tau_n = i\}} = \sum_{i \in I} (M_t^i - A_t^i) 1_{\{\tau_n = i\}}.$$ (5.23)

Now, using $1_{\{\tau_n = i\}}$ is non-negative and $\mathcal{F}_t$-measurable for $t \geq \tau_n$, the process $M_t^n$ defined on $[\tau_n, T]$ by

$$M_t^n(\omega) := \sum_{i \in I} M_t^i(\omega) 1_{\{\tau_n = i\}}(\omega), \quad \forall (\omega, t) \in [\tau_n, T]$$ (5.24)

is a uniformly integrable càdlàg martingale $\mathbb{P}$-a.s. for every $\tau_n \leq t \leq T$. Likewise, $A_t^n$ defined on $[\tau_n, T]$ by

$$A_t^n(\omega) := \sum_{i \in I} A_t^i(\omega) 1_{\{\tau_n = i\}}(\omega), \quad \forall (\omega, t) \in [\tau_n, T]$$ (5.25)

is a continuous, predictable increasing process $\mathbb{P}$-a.s. for every $\tau_n \leq t \leq T$. By equation (5.23), $M_t^n$ and $A_t^n$ provide the (unique) Meyer decomposition of $Y_t^n$ $\mathbb{P}$-a.s. for every $\tau_n \leq t \leq T$. This is the claim (5.22) and the proof is complete. $\square$

**Theorem 5.5.2 (Verification).** Suppose there exist $m$ unique processes $Y^1, \ldots, Y^m$ in $\mathcal{Q} \cap \mathcal{S}^2$ which satisfy equation (5.18). Define a sequence of times $\{\tau_n^m\}_{n \geq 0}$ and mode
Then the sequence indicators \( \{t^*_n\}_{n \geq 0} \) as follows:
\[
\begin{align*}
\tau^*_0 &= t, \quad t^*_0 = i, \\
\tau^*_n &= \inf \left\{ s \geq \tau^*_{n-1} : Y^*_{s} = \max_{j \neq j^*_{n-1}} \left( Y^*_{i} - \gamma^*_{i,j^*_{n-1}}(s) \right) \right\} \land T, \\
\iota^*_n &= \sum_{j \in I} j 1^*_{F^*_{j,i}} \\
\end{align*}
\]
for \( n \geq 1 \), where \( F^*_{j,i} \) is the event:
\[
F^*_{j,i} := \left\{ Y_{r_{j,i}}^* - \gamma_{r_{j,i},j^*_{n-1}}(\tau^*_n) = \max_{k \neq j^*_{n-1}} \left( Y_{r_{k,i}}^* - \gamma_{r_{k,i},k^*_{n-1}}(\tau^*_n) \right) \right\}.
\]
Then the sequence \( \alpha^* = (\tau^*_n, t^*_n, i)_{n \geq 0} \in \mathcal{A}_{t,i} \) and satisfies
\[
Y_i = J(\alpha^*; t, i) = \operatorname{ess} \sup_{\alpha \in \mathcal{A}_{t,i}} J(\alpha; t, i) \quad \mathsf{P} - \text{a.s.} \tag{5.27}
\]

**Proof.** Standard arguments can be used to verify that \( \tau^*_n \) is a stopping time and each \( t^*_n \) is \( F^*_{\tau^*_n} \)-measurable. Section C.1 of Appendix C confirms that \( \alpha^* \in \mathcal{A}_{t,i} \). As for the claim (5.27), it holds trivially for \( t = T \) since \( Y^*_T = \Gamma_i = V(t, i) \) a.s. for every \( i \in I \). Henceforth, we assume that \( t \in [0, T) \).

The remaining arguments of the proof are just as in [37]. The main idea is as follows: starting from an initial mode \( i \in I \) at time \( t \in [0, T] \), iteratively solve the optimal stopping problem on the right-hand-side of (5.18) using the theory of Snell envelopes. The minimal optimal stopping times characterise the switching times whilst the maximising modes are paired with them to give the switching strategy. This characterisation will eventually lead to (5.27).

Recall the process \( U_i = (U^i_t)_{0 \leq t \leq T} \) defined in equation (5.19). The assumptions on \( Y^i, \psi_i, \Gamma_i, \gamma_{i,j} \) for every \( i, j \in I \), show that \( U_i \in \mathcal{S}^2 \) and \( \left( Y^i + \int_0^t \psi_i(s)ds \right)_{0 \leq t \leq T} \) is identified formally as the Snell envelope of \( \left( U^i + \int_0^t \psi_i(s)ds \right)_{0 \leq t \leq T} \).

For \( i, j \in I \), using \( Y^i_T = \Gamma_j \) \( \mathsf{P} \)-a.s., quasi-left-continuity of \( Y^j \) and \( \gamma_{i,j} \), and condition (5.17) on the switching costs at \( T \) gives
\[
\lim_{t \uparrow T} \left( \max_{j \neq i} \{ Y^j_t - \gamma_{i,j}(t) \} \right) = \max_{j \neq i} \{ \Gamma_j - \gamma_{i,j}(T) \} \leq \Gamma_i \quad \mathsf{P} - \text{a.s.}
\]
Therefore, \( U_i \) is quasi-left-continuous on \( [0, T) \) and \( \lim_{t \uparrow T} U^i_t \leq U^i_T \) \( \mathsf{P} \)-a.s. Combining this with the continuity of the integral, the process \( \left( U^i_t + \int_0^t \psi_i(s)ds \right)_{0 \leq t \leq T} \) satisfies the hypotheses of property 3 in Proposition 5.3.2. Let \( (\tau^*_n, t^*_n, i)_{n \geq 0} \) be the pair of random times and mode indicators in the statement of the theorem and \( u^* \) be the associated
mode indicator function. In conjunction with Lemma 5.5.1, \( \{\tau^*_n\} \) defines a sequence of stopping times where, for \( n \geq 1 \), \( \tau^*_n \) is optimal for an appropriately defined optimal stopping problem. Note that each \( \iota^*_n \) is \( F_{\tau^*_n} \)-measurable since it is the (finite) sum of indicator functions of \( F_{\tau^*_n} \)-measurable sets.

Begin the iteration by using the optimality of the time \( \tau^*_1 \) for the optimal stopping problem starting at time \( \tau^*_0 = t \) with initial mode \( \iota^*_0 = i \):

\[
Y^i_t = \text{ess sup}_{\tau \in T_t} E \left[ \int_t^\tau \psi_i(s) ds + \Gamma_i \1_{\{\tau = T_t\}} + \max_{j \neq i} \{Y^j_s - \gamma_{i,j}(\tau)\} \1_{\{\tau < T_t\}} \bigg| F_t \right] \\
= E \left[ \int_t^{\tau^*_1} \psi_i(s) ds + \Gamma_i \1_{\{\tau^*_1 = T_t\}} + \max_{j \neq i} \{Y^j_s - \gamma_{i,j}(\tau^*_1)\} \1_{\{\tau^*_1 < T_t\}} \bigg| F_t \right].
\]

Then, by using the definition of \( \iota^*_1 \), one gets

\[
Y^i_t = E \left[ \int_t^{\tau^*_1} \psi_i(s) ds + \Gamma_i \1_{\{\tau^*_1 = T_t\}} + \left\{Y^i_{\tau^*_1} - \gamma_{i,i^*_1}(\tau^*_1)\right\} \1_{\{\tau^*_1 < T_t\}} \bigg| F_t \right]. \tag{5.28}
\]

Next, by Lemma 5.5.1 the following holds: P-a.s. \( \forall \tau^*_n \leq s \leq T_t \),

\[
Y^i_s = \text{ess sup}_{\tau \in T_t} E \left[ \int_s^\tau \psi_i(s) ds + \Gamma_i \1_{\{\tau = T_t\}} + \max_{j \neq i} \{Y^j_s - \gamma_{i,j}(\tau)\} \1_{\{\tau < T_t\}} \bigg| F_s \right]
\]

and using the optimality of \( \tau^*_2 \), the definition of \( \iota^*_2 \) and right-continuity of the Snell envelope gives: P-a.s.,

\[
Y^i_{\tau^*_2} = E \left[ \int_{\tau^*_1}^{\tau^*_2} \psi_i(s) ds + \Gamma_i \1_{\{\tau^*_2 = T_t\}} + \left\{Y^i_{\tau^*_2} - \gamma_{i,i^*_2}(\tau^*_2)\right\} \1_{\{\tau^*_2 < T_t\}} \bigg| F_{\tau^*_1} \right]. \tag{5.29}
\]

Substituting equation (5.29) for \( Y^i_{\tau^*_2} \) in equation (5.28) for \( Y^i_t \) then shows,

\[
Y^i_t = E \left[ \int_t^{\tau^*_2} \psi_i(s) ds + \Gamma_i \1_{\{\tau^*_1 = T_t\}} - \gamma_{i,i^*_1}(\tau^*_1) \1_{\{\tau^*_1 < T_t\}} \bigg| F_t \right] \\
+ E \left[ \left( E \left[ \int_{\tau^*_1}^{\tau^*_2} \psi_i(s) ds + \Gamma_i \1_{\{\tau^*_2 = T_t\}} - \gamma_{i,i^*_1}(\tau^*_1) \1_{\{\tau^*_1 < T_t\}} \bigg| F_{\tau^*_1} \right] \right) \1_{\{\tau^*_1 < T_t\}} \bigg| F_t \right] \\
= E \left[ \int_t^{\tau^*_2} \psi_i(s) ds + \sum_{n=1}^{2} \Gamma_{i^*_{n-1}} \1_{\{\tau^*_n < T_t\}} \1_{\{\tau^*_n = T_t\}} - \sum_{n=1}^{2} \gamma_{i,i^*_n}(\tau^*_n) \1_{\{\tau^*_n < T_t\}} \bigg| F_t \right] \\
+ E \left[ Y^i_{\tau^*_2} \1_{\{\tau^*_2 < T_t\}} \bigg| F_t \right]. \tag{5.30}
\]

where the last line was derived by using \( F_{\tau^*_1} \)-measurability of \( \1_{\{\tau^*_1 = T_t\}} \), \( \1_{\{\tau^*_1 < T_t\}} \) and \( \gamma_{i,i^*_1}(\tau^*_1) \) to bring them inside of the conditional expectation with respect to \( F_{\tau^*_1} \).
\{\tau_2 < T\} \subset \{\tau_1 < T\}; \tau_0^* = t < T so the event \{\tau_0^* < T\} occurs with probability one; and the tower property of the conditional expectation since \tau_1^* \geq t.

Since \(Y^i \in \mathcal{S}^2\) for every \(i \in \mathbb{I}\), this argument can be repeated ad infinitum which leads to: \(\forall N \geq 1\),

\[ Y^i_t = E \left[ \int_t^{\tau^*_N} \psi u^*_t(s)ds + \sum_{n=1}^{N} \Gamma_{n-1} 1_{(\tau^*_n < T)}1_{(\tau^*_n = T)} - \sum_{n=1}^{N} \gamma_{n-1,\tau^*_n} (\tau^*_n) 1_{(\tau^*_n < T)} \right| \mathcal{F}_t \]

\[ + E \left[ Y^{i,N}_{\tau^*_N} 1_{(\tau^*_N > T)} \right| \mathcal{F}_t \] \quad (5.31)

By Lemma C.1.1 and Theorem C.1.5 in the appendix respectively, the times \(\{\tau^*_n\}_{n \geq 0}\) satisfy the finiteness condition (5.1) and \(E[\sup_n |C_n^*|] < \infty\) holds for the cumulative switching costs. Appealing also to the conditional dominated convergence theorem, it is possible take the limit as \(N \to \infty\) in equation (5.31) and use the definition of \(u^*\) to get:

\[ Y^i_t = E \left[ \int_t^T \psi u^*_t(s)ds + \Gamma u^*_t - \sum_{n \geq 1} \gamma_{n-1,\tau^*_n} (\tau^*_n) 1_{(\tau^*_n < T)} \right| \mathcal{F}_t \] = \(J(\alpha^*; t, i)\) \quad (5.32)

Now, take any arbitrary admissible strategy \(\alpha = (\tau_n, \ell_n)_{n \geq 0} \in \mathcal{A}_{t,i}\). Then,

\[ Y^i_t \geq E \left[ \int_t^{\tau_1} \psi(s)ds + \Gamma_i 1_{(\tau_1 = T)} + \max_{j \neq i} \left\{ Y^j_t \right\} 1_{(\tau_1 < T)} \right| \mathcal{F}_t \]

\[ \geq E \left[ \int_t^{\tau_1} \psi(s)ds + \Gamma_i 1_{(\tau_1 = T)} + \left\{ Y^i_t \right\} 1_{(\tau_1 < T)} \right| \mathcal{F}_t \]

since \(\tau_1\) does not necessarily attain the essential supremum in equation (5.18). Just as before, this argument can be repeated ad infinitum for \(N \geq 1\) to get,

\[ Y^i_t \geq E \left[ \int_t^{\tau^*_N} \psi u^*_t(s)ds + \sum_{n=1}^{N} \Gamma_{n-1} 1_{(\tau^*_n-1 < T)}1_{(\tau^*_n = T)} - \sum_{n=1}^{N} \gamma_{n-1,\tau^*_n} (\tau^*_n) 1_{(\tau^*_n < T)} \right| \mathcal{F}_t \]

\[ + E \left[ Y^{i,N}_{\tau^*_N} 1_{(\tau^*_N < T)} \right| \mathcal{F}_t \]

By passing to the limit \(N \to \infty\) and using the conditional dominated convergence theorem,

\[ Y^i_t \geq E \left[ \int_t^T \psi u^*_t(s)ds + \Gamma u^*_t - \sum_{n \geq 1} \gamma_{n-1,\tau^*_n} (\tau^*_n) 1_{(\tau^*_n < T)} \right| \mathcal{F}_t \] = \(J(\alpha; t, i)\)

Using equation (5.32), \(\alpha^* \in \mathcal{A}_{t,i}\), and arbitrariness of \(\alpha \in \mathcal{A}_{t,i}\), we conclude that \(Y^i_t\) satisfies (5.27). \(\square\)
5.6 Existence of the candidate optimal processes

The existence of the processes $Y^1, \ldots, Y^m$ which satisfy Theorem 5.5.2 is proved in this section following the arguments of [37].

5.6.1 The case of at most $n \geq 0$ switches

For each $n \geq 0$, define process $Y_{i,n}^1, \ldots, Y_{i,n}^m$ recursively as follows: for $i \in I$ and for any $0 \leq t \leq T$, first set
\[
Y_{i,0}^i = E\left[\int_0^T \psi_i(s)ds + \Gamma_i \mid \mathcal{F}_t\right],
\]
and for $n \geq 1$,
\[
Y_{i,n}^i = \text{ess sup}_{\tau \in \mathcal{T}_t} E\left[\int_0^\tau \psi_i(s)ds + \Gamma_i 1_{\{\tau=T\}} + \max_{j \neq i} \left\{Y_{j,n-1}^j - \gamma_{i,j}(\tau)\right\} 1_{\{\tau<T\}} \mid \mathcal{F}_t\right].
\]

Define another process $\hat{U}_{i,n} = (\hat{U}_{i,n}^t)_{0 \leq t \leq T}$ by:
\[
\hat{U}_{t}^i : = \int_0^t \psi_i(s)ds + \Gamma_i 1_{\{t=T\}} + \max_{j \neq i} \left\{Y_{j,n-1}^j - \gamma_{i,j}(t)\right\} 1_{\{t<T\}}
\]

If $\hat{U}_{i,n}$ is of class $[D]$, then by Proposition 5.3.2 its Snell envelope exists and is defined by
\[
\text{ess sup}_{\tau \in \mathcal{T}_t} E[\hat{U}_{\tau}^i | \mathcal{F}_t] = Y_{i,n}^i + \int_0^t \psi_i(s)ds.
\]

Some properties of $Y_{i,n}^i$ which verify this are proved in the following lemma. In order to simplify some expressions in the proof, introduce a new process $\hat{Y}_{i,n} = (\hat{Y}_{i,n}^t)_{0 \leq t \leq T}$ which is defined by:
\[
\hat{Y}_{t}^i : = Y_{i,n}^i + \int_0^t \psi_i(s)ds.
\]

Lemma 5.6.1. For all $n \geq 0$, the processes $Y_{1,n}^1, \ldots, Y_{m,n}^m$ are in $Q \cap S^2$.

Proof. The proof follows in the same fashion as [37]. Using $\mathcal{F}_t$-measurability of the integral term, notice that
\[
\hat{Y}_{i,0}^i : = Y_{i,0}^i + \int_0^t \psi_i(s)ds = E\left[\int_0^T \psi_i(s)ds + \Gamma_i \mid \mathcal{F}_t\right].
\]
Since $\psi_i \in \mathcal{M}^2$ and $\Gamma_i \in L^2$, the conditional expectation is well-defined and $\hat{Y}_{i,0}^i$ is a uniformly integrable martingale which admits a càdlàg modification. By Doob's
maximal inequality it follows that $\hat{Y}^{i,0} \in S^2$ and therefore $Y^{i,0}$. Since the filtration is assumed to be quasi-left-continuous, Proposition 5.3.1 verifies that $\hat{Y}^{i,0} \in Q$ and therefore $Y^{i,0} \in Q$. Therefore, $Y^{i,n} \in Q \cap S^2$ for every $i \in I$ when $n = 0$.

Now, suppose by an induction hypothesis on $n \geq 0$ that for all $i \in I$, $Y^{i,n}$ is in $Q \cap S^2$. The claim $Y^{i,n+1} \in S^2$ shall be proved first. By the induction hypothesis on $Y^{i,n}$ and since $\gamma_{i,j} \in Q \cap S^2$ and $\psi_i \in M^2$, it is true that $\hat{U}^{i,n+1} \in S^2$. Therefore, by Proposition 5.3.2, $\hat{Y}^{i,n+1}$ is the Snell envelope of $\hat{U}^{i,n+1}$. For $0 \leq t \leq T$,

$$|\hat{U}^{i,n+1}_t| = E \left[ \hat{U}^{i,n+1}_{t+1} \mid F_t \right] \leq E \left[ \sup_{0 \leq s \leq T} |\hat{U}^{i,n+1}_s| \mid F_t \right]$$

and therefore

$$-E \left[ \sup_{0 \leq s \leq T} |\hat{U}^{i,n+1}_s| \mid F_t \right] \leq \hat{U}^{i,n+1}_t \leq \hat{Y}^{i,n+1}_t \leq E \left[ \sup_{0 \leq s \leq T} |\hat{U}^{i,n+1}_s| \mid F_t \right].$$

The above leads to the following bounds on $\hat{Y}^{i,n+1}$, since it is the Snell envelope of $\hat{U}^{i,n+1}$,

$$-E \left[ \sup_{0 \leq s \leq T} |\hat{Y}^{i,n+1}_s| \mid F_t \right] \leq \hat{Y}^{i,n+1}_t \leq \hat{Y}^{i,n+1}_t \leq E \left[ \sup_{0 \leq s \leq T} |\hat{Y}^{i,n+1}_s| \mid F_t \right].$$

This leads to $\hat{Y}^{i,n+1} \in S^2$ since,

$$E \left[ \left( \sup_{0 \leq t \leq T} |\hat{Y}^{i,n+1}_t| \right)^2 \right] \leq E \left[ \left( \sup_{0 \leq t \leq T} E \left[ \sup_{0 \leq s \leq T} |\hat{U}^{i,n+1}_s| \mid F_t \right] \right)^2 \right] \leq 4E \left[ \left( \sup_{0 \leq t \leq T} E \left[ |\hat{U}^{i,n+1}_s| \mid F_t \right] \right)^2 \right] = 4E \left[ \left( \sup_{0 \leq t \leq T} |\hat{U}^{i,n+1}_s| \right)^2 \right] < \infty. \quad (5.37)$$

Since $\psi_i \in M^2$ it then follows that $Y^{i,n+1} \in S^2$.

Next, to show that $Y^{i,n+1} \in Q$ an argument similar to the proof of Proposition 1.4a in [71] shall be used. First, recall that $Y^{i,n}$ is in $Q \cap S^2$ for every $i \in I$ by the induction hypothesis, and $\gamma_{i,j} \in Q \cap S^2$ for every $i, j \in I$. This means the process $(\max_{j \neq i} \{ -\gamma_{i,j}(t) + Y^{i,n}_t \})_{0 \leq t \leq T}$ is also in $Q \cap S^2$. Using $Y^{j,n}_T = \Gamma_j$, $P$-a.s. and condition (5.17) on the switching costs,

$$\lim_{t \uparrow T} \left( \max_{j \neq i} \{ Y^{j,n}_t - \gamma_{i,j}(t) \} \right) = \max_{j \neq i} \{ \Gamma_j - \gamma_{i,j}(T) \} \leq \Gamma_i.$$

Thus $\hat{U}^{i,n+1}$ is quasi-left-continuous on $[0, T)$ and has a possible positive jump at time $T$. 
Next, by Proposition 5.3.2, \( \hat{Y}^{i,n+1} \) has a unique Meyer decomposition:

\[
\hat{Y}^{i,n+1} = M - A - B,
\]

where \( M \) is a right-continuous, uniformly integrable martingale, and \( A \) and \( B \) are predictable, non-decreasing processes which are continuous and purely discontinuous respectively. Let \( \tau \in \mathcal{T} \) be any predictable time. Note that the process \( A \) is continuous so \( A_\tau = A \) holds almost surely. Moreover, the martingale \( M \) also satisfies \( M_\tau = M_{\tau^-} \) a.s. since, by Proposition 5.3.1, it is quasi-left-continuous. Predictable jumps in \( \hat{Y}^{i,n+1} \) therefore come from \( B \), and only the two events \( \{ \Delta_\tau B > 0 \} \) and \( \{ \Delta_\tau B = 0 \} \) need to be investigated since \( B \) is non-decreasing.

By property 2 of Proposition 5.3.2,

\[
\{ \Delta_\tau B > 0 \} \subset \{ \hat{Y}^{i,n+1}_\tau = \hat{U}^{i,n+1}_\tau \}
\]

and, using the dominating property of \( \hat{Y}^{i,n+1} \) and non-negativity of the predictable jumps of \( \hat{U}^{i,n+1} \), one gets

\[
E \left[ \hat{Y}^{i,n+1}_\tau 1_{\{ \Delta_\tau B > 0 \}} \right] = E \left[ \hat{U}^{i,n+1}_\tau 1_{\{ \Delta_\tau B > 0 \}} \right] \leq E \left[ \hat{U}^{i,n+1}_\tau 1_{\{ \Delta_\tau B > 0 \}} \right] \leq E \left[ \hat{Y}^{i,n+1}_\tau 1_{\{ \Delta_\tau B > 0 \}} \right]. \tag{5.38}
\]

On the other hand, the Meyer decomposition of \( \hat{Y}^{i,n+1} \) and the almost sure continuity of \( M \) and \( A \) at \( \tau \) yield the following:

\[
E \left[ \hat{Y}^{i,n+1}_\tau 1_{\{ \Delta_\tau B = 0 \}} \right] = E \left[ (M_\tau - A_\tau - B_\tau^-) 1_{\{ \Delta_\tau B = 0 \}} \right] = E \left[ (M_\tau - A_\tau - B_\tau^-) 1_{\{ \Delta_\tau B = 0 \}} \right] = E \left[ \hat{Y}^{i,n+1}_\tau 1_{\{ \Delta_\tau B = 0 \}} \right]. \tag{5.39}
\]

Equations (5.38) and (5.39) produce the inequality, \( E[\hat{Y}^{i,n+1}_\tau] \leq E[\hat{Y}^{i,n+1}_\tau] \). However, \( E[\hat{Y}^{i,n+1}_\tau] \geq E[\hat{Y}^{i,n+1}_\tau] \) since \( \hat{Y}^{i,n+1} \) is a right-continuous supermartingale (in \( S^2 \)) and \( \tau \) is predictable (Theorem VI.14 of [32]). Thus \( E[\hat{Y}^{i,n+1}_\tau] = E[\hat{Y}^{i,n+1}_\tau] \) for every predictable time \( \tau \). This means \( Y^{i,n+1} \) is a regular supermartingale (of class \([D]\)) and, by Theorem VII.10 of [32], the predictable non-decreasing component of the Meyer decomposition of \( Y^{i,n+1} \) must be continuous. Therefore, \( B \equiv 0 \) and \( Y^{i,n+1} \in \mathcal{Q} \) since the only jumps it experiences are those from the quasi-left-continuous martingale \( M \).
Lemma 5.6.2. For every $i \in \mathbb{I}$, the process $Y^{i,n}$ solves the optimal switching problem with at most $n \geq 0$ switches: $\forall t \in [0, T],$

$$Y^{i,n}_t = \text{ess sup}_{\alpha \in \mathcal{A}_{t,i}^n} \mathbb{E} \left[ \int_t^T \psi_{u_s}(s)ds + \Gamma_{u_T} - \sum_{j=1}^n \gamma_{i_{j-1},j}(\tau_j)1_{\{\tau_j < T\}} \bigg| \mathcal{F}_t \right].$$

Moreover, the sequence $\{Y^{i,n}\}_{n \geq 0}$ is increasing and converges pointwise $\mathbb{P}$-a.s. for any $0 \leq t \leq T$ to a càdlàg process $\hat{Y}^i$ satisfying: $\forall t \in [0, T],$

$$\hat{Y}^i_t = \text{ess sup}_{\alpha \in \mathcal{A}_{t,i}} J(\alpha; t, i) =: V(t, i), \quad a.s..$$

Proof. For the optimal switching problem starting from $t \in [0, T]$ with initial mode $i \in \mathbb{I}$, let $\mathcal{A}_{t,i}^n$ be the subset of admissible strategies with at most $n$ switches:

$$\mathcal{A}_{t,i}^n = \{\alpha \in \mathcal{A}_{t,i}: \tau_{n+1} = T, \mathbb{P} - a.s.\}.$$

Define a double sequence $\hat{\alpha}^{(n)} = (\hat{\tau}_k, \hat{i}_k)_{k=0}^{n+1}$ as follows

$$\hat{\tau}_0 = t, \quad \hat{i}_0 = i,$$

$$\begin{cases} 
\hat{\tau}_k = \inf \left\{ s \geq \hat{\tau}_{k-1}: Y^{i_{k-1},n-(k-1)}_s = \max_{j \neq i_{k-1}} (Y^{j,n-k}_s - \gamma_{i_{k-1},j}(s)) \right\} \wedge T, \\
\hat{i}_k = \sum_{j \in \mathbb{I}} 1_{F^{i_{k-1}}_j} 
\end{cases}$$

for $k = 1, \ldots, n$ where $F^{i_{k-1}}_j$ is the event:

$$F^{i_{k-1}}_j := \left\{ Y^{j,n-k}_{\hat{\tau}_k} - \gamma_{i_{k-1},j} (\hat{\tau}_k) = \max_{\ell \neq i_{k-1}} (Y^{\ell,n-k}_{\hat{\tau}_k} - \gamma_{i_{k-1},\ell} (\hat{\tau}_k)) \right\},$$

and set $\hat{\tau}_{n+1} = T, \hat{i}_{n+1}(\omega) = j \in \mathbb{I}$ with $j \neq \hat{i}_n(\omega)$. Since $Y^{i,n} \in \mathcal{Q} \cap \mathcal{S}^2$, one verifies that $\hat{\alpha}^{(n)} \in \mathcal{A}_{t,i}^n$ and, using the arguments of Theorem 5.5.2, that $Y^{i,n}_t = J(\hat{\alpha}^{(n)}; t, i)$ and has the representation (5.40). Furthermore, since $\mathcal{A}_{t,i}^n \subset \mathcal{A}_{t,i}^{n+1} \subset \mathcal{A}_{t,i}$, it follows that $Y^{i,n}_t$ is non-decreasing in $n$ for all $t \in [0, T]$ and $Y^{i,n}_t \leq Y^{i,n+1}_t \leq V(t, i)$ almost surely.

Recall the processes $\hat{Y}^{i,n}$ and $\hat{Y}^{i,n}$ from Lemma 5.6.1 and that $\hat{Y}^{i,n}$ is the Snell envelope of $\hat{U}^{i,n}$ for each $n \geq 0$. Then $\{\hat{Y}^{i,n}\}_{n \geq 0}$ is an increasing sequence of càdlàg supermartingales and Theorem VI.18 of [32] shows that this sequence converges to a limit $\hat{Y}^i$ defined pointwise on $[0, T]$ by

$$\hat{Y}^i_t := \sup_n \hat{Y}^{i,n}_t = \sup_n \left( Y^{i,n}_t + \int_0^t \psi_i(s)ds \right).$$
This process $\hat{Y}_t^i = (\hat{Y}_t^{i,n})_{0 \leq t \leq T}$ is indistinguishable from a càdlàg process, but is not necessarily a supermartingale since its integrability has not yet been established. Nevertheless, the sequence $\{Y_t^{i,n}\}_{n \geq 0}$ converges for every $t$ to a limit $\hat{Y}_t^i$ which, modulo indistinguishability, is càdlàg and given by

$$\hat{Y}_t^i = \sup_n Y_t^{i,n} = \hat{Y}_t^i - \int_0^t \psi_i(s)ds. \quad (5.44)$$

Next, let $\alpha = (\tau_k, \iota_k)_{k \geq 0} \in \mathcal{A}_{l,i}$ be arbitrary. Note that attention can be restricted to those strategies such that $P(\{\tau_k = \tau_{k+1}, \tau_k < T\}) = 0$ for $k \geq 1$. Indeed, if $H_k$ is the event $H_k := \{\tau_k = \tau_{k+1}, \tau_k < T\}$ and $P(H_k) > 0$ for some $k \geq 1$, then by Assumption 6:

$$\left(\gamma_{\iota_{k-1},\iota_k}(\tau_k) + \gamma_{\iota_{k},\iota_{k+1}}(\tau_{k+1})\right)1_{H_k} = \left(\gamma_{\iota_{k-1},\iota_k}(\tau_k) + \gamma_{\iota_{k},\iota_{k+1}}(\tau_{k+1})\right)1_{H_k}$$

which shows it is suboptimal to switch twice at the same time. Define $\alpha^n = (\tau^n_k, \iota^n_k)_{k \geq 0}$ to be the strategy obtained from $\alpha$ when only the first $n$ switches are kept:

$$\begin{cases} (\tau^n_k, \iota^n_k) = (\tau_k, \iota_k), & k \leq n, \\ \tau^n_k = T, & k > n. \end{cases}$$

The difference between the performance indices under $\alpha$ and $\alpha^n$ is:

$$J(\alpha; t, i) - J(\alpha^n; t, i) = \mathbb{E}\left[\int_{\tau^n}^{T} \left(\psi_{u_i}(s) - \psi_{i^n_i}(s)\right)ds + \Gamma_{u_T} - \Gamma_{i^n_i} \right. - \sum_{k>n} \gamma_{\iota_{k-1},\iota_k}(\tau_k)1_{\{\tau_k < T\}} \bigg| \mathcal{F}_t \left.\right]$$

$$= \mathbb{E}\left[\int_{\tau^n}^{T} \left(\psi_{u_i}(s) - \psi_{i^n_i}(s)\right)ds + \Gamma_{u_T} - \Gamma_{i^n_i} - (C^{\alpha} - C^n) \bigg| \mathcal{F}_t \right]$$

where $u$ is the mode indicator function associated with $\alpha$ and $\iota^n_n = \tau_{n\wedge N(\alpha)}$ is the last mode switched to before $T$ under $\alpha^n$. Since $\alpha \in \mathcal{A}_{l,i}, \psi_i \in \mathcal{M}^2$ and $\Gamma_i \in L^2$ for every $i \in \mathbb{I}$, the conditional expectation above is well-defined for every $n \geq 1$. Then, as the strategies $\alpha$ and $\alpha^n$ coincide on the event $\{N(\alpha) \leq n\}$, one gets the following integrable upper bound for $J(\alpha; t, i)$:

$$J(\alpha; t, i) \leq \mathbb{E}\left[ \left(\int_{\tau^n}^{T} \left|\psi_{u_i}(s) - \psi_{i^n_i}(s)\right|ds + |\Gamma_{u_T} - \Gamma_{i^n_i}| + |C^{\alpha} - C^n| \right)1_{\{N(\alpha) > n\}} \bigg| \mathcal{F}_t \right]$$

$$+ J(\alpha^n; t, i) \quad (5.45)$$
Using these integrability conditions again together with the observation that $N(\alpha) < \infty$ $\mathbb{P}$-a.s. and $\{\tau_k\}$ is (strictly) increasing towards $T$ shows

$$
\lim_{n \to \infty} \mathbb{E} \left[ \left( \int_{\tau_n}^T \left| \psi_{u,s}(s) - \psi_{i,n}(s) \right| ds + |\Gamma_{u} - \Gamma_{i,n}| + |C^{\alpha} - C^{\alpha}_{n}| \right) 1_{\{N(\alpha) > n\}} \right] = 0 \text{ a.s.}
$$

Therefore one can pass to the limit $n \to \infty$ in equation (5.45) to get,

$$
J(\alpha; t, i) \leq \lim_{n \to \infty} J(\alpha^n; t, i) \quad \text{a.s.} \quad (5.46)
$$

However, since $\alpha^n \in \mathcal{A}^n_{t,i}$ for each $n \geq 0$, from (5.46) and (5.44) it is true that for every $t \in [0, T]$

$$
J(\alpha; t, i) \leq \lim_{n \to \infty} J(\alpha^n; t, i) \leq \lim_{n \to \infty} Y_{i,n}^t = \tilde{Y}_i^t \quad \text{a.s.}
$$

Since $\alpha \in \mathcal{A}_{t,i}$ was arbitrary, for every $t \in [0, T]$

$$
V(t, i) := \text{ess sup}_{\alpha \in \mathcal{A}_{t,i}} J(\alpha; t, i) \leq \tilde{Y}_i^t \quad \text{a.s.}
$$

The reverse inequality holds since $Y_{i,n}^t = J(\hat{\alpha}^{(n)}; t, i) \leq V(t, i)$ almost surely for $n \geq 0$ (cf. (5.6.1)) and $\tilde{Y}_i^t$ is the pointwise supremum of the sequence $\{Y_{i,n}^t\}_{n \geq 0}$. \qed

### 5.6.2 The case of an arbitrary number of switches

The main result in this section proves that the limiting processes $\tilde{Y}^1, \ldots, \tilde{Y}^m$ satisfy the partial verification theorem 5.5.2. The main difficulty is in proving that $\tilde{Y}_i^t \in \mathcal{S}^2$, which in turn depends on the following

**Hypothesis (H1)** the sequence $\{\tilde{Y}_{i,n}^t\}_{n \geq 0}$ of Lemma 5.6.2 converges pointwise to $\tilde{Y}_i^t \in \mathcal{S}^2$.

Hypothesis (H1) can be verified a fortiori in at least two cases appearing in the literature:

**C1** there exists a constant $K > 0$ such that for almost every $\omega \in \Omega$ ([42, p. 6]):

$$
\text{Card} \{\gamma_{i,j}(\omega, t) < 0: \ i, j \in \mathbb{I}, j \neq i, \ t \in [0, T]\} \leq K. \quad (5.47)
$$

**C2** The switching costs are *time independent*, square-integrable and $\mathcal{F}_0$-measurable:

$$
\forall i, j \in \mathbb{I}, \ \gamma_{i,j}(t) := \gamma_{i,j} \in L^2(\Omega, \mathcal{F}_0, \mathbb{P}), \ \forall t \in [0, T]. \quad (5.48)
$$
In either case, one verifies directly using the arguments in Lemma 5.6.2 that the \( \mathbb{F} \)-martingale \( \zeta = (\zeta_t)_{0 \leq t \leq T} \) defined by
\[
\zeta_t := \mathbb{E} \left[ \int_0^T \max_{j \in I} |\psi_j(s)| \, ds + \max_{j \in I} |\Gamma_j| + \ell_* \cdot \max_{j_1, j_2 \in I} \left( \sup_{0 \leq s \leq t} |\gamma_{j_1, j_2}(s)| \right) \bigg\rvert \mathcal{F}_t \right] \tag{5.49}
\]
with \( \ell_* = K \) for (C1) or \( \ell_* = 1 \) for (C2) satisfies \( \zeta \in \mathcal{S}^2 \) and \( \mathbb{P} \)-a.s. for every \( t \in [0, T] \),
\[
|\hat{Y}_{t,n}^i| \leq \zeta_t, \quad \forall n \geq 0.
\]

For the case of (C2) see Lemma C.2.1 and note that the proof is easily generalised to switching costs which are martingales (see [97] for details). With the exception of non-negative switching costs, condition (C1) seems difficult check in practice. Condition (C2) is easy to check but the assumption may be considered too strong. Hypothesis (H1) can also be verified a fortiori from the results in [19, 91] obtained for Markovian models of optimal switching with signed switching costs.

**Theorem 5.6.3** (Existence). Suppose either (C1) or (C2). Then the limit processes \( \hat{Y}^1, \ldots, \hat{Y}^m \) of Lemma 5.6.2 satisfy the following: for \( i \in \mathbb{I} \),

1. \( \hat{Y}^i \in \mathcal{Q} \cap \mathcal{S}^2 \).

2. For any \( 0 \leq t \leq T \),
\[
\hat{Y}_t^i = \text{ess sup}_{\tau \geq t} \mathbb{E} \left[ \int_{\tau}^T \psi_i(s) \, ds + \Gamma_i 1_{\{\tau = T\}} + \max_{j \neq i} \left\{ \hat{Y}_\tau^j - \gamma_{i,j}(\tau) \right\} 1_{\{\tau < T\}} \bigg\rvert \mathcal{F}_\tau \right],
\]
\[
\hat{Y}_T^i = \Gamma_i.
\]

In particular, \( \hat{Y}^1, \ldots, \hat{Y}^m \) are unique and satisfy the verification theorem.

**Proof.** Recall the limit processes \( \hat{Y}^1, \ldots, \hat{Y}^m \) and \( \tilde{Y}^1, \ldots, \tilde{Y}^m \) from Lemma 5.6.1, equation (5.44). Under conditions (C1) or (C2), the martingale \( \zeta \) defined in (5.49) dominates \( \hat{Y}_{t,n}^i \) for all \( n \geq 0 \) pointwise in \( t \). Moreover, \( \hat{Y}_t^i \leq \zeta_t \) for each \( t \in [0, T] \) since \( \hat{Y}^i \) is the pointwise supremum of \( \{\hat{Y}_{t,n}^i\}_{n \geq 0} \). These observations give
\[
- \zeta_t \leq \hat{Y}_t^i \leq \zeta_t, \quad \mathbb{P} - \text{a.s.} \quad \forall 0 \leq t \leq T.
\]

Since \( \zeta \in \mathcal{S}^2 \), it follows that \( \hat{Y}^i \in \mathcal{S}^2 \) and also \( \tilde{Y}^i \in \mathcal{S}^2 \) since \( \psi_i \in \mathcal{M}^2 \). Now define a process \( \hat{U}^i = (\hat{U}_t^i)_{0 \leq t \leq T} \) for \( i = 1, \ldots, m \) similarly to \( \hat{U}_{t,n}^i \) used in Lemma 5.6.1:
\[
\hat{U}_t^i := \int_0^t \psi_i(s) \, ds + \Gamma_i 1_{\{t = T\}} + \max_{j \neq i} \left\{ \hat{Y}_t^j - \gamma_{i,j}(t) \right\} 1_{\{t < T\}}
\]
The $S^2$ processes $\hat{Y}^i$ and $\hat{U}^i$ are the respective limits of the increasing sequences of càdlàg $S^2$ processes $\{\hat{Y}^{i,n}\}_{n \geq 0}$ and $\{\hat{U}^{i,n}\}_{n \geq 0}$. Since $\hat{Y}^{i,n}$ is also the Snell envelope of $\hat{U}^{i,n}$, property 5 of Proposition 5.3.2 verifies that $\hat{Y}^i$ is the Snell envelope of $\hat{U}^i$. This leads to equation (5.50) for $\tilde{Y}^i$ and the uniqueness claim.

The final part is to show that $\tilde{Y}^i \in Q$. Let $\tau \in T$ be any predictable time. Since $\hat{Y}^i$ is the Snell envelope of $\hat{U}^i$, it has the following Meyer decomposition (cf. Proposition 5.3.2):

$$\hat{Y}^i = M - A - B,$$

where $M$ is a uniformly integrable càdlàg martingale and $A$ (resp. $B$) is non-decreasing, predictable and continuous (resp. discontinuous). Remember that $M$ is also quasi-left-continuous due to the assumption that the filtration $\mathbb{F}$ is quasi-left-continuous. This leads to the following

$$\triangle \tau \hat{Y}^i = (M_{\tau^-} - A_{\tau^-} - B_{\tau^-}) - (M_{\tau} - A_{\tau} - B_{\tau}) = -\triangle \tau B \quad \text{a.s.}$$

By property 2 of Proposition 5.3.2 concerning the jumps of $\hat{Y}^i$ (and therefore $\tilde{Y}^i$), the following holds

$$\{\triangle \tau B > 0\} \subset \{\hat{Y}^i_{\tau^-} = \hat{U}^i_{\tau^-}\}. \quad (5.52)$$

Using the definitions of $\hat{Y}^i$ and $\hat{U}^i$, the above shows:

$$\tilde{Y}^i_{\tau^-} < \tilde{Y}^i_{\tau^-} = \max_{j \neq i} \left\{ \tilde{Y}^j_{\tau^-} - \gamma_{i,j}(\tau^-) \right\} \quad \text{on} \quad \{\triangle \tau B > 0\}. \quad (5.53)$$

However, (5.52) also implies that the process $\left(\max_{j \neq i} \left\{ \hat{Y}^j_{t} - \gamma_{i,j}(t) \right\}\right)_{0 \leq t \leq T}$ jumps at time $\tau$ (since it is dominated by $\hat{Y}^i$). As the switching costs are quasi-left-continuous, we conclude that $\hat{Y}^{j^*}$ jumps at time $\tau$. Using the Meyer decomposition of $\hat{Y}^{j^*}$ and the properties of the jumps as before, this leads to

$$\tilde{Y}^{j^*}_{\tau^-} = \tilde{Y}^{j^*}_{\tau^-} - \gamma_{i,j^*}(\tau^-) = \max_{j \neq i} \left\{ \tilde{Y}^{j}_{\tau^-} - \gamma_{i,j}(\tau^-) \right\} \quad \text{on} \quad \{\triangle \tau B > 0\}. \quad (5.54)$$
Putting equations (5.53) and (5.54) together, then using the quasi-left-continuity of the switching costs and Assumption 6, the following (almost sure) inequality is obtained:

\[
\hat{Y}_i^\tau = -\gamma_{i,j^*}(\tau^-) + \hat{Y}_i^{j^*} \\
= -\gamma_{i,j^*}(\tau^-) - \gamma_{j^*,i^*}(\tau^-) + \hat{Y}_i^{i^*} \\
= -\gamma_{i,j^*}(\tau^-) - \gamma_{j^*,i^*}(\tau) + \hat{Y}_i^{i^*} \\
< -\gamma_{i,i^*}(\tau^-) + \hat{Y}_i^{i^*} \\
= -\gamma_{i,i^*}(\tau^-) + \hat{Y}_i^{i^*} \text{ on } \{\Delta_\tau B > 0\}.
\]

However, this inequality contradicts the optimality of mode \(j^*\) for \(\hat{Y}_i^\tau\). This means \(\{\Delta_\tau B = 0\}\) a.s. for every predictable time \(\tau\), and consequently \(\hat{Y}_i^\tau \in \mathcal{Q}\) for every \(i \in \mathbb{I}\).

This section is concluded with a convergence result in \(S^2\) for \(Y^1, \ldots, Y^m\) that is analogous to the one in [37].

**Proposition 5.6.4.** Suppose either (C1) or (C2). For every \(i \in \mathbb{I}\),

\[
\mathbb{E} \left[ \sup_{0 \leq s \leq T} \left| Y_s^{i,n} - Y_s^i \right|^2 \right] \to 0 \quad \text{as} \quad n \to +\infty.
\]  

**Proof.** The proof is established in the same as Proposition 4 of [37] or Proposition 2.1 of [49]. Using Theorem 5.6.3 and a weak version of Dini’s Lemma for càdlàg functions (cf. [32, p. 185]), in the \(\mathbb{P}\)-a.s. sense \(\{\hat{Y}_i^{i,n}\}_{n \geq 0}\) converges uniformly to \(Y^i\) on \([0, T]\). The dominated convergence theorem then gives (5.55). \(\square\)

## 5.7 Conclusion

This chapter investigated an extension of the multiple modes optimal switching problem in [37] to account for

1. non-zero, possibly different terminal rewards;
2. switching costs which are modelled by càdlàg, quasi-left-continuous processes;
3. signed switching costs;
4. more general filtrations, which are only assumed to satisfy the usual conditions and quasi-left-continuity.

Just as in Theorem 1 of [37], it was shown that the value function of the optimal switching problem can be defined stochastically in terms of interconnected Snell envelope-like processes. The existence of these processes was proved in a manner similar to Theorem 2 of [37], in which the sequence processes solving the optimal switching problem with at most \( n \geq 0 \) switches converges pointwise to a process with sufficient integrability. This convergence hypothesis can be verified a fortiori in some cases including:

- the number of times negative switching costs occur is limited as in [42];
- the switching costs are time-independent (and adapted to the filtration).

Further examples are given in [97]. Finally, the candidate optimal strategy defined in terms of the Snell envelope processes was shown to be admissible, and such that the cumulative switching cost is square-integrable.
Chapter 6

Pricing Cancellable Contracts for Difference via optimal switching

6.1 Introduction

This chapter provides a mathematical model for the cancellable contract for difference (CCfD) introduced in Chapter 1. The model, which is first described informally in Section 6.2 below, is an example of a stochastic game of timing known as the *Dynkin game*. This link is confirmed by a formal description of Dynkin games in Section 6.3. Given an appropriate model for the electricity market, it is then possible to determine the fair price for the premium demanded by the writer of the CCfD, and the optimal cancellation time for either party. There is also a natural connection between Dynkin games and the optimal switching problem of the previous chapter. The theory underlying this connection is presented in Section 6.5, which is then evidenced by a numerical example in Section 6.6. A representative value for the CCfD can therefore be obtained using the optimal switching formulation. This is presented in Section 6.7 for a formal mathematical model of the contract where the costs are driven by a multidimensional Markov process.

6.2 The CCfD in a local electricity market

As described in Chapter 1, the contract for difference (CfD) can be viewed as a financial contract between a generator and an electricity supplier in which certain cash flows
are traded over a finite period of time. The model introduced here assumes that these payments are made continuously, so that the contract’s duration is an interval \([0, T], 0 < T < \infty\). Section 6.7 below provides a more detailed model based on this introduction.

The terms of the CCfD described here are more realistic for generators and suppliers in a local electricity market. The generator is assumed to produce its electricity from an intermittent renewable source of energy. Let \(R(t)\) denote the \textit{non-negative} level of its renewable electricity supply at time \(t \in [0, T]\). Just as in the usual CfD, the generator can sell its electricity on the spot market at the prevailing spot price \(S(t), t \in [0, T]\), without restriction. On the other hand, the generator must cover the cost of the supplier’s shortfall whenever the latter’s demand surpasses a pre-defined level. This level may represent the total amount of supply from other contracts that the supplier has negotiated. Let \(D(t)\) and \(d^*(t), t \in [0, T]\), respectively denote the supplier’s electricity demand and the level for triggering the payments under the CCfD. The supplier’s \textit{energy shortfall} over time is given by

\[
D^*(t) := \max(D(t) - d^*(t), 0), \quad 0 \leq t \leq T.
\]  

The terms of the CCfD specify that the generator should use its own generation to cover as much of the shortfall as possible, at a rate determined by the difference between the market price \(S(\cdot)\) and a fixed strike price \(K\). The continuous payment from the generator to the supplier while the contract is active is then given by

\[
\psi(t) = \min(D^*(t), R(t)) \times (S(t) - K).
\]

The main difference between the CCfD and the regular CfD is that both parties have the option to exit (terminate) the CCfD early. The electricity supplier may decide to exit the contract early if, for instance, the market price is expected to lie below the strike price \(K\) for a prolonged period. In this case the payment from the generator to the supplier in (6.2) is likely to be negative. The supplier’s penalty for contract termination, which is denoted by \(\gamma_+\), serves as deterrent to its early exit. It therefore helps increase the likelihood of the generator recovering part of its investment cost. On the other hand, the generator may decide to exit the contract early if, for instance, market prices are significantly higher than the strike price. If it exits early then it
must pay a fine $\gamma_-$, which may be viewed as compensation for the supplier since the latter:

a) may be obliged in some way to purchase energy from renewable sources (for example, the Renewables Obligation [106]), or

b) may have to curtail the surplus demand and therefore require payment for the value of lost load.

Both termination penalties are allowed to vary over time and are assumed to satisfy $\gamma_-(t) > \gamma_+(t) \geq 0$ for all $t \in [0, T]$.

Let $\sigma$ (resp. $\tau$) denote the time that the generator (resp. supplier) opts out of the contract, the total cost from the perspective of the generator is:

$$D(\sigma, \tau) = \int_0^{\sigma \wedge \tau} \psi(t) dt + \gamma_-(\sigma) 1_{\{\sigma \leq \tau < T\}} - \gamma_+\tau+1_{\{\tau < \sigma\}}.$$  \hspace{1cm} (6.3)

A discount rate $\mu(\cdot) \geq 0$ which reflects the “time value of money” can also be included in the payoff:

$$\tilde{D}(\sigma, \tau) = \int_0^{\sigma \wedge \tau} e^{-\mu(t)} \psi(t) dt + e^{-\mu(\sigma)} \gamma_-(\sigma) 1_{\{\sigma \leq \tau < T\}} - e^{-\mu(\tau)} \gamma_+\tau+1_{\{\tau < \sigma\}}.$$  

The generator (supplier) chooses $\sigma$ (resp. $\tau$) so as to minimise (maximise) the total payment $D(\sigma, \tau)$. An appropriate concept of a solution to the game is discussed in the following section. Assumptions on the payment rate and stopping costs which ensure that such a solution exists are given for the detailed model in Section 6.7 below.

### 6.3 The Dynkin game

The CCfD is an example of the classic zero-sum Dynkin game with two players. These games have been studied extensively since Eugene Dynkin introduced them in the 1960s, and have garnered renewed interest due to the creation of Israeli Options by Yuri Kifer circa 2000 [81, pp. 1–2].

#### 6.3.1 Notation for the game

The time horizon $[0, T]$, $0 < T < \infty$, is the same as above. A probability space $(\Omega, \mathcal{F}, P)$ lies at the heart of the model, and it is equipped with a filtration $\mathcal{F} =$
The usual conditions of right-continuity and completeness are assumed for the stochastic basis \((\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})\). All adapted stochastic processes and stopping times mentioned below are assumed to be defined with respect to the given stochastic.

For a given stopping time \(\theta\), let \(\mathcal{T}_\theta\) denote the set of stopping times \(S\) which satisfy \(\theta \leq S \leq T\), \(\mathbb{P}\)-a.s., and set \(\mathcal{T} = \mathcal{T}_0\). Let \(\mathbb{E}\) denote the expectation operator and \(1_A\) be the indicator function of a set \(A\). As is customary, the dependence of a stochastic process or random function on \(\omega \in \Omega\) is often suppressed in the following.

### 6.3.2 The rules and payoff of the game

The Dynkin game between two players \(MIN\) and \(MAX\) is played until the time \(\sigma \wedge \tau := \min(\sigma, \tau)\), where \(\sigma \in \mathcal{T}\) and \(\tau \in \mathcal{T}\) are the respective strategies of \(MIN\) and \(MAX\). As long as the game progresses, \(MIN\) pays \(MAX\) at a rate of \(\psi(t)\) per unit time. If \(MIN\) exits the game prior to \(T\) and either before or at the same time that \(MAX\) exits, \(\sigma < T\) and \(\sigma \leq \tau\), then \(MIN\) pays \(MAX\) the amount \(\gamma_-(\sigma)\). Alternatively, if \(MAX\) exits the game first, \(\tau < \sigma\), then \(MAX\) pays to \(MIN\) the amount \(\gamma_+(\tau)\). If neither player exits the game before time \(T\), the rule is to set \(\sigma = \tau = T\) and have \(MIN\) pay \(MAX\) the amount \(\Gamma\). The payoff of the Dynkin game is then defined in terms of the cost to player \(MIN\):

\[
D(\sigma, \tau) = \int_0^{\sigma \wedge \tau} \psi(s)ds + \gamma_-(\sigma)1_{\{\sigma \leq \tau\}}1_{\{\sigma < T\}} - \gamma_+(\tau)1_{\{\tau < \sigma\}} + \Gamma 1_{\{\sigma = \tau = T\}}. \tag{6.4}
\]

This particular variant of the Dynkin game was studied in recent papers such as [31, 39, 65, 67]. It is a zero-sum game since costs (gains) for \(MIN\) are the gains (costs) for \(MAX\).

### 6.3.3 The players’ objectives and solution concepts

Player \(MIN\) chooses the strategy \(\sigma\) to minimise the expected value of \(D(\sigma, \tau)\) in (6.4) whereas \(MAX\) plays the strategy \(\tau\) to maximise it. This leads to upper and lower values for the game, \(W^+\) and \(W^-\) respectively, which are given by:

\[
W^+ := \inf_{\sigma \in \mathcal{T}} \sup_{\tau \in \mathcal{T}} \mathbb{E}[D(\sigma, \tau)], \quad W^- := \sup_{\tau \in \mathcal{T}} \inf_{\sigma \in \mathcal{T}} \mathbb{E}[D(\sigma, \tau)] \tag{6.5}
\]
Definition 6.3.1 (Game Value). The Dynkin game is said to be “fair” if there is equality between the upper and lower values,

$$\inf_{\sigma \in \mathcal{T}} \sup_{\tau \in \mathcal{T}} \mathbb{E}[D(\sigma, \tau)] = W = \sup_{\tau \in \mathcal{T}} \inf_{\sigma \in \mathcal{T}} \mathbb{E}[D(\sigma, \tau)]. \quad (6.6)$$

The common value $W$, when it exists, is referred to as the solution or value of the game.

Existence of a solution to the Dynkin game means neither player has an incentive to play first. When studying Dynkin games, the first course of action is to verify that the game has a solution. In fact, all that is required is to show $W^- \geq W^+$, since $W^+ \geq W^-$ follows by definition (see Lemma D.1.1 in the appendix). After establishing that the game has a solution, one searches for strategies for the players which give the game’s value or approximate it closely. This leads to the concept of a Nash equilibrium.

Definition 6.3.2 (Nash equilibrium). A pair of stopping times $(\sigma^*, \tau^*) \in \mathcal{T} \times \mathcal{T}$ is said to constitute a Nash equilibrium or a saddle point if the following property holds for any $\sigma, \tau \in \mathcal{T}$:

$$\mathbb{E}[D(\sigma^*, \tau)] \leq \mathbb{E}[D(\sigma^*, \tau^*)] \leq \mathbb{E}[D(\sigma, \tau^*)]. \quad (6.7)$$

In other words, then neither player can do better by deviating unilaterally from a Nash equilibrium point. It is readily verified (for instance, using arguments similar to Lemma D.1.1 below) that the existence of a saddle point $(\sigma^*, \tau^*)$ implies the game is fair, and its value is given by

$$\inf_{\sigma \in \mathcal{T}} \sup_{\tau \in \mathcal{T}} \mathbb{E}[D(\sigma, \tau)] = \mathbb{E}[D(\sigma^*, \tau^*)] = \sup_{\tau \in \mathcal{T}} \inf_{\sigma \in \mathcal{T}} \mathbb{E}[D(\sigma, \tau)]. \quad (6.8)$$

6.4 Notation and assumptions

6.4.1 Notation

The notation used for different classes of processes is the same as Chapter 5, and is repeated below for convenience.

1. For $p \geq 1$, let $L^p$ denote the set of random variables $Z$ satisfying $\mathbb{E}[||Z||^p] < \infty$. 

2. For \( p \geq 1 \), let \( \mathcal{M}^p \) denote the set of progressively measurable, real-valued processes \( X = (X_t)_{0 \leq t \leq T} \) satisfying,

\[
\mathbb{E} \left[ \int_0^T |X_t|^p dt \right] < \infty.
\]

3. For \( p \geq 1 \), let \( \mathcal{S}^p \) denote the set of progressively measurable processes \( X \) satisfying:

\[
\mathbb{E} \left[ \left( \sup_{0 \leq t \leq T} |X_t| \right)^p \right] < \infty.
\]

4. Let \( \mathcal{Q} \) denote the set of adapted, càdlàg processes which are quasi-left-continuous.

### 6.4.2 Assumptions

**Assumption 7.** The following integrability and measurability assumptions are made on the game parameters:

1. The instantaneous payoff satisfies \( \psi \in \mathcal{M}^1 \);

2. The switching costs satisfy \( \gamma_{\pm} \in \mathcal{S}^1 \);

3. The terminal payoff satisfies \( \Gamma \in L^1 \);

The following standing assumptions are also imposed on the stopping costs and terminal data:

**Assumption 8.** Switching cost and terminal data assumptions:

\[ i) \quad -\gamma_+(T) \leq \Gamma \leq \gamma_-(T) \quad \mathbb{P} - \text{a.s.} \quad (6.9) \]

\[ ii) \quad \forall \ t \in [0,T] : \quad \gamma_-(t) + \gamma_+(t) > 0 \quad \mathbb{P} - \text{a.s.} \quad (6.10) \]

The first condition (6.9) is standard in the literature on Dynkin games – see [41, 81] for instance. On the other hand, the second condition (6.10) is slightly stronger than the usual one for Dynkin games, where the above inequality is not necessarily strict. Nevertheless, it is often satisfied in practice as exemplified by the CCfD described above, the reversible investment problem in Section 6.6 below, and the callable (cancellable) put option in [85]. Moreover, the strict inequality in (6.10) is useful for establishing various results related to the game’s solution [67, pp. 135–144], and is typical for the optimal switching problem described in the following section [65].
6.5 Solving Dynkin games by optimal switching

6.5.1 Two-mode optimal switching

The two-mode optimal switching or “starting and stopping” problem has been studied in a variety of contexts (see [65, 68, 69] and the references therein). The two modes (also called regimes) are customarily denoted by 0 (stop) and 1 (start). Over a finite time horizon $[0, T]$, a controller earns a profit while operating in a particular mode, and this rate is determined by the random function $\Psi : \Omega \times [0, T] \times \{0, 1\} \to \mathbb{R}$. There is a cost incurred at each time when switching to a particular mode which is determined by the mapping $\gamma : \Omega \times [0, T] \times \{0, 1\} \to \mathbb{R}$. Finally, at time $T$ the controller earns a random reward $\Phi : \Omega \times \{0, 1\} \to \mathbb{R}$ which depends on the mode that is active at the time. The controller therefore influences the total profit through the operating modes it uses over $[0, T]$. Further assumptions on these rewards / costs are given below.

Definition 6.5.1 (Admissible switching controls). For a fixed time $t \in [0, T]$ and initial mode $i \in \{0, 1\}$, a switching control $\alpha = (\tau_n, \iota_n)_{n \geq 0}$ starting from $(t, i)$ consists of:

1. a non-decreasing sequence of stopping times $\{\tau_n\}_{n \geq 0} \subset T$ with $\tau_0 = t$;
2. a sequence $\{\iota_n\}_{n \geq 0}$ of modes, where $\iota_0 = i$ is the initial mode, $\iota_n : \Omega \to \{0, 1\}$ is $\mathcal{F}_{\tau_n}$-measurable and satisfies $\iota_{2n} = i$ and $\iota_{2n+1} = 1 - i$ for $n \geq 0$.

The switching control $\alpha$ is said to be admissible if it additionally satisfies:

3. the stopping times $\{\tau_n\}_{n \geq 0}$ are finite in the following sense:

   $$P\left(\{\tau_n < T, \forall n \geq 0}\right) = 0. \quad (6.11)$$

4. the sequence $\{C_n^\alpha\}_{n \geq 1}$ formed using the cumulative switching cost after $n \geq 1$ switches

   $$C_n^\alpha := \sum_{k=1}^{n} \gamma(\tau_k, \iota_k) 1_{[\tau_k < T]} \quad (6.12)$$

satisfies the condition

$$E\left[\sup_n |C_n^\alpha|\right] < \infty. \quad (6.13)$$
The set of all admissible switching controls starting from \((t, i)\) is denoted by \(A_{t,i}\). The notation \(A_i\) is used when \(t = 0\) and the subscript \(i\) is dropped when the initial mode is not important for the discussion.

Associated with each \(\alpha \in A\) is a random function \(u: \Omega \times [0,T] \to \{0,1\}\), referred to as the mode indicator function, which is defined by

\[
u_t := \nu_0 1_{\{t_0 \leq t \leq t_1\}} + \sum_{n \geq 1} \nu_n 1_{\{t_n < t \leq t_{n+1}\}}.
\]

(6.14)

The performance index for the switching control problem is given by,

\[
J(\alpha; i) := E\left[\int_0^T \Psi(s, u_s)ds + \Phi(u_T) - \sum_{n \geq 1} \gamma_\alpha(t_n, \nu_n) 1_{\{t_n < T\}}\right], \quad \alpha \in A_i.
\]

(6.15)

The optimisation problem is to maximise the performance index over all admissible switching controls. This objective is expressed using the value function \(V\), which is written in terms of the initial mode \(i \in \{0,1\}\),

\[
V(i) := \sup_{\alpha \in A_i} J(\alpha; i).
\]

(6.16)

**Remark 6.5.1.** Condition (6.13) is absent (and not necessary) in the papers [68, 69] where strictly positive switching costs are assumed. It helps avoid counterexamples where the conditional dominated convergence may fail (see [17], for instance), and provides enough integrability to arrive at conclusions similar to [68] for problems where switching costs can be negative (see Chapter 5).

It should also be noted that the admissible switching controls of [65, p. 430] satisfy a stronger integrability condition

\[
E\left[\sum_{n \geq 1} |\gamma(t_n, \nu_n)| 1_{\{t_n < T\}}\right] < \infty
\]

whilst the switching times must only satisfy \(t_n < t_{n+1}\) and \(t_n \to T\), which is weaker than (6.11) above. To see why the integrability condition in [65] is stronger, it is sufficient to consider what happens if the switching costs under \(\alpha = (t_n, \nu_n)_{n \geq 0}\) behave like the terms in a conditionally convergent series such as \(\sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n}\).

### 6.5.2 Existence of the game’s value

This section shows how to use an appropriate two-mode optimal switching problem to show that the Dynkin game has a solution. For this purpose, the rewards and costs
are assumed to satisfy the following special form: for all \( t \in [0, T] \) and \( j \in \mathbb{I} \),

\[
\Psi(t, j) := \psi(t) \cdot j \tag{6.17}
\]

\[
\gamma(t, j) := \gamma_- (t) \mathbf{1}_{\{j=1\}} + \gamma_+ (t) \mathbf{1}_{\{j=0\}} \tag{6.18}
\]

\[
\Phi(j) := \Gamma \cdot j \tag{6.19}
\]

where \( \psi, \gamma_{\pm} \) and \( \Gamma \) are the parameters in the Dynkin game’s payoff \( (6.4) \). The performance index for the auxiliary two-mode optimal switching problem associated with the Dynkin game is then given by,

\[
J(\alpha; i) = \mathbb{E} \left[ \int_0^T \psi(s) u_s ds + \Gamma u_T - \sum_{n \geq 1} \gamma(\tau_n, \iota_n) \mathbf{1}_{\{\tau_n < T\}} \right], \quad \alpha \in \mathcal{A}_i. \tag{6.20}
\]

The integrability assumptions on \( \psi, \Gamma \) and \( \gamma \), which are inherited from Section 6.4 above, ensure that the performance index is well-defined for every \( \alpha \in \mathcal{A} \).

**Definition 6.5.2** (\( \varepsilon \)-optimal strategies.). Given \( \varepsilon > 0 \), \( \alpha \in \mathcal{A}_i \) is said to be an \( \varepsilon \)-optimal strategy for the optimal switching problem with initial mode \( i \in \{0, 1\} \) if it satisfies:

\[
V(i) \leq J(\alpha; i) + \varepsilon.
\]

Note that Assumptions 7 and 8 ensure that the value function \( V(i) \) is well-defined for \( i \in \{0, 1\} \). The existence of \( \varepsilon \)-optimal strategies can be verified if \( V(i) \) is finite.

**Theorem 6.5.1.** Let \( V(i), i \in \{0, 1\} \), be the value function for the optimal switching problem with performance index \( (6.20) \). Suppose there exist \( \varepsilon \)-optimal strategies. Then the Dynkin game with payoff \( (6.4) \) is fair and its value is given by:

\[
\inf_{\sigma \in \mathcal{F}} \sup_{\tau \in \mathcal{T}} \mathbb{E} [D(\sigma, \tau)] = V(1) - V(0) = \sup_{\tau \in \mathcal{T}} \inf_{\sigma \in \mathcal{F}} \mathbb{E} [D(\sigma, \tau)]. \tag{6.21}
\]

The proof, which follows in the same manner as the one for Theorem 3.15 of [65], is given in Appendix D.

### 6.5.3 Existence of a Nash equilibrium

This section establishes the existence of a saddle point \( (\sigma^*, \tau^*) \) for the Dynkin game with payoff \( (6.4) \) using the theory for optimal switching problems developed in Chapter 5.
The martingale approach to optimal switching problems.

Assumption 9. In addition to Assumption 8 above, also suppose that

- the filtration \( \mathbb{F} \) satisfies the usual conditions and is quasi-left-continuous;
- the instantaneous payoff rate satisfies \( \psi \in \mathcal{M}^2 \);
- the early-exit stopping costs for the game satisfy \( \gamma_- , \gamma_+ \in \mathcal{S}^2 \cap \mathcal{Q} \);
- the terminal payoff satisfies \( \Gamma \in L^2 \).

The following hypothesis is crucial in what follows.

(H1) There exists a pair of processes \( (Y^0_t,Y^1_t)_{0 \leq t \leq T} \) in \( \mathcal{S}^2 \cap \mathcal{Q} \) such that for \( i \in \{0,1\} \) and \( 0 \leq t \leq T \), \( Y^i_t \) is defined by,

\[
Y^i_t = \text{ess sup}_{\theta \in \mathcal{T}_t} \mathbb{E} \left[ \int_t^\theta \psi_i(s) ds + \Gamma_i 1_{\{\theta = T\}} + \{Y^{1-i}_\theta - \gamma(\theta, 1-i)\} 1_{\{\theta < T\}} \right] \mathbb{F}_t ,
\]

\[
Y^i_T = \Gamma_i, \quad \mathbb{P} - a.s.
\]

where \( \gamma(\cdot, 1-i) \) is the switching cost defined by equation (6.18) and

\[
\psi_1(\cdot) \equiv \psi(\cdot), \quad \psi_0(\cdot) \equiv 0, \quad \Gamma_1 \equiv \Gamma; \quad \Gamma_0 \equiv 0.
\]

Under Hypothesis (H1), the random variable \( Y^i_t \) gives the optimal switching control performance for the initial data \( (t,i) \in [0,T] \times \{0,1\} \):

\[
Y^i_t = \text{ess sup}_{\alpha \in A_{t,i}} \mathbb{E} \left[ \int_t^T \psi(s) ds + \Gamma u_T - \sum_{n \geq 1} \gamma(\tau_n,t_n) 1_{\{\tau_n < T\}} \right] \mathbb{F}_t
\]

so that \( Y^i_0 = V(i) \) in particular (see Chapter 5). This hypothesis and Assumption 9 can be verified in many cases of interest, including the two examples of this chapter. See Section 5.6.2 of Chapter 5 for a more precise statement of the hypothesis.

Existence of Nash equilibria via optimal switching.

Define a pair of stopping times \( (\sigma^*, \tau^*) \) as follows

\[
\sigma^* := \inf \{ t \geq 0 : Y^0_t = -\gamma_- (t) + Y^1_t \} \wedge T
\]
\[
\tau^* := \inf \{ t \geq 0 : Y^1_t = -\gamma_+ (t) + Y^0_t \} \wedge T
\]
Then $\sigma^*$ (resp. $\tau^*$) is an optimal first switching time when starting in mode 0 (resp. 1) at time 0. It turns out that $(\sigma^*, \tau^*)$ is also a saddle point for the Dynkin game. The following lemma, which relates the pair $(Y^0, Y^1)$ to a classical assumption known as Mokobodski’s hypothesis, is a first step in the proof (see Section 5.4 of [95] for more details).

**Lemma 6.5.2.** Let $Y^i = (Y^i_t)_{0 \leq t \leq T}$, $i \in \{0, 1\}$, be the processes in Hypothesis (H1). Then $Y^0$ and $Y^1$ satisfy the following condition:

$$\forall \tau \in T : -\gamma_+(\tau) \leq Y^1_\tau - Y^0_\tau \leq \gamma_-(\tau), \quad P\text{-a.s.} \quad (6.26)$$

**Proof.** Let $U^i = (U^i_t)_{0 \leq t \leq T}$, $i \in \{0, 1\}$, be defined by

$$U^i_t = \Gamma_t \mathbf{1}_{\{t = T\}} + \{Y^1_{t-i} - \gamma(t, 1-i)\} \mathbf{1}_{\{t < T\}}. \quad (6.27)$$

Then $Y^i_t + \int^t_0 \psi_i(s)ds$ is the Snell envelope of $U^i_t + \int^t_0 \psi_i(s)ds$ on $0 \leq t \leq T$ (cf. Proposition 5.3.2). Let $\tau \in T$ be arbitrary. By the dominating property of the (right-continuous) Snell envelope, $Y^i_\tau \geq U^i_\tau$ holds $P$-a.s. Therefore, one gets

$$0 \leq Y^i_\tau - U^i_\tau = Y^i_\tau - Y^1_{\tau-i} + \gamma(\tau, 1-i) \text{ on } \{\tau < T\}$$

which means

$$-\gamma_+(\tau) \leq Y^1_\tau - Y^0_\tau \leq \gamma_-(\tau) \text{ on } \{\tau < T\}.$$

On the other hand, $Y^1_\tau - Y^0_\tau = \Gamma$ holds $P$-a.s. on the event $\{\tau = T\}$. Using this with condition (6.9) gives

$$-\gamma_+(\tau) \leq Y^1_\tau - Y^0_\tau \leq \gamma_-(\tau) \text{ on } \{\tau = T\}$$

and the claim (6.26) follows. \qed

**Proposition 6.5.3.** Let $Y^0$ and $Y^1$ be the processes in Hypothesis (H1). Then the pair of stopping times $(\sigma^*, \tau^*)$ defined in (6.25) satisfies:

$$Y^1_{\sigma^*} - Y^0_{\sigma^*} = \mathbb{E}[D(\sigma^*, \tau^*)] \quad (6.28)$$

where $D(\sigma^*, \tau^*)$ is the payoff of the Dynkin game (6.4) under the strategy $(\sigma^*, \tau^*)$. 
Proof. For $i \in \{0, 1\}$, let $U^i = (U^i_t)_{0 \leq t \leq T}$ be defined as in equation (6.27). By Proposition 5.3.2, the following stopped Snell envelopes are martingales:

\[
(Y^0 \mathbb{I}_{\tau^*})_{0 \leq t \leq T}, \quad (Y^1 \mathbb{I}_{\tau^*} + \int_0^{\tau^*} \psi(r) dr)_{0 \leq t \leq T}.
\]

The process \(\left( (Y^1_t + \int_0^t \psi(r) dr) - Y^0_t \right)_{0 \leq t \leq T}\) is therefore a martingale on \([0, \sigma^* \wedge \tau^*]\), and this yields:

\[
Y^0_1 - Y^0_0 = \mathbb{E} \left[ \int_0^{\sigma^* \wedge \tau^*} \psi(r) dr + Y^1_{\sigma^* \wedge \tau^*} - Y^0_{\sigma^* \wedge \tau^*} \right] + \mathbb{E} \left[ \psi \left( \max \left\{ \sigma^* \wedge \tau^*, \min \left\{ \tau, \sigma^* \right\} \right\} \right].
\]

The term involving the pair \((Y^0, Y^1)\) inside of the expectation may be rewritten as:

\[
\mathbb{E} \left[ Y^1_{\sigma^* \wedge \tau^*} - Y^0_{\sigma^* \wedge \tau^*} \right] = \mathbb{E} \left[ (Y^1_{\sigma^*} - Y^0_{\sigma^*}) \mathbb{1}_{\{\sigma^* \leq \tau^*\}} + (Y^1_{\tau^*} - Y^0_{\tau^*}) \mathbb{1}_{\{\tau^* < \sigma^*\}} \right].
\]

**Step 1** By equation (6.25) and conditional on the event \(\{\tau^* < T\}\), optimality of the stopping time \(\tau^*\) gives the following:

\[
Y^1_{\tau^*} \mathbb{1}_{\{\tau^* < T\}} = \left[ -\gamma_+ (\tau^*) + Y^0_{\tau^*} \right] \mathbb{1}_{\{\tau^* < T\}};
\]

Furthermore, since \(\tau^* \leq T\) and \(\sigma^* \leq T\), it is also true that:

\[
\mathbb{1}_{\{\tau^* < \sigma^*\}} = \mathbb{1}_{\{\tau^* < \sigma^*\}} \mathbb{1}_{\{\tau^* < T\}} = \mathbb{1}_{\{\tau^* < \sigma^*\}} \left( \mathbb{1}_{\{\tau^* < T\}} + \mathbb{1}_{\{\tau^* = T\}} \right) = \mathbb{1}_{\{\tau^* < \sigma^*\}} \mathbb{1}_{\{\tau^* < T\}}.
\]

Using this together with equation (6.31) gives:

\[
\mathbb{E} \left[ (Y^1_{\tau^*} - Y^0_{\tau^*}) \mathbb{1}_{\{\tau^* < \sigma^*\}} \right] = \mathbb{E} \left[ (Y^1_{\tau^*} - Y^0_{\tau^*}) \mathbb{1}_{\{\tau^* < \sigma^*\}} \mathbb{1}_{\{\tau^* < T\}} \right]
\]

\[
= \mathbb{E} \left[ (-\gamma_+ (\tau^*)) \mathbb{1}_{\{\tau^* < \sigma^*\}} \mathbb{1}_{\{\tau^* < T\}} \right]
\]

\[
= \mathbb{E} \left[ (-\gamma_+ (\tau^*)) \mathbb{1}_{\{\tau^* < \sigma^*\}} \right].
\]

**Step 2** By equation (6.25) and conditional on the event \(\{\sigma^* < T\}\), optimality of the stopping time \(\sigma^*\) gives:

\[
Y^0_{\sigma^*} \mathbb{1}_{\{\sigma^* < T\}} = \left[ -\gamma_- (\sigma^*) + Y^1_{\sigma^*} \right] \mathbb{1}_{\{\sigma^* < T\}},
\]

which is used to deduce:

\[
\mathbb{E} \left[ (Y^1_{\sigma^*} - Y^0_{\sigma^*}) \mathbb{1}_{\{\sigma^* \leq \tau^*\}} \mathbb{1}_{\{\sigma^* < T\}} \right] = \mathbb{E} \left[ \gamma_- (\sigma^*) \mathbb{1}_{\{\sigma^* \leq \tau^*\}} \mathbb{1}_{\{\sigma^* < T\}} \right].
\]
Since $\tau^* \leq T$, it follows that $\mathbf{1}_{\{\sigma^* \leq \tau^*\}} \mathbf{1}_{\{\sigma^* = T\}} = \mathbf{1}_{\{\sigma^* = \tau^* = T\}}$. Furthermore, as $Y^1_T = \Gamma$ and $Y^0_T = 0$ almost surely, one sees that
\[
\mathbb{E} \left[ (Y^1_{\sigma^*} - Y^0_{\sigma^*}) \mathbf{1}_{\{\sigma^* \leq \tau^*\}} \mathbf{1}_{\{\sigma^* = T\}} \right] = \mathbb{E} \left[ (Y^1_{\sigma^*} - Y^0_{\sigma^*}) \mathbf{1}_{\{\sigma^* = \tau^* = T\}} \right] = \mathbb{E} \left[ \Gamma \mathbf{1}_{\{\sigma^* = \tau^* = T\}} \right].
\] (6.34)

Equations (6.33) and (6.34) can now be used to assert that the following holds:
\[
\mathbb{E} \left[ (Y^1_{\sigma^*} - Y^0_{\sigma^*}) \mathbf{1}_{\{\sigma^* \leq \tau^*\}} \right] = \mathbb{E} \left[ (Y^1_{\sigma^*} - Y^0_{\sigma^*}) \mathbf{1}_{\{\sigma^* \leq \tau^*\}} (\mathbf{1}_{\{\sigma^* < T\}} + \mathbf{1}_{\{\sigma^* = T\}}) \right]
= \mathbb{E} \left[ \gamma^- (\sigma^*) \mathbf{1}_{\{\sigma^* \leq \tau^*\}} \mathbf{1}_{\{\sigma^* < T\}} + \Gamma \mathbf{1}_{\{\sigma^* = \tau^* = T\}} \right].
\] (6.35)

**Conclusion** Returning to equation (6.29) and using equations (6.30), (6.32) and (6.35), the proof is concluded since
\[
Y^1_0 - Y^0_0 = \mathbb{E} \left[ \int_0^{\sigma^* \wedge \tau^*} \psi(r)dr + Y^1_{\sigma^* \wedge \tau^*} - Y^0_{\sigma^* \wedge \tau^*} \right]
= \mathbb{E} \left[ \int_0^{\sigma^* \wedge \tau^*} \psi(r)dr \right] + \mathbb{E} \left[ (Y^1_{\sigma^*} - Y^0_{\sigma^*}) \mathbf{1}_{\{\sigma^* \leq \tau^*\}} + (Y^1_{\tau^*} - Y^0_{\tau^*}) \mathbf{1}_{\{\tau^* < \sigma^*\}} \right]
= \mathbb{E} \left[ \int_0^{\sigma^* \wedge \tau^*} \psi(r)dr \right] + \mathbb{E} \left[ \gamma^- (\sigma^*) \mathbf{1}_{\{\sigma^* \leq \tau^*\}} \mathbf{1}_{\{\sigma^* < T\}} + \Gamma \mathbf{1}_{\{\sigma^* = \tau^* = T\}} \right]
+ \mathbb{E} \left[ (-\gamma^+ (\tau^*)) \mathbf{1}_{\{\tau^* < \sigma^*\}} \right]
= \mathbb{E}[D(\sigma^*, \tau^*)].
\]

\[\Box\]

**Theorem 6.5.4.** Let $(\sigma^*, \tau^*)$ be the pair of stopping times defined in (6.25). Then for any $\sigma, \tau \in \mathcal{T}$:
\[
\mathbb{E} \left[ D(\sigma^*, \tau) \right] \leq \mathbb{E} \left[ D(\sigma^*, \tau^*) \right] \leq \mathbb{E} \left[ D(\sigma, \tau^*) \right],
\] (6.36)
The pair $(\sigma^*, \tau^*)$ is therefore a saddle point for the Dynkin game with payoff (6.4).

**Proof.** Proposition 6.5.3 showed that $Y^1_0 - Y^0_0 = \mathbb{E}[D(\sigma^*, \tau^*)]$, where $Y^0$ and $Y^1$ are the processes from Hypothesis (H1). Let $\sigma, \tau \in \mathcal{T}$ be arbitrary. Define a process $\hat{Y}^1 = (\hat{Y}^1_t)_{0 \leq t \leq T}$ by $\hat{Y}^1_t := Y^1_t + \int_0^t \psi(r)dr$. By Theorem II.77.4 of [117], a stopped supermartingale is also a supermartingale. The following stopped Snell envelopes are therefore supermartingales:
\[
(Y^0_{t \wedge (\sigma \wedge \tau^*)})_{0 \leq t \leq T}, \quad (\hat{Y}^1_{t \wedge (\sigma^* \wedge \tau^*)})_{0 \leq t \leq T}.
\]

Then, using the martingale property of the stopped Snell envelope as in Proposition 6.5.3, $\hat{Y}^1 - Y^0$ satisfies the following: for any $\sigma, \tau \in \mathcal{T}$,
1. \( \left( \hat{Y}^1_t - Y^0_t \right)_{0 \leq t \leq (\sigma^* \wedge \tau)} \) is a supermartingale;

2. \( \left( \hat{Y}^1_t - Y^0_t \right)_{0 \leq t \leq (\sigma \wedge \tau^*)} \) is a submartingale.

**Case (A)**  The following lines deal with the first inequality in the claim (6.36). The supermartingale property of \( \hat{Y}^1 - Y^0 \) on \([0, \sigma^* \wedge \tau]\) and the definition of \( \hat{Y}^1 \) gives:

\[
Y^1_0 - Y^0_0 \geq E \left[ \int_0^{\sigma^* \wedge \tau} \psi(r)dr + Y^1_{\sigma^* \wedge \tau} - Y^0_{\sigma^* \wedge \tau} \right] \\
= E \left[ \int_0^{\sigma^* \wedge \tau} \psi(r)dr \right] + E \left[ (Y^1_{\sigma^*} - Y^0_{\sigma^*}) 1_{\{\sigma^* \leq \tau^*\}} \right] \\
+ E \left[ (Y^1_{\tau^*} - Y^0_{\tau^*}) 1_{\{\tau^* \leq \sigma^*\}} \right]. \\
(6.37)
\]

Then, using the same arguments as in the proof of Proposition 6.5.3 leading up to equation (6.35), the following can be shown:

\[
E \left[ (Y^1_{\sigma^*} - Y^0_{\sigma^*}) 1_{\{\sigma^* \leq \tau\}} \right] = E \left[ \gamma_-(\sigma^*) 1_{\{\sigma^* \leq \tau\}} 1_{\{\sigma^* < T\}} + \Gamma 1_{\{\sigma^* = \tau = T\}} \right]. \\
(6.38)
\]

On the other hand, \( Y^1_{\tau^*} - Y^0_{\tau^*} \geq -\gamma_+(\tau) \) almost surely by Lemma 6.5.2. Using this together with equation (6.38) gives the conclusion:

\[
Y^1_0 - Y^0_0 \geq E \left[ \int_0^{\sigma^* \wedge \tau} \psi(r)dr \right] + E \left[ (Y^1_{\sigma^*} - Y^0_{\sigma^*}) 1_{\{\sigma^* \leq \tau^*\}} + (Y^1_{\tau^*} - Y^0_{\tau^*}) 1_{\{\tau^* \leq \sigma^*\}} \right] \\
\geq E \left[ \int_0^{\sigma^* \wedge \tau} \psi(r)dr + \gamma_-(\sigma^*) 1_{\{\sigma^* \leq \tau\}} 1_{\{\sigma^* < T\}} - \gamma_+(\tau) 1_{\{\tau^* \leq \sigma^*\}} \right] \\
+ E \left[ D(\sigma^*, \tau^*) \right]. \\
(6.39)
\]

**Case (B)**  Now on to the second inequality in the claim (6.36). The submartingale property of \( \hat{Y}^1 - Y^0 \) on \([0, \sigma \wedge \tau^*]\) and the definition of \( \hat{Y}^1 \) leads to:

\[
Y^1_0 - Y^0_0 \leq E \left[ \int_0^{\sigma \wedge \tau^*} \psi(r)dr + Y^1_{\sigma \wedge \tau^*} - Y^0_{\sigma \wedge \tau^*} \right] \\
= E \left[ \int_0^{\sigma \wedge \tau^*} \psi(r)dr \right] + E \left[ (Y^1_{\sigma} - Y^0_{\sigma}) 1_{\{\sigma \leq \tau^*\}} + (Y^1_{\tau^*} - Y^0_{\tau^*}) 1_{\{\tau^* \leq \sigma\}} \right]. \\
(6.40)
\]

Then, using the same arguments as in the proof of Proposition 6.5.3 leading up to equation (6.32), the following can be shown:

\[
E \left[ (Y^1_{\tau^*} - Y^0_{\tau^*}) 1_{\{\tau^* \leq \sigma\}} \right] = E \left[ (-\gamma_+(\tau^*)) 1_{\{\tau^* \leq \sigma\}} \right]. \\
(6.41)
\]
By the arguments establishing equation (6.34) in Proposition 6.5.3 and using Lemma 6.5.2 to assert that $Y_1^\sigma - Y_0^\sigma \leq \gamma_-(\sigma)$ almost surely, the following is derived:

$$
E \left[ (Y_1^\sigma - Y_0^\sigma) 1_{\{\sigma \leq \tau^*\}} \right] = E \left[ (Y_1^\sigma - Y_0^\sigma) 1_{\{\sigma < T\}} + 1_{\{\sigma = T\}} \right] 
\leq E \left[ \gamma_-(\sigma) 1_{\{\sigma \leq \tau^*\}} 1_{\{\sigma < T\}} + \Gamma 1_{\{\sigma = \tau^* = T\}} \right].
$$

(6.42)

Using equations (6.41) and (6.42) in equation (6.40) above concludes the proof:

$$
Y_1^\sigma - Y_0^\sigma \leq E \left[ \int_0^{\sigma \wedge \tau^*} \psi(r)dr \right] + E \left[ (Y_1^\sigma - Y_0^\sigma) 1_{\{\sigma \leq \tau^*\}} + (Y_1^\tau - Y_0^\tau) 1_{\{\tau^* < \sigma\}} \right] 
\leq E \left[ \int_0^{\sigma \wedge \tau^*} \psi(r)dr + \gamma_-(\sigma) 1_{\{\sigma \leq \tau^*\}} 1_{\{\sigma < T\}} - \gamma_+(\tau^*) 1_{\{\tau^* < \sigma\}} + \Gamma 1_{\{\sigma = \tau^* = T\}} \right] 
= E[D(\sigma, \tau^*)].
$$

\[\square\]

**Remark 6.5.2.** The results of Theorem 6.5.4 were obtained in a similar fashion to existing studies of Dynkin games which use probabilistic approaches. In particular, [100] (particularly Theorem 1) which uses martingale methods for Dynkin games; [110] (particularly Theorem 2.1) which has a semi-harmonic characterisation of the value function for the Dynkin game in a Markovian setting; and [39, 67] which use the concept of doubly reflected backward stochastic differential equations.

**Remark 6.5.3.** Although this chapter began with a Dynkin game and subsequently formulated an optimal switching problem which gave its solution, one could have arrived at a similar conclusion by doing the reverse. More precisely, take any two-mode optimal switching problem (satisfying the assumptions outlined earlier) with the given profit rate processes $\psi_1(\cdot) \equiv \Psi(\cdot, 1)$ and $\psi_0(\cdot) \equiv \Psi(\cdot, 0)$, switching costs $\gamma(\cdot, 1)$ and $\gamma(\cdot, 0)$, and terminal reward data $\Gamma_1 \equiv \Phi(1)$ and $\Gamma_0 \equiv \Phi(0)$, then construct the payoff (6.4) for the corresponding Dynkin game by setting

$$
\Gamma := \Gamma_1 - \Gamma_0, \quad \psi := \psi_1 - \psi_0, \quad \gamma_-(\cdot) := \gamma(\cdot, 1), \quad \gamma_+(\cdot) := \gamma(\cdot, 0).
$$

The above generalises the form of the parameters assumed in Section 6.5.2, and still ensures that Theorem 6.5.4 and the other results in this section remain valid.
6.6 Numerical example: A partially reversible investment problem

An important aspect of economic theory is the study of how firms or similar agents can investment optimally in projects within an uncertain economic environment. For example, a firm’s profit may depend on the strategy it uses to invest in the production capacity of a particular good that it sells on the market. This problem is made more difficult if one also allows the firm to recover part of its investment through some form of disinvestment. The problem then becomes one of (partially) reversible investment, and this has been studied in a variety of contexts including optimal switching [90, 93] and singular stochastic control [31, 65].

The connection between Dynkin games, optimal stopping and singular stochastic control problems is already well known as [31, 65] and their references show. This connection enabled the authors in [31] to solve a partially reversible investment problem on a finite time horizon. Their approach included a thorough analysis of a zero-sum Dynkin game with the following payoff:

\[
D (\sigma, \tau; t, x_0) = \int_0^{\sigma \wedge \tau} e^{-\bar{\mu} s} F \left( Z_{x_0}^s \right) ds + \frac{\hat{c}}{\eta} e^{-\bar{\mu} \sigma} 1_{\{\sigma \leq \tau\}} 1_{\{\sigma < T - t\}} + \frac{\check{c}}{\eta} e^{-\bar{\mu} \tau} 1_{\{\tau < \sigma\}} + \frac{\hat{c}}{\eta} e^{-\bar{\mu} (T - t)} 1_{\{\sigma = \tau = T - t\}}
\]

(6.43)

where \( T > 0, t \in [0, T] \) is the fixed initial time and \( x_0 > 0 \) is the initial capacity level. Strategies \( \sigma \in \mathcal{T} \) and \( \tau \in \mathcal{T} \) are admissible in this case if \( \sigma, \tau \leq T - t \).

In equation (6.43), \( Z^z = (Z^z_t)_{0 \leq t \leq T} \) is a geometric Brownian motion with explicit representation:

\[
\begin{align*}
Z_0^z &= z \in \mathbb{R}_+; \\
Z_t^z &= z \cdot \exp \left( \hat{\mu} t + \pi B_t \right), \quad t \in [0, T]
\end{align*}
\]

(6.44)

where \( \hat{\mu} \) and \( \pi \) are positive constants and \( B = (B_t)_{0 \leq t \leq T} \) is a standard Brownian motion with respect a filtered probability space \((\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})\). The parameters \( \hat{c} \) and \( \check{c} \), which represent the constant cost/benefit of investment and disinvestment respectively, satisfy the relation \( \hat{c} > \check{c} > 0 \). The constant \( \eta > 0 \) is a factor for converting units of investment into production capacity, and \( \bar{\mu} > 0 \) is an adjusted discount rate. The mapping \( Z \mapsto F(Z) \), which represents the marginal rate of production, is continuously
differentiable on \((0, \infty)\), strictly decreasing and satisfies the Inada conditions:

\[
\lim_{Z \to 0} F(Z) = \infty, \quad \lim_{Z \to \infty} F(Z) = 0. \tag{6.45}
\]

### 6.6.1 An exact solution for the Dynkin game.

It was verified in [31, p. 4100] that the game with payoff (6.43) has a solution \(v(t, x_0)\) given by the following semi-explicit representation:

\[
\begin{align*}
v(t, x_0) &= e^{-\hat{\mu}(T-t)} \frac{\hat{c}}{\eta} + \int_0^{T-t} e^{-\hat{\mu} s} \mathbb{E} \left[ F(Z_s^{x_0}) 1_{\{\hat{x}(t+s) < Z_s^{x_0} < \hat{x}(t+s)\}} \right] ds \\
&\quad + \frac{\hat{\mu}}{\eta} \int_0^{T-t} e^{-\hat{\mu} s} \left[ \hat{c} \mathbb{P}(Z_s^{x_0} < \hat{x}(t+s)) + \hat{c} \mathbb{P}(Z_s^{x_0} > \hat{x}(t+s)) \right] ds \tag{6.46}
\end{align*}
\]

The functions \(\hat{x}(\cdot)\) and \(\hat{x}(\cdot)\) appearing in equation (6.46) are continuous, decreasing and solve coupled integral equations of a similar form to (6.46):

\[
\begin{align*}
\frac{\hat{c}}{\eta} &= e^{-\hat{\mu}(T-t)} \frac{\hat{c}}{\eta} + \int_0^{T-t} e^{-\hat{\mu} s} \mathbb{E} \left[ F(Z_s^{\hat{x}(t)}) 1_{\{\hat{x}(t+s) < Z_s^{\hat{x}(t)} < \hat{x}(t+s)\}} \right] ds \\
&\quad + \frac{\hat{\mu}}{\eta} \int_0^{T-t} e^{-\hat{\mu} s} \left[ \hat{c} \mathbb{P}(Z_s^{\hat{x}(t)} < \hat{x}(t+s)) + \hat{c} \mathbb{P}(Z_s^{\hat{x}(t)} > \hat{x}(t+s)) \right] ds, \tag{6.47}
\end{align*}
\]

\[
\begin{align*}
\hat{x}(T) &= 0
\end{align*}
\]

\[
\begin{align*}
\frac{\hat{c}}{\eta} &= e^{-\hat{\mu}(T-t)} \frac{\hat{c}}{\eta} + \int_0^{T-t} e^{-\hat{\mu} s} \mathbb{E} \left[ F(Z_s^{\hat{x}(t)}) 1_{\{\hat{x}(t+s) < Z_s^{\hat{x}(t)} < \hat{x}(t+s)\}} \right] ds \\
&\quad + \frac{\hat{\mu}}{\eta} \int_0^{T-t} e^{-\hat{\mu} s} \left[ \hat{c} \mathbb{P}(Z_s^{\hat{x}(t)} < \hat{x}(t+s)) + \hat{c} \mathbb{P}(Z_s^{\hat{x}(t)} > \hat{x}(t+s)) \right] ds, \tag{6.48}
\end{align*}
\]

where \(F^{-1}\) is the inverse of \(F\).

### 6.6.2 Derivation of the auxiliary optimal switching problem.

The next few lines verify that the value of the Dynkin game with payoff (6.43) can be found using optimal switching. For a fixed \(t \in [0, T]\), let \(\tilde{T} = T - t\) be the terminal time for the Dynkin game with parameters:

\[
\psi(s) = F(Z_s^{x_0}), \quad \gamma_-(s) = \frac{\hat{c}}{\eta}, \quad \gamma_+(s) = -\frac{\hat{c}}{\eta}, \quad \Gamma = \frac{\hat{c}}{\eta}, \tag{6.49}
\]
where \( s \in [0, \tilde{T}] \). The payoff in equation (6.43) has the same form as the Dynkin game with payoff (6.4) determined by the parameters in (6.49), and the new time horizon \([0, \tilde{T}]\). Since \( \hat{c} > \bar{c} \) and \( \eta > 0 \), the switching costs satisfy

\[
\gamma_-(s) + \gamma_+(s) = \frac{\xi}{\eta} - \frac{\xi}{\eta} > 0
\]

for all \( s \in [0, \tilde{T}] \). Similarly one verifies that equation (6.9) holds. Note that the geometric Brownian motion \( Z^z = (Z^z_t)_{0 \leq t \leq T} \) almost surely never hits 0 in finite time when starting from \( z > 0 \), and is well known to be in \( \mathcal{S}^2 \) (by the Burkholder-Davis-Gundy inequality or moment estimates for solutions to SDEs, for instance). A suitable choice of the mapping \( Z \to F(Z) \) therefore ensures that either Assumption 7 or 9 holds.

The value of the game with payoff (6.43) can therefore be obtained using optimal switching. In order to account for the discount rate \( \bar{\mu} \), the pair of processes \((Y^0, Y^1)\) in Hypothesis (H1) are discounted so that, \( P - a.s., \)

\[
\left\{ \begin{array}{l}
e^{-\bar{\mu}r}Y^i_r = \operatorname{ess sup}_{\theta \in \tilde{\mathcal{T}}_r} E \left[ \int_r^\theta e^{-\bar{\mu}s} \psi_i(s) ds + e^{-\bar{\mu}\tilde{T}} \Gamma_i 1_{\{\theta = \tilde{T}\}} + e^{-\bar{\mu}r} \{Y^1_{\theta} - \gamma(\theta, 1 - i)\} 1_{\{\theta < T\}} \bigg| \mathcal{F}_r \right] \\
Y^i_{\tilde{T}} = \Gamma_i
\end{array} \right.
\]

where \( r \in [0, \tilde{T}] \) and \( \tilde{\mathcal{T}}_r \) is the set of stopping times \( \theta \) satisfying \( r \leq \theta \leq \tilde{T} \).

6.6.3 Numerical results

Free-boundary method.

For initial values \( x_0 > 0 \) and \( t \in [0, T] \), the values \( v(t, x_0) \), \( \hat{x}(t) \) and \( \bar{x}(t) \) were computed simultaneously using the following steps:

1. use a basic quadrature routine such as the rectangle rule to approximate the integral terms in (6.47) and (6.48) on a (uniform) partition of \([t, T]\);

2. starting with the boundary conditions in (6.47) and (6.48), use a root-finding algorithm to solve for \( \hat{x}(\cdot) \) and \( \bar{x}(\cdot) \) backwards in time on the partition of \([t, T]\);

3. use the values obtained recursively for \( \hat{x}(\cdot) \) and \( \bar{x}(\cdot) \) to calculate \( v(t, x_0) \) according to equation (6.46).
Least squares Monte Carlo regression.

The least squares Monte Carlo regression (LSMC) method was used to approximate the value of the Dynkin game via a numerical solution of its auxiliary optimal switching problem. More details on the algorithm can be found in Appendix B. The numerical experiment used 10,000 simulated sample paths of the geometric Brownian motion (6.44), with half of these paths obtained via antithetic sampling. Simple monomials were used as the basis functions for approximating the conditional expectations.

Comparison of results and discussion.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$T$</td>
<td>1.0</td>
<td>Terminal time</td>
</tr>
<tr>
<td>$x_0$</td>
<td>2.5</td>
<td>Initial value of the state variable</td>
</tr>
<tr>
<td>$\hat{\mu}, \pi$</td>
<td>0.2, 1.0</td>
<td>Parameters for stochastic process</td>
</tr>
<tr>
<td>$F(\cdot)$</td>
<td>$F(x) = \frac{1}{\sqrt{x}}$</td>
<td>Instantaneous payoff</td>
</tr>
<tr>
<td>$\hat{c}, \hat{c}, \eta, \bar{\mu}$</td>
<td>1.0, 0.8, 1.0, 0.8</td>
<td>Parameters in the game’s payoff</td>
</tr>
<tr>
<td>$N$</td>
<td>200</td>
<td>Number of time intervals</td>
</tr>
<tr>
<td>$N_S$</td>
<td>10,000</td>
<td>Number of sample paths</td>
</tr>
</tbody>
</table>

Table 6.1: Table of parameter values used in numerical solution of the Dynkin game.

The parameters used for the LSMC approximation are summarised in Table 6.1 above, whilst Figure 6.1 below shows the Dynkin game’s solution obtained via the free boundary approach and the LSMC approximation with different degrees of the basis polynomials.

![Figure 6.1: Comparison of free-boundary and Least-Squares Monte Carlo solutions to the Dynkin game.](image)

Figure 6.1: Comparison of free-boundary and Least-Squares Monte Carlo solutions to the Dynkin game.

The plot confirms that the solution of the auxiliary optimal switching indeed leads
to the solution of the Dynkin game. For an individual run (fixed polynomial basis
degree), the LSMC method took about 30 minutes to compute the game value for all
of the 200 (discrete) time points between 0 and $T$. On the other hand, solving the
non-linear integral equations for the game value at $t = 0$ took over 4 hours on the
same computer. Considering the accuracy of the results obtained, there was a major
benefit to using LSMC in terms of computational performance.

The vast difference in running times might be explained by certain features of the
problem. In particular, the non-linear integral equations arising in the free boundary
method can be difficult to solve numerically since the integrands become unbounded
near $t = T$ (recall the Inada conditions (6.45), $\hat{x}(T) = 0$ and $\hat{x}(\cdot)$ is continuous).
Nevertheless, many computational software packages can cope with such singularities
in numerical integration, albeit at additional computational cost.

6.7 Pricing cancellable contracts for difference

Section 6.6 verified that the standard LSMC algorithm for solving optimal switching
problems gives accurate approximations to the Dynkin game’s value. One of the
advantages of this numerical method is its suitability for high-dimensional problems,
and [2] gives an example of an optimal investment problem in electricity generation
with eight underlying factors. This section provides a sufficiently general multi-factor
model for the CCfD for which the LSMC algorithm may be appropriate. It formalises
the previous model given in Section 6.2 above.

6.7.1 Probabilistic model and assumptions

Let $0 < T < \infty$ be given and $X^x = (X^x(t))_{0 \leq t \leq T}$ be an $\mathbb{R}^n$-valued càdlàg strong
Markov process that is defined on a stochastic basis $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{0 \leq t \leq T}, P_x)$ satisfying
the usual conditions. The probability measure $P_x$ is such that $P_x(\{X^x(0) = x\}) = 1$
for a given initial state $x \in \mathbb{R}^n$. It is further assumed that $X$ is quasi-left-continuous
and the filtration $\mathbb{F} = (\mathcal{F}_t)_{0 \leq t \leq T}$ is the natural completed filtration of $X^x$.

The process $X^x$ represents the dynamics of $n \geq 1$ economic / environmental factors
that influence the electricity market. Suppose that four functions are given, $f_1: [0, T] \times
\mathbb{R}^n \to \mathbb{R}$, and $f_i: [0, T] \times \mathbb{R}^n \to \mathbb{R}_+$ for $i = 2, 3, 4$. For $i \in \{1, 2, 3, 4\}$, the function
$f_i(\cdot, \cdot)$ is assumed to be continuous on $[0, T] \times \mathbb{R}^n$ and satisfies $(f_i(t, X^x(t)))_{0 \leq t \leq T} \in \mathcal{S}^2$. These functions are assumed to define the electricity market variables as follows: for $t \in [0, T]$,

1. the market price of electricity is given by $S(t) = f_1(t, X^x(t))$;

2. the supplier’s demand is defined by $D(t) = f_2(t, X^x(t))$;

3. the trigger level for the CCfD is defined by $d^*(t) = f_3(t, X^x(t))$;

4. the generator’s supply is given by $R(t) = f_4(t, X^x(t))$.

The stopping costs $\gamma_{\pm}$ are also assumed to satisfy $\gamma_{\pm}(\cdot) = \tilde{\gamma}_{\pm}(\cdot, X^x(\cdot))$, where the maps $\tilde{\gamma}_{\pm}: [0, T] \times \mathbb{R}^n \to \mathbb{R}_+$ are continuous and such that the processes $\tilde{\gamma}_{\pm}(\cdot, X^x(\cdot))$ are in $\mathcal{S}^2$. To avoid excessive notation, $\gamma_{\pm}(\cdot, X^x(\cdot))$ is used instead of $\tilde{\gamma}_{\pm}(\cdot, X^x(\cdot))$ below.

**Remark 6.7.1.** Suitable models for the market variables which satisfy the above assumptions can come from the reduced-form and structural approaches to modelling electricity markets for instance – see the survey [29] for more details and examples.

### 6.7.2 The fair value for the CCfD

Continuing from Section 6.2 above, with additional emphasis on the dependence on $X^x$, the payoff of the CCfD in terms of the generator is given by (cf. (6.3)):

$$D(\sigma, \tau; x) = \int_{0}^{\sigma \land \tau} \psi(t, X^x_t) dt + \gamma_-(\sigma, X^x(\sigma)) 1_{\{\sigma \leq \tau < T\}} - \gamma_+(\tau, X^x(\tau)) 1_{\{\tau < \sigma\}}. \quad (6.50)$$

Introduce the following upper and lower values of the Dynkin game

$$W^+(x) = \inf_{\sigma \in T} \sup_{\tau \in T} \mathbb{E}_x[D(\sigma, \tau; x)] \quad \text{and} \quad W^-(x) = \sup_{\tau \in T} \inf_{\sigma \in T} \mathbb{E}_x[D(\sigma, \tau; x)].$$

These values are well-defined based on the earlier assumptions, and satisfy

$$W^-(x) \leq W^+(x), \quad \forall x \in \mathbb{R}^n.$$

If $W^-(x) = W^+(x), \forall x \in \mathbb{R}^n$, the game is said to be “fair” and the common value $V = W^- = W^+$ is its solution [41, p. 684]. The value $V(x)$ represents the fair price of the generator’s premium when it accepts the contract’s terms and the initial state is $x \in \mathbb{R}^n$. 
The game’s solution.

Under Hypothesis (H1) (which holds since $\gamma_- > \gamma_+ \geq 0$), for every $x \in \mathbb{R}^n$, there exist two adapted processes, $Y^{0,x} = (Y^{0,x}(t))_{0 \leq t \leq T}$ and $Y^{1,x} = (Y^{1,x}(t))_{0 \leq t \leq T}$, which are defined by:

$$Y^{1,x}(t) = \text{ess sup}_{\theta \in \mathcal{T}} E_x \left[ \int_{t}^{\theta} \psi(s, X^x_s) ds + \{Y^{0,x}(\theta) - \gamma_+(\theta, X^x_{\theta})\} \mathbf{1}_{\{\theta < T\}} \right| \mathcal{F}_t],$$

$$Y^{0,x}(t) = \text{ess sup}_{\theta \in \mathcal{T}} E_x \left[ \{Y^{1,x}(\theta) - \gamma_-(\theta, X^x_{\theta})\} \mathbf{1}_{\{\theta < T\}} \right| \mathcal{F}_t].$$

These processes have paths that are right-continuous and quasi-left-continuous, are in $\mathcal{S}^2$, and also satisfy

$$Y^{1,x}(0) - Y^{0,x}(0) = W^-(x) = W^+(x), \quad x \in \mathbb{R}^n.$$

The value of the game can therefore be computed for each initial value of the state using suitable methods for (numerically) solving high dimensional optimal switching problems such as the least squares Monte Carlo algorithm.

### 6.8 Conclusion

This chapter introduced a mathematical model for the cancellable contract for difference (CCfD), which is a variant of the Contract for Difference that allows both parties to exit the arrangement before its maturity. The payoff of the contract is a special case of the classic zero-sum Dynkin game, which was subsequently analysed using the theory of continuous-time optimal switching problems studied in Chapter 5. This chapter gave a formulation of a two-mode optimal switching problem that not only solves the Dynkin game, but also provides a saddle point strategy. Utilising an example of a finite time horizon Dynkin game with an analytic solution, a numerical experiment provided ample evidence that the optimal switching formulation works. In this case, the least squares Monte Carlo method, which is a popular simulation-based numerical approach for optimal switching problems, gave quick and accurate approximations when compared to the analytic method.

One of the advantages of the Least-Squares Monte Carlo method is its ability to handle complicated high-dimensional problems for which analytic solutions may be out of reach. It is therefore a suitable candidate for valuing a complex instrument.
as the CCfD. To this end, an abstract Markovian model for the variables underlying the CCfD was put forward, and the formulation of its solution via optimal switching was then given. The results of this chapter give the necessary confidence that this approach will work.
Chapter 7

A stochastic game of load balancing in a power system

7.1 Introduction

This chapter develops a mathematical model for a balancing services contract between an electricity transmission system operator (SO) and the owner of an electric energy storage device (battery operator or BO). The SO chooses times (up to a finite number) and amounts of electricity (up to a maximum amount) to instantaneously affect the imbalance between demand and supply on the power system. This description is similar to that of a swing option, which is a popular contract in energy markets that grants the holder multiple conditional rights to alter the amount of energy it receives [77]. However, an important difference in the balancing services contract is that the energy is supplied by the BO’s battery, therefore putting a dynamic constraint on the actual amount of energy delivered to the SO. The BO is assumed to use the battery only for the balancing services contract, and is allowed to charge it between request times with electricity purchased at the spot market rate. Battery charging is assumed to have negligible effect on the system imbalance.

The problem for the SO is to choose appropriate delivery times and energy amounts that minimise the total electricity imbalance and cost of the balancing services over a prescribed time interval. Simultaneously, the battery operator would like to minimise the cost of its energy purchases and any additional penalties for partial delivery at the request times. These optimisation problems are combined into a non-zero sum
stochastic differential game where one player (BO) uses a classic (continuous) control (like Chapter 3) and the other player (SO) uses an impulse control. A Nash equilibrium solution to this game gives a “fair” value for the balancing services contract.

The following section further motivates the mathematical problem in the context of local electricity markets and proposes a programme for its solution. Section 7.3 then provides some background material on Markov diffusion processes that are controlled continuously through their drift and diffusion coefficients, and discretely via impulses. A more precise mathematical formulation of the stochastic game is presented in Section 7.4. Section 7.5 then presents a verification theorem which serves as a basis for the game’s solution. Section 7.6 studies the verification theorems in the context of the balancing services contract and discusses the findings. A conclusion and closing remarks for the thesis are given in Section 7.7.

7.2 The load balancing game

The balancing services contract is investigated on the time horizon \([0, T]\) where \(0 < T < \infty\). For expository purposes, introduce a probability space \((\Omega, \mathcal{F}, P)\) equipped with a filtration \(\mathbb{F} = (\mathcal{F}_t)_{0 \leq t \leq T}\). More details on this space are given below in Section 7.3. Let \(s \in [0, T)\) be an initial time.

7.2.1 A model for the local electricity market

Electricity storage variable.

The BO’s battery is assumed to have a finite capacity \(B_{max} \in (0, \infty)\). Let \(u_{max} \in (0, \infty)\) and define a set \(U = [0, u_{max}]\). The battery’s effective charging rate (including losses) is described by a given deterministic function \(b^{X_1} : [0, B_{max}] \times U \rightarrow \mathbb{R}_+\) satisfying the following:

- the mapping \((x_1, v) \mapsto b^{X_1}(x_1, v)\) is continuous on \([0, B_{max}] \times U\);
- furthermore, \(b^{X_1}(B_{max}, \cdot) = 0\) and for each \(x_1 \in [0, B_{max}]\), \(b^{X_1}(x_1, 0) = 0\) and \(v \mapsto b^{X_1}(x_1, v)\) is non-decreasing.

The battery is charged over time according to a \(U\)-valued control variable \((u(t))_{t \geq s}\). This control variable may be random, \(u(t) = u(\omega, t)\) for \(\omega \in \Omega\), in which case \(u(\cdot)\) is
assumed to be \( \mathbb{F} \)-progressively measurable. The evolution of the battery level \( X_1 \) over \([s, T]\) without interventions is given by the following equation: \( \mathbb{P} \)-a.s.,

\[
\begin{align*}
X_1(s) &= x_1 \in [0, B_{\text{max}}]; \\
X_1(t) &= x_1 + \int_s^t b^{X_1}(X_1(\theta), u(\theta)) d\theta, \quad s \leq t \leq T.
\end{align*}
\]

An admissible charging control \( (u(t))_{t \geq s} \) ensures the existence of a solution \( X_1(\cdot) \) to (7.1) which is \( \mathbb{F} \)-adapted and has continuous paths almost surely \( \mathbb{P} \). This concept of admissibility is made more precise in Section 7.3 below.

**Electricity demand-supply imbalance.**

The demand-supply imbalance on the electricity grid without interventions is described by an adapted stochastic process \( X_2(\cdot) \). Positive values of \( X_2 \) indicate over-supply on the grid whereas negative values mean there is an under-supply. The imbalance process is assumed to evolve according to the following equation: \( \mathbb{P} \)-a.s.,

\[
\begin{align*}
X_2(s) &= x_2 \in \mathbb{R}; \\
X_2(t) &= x_2 + \int_s^t b^{X_2}(\theta, X_2(\theta)) d\theta + \int_s^t \sigma^{X_2}(\theta, X_2(\theta)) dW^2(\theta), \quad s \leq t \leq T.
\end{align*}
\]

where \( W^2 \) is a standard Brownian motion adapted to \( \mathbb{F} \) and \( b^{X_2}: \mathbb{R}_+ \times \mathbb{R} \rightarrow \mathbb{R} \) and \( \sigma^{X_2}: \mathbb{R}_+ \times \mathbb{R} \rightarrow \mathbb{R} \) are given measurable functions satisfying: uniformly in \( t \in [0, T] \),

- \( x_2 \mapsto (b^{X_2}(t, x), \sigma^{X_2}(t, x_2)) \) is locally Lipschitz;
- \( x_2 \mapsto (b^{X_2}(t, x), \sigma^{X_2}(t, x_2)) \) has at most linear growth.

For example, with \( b^{X_2}(t, x_2') = -\alpha x_2' \) and \( \sigma^{X_2}(t, x_2') = \beta \), where \( \alpha \) and \( \beta \) are positive constants, \( X_2 \) becomes an Ornstein-Uhlenbeck process with explicit solution:

\[
X_2(t) = x_2 e^{-\alpha(t-s)} + \int_s^t \beta e^{\alpha(r-t)} dW^2(r), \quad s \leq t \leq T.
\]

**Market price of electricity.**

Let \( X_3 = (X_3(t))_{0 \leq t \leq T} \) denote the random price of electricity in the spot market. This process is assumed to satisfy the equation:

\[
\begin{align*}
X_3(s) &= x_3 \in \mathbb{R}; \\
X_3(t) &= x_3 + \int_s^t b^{X_3}(\theta, X_3(\theta)) d\theta + \int_s^t \sigma^{X_3}(\theta, X_3(\theta)) dW^3(\theta), \quad s \leq t \leq T.
\end{align*}
\]
where $W^3(\cdot)$ is a standard Brownian motion adapted to $\mathbb{F}$ and $b^X_3: \mathbb{R}_+ \times \mathbb{R} \to \mathbb{R}$ and $\sigma^X_3: \mathbb{R}_+ \times \mathbb{R} \to \mathbb{R}$ are measurable functions satisfying the same Lipschitz and growth assumptions as $b^X_2$ and $\sigma^X_2$. The Brownian motions $W^2$ and $W^3$ are assumed to be independent for simplicity. However, correlation may be introduced by an appropriate modification of the diffusion coefficient $\sigma$ in (7.4) below.

The multidimensional state variable.

Let $X_1$, $X_2$ and $X_3$ denote the battery level, system imbalance and spot price state variables respectively. Recall the constraint set $[0, B_{max}]$ for $X_1$. Set $X = [0, B_{max}] \times \mathbb{R} \times \mathbb{R}$. Define functions $b: [0, T] \times X \times U \to \mathbb{R}$ and $\sigma: [0, T] \times X \to \mathbb{R}^{3 \times 2}$ as follows: for $x = (x_1, x_2, x_3) \in X$, $t \in [0, T]$, $v \in U$,

$$b(t, x, v) = \begin{pmatrix} b^X_1(x_1, v) \\ b^X_2(t, x_2) \\ b^X_3(t, x_3) \end{pmatrix}, \quad \sigma(t, x) = \begin{pmatrix} 0 & 0 \\ \sigma^X_2(t, x_2) & 0 \\ 0 & \sigma^X_3(t, x_3) \end{pmatrix} \quad (7.4)$$

The multidimensional state variable of interest in this case is a 3-dimensional process $X$ which evolves according to the following dynamics:

$$\begin{cases} X(s) = x \in X; \\
X(t) = x + \int_s^t b(\theta, X(\theta), u(\theta)) d\theta + \int_s^t \sigma(\theta, X(\theta)) \cdot dW(\theta), \quad s \leq t \leq T 
\end{cases} \quad (7.5)$$

where $(W(t))_{s \leq t \leq T}$ is a 2-dimensional standard Brownian motion and $(u(t))_{s \leq t \leq T}$ is a progressively measurable $U$-valued process with respect to $(\Omega, \mathcal{F}, \mathbb{F}, P)$. Thus (7.5) is a stochastic differential equation with random coefficients. If $(\Omega, \mathcal{F}, \mathbb{F}, P)$ and the Brownian motion $(W(t))_{s \leq t \leq T}$ are given a priori, then Theorem 1.6.16 of [138] proves under certain Lipschitz and growth conditions that there exists a unique $\mathbb{F}$-adapted continuous solution $X$ to (7.5).

In some instances it is also possible to obtain a solution $X$ to (7.5) which is in feedback form: for $s \in [0, T)$ given and $s \leq t \leq T$,

$$\begin{cases} X(s) = x \in X; \\
X(t) = x + \int_s^t b(\theta, X(\theta), u(\theta, X(\theta))) d\theta + \int_s^t \sigma(\theta, X(\theta)) \cdot dW(\theta) 
\end{cases} \quad (7.6)$$

where $u: [0, T] \times X \to U$ is Borel-measurable, so that $u(t) = u(t, X(t))$, $s \leq t \leq T$, is the control variable appearing in (7.5). This is true when $x \mapsto (b^u(t, x), \sigma(t, x))$,
b^n(t, x) \equiv b(t, x, u(t, x)), satisfies particular Lipschitz and growth conditions uniformly in t. In a canonic setup where (Ω, ℱ, ℱ_t) is given (for instance, in Section 7.3 below) and under slightly weaker assumptions on x \mapsto (b^n(t, x), σ(t, x)), it is possible to obtain a family of probability measures \{P_{sz}: (s, x) \in [0, T] \times X\} on (Ω, ℱ) forming a continuous Markov process – see Section 5 of [73] and Chapter 12 of [125].

7.2.2 Balancing actions as system state interventions

This section describes informally the way in which the SO can influence the battery level, imbalance and spot price processes through balancing actions. For this purpose, introduce another measurable space (\tilde{Ω}, \tilde{ℱ}) which extends (Ω, ℱ) in a particular way (discussed in Section 7.3). Let \tilde{ℱ} = (\tilde{ℱ}_t)_{0 \leq t \leq T} denote a filtration on this extended space.

Impulse control: an informal description.

The SO requests impulses \{ζ_1, ζ_2, ...\} of electricity from the BO and uses them to intervene in the evolution of the imbalance process X_2. Each impulse has a value in the fixed range \mathcal{I} := [i_{\min}, i_{\max}] \subseteq [0, B_{\max}] since the BO supplies the energy from its store. Let \{τ_1, τ_2, ...\} denote a sequence of \tilde{ℱ}-stopping times, the corresponding impulse times, satisfying s \leq τ_1(\tilde{ω}) \leq ... \leq τ_n(\tilde{ω}) \leq ... \leq T for every \tilde{ω} \in \tilde{Ω}. Each impulse ζ_n is then taken to be an \tilde{ℱ}_{τ_n}-measurable \mathcal{I}-valued random variable. The sequence of impulse time / amount pairs π = (τ_n, ζ_n)_{n \geq 1} is called an impulse control or intervention strategy. It is assumed that an upper limit 0 \leq N < \infty has been imposed on the total number of impulses. In this case one formally sets τ_{N+1} = T and it is also commonplace and notationally convenient to define τ_0 \equiv s. A more formal description of these strategies is given in Definition 7.3.3 below.

State interventions: an informal description.

Let N \geq 1 be a positive integer and \{s_n\}_{n=0}^N \subset T be an arbitrary non-decreasing sequence of times satisfying s_0 = s. Suppose there are U-valued processes u^{(0)}, u^{(1)}, ..., u^{(N)} given such that u^{(n)} is a progressively measurable with respect to (Ω, ℱ, ℱ_t) from times s_n onwards (these processes may be obtained from feedback controls u^{(0)}, u^{(1)}, ..., u^{(N)}).
For $0 \leq n \leq N$, let $X^{(n)}(\cdot; s_n, x_n)$ denote a process with similar dynamics to $X$ in (7.5) but starting from position $x_n \in \mathcal{X}$ at time $s_n \in [s, T]$. Let $\mathbf{u} = (u^{(0)}, \ldots, u^{(n)})$ denote the vector of controls and $\pi = (\tau_n, \zeta_n)_{n=1}^{N}$ be an impulse control strategy. Associated with the pair $(\pi, \mathbf{u})$ is a probability measure $P_{sx}^{\pi,\mathbf{u}}$ on $(\tilde{\Omega}, \mathcal{F})$. This probability measure is described in more detail in Section 7.3.2 below.

Let $\Gamma_{X_1}: \mathcal{X} \times \mathcal{I} \rightarrow [0, B_{\max}]$, $\Gamma_{X_2}: \mathcal{X} \times \mathcal{I} \rightarrow \mathbb{R}$ and $\Gamma_{X_3}: \mathcal{X} \times \mathcal{I} \rightarrow \mathbb{R}$ be given deterministic functions determining the respective positions of the components $X_1, X_2$ and $X_3$ after an impulse. Let $x = (x_1, x_2, x_3) \in \mathcal{X}$ and $\zeta \in \mathcal{I}$. Using the notation $\Gamma(x, \zeta) = (\Gamma_{X_1}(x, \zeta), \Gamma_{X_2}(x, \zeta), \Gamma_{X_3}(x, \zeta))^T$ where $T$ is the transpose operator, the state variable $X^{(\pi)}$ under an impulse control $\pi = (\tau_n, \zeta_n)_{n \geq 1}$ evolves informally as follows: for $0 \leq n \leq N$, $P_{sx}^{\pi,\mathbf{u}}$ almost surely

$$
\begin{cases}
X^{(\pi)}(s^-) = X^{(n)}(s); \\
X^{(\pi)}(t) = X(t; \tau_n, X^{(\pi)}(\tau_n)), & t \in [\tau_n, \tau_{n+1}); \\
X^{(\pi)}(\tau_{n+1}) = \Gamma(X^{(\pi)}(\tau_n^-), \zeta_{n+1}) \text{ on } \{\tau_{n+1} < T\}.
\end{cases} \tag{7.7}
$$

Intuitively, on each of the $N+1$ copies of $(\tilde{\Omega}, \mathcal{F}, \mathcal{F})$ there is a version of the continuously controlled state $X$ which evolves without intervention. The process $X^{(\pi)}$ evolves according to the law of one of these versions until an impulse time $\tau_{n+1}$, $0 \leq n \leq N$. It then jumps to a new position $X^{(\pi)}(\tau_{n+1})$ on $\{\tau_{n+1} < T\}$ and thereafter evolves according to the law of another version of $X$ until the next impulse time. This procedure is repeated from $n = 0$ until $n = N$. Details on the exact probabilistic construction used for Markov processes under an impulse control are given in Section 7.3.2 below.

### 7.2.3 Contract cash flows and performance criteria

Based on the physical description of the balancing services contract, it is reasonable to assume that the intervention position functions $\Gamma_{X_1}$ and $\Gamma_{X_2}$, which are associated with the battery level and system imbalance respectively, are defined as follows: for $x = (x_1, x_2, x_3) \in \mathcal{X}$ and $\zeta \in \mathcal{I}$,

$$
\Gamma_{X_1}(x, \zeta) := x_1 - \min(x_1, \zeta), \quad \Gamma_{X_2}(x, \zeta) := x_2 + \min(x_1, \zeta) \tag{7.8}
$$
If interventions have negligible effect on the spot price then one may take $\Gamma_{X^3}(\cdot, \cdot) \equiv 0$. This is a reasonable assumption for a model where the system operator is replaced by a small electricity supplier.

Let $s \in [0, T)$ and $x \in \mathcal{X}$ be given initial conditions and $\mathcal{U}^{(N)}$ and $\Pi^{(N)}$ denote suitable classes of controls used by the BO and SO respectively. An arbitrary element in $\mathcal{U}^{(N)}$ is denoted by $\vec{u}$ whilst the impulse control $\pi = (\tau_n, \zeta_n)_{1 \leq n \leq N}$ represents an arbitrary element in $\Pi^{(N)}$. The process $\tilde{u} = (\tilde{u}_t)_{s \leq t \leq T}$ below is an $\tilde{F}$-adapted process associated with $\vec{u}$ and $\pi$ which describes how the controls in $\vec{u}$ are applied. Expectations with respect to $(\tilde{\Omega}, \tilde{F}; \mathbb{P}_{s \vec{u} \pi \mathbf{x}})$ are denoted by $E_{s \vec{u} \pi \mathbf{x}}$. Section 7.3.2 below explains these in greater detail.

**The battery operator’s perspective.**

The BO is paid for the energy that it delivers to the SO at each intervention time. If the battery has $x_1 \in [0, B_{\text{max}}]$ units of energy and the SO requests $\zeta \in \mathcal{I}$ units from it, then it is assumed that the BO delivers the amount $\min(\zeta, x_1)$ to the SO. The payment rate for this delivered energy is assumed to be a positive constant which is denoted by $h$. The intervention cost from the perspective of the BO is therefore,

$$\hat{G}(x, \zeta) := -h \min(\zeta, x_1), \quad h > 0; \quad \zeta \in \mathcal{I}; \quad x = (x_1, x_2, x_3) \in \mathcal{X}; \quad y \in \mathbb{R}. \quad (7.9)$$

However, the BO has to purchase electricity from the spot market in order to charge the battery to a desired level. The cost rate for charging the battery if the prevailing spot price is $x_3 \in \mathbb{R}$ is then given by $x_3 \cdot v, \quad x_3 \in \mathbb{R}; \quad v \in \mathcal{U}$. Since the BO is assumed to use the battery only for the purposes of the balancing services contract, the following assumption is reasonable.

**Assumption 10.** The BO has a running cost, $\hat{L}: [0, T] \times \mathcal{X} \times \mathcal{U} \to \mathbb{R}$, which is determined only by a function $\hat{f}$ of the charging cost: for $t \in [0, T], \quad x = (x_1, x_2, x_3) \in \mathcal{X}$ and $v \in \mathcal{U},$

$$\hat{L}(t, x, v) = \hat{f}(x_3 \cdot v) \quad (7.10)$$

where $\hat{f}: \mathbb{R} \to \mathbb{R}$ is a convex, increasing function with at most polynomial growth.

Note that this assumption covers the linear case $\hat{f}(z) = z$. If the BO uses the control $\vec{u} \in \mathcal{U}^{(N)}$, its total expected cost over the interval $[s, T]$ under the SO’s impulse
control $\pi \in \Pi_s^{(N)}$ is:

$$J(s, x; \pi, \bar{u}) = \mathbb{E}_{x, u}^{\pi}[\int_s^T \tilde{L}(t, X^{(\pi)}(t^-), \bar{u}(t)) dt + \sum_{n=1}^N \tilde{G}(X^{(\pi)}(\tau_n^-), \zeta_n) 1_{\tau_n < T}].$$ (7.11)

Notice that the impact of the impulses is only considered on the interval $[s, T)$. The objective for the BO is to minimise the cost (7.11) over all controls $\bar{u} \in U^{(N)}$ when the SO uses the control $\pi$:

**Problem (BO).** For each given $\pi \in \Pi_s^{(N)}$, find $\bar{u}^* \in U^{(N)}$ such that:

$$J(s, x; \pi, \bar{u}^*) = \inf_{\bar{u} \in U^{(N)}} J(s, x; \pi, \bar{u})$$

$$=: \hat{V}(s, x; \pi).$$ (7.12)

**Remark 7.2.1.** The domain of integration in equation (7.11) involving $X^{(\pi)}$ is interpreted as being open on the left, that is on $(s, T]$. This is the convention followed throughout this chapter.

**The system operator’s perspective.**

The SO monitors the imbalance on the power system by observing the process $X_2$. For the purpose of grid stability, it prefers the imbalance to be near 0 at all times, and uses the strategy $\pi$ to help accomplish this. Suppose that the SO assigns a positive cost to the imbalance regardless of its sign. If the SO views under-supply and over-supply conditions symmetrically, then it is reasonable to assume the following:

**Assumption 11.** The SO has a running cost, $\tilde{L}: [0, T] \times \mathcal{X} \times U \rightarrow \mathbb{R}$, which is determined only by a function $\tilde{f}$ of the system imbalance: for $t \in [0, T]$, $x = (x_1, x_2, x_3) \in \mathcal{X}$ and $v \in U$,

$$\tilde{L}(t, x, v) = \tilde{f}(x_2)$$

where $\tilde{f}: \mathbb{R} \rightarrow \mathbb{R}_+$ is a strictly convex, increasing function with at most polynomial growth.

An example of such a specification for $\tilde{f}$ is $\tilde{f}(y) := \kappa y^2$, $y \in \mathbb{R}$ where $\kappa > 0$ is a constant. Recall that the SO pays the battery operator each time it intervenes on the system. The impulse cost from the SO’s perspective is then given by

$$G(x, \zeta) := -\tilde{G}(x, \zeta)$$

$$= h \min(\zeta, x_1), \quad h > 0; \quad \zeta \in \mathcal{I}; \quad x = (x_1, x_2, x_3) \in \mathcal{X}.$$ (7.14)
The performance of the SO’s strategy \( \pi \in \Pi_s(N) \) over the interval \([s, T]\) when the BO uses the control \( \tilde{\mathbf{u}} \in U^{(N)} \) is:

\[
\hat{J}(s, x; \pi, \tilde{\mathbf{u}}) = \mathbb{E}_x \left[ \int_s^T \hat{L}(t, X^{(\pi)}(t), \tilde{\mathbf{u}}(t)) \, dt + \sum_{n=1}^N \hat{G}(X^{(\pi)}(\tau_n^-), \zeta_n) \mathbf{1}_{\{\tau_n < T\}} \right].
\] (7.15)

The SO’s objective is to minimise the cost (7.15) over all intervention strategies \( \pi \in \Pi_s(N) \) given the BO’s strategy \( \tilde{\mathbf{u}} \):

**Problem (SO).** For each given \( \tilde{\mathbf{u}} \in U^{(N)} \), find \( \pi^* \in \Pi_s(N) \) such that:

\[
\hat{J}(s, x; \pi^*, \tilde{\mathbf{u}}) = \inf_{\pi \in \Pi_s(N)} \hat{J}(s, x; \pi, \tilde{\mathbf{u}}) =: V(s, x; \tilde{\mathbf{u}}).
\] (7.16)

### 7.2.4 A non-zero sum game of impulse and classic control

The control problems for the BO and SO are coupled through the intervention strategy \( \pi \) and control \( \tilde{\mathbf{u}} \) (due to \( X^{(\pi)} \)). Due to the asymmetry in their objectives, these problems constitute a *non-zero sum game* when viewed simultaneously. One may then search for strategies \( \tilde{\mathbf{u}}^* \in U^{(N)} \) and \( \pi^* \in \Pi_s(N) \) which jointly minimise equations (7.11) and (7.16). This leads to the concept of a *Nash equilibrium*, which is a pair \((\tilde{\mathbf{u}}^*, \pi^*) \in U^{(N)} \times \Pi_s(N)\) such that for all \( \tilde{\mathbf{u}} \in U^{(N)} \) and \( \pi \in \Pi_s(N) \):

\[
\begin{align*}
\hat{J}(s, x; \pi^*, \tilde{\mathbf{u}}^*) &\leq \hat{J}(s, x; \pi^*, \tilde{\mathbf{u}}); \\
\hat{J}(s, x; \pi^*, \tilde{\mathbf{u}}^*) &\leq \hat{J}(s, x; \pi, \tilde{\mathbf{u}}^*).
\end{align*}
\]

A Nash equilibrium solution \((\tilde{\mathbf{u}}^*, \pi^*)\) is such that neither player can do better by deviating unilaterally from the equilibrium strategy. This concept is revisited below.

### 7.3 Controlled Markov processes

This chapter explores a PDE approach for verifying the existence of a Nash equilibrium and therefore relies on concepts related to controlled Markov diffusions. The definitions used below mainly follow the books [57, 58]. However, wherever appropriate they are supplemented by results from [73, 84, 138].

Throughout this section \([0, T]\) is a fixed finite time horizon. Let \( Q = [0, T] \times \mathbb{R}^d \) and \( \bar{Q} = [0, T] \times \mathbb{R}^d \) denote its closure. For a set \( A \), let \( \mathcal{B}(A) \) denote the Borel \( \sigma \)-algebra.
on $A$. The set of sample paths $\Omega$ is the canonic one for continuous trajectories from $[0, T]$ to $\mathbb{R}^d$, $\Omega = C([0, T]; \mathbb{R}^d)$. Define the canonical process $(\xi_t)_{0 \leq t \leq T}$ on $(\Omega, \mathcal{B}(\Omega))$ by $\xi_t(\omega) := \xi(\omega, t) := \omega(t)$ for every $\omega \in \Omega$. Define the $\sigma$-algebras $\mathcal{F}_t = \sigma(\{\xi_s : 0 \leq s \leq t\})$, $\mathcal{F} = \mathcal{F}_T$ and set $\mathcal{F} = (\mathcal{F}_t)_{0 \leq t \leq T}$.

7.3.1 Controlled Markov diffusions

Let $\mathbb{S}^d$ denote the set of $d \times d$ symmetric real-valued matrices and $U \subset \mathbb{R}^k$ be compact. Suppose two measurable functions $a : [0, T] \times \mathbb{R}^d \to \mathbb{S}^d$ and $b : [0, T] \times \mathbb{R}^d \times U \to \mathbb{R}^d$ are given. These functions are assumed to satisfy:

Assumption 12. The maps $x \mapsto a(t, x)$ and $(x, v) \mapsto b(t, x, v)$ are continuous uniformly in $t$. Moreover, there exist non-negative constants $C, \gamma$ with $0 \leq \gamma \leq 2$ such that for all $t \in [0, T]$, $x \in \mathbb{R}^d$ and $v \in U$:

$$|a(t, x)| \leq C(1 + |x|)^\gamma$$

$$|b(t, x, v)| \leq C(1 + |x|).$$

For each fixed $v \in U$, let $\mathcal{A}^v$ denote the backward operator for a controlled $d$-dimensional Markov diffusion which acts on all functions $\varphi \in C^{1,2}(\bar{Q})$ as follows:

$$\mathcal{A}^v[\varphi](t, x) := \frac{\partial}{\partial t}\varphi(t, x) + \left[ \frac{1}{2} \sum_{i,j=1}^d a_{ij}(t, x) \frac{\partial^2}{\partial x_i \partial x_j} + \sum_{i=1}^d b_i(t, x, v) \frac{\partial}{\partial x_i} \right] \varphi(t, x). \quad (7.17)$$

The matrices $(a_{ij}(t, x))$ are symmetric, non-negative definite and, in keeping with Section 7.2.1 above, they are assumed to be uncontrolled. Thus, there exists a constant $c \geq 0$ such that for all $(t, x) \in Q$ and $\lambda \in \mathbb{R}^d$

$$\sum_{i,j=1}^d a_{ij}(t, x) \lambda_i \lambda_j \geq c|\lambda|^2.$$

If $c > 0$ then $\{\mathcal{A}^v, \ v \in U\}$ is said to be a family of uniformly parabolic operators, otherwise it is said to be degenerate parabolic.

Feedback controlled Markov diffusions.

A feedback control $u$ is a Borel measurable map $u : [0, T] \times \mathbb{R}^d \to U$. Define the backward operator $\mathcal{A}^u$ for a feedback controlled $d$-dimensional Markov diffusion from
(7.17) as follows:

\[
\mathcal{A}^u[\varphi](t,x) = \mathcal{A}^{u(t,x)}[\varphi](t,x) = \mathcal{A}^v[\varphi](t,x), \quad \text{if } u(t,x) = v.
\]  

(7.18)

For a given feedback control \( u \), define the function \( b^u : [0,T] \times \mathbb{R}^d \rightarrow \mathbb{R}^d \) by \( b^u(t,x) \). The feedback controlled Markov diffusion associated with \( \mathcal{A}^u \) is assumed to satisfy the following state equation: for some probability measure \( P_{sx}^u \) on \((\Omega, \mathcal{F})\),

\[
\begin{cases}
X(s) = x \in \mathbb{R}^d; \\
X(t) = x + \int_s^t b^u(\theta,X(\theta))d\theta + \int_s^t \sigma(\theta,X(\theta)) \cdot dW(\theta), \quad s \leq t \leq T
\end{cases}
\]

(7.19)

\( P_{sx}^u \)-almost surely, where \( \sigma(t,x) \) is a \( d \times d' \)-dimensional real-valued matrix satisfying \( \sigma(t,x)\sigma(t,x)^T = a(t,x) \) and \((W(t))_{s \leq t \leq T}\) is a \( d' \)-dimensional standard Brownian motion with respect to \((\Omega, \mathcal{F}, P_{sx}^u)\). Note that neither \( P_{sx}^u \) nor \((W(t))_{s \leq t \leq T}\) are specified apriori, which means a solution \( X \) to (7.19) may only exist in a weak sense (see also [73]).

**Definition 7.3.1.** A Borel measurable map \( u : [0,T] \times \mathbb{R}^d \rightarrow U \) is said to be an admissible feedback control if for each \((s,x) \in \bar{Q}\):

1. there exists a probability measure \( P_{sx}^u \) on the canonic space \((\Omega, \mathcal{F})\) such that, with respect to \((\Omega, \mathcal{F}, \mathcal{F}, P_{sx}^u)\), \( X^u \equiv \xi \) is a continuous \( \mathbb{R}^d \)-valued Markov process starting from \( X^u(s) = x \) with generator (backward evolution operator) \( \mathcal{A}^u \);

2. the process \((X^u(t))_{s \leq t \leq T}\) has bounded moments (with respect to \( P_{sx}^u \)) of all orders, with bound possibility depending on \((s,x)\).

Let \( \mathcal{U} \) denote the class of all admissible feedback controls.

**Remark 7.3.1.** Suppose \( \sigma : [0,T] \times \mathbb{R}^d \rightarrow \mathbb{R}^{d \times d'} \) is given and such that \( \sigma(t,x)\sigma(t,x)^T = a(t,x) \) (which is the case in Section 7.2.1). If \( u : [0,T] \times \mathbb{R}^d \rightarrow U \) is such that, uniformly in \( t, x \mapsto (b^u, \sigma) \) is locally Lipschitz and has at most linear growth, then \( u \in \mathcal{U} \) [57, p. 156].

**Definition 7.3.2.** Let \( \mathcal{D} = C_p^{1,2}(\bar{Q}) \) be the space of functions \((t,x) \mapsto \varphi : \bar{Q} \rightarrow \mathbb{R}\) such that the first (respectively first and second) derivative of \( \varphi \) with respect to \( t \) (respectively \( x \)) is continuous, and \( \varphi \) has at most polynomial growth on \( \bar{Q} \).
Note that under Assumption 12, every $\varphi \in \mathcal{D}$ satisfies the following: for every $u \in \mathcal{U}$,

- the mapping $(t, x) \mapsto \mathcal{A}^u[\varphi](t, x)$ is well-defined;
- $E_{sx}^u[|\varphi(t, X^u(t))|] + E_{sx}^u \left[ \int_s^t |\mathcal{A}^u(r)[\varphi](r, X^u(r))|dr \right] < \infty, \ \forall s < t \leq T.$
- the process $M^u = (M^u(t))_{s \leq t \leq T}$ defined by
  \[ M^u(t) = \varphi(t, X^u(t)) - \varphi(s, x) - \int_s^t \mathcal{A}^u[\varphi](r, X^u(r))dr \quad (7.20) \]
  is a martingale on $[s, T]$ with respect to $(\Omega, \mathcal{F}, \mathbb{P}, \mathcal{P}_{sx}).$

Equation (7.20), which can be verified using Itô’s formula, implies the following form of Dynkin’s formula for controlled diffusions: for $\varphi \in \mathcal{D}$ and $u \in \mathcal{U},$

\[ E_{sx}^u[\varphi(t, X^u(t))] = \varphi(s, x) + E_{sx}^u \left[ \int_s^t \mathcal{A}^u[\varphi](r, X^u(r))dr \right], \quad s \leq t \leq T. \quad (7.21) \]

It is these properties and the Markovian nature of weak solutions to (7.19) under admissible feedback controls that are essential for the verification theorems in Section 7.4 below.

### 7.3.2 Controlled Markov diffusions with interventions

Let $\mathcal{I} \subset \mathbb{R}^k$ be a compact set and $\Gamma: \mathbb{R}^d \times \mathcal{I} \to \mathbb{R}^d$ be a given Borel-measurable map. For each $N \in \mathbb{N}$, define the product space $\Omega^{(N)} = \times_{n=1}^N \Omega$ and corresponding $\sigma$-algebras $\mathcal{F}^{(N)} = \times_{n=1}^N \mathcal{F}$ and $\mathcal{F}_t^{(N)} = \times_{n=1}^N \mathcal{F}_t$. Note that $\Omega^{(1)} = \Omega$, $\mathcal{F}^{(1)} = \mathcal{F}$ and $\mathcal{F}_t^{(1)} = \mathcal{F}_t$. Now let $N \in \mathbb{N}$, $s \in [0, T)$ and $x \in \mathbb{R}^d$ be fixed.

**Remark 7.3.2.** Measurable functions $\xi$ and $\xi'$ defined on measurable spaces $(\Omega, \mathcal{F})$ and $(\Omega', \mathcal{F}')$ can be extended to measurable functions $\tilde{\xi}$ and $\tilde{\xi}'$ defined on the product space $(\Omega \times \Omega', \mathcal{F} \times \mathcal{F}')$ by setting, for example, $\tilde{\xi}(\tilde{\omega}) = \xi(\omega)$ and $\tilde{\xi}'(\tilde{\omega}) = \xi'(\omega')$ for each $\tilde{\omega} = (\omega, \omega') \in \tilde{\Omega} := \Omega \times \Omega'$.

Let $\mathcal{I} \subset \mathbb{R}^k$ be a compact set and $\Gamma: \mathbb{R}^d \times \mathcal{I} \to \mathbb{R}^d$ be a given Borel-measurable map.
Definitions of control strategies for the players.

Definition 7.3.3. Let \( \pi = (\tau_n, \zeta_n)_{n=1}^N \) be a sequence of pairs of impulse times \( \tau_n : \Omega^{(N+1)} \to [0, T] \) and impulse amounts \( \zeta_n : \Omega^{(N+1)} \to \mathcal{I} \) satisfying

- \( \tau_n \) is an \( \left( \mathcal{F}^{(n)}_t \times (X^{(n)}_j = \emptyset, \Omega) \right) \)-stopping time;
- \( \zeta_n \) is \( \left( \mathcal{F}^{(n)}_t \times (X^{(n)}_j = \emptyset, \Omega) \right) \)-measurable;
- For every \( \tilde{\omega} \in \Omega^{(N+1)} \), \( s \leq \tau_1(\tilde{\omega}) \leq \ldots \leq \tau_n(\tilde{\omega}) \leq \ldots \leq \tau_N(\tilde{\omega}) \leq T \);
- For each \( \tilde{\omega} \in \Omega^{(N+1)} \) with \( \tilde{\omega} = (\omega_1, \ldots, \omega_{N+1}) \):

\[
\tau_n(\tilde{\omega}) = \tau_n(\omega_1, \ldots, \omega_n), \quad \zeta_n(\tilde{\omega}) = \zeta_n(\omega_1, \ldots, \omega_n);
\]

The double sequence \( \pi \) is called an impulse control (or intervention) strategy with at most \( N \) interventions starting from \( s \). Let \( \Pi_\pi^{(N)} \) denote the class of all such controls.

To every \( \pi \in \Pi_\pi^{(N)} \) the pairs \( (\tau_0, \zeta_0) \) and \( (\tau_{N+1}, \zeta_{N+1}) \) may be adjoined, where \( \tau_0 \equiv s \), \( \tau_{N+1} \equiv T \) and \( \zeta_0 \in \mathcal{I} \), \( \zeta_{N+1} \in \mathcal{I} \) are arbitrary. In the case \( N = 0 \), let \( \Pi_\pi^{(0)} \) denote the singleton set with member \( \pi_0 = ((\tau_0, \zeta_0), (\tau_1, \zeta_1)) \) consisting of the adjoined variables defined in the previous line.

Definition 7.3.4. For each \( \pi \in \Pi_\pi^{(N)} \), define an \( (\mathcal{F}^{(N+1)}_t) \)-adapted càdlàg process \( X(\pi) \) on \( (\Omega^{(N+1)}, \mathcal{F}^{(N+1)}) \) as follows: for \( n = 0, \ldots, N \) and each \( \tilde{\omega} \in \Omega^{(N+1)} \) with \( \tilde{\omega} = (\omega_1, \ldots, \omega_{N+1}) \),

\[
\begin{align*}
X^{(\pi)}(\tilde{\omega}, s^-) &= \xi^{(0)}(\omega_1, s); \\
X^{(\pi)}(\tilde{\omega}, t) &= \xi^{(n)}(\omega_{n+1}, t), \quad t \in [\tau_n, \tau_{n+1}); \quad \tau_0 = s, \quad \tau_{N+1} = T; \\
X^{(\pi)}(\tilde{\omega}, \tau_{n+1}(\tilde{\omega})) &= \Gamma(\xi^{(n)}(\omega_{n+1}, \tau_{n+1}(\tilde{\omega})), \zeta_{n+1}(\tilde{\omega})) \quad \text{on} \quad \{\tau_{n+1} < T\};
\end{align*}
\]

In equation (7.22), \( \xi^{(0)}, \ldots, \xi^{(N)} \) are the canonical processes on the \( N + 1 \) copies of the space \( \Omega \). The process \( X(\pi) \) is referred to as the state process for the non-zero sum game of classic and impulse control.

Definition 7.3.5. A strategy for the classic controller in the non-zero sum game of classic and impulse control (with at most \( N \) interventions) is an \( (N + 1) \)-tuple \( \bar{u} = (u^{(0)}, \ldots, u^{(N)}) \), where \( u^{(n)} \in \mathcal{U} \) for \( n = 0, \ldots, N \). Let \( \mathcal{U}^{(N)} \) denote this class of
strategies. For every \( \pi \in \Pi_s^{(N)} \), \( \pi = (\tau_n, \zeta_n)_{n=1}^N \), and \( \tilde{u} \in U^{(N)} \), define an \( (F_t^{(N+1)}) \)-progressively measurable process \( \tilde{u} : \Omega^{(N+1)} \times [s, T] \to U \) in the following way: for each \( \tilde{\omega} \in \Omega^{(N+1)} \), \( \tilde{\omega} = (\omega_1, \ldots, \omega_{N+1}) \),

\[
\tilde{u}(\tilde{\omega}, t) = u^{(n)}(t, \xi^{(n)}(\omega_{n+1}, t)), \quad t \in [\tau_n, \tau_{n+1}); \quad \tau_0 = s, \quad \tau_{N+1} = T. \tag{7.23}
\]

The process \( \tilde{u} \) is called the admissible classic control with interventions corresponding to \( \pi \in \Pi_s^{(N)} \) and \( \tilde{u} \in U^{(N)} \).

**Remark 7.3.3.** Intuitively, \( \tilde{u} \) behaves like a feedback control which, after \( n = 0, 1, \ldots, N \) impulses, depends on only the \( (n + 1) \)-st controlled Markov process associated with the tuple \( \tilde{u} \).

**Definition of the probability measure generated by control strategies.**

The above definitions for the intervention strategy and state process with interventions are closest to [124], where they were used for a time-homogeneous Markov process without classic, continuous controls. Chapter 8 of the book [108] defines a time-homogeneous state process which is influenced by impulse and classic controls, but without a detailed description of the probabilistic setup. There are, however, other related constructions available in the literature which are valid for the time-inhomogeneous case. One example is the article [116], which studied an inventory control problem that involved switching amongst finitely many Markov processes. Another example is the paper [141], which studied combined impulse and classic control in the framework of Markov decision processes. The properties of the probability measure generated by classic and impulse controls described below uses a combination of the constructions in [116, 124].

To deal with the time-inhomogeneous case, introduce the set \( \tilde{\Omega} := [0, T] \times \Omega \) with the corresponding canonical time-space process \( (\tilde{\xi}_t)_{t \geq 0}, \tilde{\xi}_t := (t, \xi_t) \). The \( \sigma \)-algebras \( \mathcal{F}_t \) and \( \mathcal{F} \) are defined analogously to \( \mathcal{F}_t \) and \( \mathcal{F} \) respectively, noting that \( \mathcal{F} = \mathcal{B}(\Omega) = \mathcal{B}([0, T]) \times \mathcal{F} \). Define \( \tilde{P}_{sx}^{(n)} := \varepsilon_s \times P_{sx}^{(n)} \), where \( \varepsilon_s \) is the unit mass at \( s \in [0, T] \). The product set \( \tilde{\Omega}^{(N+1)} \) with corresponding \( \sigma \)-algebras \( \tilde{\mathcal{F}}^{(N+1)} \) and \( \tilde{\mathcal{F}}_{t}^{(N+1)} \) are defined just as before. Recalling Remark 7.3.2, the intervention strategy \( \pi \) can be extended from \( (\Omega^{(N+1)}, F^{(N+1)}) \) to \( (\tilde{\Omega}^{(N+1)}, \tilde{F}^{(N+1)}) \) without having to change Definition 7.3.3. Let \( \tilde{\xi}_t^{(n)} = (t, \xi_t^{(n)}) \) be the associated canonical time-space processes. For \( n = 1, \ldots, N \) and
\( \tilde{\omega} = (\omega_1, \ldots, \omega_{N+1}) \in \Omega^{(N+1)} \), set \( [\tilde{\omega}]_n = (\omega_i, \ 1 \leq i \leq n) \) and

\[
\tilde{\Gamma}(\tau_n, [\tilde{\omega}]_n) := \left( \tau_n([\tilde{\omega}]_n), \Gamma(\xi^{(n-1)}(\omega_n, \tau_n([\tilde{\omega}]_n)), \zeta_n([\tilde{\omega}]_n)) \right).
\]

For \( y \in \mathbb{R}^d \), let \( \varphi_y(\cdot) \) denote the function \( \varphi_y(t) = (t, y) \) for \( t \geq 0 \), and let \( \delta \) denote the Dirac measure on \( (\tilde{\Omega}, \tilde{\mathcal{F}}) \).

**Definition 7.3.6.** Let \( \pi \in \Pi_s^{(N)} \), \( \vec{u} \in \mathcal{U}^{(N)} \) be given and set \( P^{(0)}_{\pi, \vec{u}} = \tilde{P}^{(0)}_{\pi, \vec{u}} \). The probability measure \( P^{(n)}_{\pi, \vec{u}} \) on \( (\tilde{\Omega}^{(N+1)}, \tilde{\mathcal{F}}^{(N+1)}) \) generated by \( \pi \) and \( \vec{u} \) is one such that its projections \( P^{(n)} \) on the spaces \( \tilde{\Omega}^{(n)} \), \( n = 1, \ldots, N \), satisfy:

\[
P^{(n)}_{\pi, \vec{u}} = \delta_{\varphi_y(\tau_{n-1})} \times \delta_{\varphi_y(\tau_n)} \quad \text{on} \quad \sigma(\tilde{\mathcal{F}}^{(n+1)}_{\tau_n}, \tilde{\mathcal{F}}^{(n)}_{\tau_n} \times \{\emptyset, \tilde{\Omega}\}) =: \tilde{G}_{\tau_n}
\]

\[
\forall t \geq 0, \quad P^{(n)}_{\pi, \vec{u}}(\tilde{\xi}^{(n)}_{\tau_n+t} \in B_1, \ldots, \tilde{\xi}^{(n)}_{\tau_n+t} \in B_{n+1} | \tilde{G}_{\tau_n})
\]

\[
= \delta_{\varphi_{y|(\tau_1)}(\tau_1)}(B_1) \times \ldots \times \delta_{\varphi_{y|(\tau_n)}(\tau_n)}(B_n) \times \tilde{P}^{(n)}_{\pi, \vec{u}}(\tilde{\xi}^{(n)}_{\tau_n+t} \in B_{n+1}),
\]

\[
(B_1, \ldots, B_{n+1}) \in \tilde{\mathcal{F}}^{(n+1)}, \quad \text{on} \quad \{\tau_n + t < \tau_{n+1}\}.
\]

where \( y \in \mathbb{R}^d \) is arbitrary.

For further details on the construction of \( P^{\pi, \vec{u}}_{\pi, \vec{u}} \) see [124, p. 290], [116, p. 158] and the references therein. Its existence and uniqueness are related to Tulcea’s extension theorem (also see [124, p. 280] and [141, p. 478]).

### 7.4 A non-zero sum game of classic and impulse control

This section presents a more formal mathematical treatment of the non-zero sum game of Section 7.2. For ease of presentation, it avoids the use of the space-time process and the related notation in the time-inhomogeneous case. Let \( (s, x) \in Q \) be a given initial condition and \( N \in \{0, 1, 2, \ldots\} \) be given. Set \( \tilde{\Omega} = \Omega^{(N+1)}, \tilde{\mathcal{F}} = \mathcal{F}^{(N+1)}, \tilde{\mathcal{F}}_t = \mathcal{F}^{(N+1)}_t \) and \( \tilde{\mathcal{F}} = (\tilde{\mathcal{F}}_t)_{0 \leq t \leq T} \).

#### 7.4.1 Performance criteria

Let \( \hat{L}: \tilde{Q} \times U \to \mathbb{R}, \hat{M}: \mathbb{R}^d \to \mathbb{R} \) and \( \hat{G}: \mathbb{R}^d \times \mathcal{I} \to \mathbb{R} \) be functions which denote the respective running, terminal and impulse costs to the classic controller. Analogous
CHAPTER 7. LOAD BALANCING GAME IN A POWER SYSTEM

192
costs $\hat{L}$, $\hat{M}$ and $\hat{G}$ are defined for the impulse controller. The assumptions on these
costs are given in Section 7.4.2 below.

Let $\Pi_s^{(N)}$ and $\mathcal{U}^{(N)}$ denote the respective classes of controls for the impulse and
classic controller according to Definitions 7.3.3 and 7.3.5. Let $\pi \in \Pi_s^{(N)}$, $\bar{u} \in \mathcal{U}^{(N)}$
and $P_{xx}^{\pi,\bar{u}}$ be the probability measure on $(\tilde{\Omega}, \tilde{\mathcal{F}})$ generated by the pair $(\pi, \bar{u})$ as per
Definition 7.3.6. The BO’s expected total cost with respect to
$P_{xx}^{\pi,\bar{u}}$ over $[s, T]$ when
the SO intervenes according to its strategy $\pi$ is:

$$\hat{J}(s, x; \pi, \bar{u}) = \mathbb{E}_{xx}^{\pi,\bar{u}} \left[ \int_s^T \hat{L}(t, X^{(\pi)}(t^-), \bar{u}(t)) dt + \hat{M}(X^{(\pi)}(T)) \right. \left. + \sum_{n=1}^N \hat{G}(X^{(\pi)}(\tau^-_n), \zeta_n) 1_{\{\tau^-_n < T\}} \right].$$

(7.26)

Simultaneously, the expected total cost to the SO is:

$$\check{J}(s, x; \pi, \bar{u}) = \mathbb{E}_{xx}^{\pi,\bar{u}} \left[ \int_s^T \check{L}(t, X^{(\pi)}(t^-), \bar{u}(t)) dt + \check{M}(X^{(\pi)}(T)) \right. \left. + \sum_{n=1}^N \check{G}(X^{(\pi)}(\tau^-_n), \zeta_n) 1_{\{\tau^-_n < T\}} \right].$$

(7.27)

The classic and impulse controllers both try to minimise their respective total costs,
and this leads to the following concept of a solution.

**Definition 7.4.1.** A pair $(\bar{u}^*, \pi^*) \in \mathcal{U}^{(N)} \times \Pi_s^{(N)}$ is called a *Nash equilibrium*
solution to the non-zero sum game with payoffs (7.26)–(7.27), if for all $\bar{u} \in \mathcal{U}^{(N)}$ and $\pi \in \Pi_s^{(N)}$:

$$\begin{cases}
\hat{J}(s, x; \pi^*, \bar{u}^*) \leq \hat{J}(s, x; \pi^*, \bar{u}); \\
\check{J}(s, x; \pi^*, \bar{u}^*) \leq \check{J}(s, x; \pi, \bar{u}^*).
\end{cases}$$

(7.28)

7.4.2 Assumptions

This section collects standing assumptions on the data for the stochastic game.

**Assumption 13.** The intervention position function, $\Gamma: \mathbb{R}^d \times \mathcal{I} \to \mathbb{R}^d$, is $C(\mathbb{R}^d \times \mathcal{I})$
and $x \mapsto \Gamma(x, i)$ has at most polynomial growth for each $i \in \mathcal{I}$.

**Assumption 14.** The cost functions for the classic controller satisfy the following:

- the running cost $\hat{L}: \bar{Q} \times U \to \mathbb{R}$ is $C(\bar{Q} \times U)$ and such that $x \mapsto \hat{L}(t, x, v)$ has
  at most polynomial growth for each $(t, v) \in [0, T] \times U$;
• the terminal cost $\hat{M} : \mathbb{R}^d \to \mathbb{R}$ is continuous and has at most polynomial growth for each $t \in [0, T]$;

• the impulse cost $\hat{G} : \mathbb{R}^d \times \mathcal{I} \to \mathbb{R}$ is $C(\mathbb{R}^d \times \mathcal{I})$ and such that $x \mapsto \hat{G}(x, i)$ has at most polynomial growth for each $i \in \mathcal{I}$.

The cost functions for the impulse controller satisfy similar assumptions as their counterparts above, with the exception that the impulse cost is non-negative, $\hat{G} : \mathbb{R}^d \times \mathcal{I} \to \mathbb{R}_+$. These assumptions along with the existence of finite moments for the controlled state variable (Definition 7.3.1) ensure that the performance criteria for both players are well defined and finite for any pair of admissible strategies.

### 7.5 Solution via the Hamilton-Jacobi-Bellman PDE

This section proposes a method for verifying a fortiori the existence of a Nash equilibrium solution (7.28).

#### 7.5.1 Solution with $N = 0$ interventions

Let $(s, x) \in Q$ be given and suppose $\mathcal{U}^{(0)} \equiv \mathcal{U} \neq \emptyset$. Let $\pi_0 \in \Pi_s^{(0)}$ be the strategy with 0 interventions according to Definition 7.3.3. The performance criteria for the players using classic and impulse controls are respectively,

$$
\hat{J}(s, x; \pi_0, \bar{u}) = \mathbb{E}^{\pi_0, \bar{u}}_{s,x} \left[ \int_s^T \hat{L}(t, X^{(\pi_0)}(t^-), \hat{u}(t)) dt + \hat{M}(X^{(\pi_0)}(T)) \right]
= \mathbb{E}^{\bar{u}}_{s,x} \left[ \int_s^T \hat{L}(t, X^{(\pi_0)}(t^-), \hat{u}(t)) dt + \hat{M}(X^{(\pi_0)}(T)) \right],
\tag{7.29}
$$

and

$$
\check{J}(s, x; \pi_0, \bar{u}) = \mathbb{E}^{\pi_0, \bar{u}}_{s,x} \left[ \int_s^T \check{L}(t, X^{(\pi_0)}(t^-), \hat{u}(t)) dt + \check{M}(X^{(\pi_0)}(T)) \right]
= \mathbb{E}^{\bar{u}}_{s,x} \left[ \int_s^T \check{L}(t, X^{(\pi_0)}(t^-), \hat{u}(t)) dt + \check{M}(X^{(\pi_0)}(T)) \right].
\tag{7.30}
$$

Note that the classic controller is the only one with the ability to influence the performance criteria.
Using the dynamic programming method, the problem may be solved by analysing the value function

\[ \hat{V}^{(0)}(s, x; \pi_0) = \inf_{u \in \mathcal{U}} \hat{J}(s, x; \pi_0, u) \]

\[ = \inf_{u \in \mathcal{U}} \mathbb{E}^u_{sx} \left[ \int_s^T \hat{L}(t, X^u(t), \tilde{u}(t)) dt + \hat{M}(X^u(T)) \right] \quad (7.31) \]

where the infimum is taken over all admissible feedback controls \( u \in \mathcal{U} \). Also define the following "value function" for the impulse controller:

\[ \check{V}^{(0)}(s, x; u) = \inf_{\pi \in \Pi_s(0)} \check{J}(s, x; \pi_0, u) \]

\[ = \mathbb{E}^u_{sx} \left[ \int_s^T \hat{L}(t, X^u(t), \tilde{u}(t)) dt + \hat{M}(X^u(T)) \right] \quad (7.32) \]

**Theorem 7.5.1.** Suppose there exist two functions \( \varphi^{(0)} \) and \( \psi^{(0)} \) belonging to \( \mathcal{D} \), together with a feedback control \( u^* \in \mathcal{U} \), which satisfy the following system of equations:

\[ i. \quad \forall v \in \mathcal{U}, \mathcal{A}^v[\varphi^{(0)}](s, x) + \hat{L}(s, x, v) \geq 0 \quad \forall (s, x) \in Q \quad (7.33) \]

\[ \text{ii.} \begin{cases} \inf_{v \in \mathcal{U}} \{ \mathcal{A}^v[\varphi^{(0)}](s, x) + \hat{L}(s, x, v) \} = 0 & \forall (s, x) \in Q \\ \varphi^{(0)}(T, x) = \hat{M}(x) & \forall x \in \mathbb{R}^d \end{cases} \quad (7.34) \]

\[ \text{iii.} \quad u^*(s, x) \in \arg \min_{v \in \mathcal{U}} \{ \mathcal{A}^v[\varphi^{(0)}](s, x) + \hat{L}(s, x, v) \} & \forall (s, x) \in Q \quad (7.35) \]

\[ \text{iv.} \begin{cases} \mathcal{A}^u[\psi^{(0)}](s, x) + \hat{L}(s, x, u^*(s, x)) = 0 & \forall (s, x) \in Q \\ \psi^{(0)}(T, x) = \hat{M}(x) & \forall x \in \mathbb{R}^d \end{cases} \quad (7.36) \]

Then for every \((s, x) \in Q\), \( \varphi^{(0)}(s, x) = \hat{V}^{(0)}(s, x; \pi_0) \), \( \psi^{(0)}(s, x) = \check{V}^{(0)}(s, x; u^*) \) and \((u^*, \pi_0)\) is a Nash equilibrium solution to the game with \( N = 0 \) interventions.

**Sketch of proof.** The proof involves repeated applications of Dynkin’s formula (7.21) and is more or less standard (see [58] for instance). It is therefore only sketched here. Let \((s, x) \in Q\) and \( u \in \mathcal{U} \) be arbitrary. Since \( \varphi^{(0)} \in \mathcal{D} \), Dynkin’s formula is valid and one can use equation (7.33) to obtain

\[ \mathbb{E}^u_{sx}[\varphi^{(0)}(T, X^u(T))] = \varphi^{(0)}(s, x) + \mathbb{E}^u_{sx} \left[ \int_s^T \mathcal{A}^u[\varphi^{(0)}](t, X^u(t)) dt \right] \]

\[ \geq \varphi^{(0)}(s, x) - \mathbb{E}^u_{sx} \left[ \int_s^T \hat{L}(t, X^u(t), u(t, X^u(t))) dt \right] . \]
Then, using the boundary condition at time $T$, the above leads to
\[
\varphi(0)(s, x) \leq E^u_{sx} \left[ \tilde{M}(X^u(T)) + \int_s^T \tilde{L}(t, X^u(t), u(t, X^u(t)))dt \right] \\
= \hat{J}(s, x; \pi_0, u)
\] (7.37)
and, by arbitrariness of $u$, it holds for every $u \in U$. Equality in (7.37) holds when $u \equiv u^*$ and therefore shows $\varphi(0)(s, x) = \hat{V}(0)(s, x; \pi_0)$. The other claims follow in a similar manner.

Remark 7.5.1. Equation (7.33) is called the Hamilton-Jacobi-Bellman partial differential equation. Equation (7.34) shows that a candidate optimal control minimises the Hamiltonian $\mathcal{H}$,
\[
\mathcal{H}[\varphi(0)](s, x) = \inf_{v \in U} \left\{ \sum_{i=1}^d b_i(s, x, v) \frac{\partial}{\partial x_i} \varphi(0)(s, x) + \hat{L}(s, x, v) \right\}.
\]
If $\varphi(0): \tilde{Q} \to \mathbb{R}$ is such that its gradient is continuous (as in the hypotheses of Theorem 7.5.1), then a selection lemma can be used to obtain a Borel-measurable map $u^*: \tilde{Q} \to U$ which achieves the minimum for $\mathcal{H}[\varphi](s, x)$ for almost every in $(s, x) \in Q$.

However, one still needs to show that $u^*$ is admissible (particularly, $\mathcal{A}u^*$ is the generator of a continuous Markov process).

Alternatively, the other hand the proof of the verification theorem 7.5.1 provides a stochastic interpretation of the value functions $\varphi(0), \psi(0)$ for the non-zero sum game. This alternative formulation can be useful when the functions comprising the problem data are not smooth. Specifically, the candidate $\varphi(0): \tilde{Q} \to \mathbb{R}$ is a measurable function such that for every $(s, x) \in Q$ and $u \in U$, the process $J^u = (J^u(t))_{s \leq t \leq T}$ defined by
\[
J^u(t) = \varphi(0)(t, X^u(t)) + \int_s^T \tilde{L}(t, X^u(t), u(t, X^u(t)))dt
\]
is a:

- sub-martingale with respect to $(\Omega, \mathcal{F}, \mathbb{F}, P^u_{sx})$ on $[s, T]$;

- martingale with respect to $(\Omega, \mathcal{F}, \mathbb{F}, P^u_{sx})$ on $[s, T]$ whenever $u \equiv u^*$ is an optimal (minimising) control.

In the latter case, there is a joint martingale formulation for the value function $\psi(0)$. Therefore, the Nash equilibrium solution $(u^*, \pi_0)$ is such that the two candidate value
functions $\varphi^{(0)}$ and $\psi^{(0)}$ are measurable functions that are, up to an integral term, martingales along the path of the controlled process $X^{u^*}$ on $[s, T]$. This characterisation of the Nash equilibrium also features in the case of $N \geq 1$ impulses as shown below.

For an in depth discussion of the martingale approach to stochastic control when the family of operators $\{A^v, \ v \in U\}$ is uniformly parabolic see Chapter 16 of [51]. In the degenerate case see [44, 73].

### 7.5.2 Solution with $N = 1$ interventions

An intermediate game of control and stopping.

Consider again the canonic space $(\Omega, \mathcal{F}, \mathbb{F})$ of Section 7.3. Recall $\Omega^{(2)} = \Omega \times \Omega$, $\mathcal{F}^{(2)} = \mathcal{F} \times \mathcal{F}$ and $\mathcal{F}^{(2)}_t = \mathcal{F}_t \times \mathcal{F}_t$.

**Definition 7.5.1.** An intermediate intervention $\nu$ is a pair $\nu = (\tau, \zeta)$ where $\tau: \Omega \to [0, T]$ is an $\mathbb{F}$-stopping time and $\zeta: \Omega \to \mathcal{I}$ is $\mathcal{F}_\tau$-measurable. For a given initial condition $(s, x) \in Q$, $\mathcal{I}_s$ denotes the class of intermediate interventions $\nu = (\tau, \zeta)$ satisfying $\tau(\omega) \geq s$ for every $\omega \in \Omega$.

**Remark 7.5.2.** Using this definition of an intermediate intervention, an impulse control strategy $\pi \in \Pi^{(1)}_s$, $\pi = (\tau_1, \zeta_1)$ (cf. Definition 7.3.3) could also be viewed as an intermediate intervention $\nu \in \mathcal{I}_s$ which has been extended from the canonic space $(\Omega, \mathcal{F}, \mathcal{F}_t)$ to $(\Omega^{(2)}, \mathcal{F}^{(2)}, \mathcal{F}^{(2)}_t)$ in a specific measurable way.

Let $\Phi: \bar{Q} \to \mathbb{R}$ and $\Psi: \bar{Q} \to \mathbb{R}$ be two given measurable functions belonging to $\mathcal{D}$. For $u \in \mathcal{U}$, the notation $u(t) := u(t, X^u(t))$, $t \geq 0$ is used below for convenience. For each $u \in \mathcal{U}$ and $\nu \in \mathcal{I}_s$, define two performance criteria: one for a controller,

$$\tilde{J}(s, x; \nu, u) := \mathbb{E}^u_{sx} \left[ \int_s^T \tilde{L}(t, X^u(t), u(t)) \, dt + \tilde{M}(X^u(T)) \mathbf{1}_{\{\tau = T\}}^{(7.38)} + \{ \Phi(\tau, \Gamma(X^u(\tau), \zeta)) + \tilde{G}(X^u(\tau), \zeta) \} \mathbf{1}_{\{\tau < T\}} \right],$$

and the other for a stopper,

$$\tilde{J}(s, x; \nu, u) := \mathbb{E}^u_{sx} \left[ \int_s^T \tilde{L}(t, X^u(t), u(t)) \, dt + \tilde{M}(X^u(T)) \mathbf{1}_{\{\tau = T\}}^{(7.39)} + \{ \Psi(\tau, \Gamma(X^u(\tau), \zeta)) + \tilde{G}(X^u(\tau), \zeta) \} \mathbf{1}_{\{\tau < T\}} \right].$$
Given that the stopper uses $\nu \in \mathcal{I}_s$, the controller seeks a strategy $u \in \mathcal{U}$ which minimises (7.38). Simultaneously, given that the controller uses $u \in \mathcal{U}$, the stopper seeks $\nu \in \mathcal{I}_s$ which minimises (7.39). This joint optimisation problem leads to a non-zero sum controller-stopper game (also known as a game of control and stopping). Similar to Definition 7.4.1 above, one can define the following concept of a solution to the game.

**Definition 7.5.2.** A pair $(u^*, \nu^*) \in \mathcal{U} \times \mathcal{I}_s$ is called a Nash equilibrium solution to the intermediate non-zero sum controller-stopper game with initial condition $(s, x) \in Q$, if for all $u \in \mathcal{U}$ and $\nu \in \mathcal{I}_s$:

$$\begin{align*}
\hat{J}(s, x; \nu^*, u^*) &\leq \hat{J}(s, x; \nu^*, u), \\
\hat{J}(s, x; u^*, \nu^*) &\leq \hat{J}(s, x; \nu, u^*).
\end{align*}$$

(7.40)

Introduce value functions $\hat{\mathcal{V}}$ and $\hat{\mathcal{V}}$ for the controller and stopper respectively

$$\begin{align*}
\hat{\mathcal{V}}(s, x; \nu) &:= \inf_{u \in \mathcal{U}} \hat{J}(s, x; \nu, u), \quad \nu \in \mathcal{I}_s; \\
\hat{\mathcal{V}}(s, x; u) &:= \inf_{\nu \in \mathcal{I}_s} \hat{J}(s, x; \nu, u), \quad u \in \mathcal{U}.
\end{align*}$$

(7.41) (7.42)

**Proposition 7.5.2.** Let $\Phi, \Psi \in \mathcal{D}$ be given. Suppose there exist two functions $\varphi^{(1)}$ and $\psi^{(1)}$ belonging to $\mathcal{D}$, together with an admissible feedback control $u^* \in \mathcal{U}$ and a Borel-measurable map $\tilde{\zeta}^*: [0, T] \times \mathbb{R}^d \rightarrow \mathcal{S}$, which satisfy the following system of equations:

i. $\mathcal{A}^{u^*}[\psi^{(1)}](s, x) + \hat{L}(s, x, u^*(s, x)) \geq 0, \quad \forall (s, x) \in Q$ (7.43)

ii. $\psi^{(1)}(s, x) \leq \inf_{i \in \mathcal{I}} \{ \Psi(s, \Gamma(x, i)) + \hat{G}(x, i) \}, \quad \forall (s, x) \in Q$ (7.44)

set $\mathcal{C}^{(1)} := \left\{ (s, x) \in Q: \psi^{(1)}(s, x) < \inf_{i \in \mathcal{I}} \{ \Psi(s, \Gamma(x, i)) + \hat{G}(x, i) \} \right\}$ (7.45)

iii. $\forall (s, x) \in \mathcal{C}^{(1)}:
\begin{align*}
\forall v \in \mathcal{U}, & \quad \mathcal{A}^v[\varphi^{(1)}](s, x) + \hat{L}(s, x, v) \geq 0 \\
\inf_{v \in \mathcal{U}} \{ \mathcal{A}^v[\varphi^{(1)}](s, x) + \hat{L}(s, x, v) \} &\leq 0 \\
u^*(s, x) &\in \arg \min_{v \in \mathcal{U}} \{ \mathcal{A}^v[\varphi^{(1)}](s, x) + \hat{L}(s, x, v) \}
\end{align*}$

(7.46)

iv. $\forall (s, x) \in Q \setminus \mathcal{C}^{(1)}$:
\[
\begin{align*}
\psi^{(1)}(s, x) &= \inf_{i \in \mathcal{I}} \{ \Psi(s, \Gamma(x, i)) + \hat{G}(x, i) \} \\
\bar{\zeta}^{*}(s, x) &= \arg \min_{i \in \mathcal{I}} \{ \Psi(s, \Gamma(x, i)) + \hat{G}(x, i) \} \\
\varphi^{(1)}(s, x) &= \Phi(s, \Gamma(x, \bar{\zeta}^{*}(s, x))) + \hat{G}(x, \bar{\zeta}^{*}(s, x))
\end{align*}
\]

\[\tag{7.47}\]

\[
\forall x \in \mathbb{R}^{d}, \quad \begin{cases} 
\varphi^{(1)}(T, x) = \hat{M}(x) \\
\psi^{(1)}(T, x) = \hat{M}(x) 
\end{cases}
\]

\[\tag{7.48}\]

Let \(\xi\) denote the canonical process on \((\Omega, \mathcal{F})\). Define \(\tau^{*}: \Omega \to [0, T]\) and \(\zeta^{*}: \Omega \to \mathcal{I}\) by

\[
\tau^{*} = \inf \{ t \geq s : (t, \xi_t) \notin C^{(1)} \} \wedge T \\
\zeta^{*} = \begin{cases} 
\bar{\zeta}^{*}(\tau^{*}, \xi_{\tau^{*}}) & \text{on } \{ \tau^{*} < T \} \\
i \in \mathcal{I} & \text{otherwise}
\end{cases}
\]

and set \(\nu^{*} = (\tau^{*}, \zeta^{*})\). Then \(\varphi^{(1)}(s, x) = \hat{V}(s, x; \nu^{*})\), \(\psi^{(1)}(s, x) = \hat{V}(s, x; u^{*})\), and \((u^{*}, \nu^{*})\) is a Nash equilibrium solution to the intermediate controller-stopper game (7.38)–(7.39).

Proof of Proposition 7.5.2 (sketch). Let \((s, x) \in Q\), \(\nu^{*} = (\tau^{*}, \zeta^{*})\) and be as stated above. Note that the compactness of \(\mathcal{I}\) and the assumptions on the functions \(\Psi, \Gamma\) and \(\hat{G}\) ensure that \((s, x) \mapsto \inf_{i \in \mathcal{I}} \{ \Psi(s, \Gamma(x, i)) + \hat{G}(x, i) \}\) is continuous on \(Q\). Since \(\psi^{(1)}\) is continuous on \(\bar{Q}\) by assumption, the continuation region \(C^{(1)}\) defined in (7.45) is an open subset of \(Q\). Therefore \(\tau^{*}\) is a stopping time with respect to \((\Omega, \mathcal{F}, \mathcal{F})\), and it also satisfies \(\tau^{*}(\omega) \geq s\) for all \(\omega \in \Omega\). One can also readily verify that \(\zeta^{*}: \Omega \to \mathcal{I}\) is \(\mathcal{F}_{\tau^{*}}\)-measurable, and therefore \(\nu^{*} = (\tau^{*}, \zeta^{*}) \in \mathcal{I}_{s}\).

Step 1: \(\varphi^{(1)}(s, x) = \hat{V}(s, x; \nu^{*})\).

Let \(u \in \mathcal{U}\) be arbitrary, then apply Dynkin’s formula to \(\varphi^{(1)}\) along \((t, X^{u}(t))\) up to \(\tau^{*}\) and use (7.46) to get:

\[
\mathbb{E}_{sx}^{u} \left[ \varphi^{(1)}(\tau^{*}, X^{u}(\tau^{*})) \right] = \varphi^{(1)}(s, x) + \mathbb{E}_{sx}^{u} \left[ \int_{s}^{\tau^{*}} \mathcal{A}_{u}[\varphi^{(1)}](t, X^{u}(t))dt \right] \\
\geq \varphi^{(1)}(s, x) - \mathbb{E}_{sx}^{u} \left[ \int_{s}^{\tau^{*}} \hat{L}(t, X^{u}(t), u(t, X^{u}(t)))dt \right]. \tag{7.50}\]
Conditioning on the events \( \{ \tau^* < T \} \) and \( \{ \tau^* = T \} \) then using the boundary conditions (7.48) and (7.48) gives

\[
E_{sx}^u [\varphi^{(1)}(\tau^*, X^u(\tau^*))] = E_{sx}^u [1_{\{\tau^* < T\}} \varphi^{(1)}(\tau^*, X^u(\tau^*))] + E_{sx}^u [1_{\{\tau^* = T\}} \varphi^{(1)}(T, X^u(T))]
\]

\[
= E_{sx}^u [1_{\{\tau^* < T\}} \{ \Phi(\tau^*, \Gamma(X^u(\tau^*), \zeta^*)) + \hat{G}(X^u(\tau^*), \zeta^*) \}]
\]

\[
+ E_{sx}^u [1_{\{\tau^* = T\}} \hat{M}(X^u(T))]. \tag{7.51}
\]

Since equations (7.50) and (7.51) together yield

\[
\varphi^{(1)}(s, x) \leq E_{sx}^u [1_{\{\tau^* < T\}} \{ \Phi(\tau^*, \Gamma(X^u(\tau^*), \zeta^*)) + \hat{G}(X^u(\tau^*), \zeta^*) \}]
\]

\[
+ E_{sx}^u [1_{\{\tau^* = T\}} \hat{M}(X^u(T)) + \int_s^\tau \hat{L}(t, X^u(t), u(t, X^u(t)))dt]
\]

\[
= \hat{J}(s, x; \nu^*, u) \tag{7.52}
\]

and \( u \in \mathcal{U} \) was arbitrary, this means \( \varphi^{(1)}(s, x) \leq \hat{J}(s, x; \nu^*, u) \) for all \( u \in \mathcal{U} \). Equality in (7.52) holds for \( u \equiv u^* \), that is \( \varphi^{(1)}(s, x) = \hat{V}(s, x; \nu^*) \), due to condition (7.46) above.

**Step 2:** \( \psi^{(1)}(s, x) = \hat{V}(s, x; u^*) \).

Let \( \nu = (\tau, \zeta) \in \mathcal{I}_s \) be arbitrary and \( u^* \in \mathcal{U} \) be as stated in the hypothesis. Applying Dynkin’s formula to \( \psi^{(1)} \) along \( (t, X^{u^*}(t)) \) up to \( \tau \) and using (7.43) shows that

\[
E_{sx}^{u^*} [\psi^{(1)}(\tau, X^{u^*}(\tau))] = \psi^{(1)}(s, x) + E_{sx}^{u^*} \left[ \int_s^\tau s\psi^{u^*}(t)[\psi^{(1)}](t, X^{u^*}(t))dt \right]
\]

\[
\geq \psi^{(1)}(s, x) - E_{sx}^{u^*} \left[ \int_s^\tau \hat{L}(t, X^{u^*}(t), u^*(t, X^{u^*}(t)))dt \right]. \tag{7.53}
\]

Similar to before, conditioning on the events \( \{ \tau < T \} \) and \( \{ \tau = T \} \), utilising the constraint (7.44) and boundary condition (7.48) gives:

\[
E_{sx}^{u^*} [\psi^{(1)}(\tau, X^{u^*}(\tau))] \leq E_{sx}^{u^*} [1_{\{\tau < T\}} \psi^{(1)}(\tau, X^{u^*}(\tau))] + E_{sx}^{u^*} [1_{\{\tau = T\}} \psi^{(1)}(T, X^{u^*}(T))]
\]

\[
\leq E_{sx}^{u^*} [1_{\{\tau < T\}} \{ \Psi(\tau, \Gamma(X^{u^*}(\tau), \zeta)) + \hat{G}(X^{u^*}(\tau), \zeta) \}]
\]

\[
+ E_{sx}^{u^*} [1_{\{\tau = T\}} \hat{M}(X^{u^*}(T))]. \tag{7.54}
\]

Then, putting equations (7.53) and (7.54) together leads to

\[
\psi^{(1)}(s, x) \leq E_{sx}^{u^*} [1_{\{\tau < T\}} \{ \Psi(\tau, \Gamma(X^{u^*}(\tau), \zeta)) + \hat{G}(X^{u^*}(\tau), \zeta) \}]
\]

\[
+ E_{sx}^{u^*} [1_{\{\tau = T\}} \hat{M}(X^{u^*}(T)) + \int_s^\tau \hat{L}(t, X^{u^*}(t), u^*(t, X^{u^*}(t)))dt]
\]

\[
= \hat{J}(s, x; \nu, u^*). \tag{7.55}
\]
Hence $\psi^{(1)}(s, x) \leq \tilde{J}(s, x; \nu, u^*)$ for all $\nu \in \mathcal{I}_s$. Equality in (7.52) holds for $\nu \equiv \nu^*$ due to the conditions (7.44) and (7.46) and the definition of $\nu^*$, which means $\psi^{(1)}(s, x) = \tilde{V}(s, x; u^*)$.

Conclusion.

The preceding arguments established the following: for all $u \in \mathcal{U}$,

$$\hat{J}(s, x; \nu^*, u^*) = \varphi^{(1)}(s, x) \leq \hat{J}(s, x; \nu^*, u)$$

and for all $\nu \in \mathcal{I}_s$,

$$\tilde{J}(s, x; \nu^*, u^*) = \psi^{(1)}(s, x) \leq \tilde{J}(s, x; \nu, u^*).$$

The pair $(u^*, \nu^*)$ is therefore a Nash equilibrium solution to the intermediate controller-stopper game.

A solution to the game of classic and impulse control.

The solutions to the game of classic and impulse control with zero impulses in Section 7.5.1 and the non-zero sum controller-stopper game can be combined to solve the game of classic and impulse control with at most $N = 1$ interventions.

**Theorem 7.5.3.** Suppose the hypotheses of Theorem 7.5.1 are true and let $\varphi^{(0)}, \psi^{(0)} \in \mathcal{D}$ and $u_0^* \in \mathcal{U}$ be the functions and feedback control obtained therewith. Suppose furthermore that Proposition 7.5.2 also holds true with $\varphi \equiv \varphi^{(0)}$ and $\psi \equiv \psi^{(0)}$ and let $\varphi^{(1)}, \psi^{(1)} \in \mathcal{D}$ and $u_1^*$ be the functions and feedback control obtained therewith. Let $(s, x) \in Q$ be given and $\nu_1^* = (\tau_1^*, \zeta_1^*) \in \mathcal{I}_s$ denote the intermediate intervention from (7.49) in Proposition 7.5.2. Let $\pi^* = (\tilde{\tau}_1^*, \tilde{\zeta}_1^*)$ be an extension of $\nu_1^* = (\tau_1^*, \zeta_1^*)$ from $(\Omega, \mathcal{F}, (\mathcal{F}_t))$ to $(\Omega^{(2)}, \mathcal{F}^{(2)}(\tau_1^*) temperament)$ satisfying:

- $\tilde{\tau}_1^*$ is an $(\mathcal{F}_t \times \{\emptyset, \Omega\})$-stopping time;
- $\tilde{\zeta}_1^*$ is $(\mathcal{F}_{\tilde{\tau}_1^*} \times \{\emptyset, \Omega\})$-measurable;
- For each $\tilde{\omega} \in \Omega^{(2)}$, with $\tilde{\omega} = (\omega_1, \omega_2)$, $\tilde{\tau}_1^*(\tilde{\omega}) = \tau_1^*(\omega_1)$ and $\tilde{\zeta}_1^*(\tilde{\omega}) = \zeta_1^*(\omega_1)$.

Then, setting $\bar{u}^* = (u_1^*, u_0^*)$, the functions $\varphi^{(1)}$ and $\psi^{(1)}$ satisfy $\varphi^{(1)}(s, x) = \tilde{J}(s, x; \pi^*, \bar{u}^*)$ and $\psi^{(1)}(s, x) = \tilde{J}(s, x; \pi^*, \bar{u}^*)$. Furthermore, the pair $(\bar{u}^*, \pi^*)$ is a Nash Equilibrium solution to the game of classic and impulse control (7.26)–(7.27) with at most $N = 1$ interventions.
Sketch of proof. Note that \( \pi^* \in \Pi_s^{(1)} \) as per Definition 7.3.3 and setting \( \hat{\mathbf{u}}^* = (u^*_1, u^*_0) \) leads to \( \hat{\mathbf{u}}^* \in \mathcal{U}^{(1)} \). The corresponding state process \( X^{(\pi^*)} \), admissible classic control with interventions \( u^* \), and probability measure \( P_{sx}^{\pi^*,d^*} \) on \( (\Omega^{(2)}, \mathcal{F}^{(2)}) \) can all be defined as in Section 7.3.2 above. The proof proceeds similarly to that of Theorem 7.5.1 and Proposition 7.5.2 above and is therefore sketched.

**Step 1:** \( \phi^{(1)}(s, x) = \hat{J}(s, x; \pi^*, \hat{\mathbf{u}}^*) \).

Let \( \hat{\mathbf{u}} \in \mathcal{U}^{(1)} \), \( \hat{\mathbf{u}} = (u^{(0)}, u^{(1)}) \), be arbitrary. The following expression holds for \( \phi^{(1)}(s, x) \) (compare with equation (7.52) in the proof of Proposition 7.5.2):

\[
\phi^{(1)}(s, x) \leq E_{sx}^{(0)} \left[ 1_{\{\tau^*_1 = T\}} \hat{M}(X^{(0)}(T)) + \int_s^{\tau^*_1} \hat{L}(t, X^{(0)}(t), u(t, X^{(0)}(t)))dt \right]
+ \mathbb{E}_{sx}^{(0)} [1_{\{\tau^*_1 < T\}} \{\phi^{(0)}(\tau^*_1, \Gamma(X^{(0)}(\tau^*_1), \zeta^*_1)) + \hat{G}(X^{(0)}(\tau^*_1), \zeta^*_1)\}] \tag{7.56}
\]

On the other hand, the process \( X^{\pi^*} \) agrees with \( X^{(0)} \) when evaluated under the measure \( P_{sx}^{\pi^*,d^*} \) for times before \( \tau^*_1 \). Using the properties outlined in Section 7.3.2 on controlled Markov processes with interventions, one can rewrite the right-hand side of equation (7.56) in the following way:

\[
\phi^{(1)}(s, x) \leq E_{sx}^{\pi^*,d^*} \left[ 1_{\{\tau^*_1 = T\}} \{\phi^{(0)}(\tau^*_1, X^{\pi^*}(\tau^*_1)) + \hat{G}(X^{\pi^*}(\tau^*_1), \zeta^*_1)\} \right]
+ \mathbb{E}_{sx}^{\pi^*,d^*} \left[ 1_{\{\tau^*_1 > T\}} \hat{M}(X^{\pi^*}(T)) + \int_s^{\tau^*_1} \hat{L}(t, X^{\pi^*}(t), u(t))dt \right] \tag{7.57}
\]

Consider those \( \tilde{\omega} \in \Omega^{(2)}, \tilde{\omega} = (\omega, \omega_2) \), such that \( \omega \in \Omega \) is fixed. Then \( \tau^*_1(\tilde{\omega}) := \tau^*_1(\omega) = r \in [s, T] \) and \( \zeta^*_1(\tilde{\omega}) := \zeta^*_1(\omega) = i \in \mathcal{I} \) is fixed for all such \( \tilde{\omega} \). Denote \( \Gamma(X^{(0)}(r), i) = y \in \mathbb{R}^d \) and note that, with respect to \( (\Omega, \mathcal{F}, \mathbb{P}, P_{ry}^{u^{(1)}}) \), the process \( X^{(1)} \) is actually the canonical one \( \xi^{(1)} \) and is evaluated using \( \omega_2 \in \Omega \) (not fixed). Using similar arguments as in Theorem 7.5.1, apply Dynkin’s formula to \( \phi^{(0)} \) along \( (t, X^{(1)}(t)) \) with respect to \( (\Omega, \mathcal{F}, \mathbb{P}, P_{ry}^{u^{(1)}}) \) to get:

\[
E_{ry}^{u^{(1)}} [\phi^{(0)}(r, y)] \leq E_{ry}^{u^{(1)}} \left[ \hat{M}(X^{(1)}(T)) + \int_r^T \hat{L}(t, X^{(1)}(t), u^{(1)}(t, X^{(1)}(t)))dt \right] \tag{7.58}
\]

Equation (7.58) is true for every \( \omega \in \Omega \) such that \( \tau^*_1(\omega) = r \) and \( X^{(1)}(r) = y \) almost surely \( P_{ry}^{u^{(1)}} \). Since the state process \( X^{\pi^*} \) agrees with \( X^{(1)} \) under \( P_{sx}^{\pi^*,d^*} \) from time \( \tau^*_1 \) onwards on the event \( \{\tau^*_1 < T\} \), applying the properties of \( P_{sx}^{\pi^*,d^*} \) from Definition 7.3.6...
and (7.58) to equation (7.57) gives:
\[
\psi^{(1)}(s, x) \leq E_{sx}^{\pi, \tilde{u}} \left[ \mathbf{1}_{\{\tau_1 < T\}} \left\{ \psi^{(0)}(\tilde{\tau}_1, X^\pi(\tilde{\tau}_1)) + G(X^\pi(\tilde{\tau}_1 -), \tilde{\zeta}_1) \right\} \right] \\
+ E_{sx}^{\pi, \tilde{u}} \left[ \mathbf{1}_{\{\tau_1 = T\}} \tilde{M}(X^\pi(T)) + \int_{\tilde{\tau}_1}^T \tilde{L}(t, X^\pi(t^-), u(t))dt \right] \\
\leq E_{sx}^{\pi, \tilde{u}} \left[ \int_{\tilde{s}}^T \tilde{L}(t, X^\pi(t^-), u(t))dt + \tilde{M}(X^\pi(T)) \right] \\
+ \mathbf{1}_{\{\tau_1 < T\}} \tilde{G}(X^\pi(\tilde{\tau}_1 -), \tilde{\zeta}_1) \\
= \hat{J}(s, x; \pi^*, \tilde{u})
\]
(7.59)

Following these same arguments shows that equality in (7.59) holds for \( \tilde{u} \equiv \tilde{u}^* \).

**Step 2:** \( \psi^{(1)}(s, x) = \hat{J}(s, x; \pi^*, \tilde{u}^*) \).

Let \( \pi \in \Pi_x^{(1)}, \pi = (\tau_1, \zeta), \) be arbitrary. Under \( P_{sx}^{\pi, \tilde{u}^*} \), the state process \( X^\pi \) agrees with \( X^{\tilde{u}^*} \) for times before \( \tau_1 \), and agrees with \( X^{\bar{u}_0} \) from time \( \tau_1 \) onwards on the event \( \{ \tau_1 < T \} \). Following similar arguments to those above, one can establish (compare with (7.54) above):
\[
\psi^{(1)}(s, x) \leq E_{sx}^{\pi, \bar{u}^*} \left[ \mathbf{1}_{\{\tau_1 < T\}} \left\{ \psi^{(0)}(\tau_1, X^\pi(\tau_1)) + \bar{G}(X^\pi(\tau_1 -), \zeta_1) \right\} \right] \\
+ E_{sx}^{\pi, \bar{u}^*} \left[ \mathbf{1}_{\{\tau_1 = T\}} \bar{M}(X^\pi(T)) + \int_{\tau_1}^T \bar{L}(t, X^\pi(t^-), u(t))dt \right] \\
\leq E_{sx}^{\pi, \bar{u}^*} \left[ \int_{\tau_1}^T \bar{L}(t, X^\pi(t^-), u(t))dt + \bar{M}(X^\pi(T)) \right] \\
+ \mathbf{1}_{\{\tau_1 < T\}} \bar{G}(X^\pi(\tau_1 -), \zeta_1) \\
= \hat{J}(s, x; \pi, \bar{u}^*)
\]
(7.60)

Again, equality holds in equation (7.60) when \( \pi \equiv \pi^* \).

**Conclusion.**

The previous arguments have just shown that for all \( \tilde{u} \in U^{(1)} \),
\[
\hat{J}(s, x; \pi^*, \tilde{u}) = \psi^{(1)}(s, x) \leq \hat{J}(s, x; \pi^*, \tilde{u})
\]
and for all \( \pi \in \Pi_x^{(1)} \),
\[
\hat{J}(s, x; \pi^*, \tilde{u}^*) = \psi^{(1)}(s, x) \leq \hat{J}(s, x; \pi, \tilde{u}^*)
\]
The pair \( (\tilde{u}^*, \pi^*) \) is therefore a Nash equilibrium solution to the game of classic and impulse control (7.26)–(7.27) with at most 1 intervention. \( \Box \)
7.5.3 Solution with $N \geq 1$ interventions

This section shows that the game (7.26)–(7.27) with at most $N \geq 1$ interventions is solved in a similar way to the game with at most 1 intervention: solve the intermediate controller-stopper game which has the solution to the game of classic and impulse control with at most $N - 1$ interventions as a boundary condition.

**Theorem 7.5.4.** Let $\varphi^{(0)}$, $\psi^{(0)}$ and $u_0^*$ be the respective value functions and optimal admissible strategy from Theorem 7.5.1. Suppose for $n = N, \ldots, 1$ there exist functions $\varphi^{(n)}$ and $\psi^{(n)}$ belonging to $\mathcal{D}$, together with an admissible feedback control $u_n^* \in \mathcal{U}$ and a Borel-measurable map $\bar{\zeta}_n^* : [0, T] \times \mathbb{R}^d \to \mathcal{I}$, which satisfy the following system of equations:

1. $\mathcal{A}u_n^*[\varphi^{(n)}](s, x) + \hat{L}(s, x, u_n^*(s, x)) \geq 0 \quad \forall (s, x) \in Q$ \hfill (7.61)
2. $\psi^{(n)}(s, x) \leq \inf_{i \in \mathcal{I}} \left\{ \psi^{(n-1)}(s, \Gamma(x, i)) + \hat{G}(x, i) \right\}, \quad \forall (s, x) \in Q$ \hfill (7.62)
3. $\mathcal{C}(n) := \left\{ \psi^{(n)}(s, x) < \inf_{i \in \mathcal{I}} \left\{ \psi^{(n-1)}(s, \Gamma(x, i)) + \hat{G}(x, i) \right\} \right\}$ \hfill (7.63)
4. $\forall (s, x) \in Q \setminus \mathcal{C}(n)$:
   
   \[
   \begin{aligned}
   &\forall v \in \mathcal{U}, \quad \mathcal{A}v[\varphi^{(n)}](s, x) + \hat{L}(s, x, v) \geq 0 \\
   &\inf_{v \in \mathcal{U}} \{\mathcal{A}v[\varphi^{(n)}](s, x) + \hat{L}(s, x, v)\} = 0 \\
   &u_n^*(s, x) \in \arg\min_{v \in \mathcal{U}} \{\mathcal{A}v[\varphi^{(n)}](s, x) + \hat{L}(s, x, v)\} \\
   &\mathcal{A}u_n^*[\psi^{(n)}](s, x) + \hat{L}(s, x, u_n^*(s, x)) = 0
   \end{aligned}
   \] \hfill (7.64)
5. $\forall x \in \mathbb{R}^d$:
   
   \[
   \begin{aligned}
   &\varphi^{(n)}(s, x) = \varphi^{(n-1)}(s, \Gamma(x, i)) + \hat{G}(x, i) \\
   &\bar{\zeta}_n^*(s, x) = \arg\min_{i \in \mathcal{I}} \left\{ \psi^{(n-1)}(s, \Gamma(x, i)) + \hat{G}(x, i) \right\} \\
   &\psi^{(n)}(s, x) = \varphi^{(n-1)}(s, \Gamma(x, \bar{\zeta}_n^*(s, x))) + \hat{G}(x, \bar{\zeta}_n^*(s, x))
   \end{aligned}
   \] \hfill (7.65)
6. $\forall x \in [0, T)$ be fixed and $\xi^{(0)}, \ldots, \xi^{(N-1)}$ denote the canonical processes on $N$ copies of the space $(\Omega, \mathcal{F})$. Let $(s, x) \in Q$ be given, set $\tau_0^* = s$ and for $n = 1, \ldots, N$, define
\( \tau_n^* : \Omega^{(n)} \to [0, T] \) and \( \zeta_n^* : \Omega^{(n)} \to \mathcal{I} \) by

\[
\tau_n^*(\omega_1, \ldots, \omega_n) = \inf \left\{ t \geq \tau_{n-1}^*(\omega_1, \ldots, \omega_{n-1}) : (t, \xi_t^{(n-1)}(\omega_n)) \notin \mathcal{C}^{(N-n+1)} \right\} \land T
\]

\[
\zeta_n^*(\omega_1, \ldots, \omega_n) = \begin{cases} 
\tilde{\zeta}_n^*(\tau_n^*, \xi_{\tau_n^*}^{(n-1)}(\omega_n)) \text{ on } \{ \tau_n^* \leq T \}, \\
i \in \mathcal{I} \text{ otherwise}.
\end{cases}
\] (7.67)

Define \( \pi^* = (\tau_n^*, \xi_n^*)_{n=1}^N \) and \( \bar{u}^* = (u_N^*, \ldots, u_0^*) \). Then, \( \varphi^{(N)}(s, x) = \hat{J}(s, x; \pi^*, \bar{u}^*) \), \( \psi^{(N)}(s, x) = \check{J}(s, x; \pi^*, \bar{u}^*) \), and \( (\bar{u}^*, \pi^*) \) is a Nash equilibrium solution to the game of classic and impulse control with at most \( N \geq 1 \) impulses.

**Remark 7.5.3.** Intuitively, for \( n = N, N-1, \ldots, 0 \) the feedback control \( u_n^* \) is optimal for the classic controller and the sets \( \mathcal{C}^{(n)} \) for \( n = N, \ldots, 1 \) defined using (7.63) are the continuation regions for the impulse controller when there are at most \( n \) impulses remaining (in the case \( n = 0 \) set \( \mathcal{C}^{(0)} = Q \)).

**Proof of Theorem 7.5.4 (sketch).** One readily verifies that \( \pi^* \in \Pi^{(N)}_s \) and \( \bar{u}^* \in \mathcal{U}^{(N)} \) by construction, recalling the discussion in the proof of Proposition 7.5.2. The proof of this theorem is essentially the same as for Theorem 7.5.3 above and only the main points are summarised.

Let \( X^{u}_j \) for \( j \in \{0, 1, \ldots, N\} \) denote the continuous Markov process associated with the feedback control \( u_n^* \). Iteratively for \( n = 0, 1, \ldots, N \), a stochastic representation for \( \varphi^{(N-n)} \) in terms of the running, impulse and terminal costs can be obtained by applying Dynkin’s formula to \( \varphi^{(N-n)} \) along the path of \( X^{u_{N-n}} \) on the interval \([\tau_n^*, \tau_{n+1}^*] \), and by using the various boundary conditions for \( \varphi^{(N-n)} \) at \( \tau_{n+1}^* \) in the hypothesis (recall \( \tau_{n+1}^* \equiv T \)). Under the measure \( \mathbb{P}^{\pi^*, a^*}_{s_x} \), the process \( X^{u_N} \) starts from \( x \in \mathbb{R}^d \) at time \( s \) and for \( n \geq 1 \), conditional on the event \( \{ \tau_n^* < T \} \) the process \( X^{u_{N-n}} \) starts from \( \Gamma(X^{u_{N-n}}, \zeta_n^*) \) at time \( \tau_n^* \). Since under \( \mathbb{P}^{\pi^*, a^*}_{s_x} \) the state process \( X^{\pi^*} \) agrees (in a certain sense) with \( X^{u_{N-n}} \) on each interval \([\tau_n^*, \tau_{n+1}^*] \), one can then show that

\[
\varphi^{(N)}(s, x) = \hat{J}(s, x; \pi^*, \bar{u}^*).
\]

On the other hand, there are inequalities (\( \leq \)) in the stochastic representation for each \( \varphi^{(N-n)} \) when an arbitrary control \( \bar{u} \in \mathcal{U}^{(N)} \) is used together with \( \pi^* \). This leads to

\[
\hat{J}(s, x; \pi^*, \bar{u}^*) = \varphi^{(N)}(s, x) \leq \hat{J}(s, x; \pi^*, \bar{u})
\]
for every $\bar{u} \in U^{(N)}$. In a similar fashion, one can show that for all $\pi \in \Pi^{(N)}_{s}$,

$$
\bar{J}(s, x; \pi^{*}, \bar{u}^{*}) = \psi^{(N)}(s, x) \leq \bar{J}(s, x; \pi, \bar{u}^{*})
$$

and conclude that $(\bar{u}^{*}, \pi^{*})$ is a Nash equilibrium solution.

7.6 The balancing services contract revisited

This section investigates the application of the verification theorems to the balancing services contract example of Section 7.2. The following assumption is made so that the investigation can proceed without changing the form of the payoff or PDEs in Section 7.5 (also see Remark 2.3 of [11]).

**Assumption 15.** If $X(\cdot)$ is a continuous stochastic process solving (7.5) relative to some setup $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ with $X(s) = x \in \mathcal{X} := [0, B_{\text{max}}] \times \mathbb{R} \times \mathbb{R}$, then $X(t) \in \mathcal{X}$ for all $s \leq t \leq T \mathbb{P}$-almost surely.

The operators $\mathcal{A}^{v}, v \in U$, associated with the process $X$ of (7.5) are given by

$$
\mathcal{A}^{v}[\varphi](t, x) = \frac{\partial}{\partial t}\varphi(t, x) + \left[ \frac{1}{2} \sum_{i,j=1}^{3} a_{ij}(t, x) \frac{\partial^2}{\partial x_i \partial x_j} + \sum_{i=1}^{3} b_{i}(t, x, v) \frac{\partial}{\partial x_i} \right] \varphi(t, x)
$$

where $\varphi \in \tilde{D} := C^{1,2}_{\mathbb{P}}([0, T] \times \mathcal{X})$ and $a(t, x) = \sigma(t, x)\sigma(t, x)^{T}$ is the covariance matrix

$$
a(t, x) = \begin{pmatrix}
0 & 0 & 0 \\
0 & (\sigma^{X_2}(t, x_2))^2 & 0 \\
0 & 0 & (\sigma^{X_3}(t, x_3))^2
\end{pmatrix}
$$

Note that the family of operators $\{\mathcal{A}^{v}, v \in U\}$ in this case is degenerate parabolic.

7.6.1 The case of $N = 0$ interventions

According to Theorem 7.5.1 and using the model of Section 7.2.1, the game of classic and impulse control with $N = 0$ interventions is solved if one can find functions
\( \varphi^{(0)}, \psi^{(0)} \in \tilde{D} \) and \( u^* \in \mathcal{U} \) such that

\[
\begin{align*}
\varphi^{(0)}(T, x) &= 0 \quad \forall x \in \mathcal{X} \\
\psi^{(0)}(T, x) &= 0 \quad \forall x \in \mathcal{X}
\end{align*}
\quad (7.68)
\]

\[ u^*(s, x) = \arg \min_{v \in \mathcal{U}} \{ A v[\varphi^{(0)}](s, x) + \hat{f}(x_3 \cdot v) \} \quad \forall (s, x) \in [0, T) \times \mathcal{X} \quad (7.69) \]

\[ A u^*[\psi^{(0)}](s, x) + \check{f}(x_2) = 0 \quad \forall (s, x) \in [0, T) \times \mathcal{X} \quad (7.70) \]

Ignore the function \( \psi^{(0)} \) and (7.70) momentarily. Then (7.68) is the Hamilton-Jacobi-Bellman equation associated with a finite time horizon stochastic control problem, which is discussed in [58]. Chapter IV of [58] discusses the existence of solutions to (7.68) in the class \( \tilde{D} \). The authors show that, given sufficiently smooth data for (7.68), there exists a solution \( \varphi^{(0)} \) to (7.68) in the class \( \tilde{D} \) when the operators \( \{ \mathcal{A}^v \} \) are uniformly parabolic. However, in the degenerate case and even with the same degree of smoothness, the authors are only able to prove the existence of a solution to (7.68) in some generalised sense (see their Theorem IV.10.2). Moreover, it is not guaranteed that a measurable function \( u^* \) achieving (7.69) in some generalised sense also admits a solution \( X u^* \) to (7.6) (see Chapter 5, Section 6 of [138]).

Approximation by non-degenerate controlled processes.

The previous discussion highlighted the degeneracy of the controlled diffusion as an impediment to applying the results of Section 7.5 to the balancing services contract. First, extend the domain of the drift coefficient \( b \) from \([0, T] \times \mathcal{X} \times \mathcal{U} \) to \([0, T] \times \mathbb{R}^3 \times \mathcal{U} \) in a continuous way, truncating it so that it retains the same bounds (see [58, p. 401] for Lipschitz functions). The domain of the costs \( \hat{L} \) and \( \check{L} \) as well as the intervention position function \( \Gamma \) should also be extended in an appropriate manner to account for an enlargement of the state space from \( \mathcal{X} \) to \( \mathbb{R}^3 \). Let \( \varepsilon > 0 \) be fixed and suppose that the diffusion coefficient \( \sigma \) has been replaced by \( \sigma^\varepsilon \) which is defined on all of \([0, T] \times \mathbb{R}^3 \) as follows: for \((t, x) \in \bar{Q}, x = (x_1, x_2, x_3), \)

\[
\sigma^\varepsilon(t, x) = \begin{pmatrix}
0 & 0 & \sqrt{2\varepsilon} & 0 & 0 \\
\sigma^{x_2}(t, x_2) & 0 & 0 & \sqrt{2\varepsilon} & 0 \\
0 & \sigma^{x_3}(t, x_3) & 0 & 0 & \sqrt{2\varepsilon}
\end{pmatrix}
\]
Then, the covariance matrix \( a^\varepsilon(t, x) = \sigma^\varepsilon(t, x)\sigma^\varepsilon(t, x)^T, \) \((t, x) \in \bar{Q}, \) is

\[
a^\varepsilon(t, x) = \begin{pmatrix}
2\varepsilon & 0 & 0 \\
0 & (\sigma X_2(t, x_2))^2 + 2\varepsilon & 0 \\
0 & 0 & (\sigma X_3(t, x_3))^2 + 2\varepsilon
\end{pmatrix},
\]

which satisfies for all \( \lambda \in \mathbb{R}^3: \)

\[
\sum_{i,j=1}^3 a^\varepsilon_{ij}(t, x_2)\lambda_i\lambda_j \geq 2\varepsilon|\lambda|^2.
\]

One can then define a family of uniformly parabolic partial differential operators \( \{\mathcal{A}^v, \ v \in U\} \) by

\[
\mathcal{A}^v_\varepsilon[\varphi](t, x) := \frac{\partial}{\partial t}\varphi(t, x) + \left[ \frac{1}{2} \sum_{i,j=1}^3 a^\varepsilon_{ij}(t, x) \frac{\partial^2}{\partial x_i \partial x_j} + \sum_{i=1}^3 b_i(t, x, v) \frac{\partial}{\partial x_i} \right] \varphi(t, x)
\]

\[
= \mathcal{A}^v[\varphi](t, x) + \varepsilon\Delta[\varphi](t, x) \quad (7.71)
\]

where \( \varphi \in \mathcal{D} \) and \( \Delta = \sum_{i=1}^3 \frac{\partial^2}{\partial (x_i)^2} \) is the Laplace operator. The operators \( \{\mathcal{A}^v, \ v \in U\} \) are associated with a process \( (X^\varepsilon(t))_{s \leq t \leq T} \) which is a weak solution to the following stochastic equation: for \( s \in [0, T) \) given and \( s \leq t \leq T, \)

\[
\begin{align*}
X^\varepsilon(s) &= x \in \mathcal{X}; \\
X^\varepsilon(t) &= x + \int_s^t b(\theta, X^\varepsilon(\theta), u(\theta))) \, d\theta + \int_s^t \sigma^\varepsilon(\theta, X^\varepsilon(\theta)) \cdot d\hat{W}(\theta)
\end{align*}
\]

(7.72)

where \( (\hat{W}(t))_{s \leq t \leq T} \) is a 5-dimensional Brownian motion and \( (u(t))_{s \leq t \leq T} \) is a \( U \)-valued progressively measurable process, both with respect to some stochastic basis \( (\hat{\Omega}, \hat{\mathcal{F}}, \hat{\mathbb{F}}, \hat{\mathbb{P}}) \) \((\hat{\Omega}, \hat{\mathcal{F}}) \) may be an enlargement of the canonic space \( (\Omega, \mathcal{F}) \)). There is an analogous equation for a process \( X^\varepsilon \) in feedback form (cf. (7.6)).

A non-degenerate controlled process \( X^\varepsilon \) satisfying (7.72), if it exists, approximates the degenerate one \( X \) of (7.5) as \( \varepsilon \downarrow 0 \). However, it is no longer true that \( X^\varepsilon \) satisfies the state constraint (cf. Assumption 15). Nevertheless, it is possible to analyse the HJB equation with the operators \( \{\mathcal{A}^v_\varepsilon, \ v \in U\} \) defined in (7.71) and investigate what happens as \( \varepsilon \downarrow 0 \). A similar concept was discussed briefly in Section 3.6.4 of Chapter 3 (see also Chapter 4, Section 6 of [84] and Chapter 4, Section 4.1 of [138]). It is highlighted again in Section 7.7 below.
An attempt at solving the problem for the classic controller.

For completeness, the HJB formulation (7.68)–(7.69) for the non-degenerate case is given by

\[
\begin{align*}
&i. \inf_{v \in U} \{ \mathcal{A}^v[\varphi_0^0](s, x) + \hat{L}(s, x, v) \} + \varepsilon \Delta[\varphi_0^0](s, x) = 0 \quad \forall (s, x) \in Q \\
&\varphi_0^0(T, x) = 0 \quad \forall x \in \mathbb{R}^3
\end{align*}
\] (7.73)

\[
\begin{align*}
&ii. \exists \mathbf{u}_*^\varepsilon \in U \text{ such that} \\
&\mathbf{u}_*^\varepsilon(s, x) \in \arg \min_{v \in U} \{ \mathcal{A}^v[\varphi_0^0](s, x) + \hat{f}(x \cdot v) \} \quad \forall (s, x) \in Q
\end{align*}
\] (7.74)

If \(\sigma(t, x), \hat{L}(t, x, v) = \hat{f}(x \cdot v)\) and \(b(t, x, v)\) are continuous functions on their domains, and furthermore admit partial derivatives with respect \(t\) and \(x\) which are continuous and bounded, then Theorem IV.4.2 of [58] proves that a solution \(\varphi_0^0\) to (7.73) in the class \(\mathcal{D}\) exists and is unique. It is also possible to prove the existence of a feedback control \(\mathbf{u}_*^\varepsilon\) satisfying (7.74) and such that (7.72) has a weak solution – see [58, pp.163–164] or [138, pp. 277-278] for example.

Remark 7.6.1. Note that as \(\{\mathcal{A}^v, v \in U\}\) is uniformly parabolic and \(a^\varepsilon\) is independent of \(v \in U\), existence of a weak solution \(X^{\mathbf{u}_*^\varepsilon}\) to (7.72) under a feedback control \(\mathbf{u}_*^\varepsilon\) is guaranteed if the diffusion coefficient \(a^\varepsilon\) and drift \(b\) satisfy Assumption 12 above (note that \((t, x) \mapsto b(t, x, \mathbf{u}_*^\varepsilon(t, x))\) only needs to be locally bounded and measurable). One can also show in this case that \(\mathbf{u}_*^\varepsilon\) is admissible according to Definition 7.3.1 above – see Chapter 10 of [125] and Lemma 2.3 of [73].

An attempt at solving the problem for the impulse controller.

Suppose the criteria for obtaining a function \(\varphi_0^0 \in \mathcal{D}\) which solves (7.73) and a feedback control \(\mathbf{u}_*^\varepsilon\): \(\tilde{Q} \rightarrow U\) which satisfies (7.74) and guarantees the existence of a weak solution \(X^\varepsilon\) to (7.72). Then, one can look at the secondary problem of finding a suitable function \(\psi_0^0\) satisfying:

\[
\begin{align*}
&\mathcal{A}^{\mathbf{u}_*^\varepsilon}[\psi_0^0](s, x) + \varepsilon \Delta[\psi_0^0](s, x) + \kappa(x_2)^2 = 0 \quad \forall (s, x) \in Q \\
&\psi_0^0(T, x) = 0 \quad \forall x \in \mathbb{R}^3
\end{align*}
\] (7.75)

However, there is no guarantee that there exists a solution \(\psi_0^0\) to (7.75) which is in \(\mathcal{D}\). The difficulty is that the feedback control \(\mathbf{u}_*^\varepsilon\) is not necessarily smooth on \(\tilde{Q}\). In
fact, Lemma VI.6.3 of [57] was only able to show that $u^*_\varepsilon$ is locally Lipschitz under more stringent assumptions than those leading to the existence of $\varphi^{(0)}_\varepsilon \in \mathcal{D}$. Nevertheless, under Lipschitz and growth assumptions on the coefficients for the Cauchy problem (7.75), it is still possible to obtain the stochastic representation

$$
\psi^{(0)}_\varepsilon(s, x) = \mathbb{E}_{s,x}^u \left[ \int_s^T \tilde{L}(t, X^{u^*_\varepsilon}(t), u^*_\varepsilon(t)) dt + \tilde{M}(X^{u^*_\varepsilon}(T)) \right], \quad \forall (s, x) \in [0, T) \times \mathcal{X}
$$

where $u^*_\varepsilon(t) = u^*_\varepsilon(t, X^{u^*_\varepsilon}(t))$ for $t \geq s$. This follows from a generalised version of the Feynman-Kac formula, Theorem 7.4.1 of [138].

### 7.6.2 Discussion

Section 7.6.1 showed that, contrary to the hypotheses of the verification theorems in Section 7.5, the candidate value functions for the balancing services contract are unlikely to solve the proposed PDEs in a classical sense. This observation followed first and foremost from the degeneracy of the parabolic partial differential operators. This particular difficulty arose because the state process for the storage level satisfies an ordinary differential equation, whereas the system imbalance and spot price are described by diffusions.

By introducing another Brownian component to the state equation, it was possible to alleviate the aforementioned problem of degenerate operators. Under rather strong assumptions on the payoff function for the classic controller, it is also possible to partially satisfy the hypotheses for the verification theorem when there are no interventions. However, these assumptions are still insufficient to confirm the hypotheses on the candidate value function for the impulse controller. This problem then propagates through the verification theorems for $N \geq 1$ interventions due to their dependence on the solutions for $N - 1$ interventions.

### 7.7 Conclusion

#### 7.7.1 Conclusion for this chapter

The preceding discussion showed that the hypotheses of the verification theorems in Section 7.5 must be weakened for them to be applicable to the balancing services contract. The following issues therefore need to be addressed:
1. the existence (and uniqueness) of appropriate, possibly non-smooth solutions to the coupled PDEs which arise;

2. the existence of trajectories for the state variable synthesised according to the optimality criteria associated with the PDE for the classic controller's value function;

3. the existence of Markovian optimal controls and a Markovian selector for a single optimal impulse;

4. the ability to obtain stochastic representations for the candidate value functions along the path of the synthesized state variable.

Section 7.6 explained that a pair of smooth solutions to the coupled PDEs need not exist even in the non-degenerate case. Nevertheless, it might be possible to show that there exist solutions to those PDEs in the *viscosity sense*. In this case there is a wide literature available including the references [58, 138] whose results were used previously. The paper [11] recently proposed a framework for verification theorem for continuous viscosity solutions to Hamilton-Jacobi-Bellman equations. Unfortunately, their approach does not furnish the existence of an optimal feedback control strategy. On the other hand, the paper [64] provides an alternative framework for establishing verification theorems and the existence of optimal feedback controls in stochastic control problems where the diffusion term may degenerate. The approach in [64] relies on a version of Itô's formula for $C^{0,1}$ functions, assuming either

1. boundedness of the pseudo-inverse of the diffusion matrix;

2. the existence of an approximating sequence of classical solutions to PDEs related to the HJB equation for the stochastic control problem of interest.

This author also notes the recent partially related work of [55] on proving existence and uniqueness of stochastic representations for solutions to degenerate elliptic and parabolic PDEs with constraints.

An alternative approach to the PDE method may combine the *compactification method* of [73] for the existence of (Markov) optimal controls with a *fixed point theorem* for the existence of Nash equilibria. A similar programme was carried out in
[18] for non-zero sum games of classic control in the setting of relaxed controls. However, the approach requires some topological and geometric properties to be satisfied in order to utilise the aforementioned fixed point theorem. Such properties are quite difficult to verify in general, and were obtained for the games in [18] via an intermediary PDE formulation for the associated value functions. In the absence of such a PDE formulation, the compactification method appears to be more difficult than the previously proposed approach using PDEs only. It should also be noted that the compactification method only furnishes the existence of a Nash equilibrium and does not characterise the optimal strategies (similar to Section 3.4 of Chapter 3).

7.7.2 Closing remarks for the thesis

This chapter studied a generalisation of the classic control problem in Chapter 3 within a stochastic setting and with another decision maker. This new agent utilised strategies which, just as in Chapters 4 and 5, are generalisations of the kind found in optimal stopping problems. While it was possible to obtain fairly general and satisfactory results for the single agent optimisation problems of Chapters 3–5, it was difficult to get similar results for the games studied thereafter. The cancellable contract for difference could be described in purely financial terms and the resulting zero-sum game was studied in Chapter 6 within a very general setting. On the other hand, differing financial and physical objectives of the system and battery operators led to a non-zero sum game for the balancing services contract. In this case, one needs to analyse two value functions simultaneously and also ensure the controlled state variables are well-defined. This made the problem much more difficult than the one in Chapter 6, and only partial results could be obtained. Nevertheless, this thesis has exemplified the utility of stochastic control theory in analysing optimisation problems within local electricity markets, and this author hopes that it spurs on further research in theory and applications.
Bibliography


Appendix A

Appendix to Chapter 3

A.1 Supplementary proofs

Proposition A.1.1 ([57], Problem III.7). Let \( f: [0, T] \times \mathbb{R} \times U \) be a measurable function and of class \( C^1 \) in \( x \) and \( u \) for every \( t \). Suppose \( 0 \in U \), \( |f(t, 0, 0)| \leq C \), \( |f_x(t, x, u)| \leq C(1 + |u|) \) and \( |f_u(t, x, u)| \leq C \) for some (positive) constant \( C \). Then for all \( t \in [0, T] \), \( x, x' \in \mathbb{R} \) and \( u \in U \):

\[
|f(t, x, u)| \leq C(1 + |x| + |u|) \tag{A.1}
\]

\[
|f(t, x', u) - f(t, x, u)| \leq C|x' - x|(1 + |u|) \tag{A.2}
\]

Proof. Since \( f(t, x, u) \) is continuously differentiable in \( u \), by the Mean Value Theorem (Theorem V.19.6, [8]) there exists \( \hat{u} \in \text{int}(U) \) (where \( \text{int}(U) \) is the interior of the control set \( U \)) such that,

\[
f(t, x, u) - f(t, x, 0) = f_u(t, x, \hat{u})(u - 0)
\]

and taking absolute values then applying the (reverse) triangle inequality shows

\[
|f(t, x, u)| - |f(t, x, 0)| \leq |f_u(t, x, \hat{u})||u|.
\]

Using the inequality \( |f_u(t, x, u)| \leq C \) in the hypothesis gives,

\[
|f(t, x, u)| - |f(t, x, 0)| \leq C|u| \tag{A.3}
\]

Similarly, since \( f(t, x, u) \) is continuously differentiable in \( x \), there exists \( \hat{x} \in \mathbb{R} \) such that,

\[
f(t, x, 0) - f(t, 0, 0) = f_x(t, \hat{x}, 0)(x - 0)
\]
and also
\[ |f(t, x, 0)| - |f(t, 0, 0)| \leq |f_x(t, \hat{x}, 0)||x|. \]

Using the inequality \(|f_x(t, \hat{x}, u)| \leq C(1 + |u|)\) with \(u = 0\) and \(|f(t, 0, 0)| \leq C\) yields,
\[ |f(t, x, 0)| - C \leq C|x| \]
which is used in (A.3) to get (A.1). Finally, by the Mean Value Theorem there exists \(\hat{x} \in \mathbb{R}\) such that for \(x, x' \in \mathbb{R}\)
\[ f(t, x', u) - f(t, x, u) = f_x(t, \hat{x}, u)(x' - x) \]
and using the inequality \(|f_x(t, x, u)| \leq C(1 + |u|)\) gives (A.2).

**Lemma A.1.2** (Continuous dependence on initial conditions). Let \(x(\cdot)\) denote the state variable which evolves according to
\[ \frac{d}{dt}x(t) = f(t, x(t), u(t)), \quad t \in [s, T] \]  
(A.4)
where \(s \in [0, T]\) and \(f\) satisfies:
\[ |f(t, x, v) - f(t, y, v)| \leq C|x - y|, \quad \forall (t, x), (t, y) \in \bar{Q}, \ v \in U \]  
(A.5)
\[ |f(t, x, v)| \leq C(1 + |x|), \quad \forall (t, x) \in \bar{Q}, \ v \in U \]  
(A.6)
for some constant \(C > 0\).

i). Let \(u \in U_d(s, y)\) and \(x = x(\cdot; s, y)\) be the associated (unique) solution to (A.4) with initial condition \(x(s; s, y) = y\). Then, for every \(t \in [s, T]\),
\[ |x(t) - y| \leq C_2|t - s| \]  
(A.7)
where \(C_2\) is a constant depending on \(|y|\) and \(C\) in equation (A.6).

ii). Let \(u : [0, T] \to U\) be an arbitrary Lebesgue-measurable function, \(s_1, s_2 \in [0, T]\) and \(y_1, y_2 \in K\) with \(K \subset \mathbb{R}\) compact. For \(i \in \{1, 2\}\), let \(x^i = x(\cdot; s_i, y_i)\) denote the solution to (A.4) corresponding to the restriction of \(u(\cdot)\) to \([s_i, T]\) with initial condition \(x^i(s_i) = y_i\). Then these solutions satisfy
\[ |x^2(t) - x^1(t)| \leq K[|y_2 - y_1|^2 + |s_2 - s_1|^2]^{1/2}, \quad (s_1 \vee s_2) \leq t \leq T \]  
(A.8)
for some constant \(K\) which depends on \(T, C\) in equations (A.5) and (A.6), but not on the choice of control \(u(\cdot)\).
Proof. Fix any \( t \in [s, T] \). Let \( u \in \mathcal{U}_d(s, y) \) be arbitrary and express the solution \( x = x(\cdot; s, y) \) to (A.4) in integral form

\[
x(t) = y + \int_s^t f(r, x(r), u(r)) \, dr.
\]

(A.9)

From equation (A.6) and the triangle inequality one gets

\[
|f(t, x, u)| \leq C(1 + |x|) \leq C(1 + |y| + |x - y|)
\]

which is used with (A.9) to deduce

\[
|x(t) - y| \leq D + \int_s^t Cy(r) dr
\]

(A.10)

with \( D = (t-s)[C(1+|y|)] \). The condition (A.7) holds since, upon applying Gronwall’s inequality to (A.10),

\[
|x(t) - y| \leq D \exp (C(t - s)) \leq C_2 |t - s|.
\]

with \( C_2 := C(1 + |y|) \exp (CT) \).

Let \( \mathcal{O}_a \) (resp. \( \mathcal{C}_a \)) be the open (resp. closed) “ball” of radius \( a > 0 \) (cf. (3.28)) and \( u: [0, T] \to U \) be an arbitrary Lebesgue-measurable function. Take any \( s_1, s_2 \in [0, T] \) and suppose without loss of generality that \( s_2 > s_1 \). For \( i \in \{1, 2\} \) and \( y_i \in \mathcal{O}_a \), one obtains a solution \( x^i = x(\cdot; s_i, y_i) \) to (A.4) corresponding to the restriction of \( u(\cdot) \) to \([s_i, T] \), and this solution is an absolutely continuous function. One then readily verifies for every \( t \in [s_2, T] \) that \( x^1 \) satisfies

\[
x^1(t) = y_1 - y_1 + x^1(s_2) + \int_{s_2}^t f(r, x^1(r), u(r)) dr
\]

Therefore, writing the \( x^i \) in integrated form and applying elementary inequalities shows

\[
|x^2(t) - x^1(t)| \leq |y_2 - y_1| + \int_{s_2}^t |f(r, x^2(r), u(r)) - f(r, x^1(r), u(r))| dr
\]

\[
+ |x^1(s_2) - y_1|.
\]

(A.11)

Now, from equation (A.5) it is true that

\[
\int_{s_2}^t |f(r, x^2(r), u(r)) - f(r, x^1(r), u(r))| dr \leq \int_{s_2}^t C|x^2(r) - x^1(r)| dr.
\]

On the other hand, equation (A.7) shows that

\[
|x^1(s_2) - y_1| \leq K_1 |s_2 - s_1|
\]
where $K_1$ is a constant that depends on $C$ and $|y_1|$. Using these inequalities in (A.11) shows that

$$|x^2(t) - x^1(t)| \leq |y_2 - y_1| + \int_{s_2}^t C|x^2(r) - x^1(r)|dr + K_1|s_2 - s_1|$$

and an application of Gronwall’s inequality shows that on $C_a$:

$$|x^2(t) - x^1(t)| \leq [|y_2 - y_1| + K_1|s_2 - s_1|] \exp\left(C(t - s_2)\right)$$

$$\leq K\left[|y_2 - y_1|^2 + |s_2 - s_1|^2\right]^{1/2}$$

for some constant $K$ depending on $C, T$ and $y_1 \in O_a$. Recalling that $y_1 \in O_a$ and $u(\cdot)$ were arbitrary, the proof is completed by letting $a \to \infty$ in (3.28).

Proposition A.1.3. For every $s \in [0, T), y \in \mathbb{R}$ and Borel-measurable map $u: [0, T] \times \mathbb{R} \to U$, there exists a solution $x^\varepsilon$ to

$$\begin{cases} x^\varepsilon(s) = y; \\ dx^\varepsilon(t) = f\left(t, x^\varepsilon(t), u(t, x^\varepsilon(t))\right) dt + \sqrt{2\varepsilon} dW(t), & t \geq s. \end{cases}$$

(A.12)

Moreover, this solution is unique in probability law.

Proof. Suppose the drift $f^u$ can be expressed as

$$f^u(t, x) = b(t, x) + \sigma(t, x)\Theta^u(t, x)$$

where $b, \sigma: [0, T] \times \mathbb{R} \to \mathbb{R}$ are deterministic functions satisfying the \textit{Itô conditions} (see (A.14) and (A.15) below) with $\sigma$ bounded, and $\Theta^u: [0, T] \times \mathbb{R} \to \mathbb{R}$ is bounded and Borel-measurable. Then the existence of the solution $x^\varepsilon$ and its uniqueness in probability law follow from Theorems V.10.2 and V.10.3 of [57] respectively. This hypothesis is verified below.

For any Borel-measurable map $u: [0, T] \times \mathbb{R} \to U$, $f^u$ is defined as (cf. (3.2))

$$f^u(t, x) := -\frac{1}{\tau} (x - T^u - T^y u(t, x))$$

where $\tau > 0$, $T^u, T^y \in \mathbb{R}$ are constants. Let $\varepsilon > 0$ be given and define $\sigma, b$ and $\Theta^u$ as follows

$$\sigma(t, x) = \sqrt{2\varepsilon}, \quad b(t, x) = -\frac{1}{\tau} [x - T^u], \quad \Theta^u(t, x) = \frac{1}{\sqrt{2\varepsilon}} \cdot \frac{T^y}{\tau} u(t, x).$$

(A.13)
Then \( f^\nu(t, x) = b(t, x) + \sigma(t, x)\Theta^\nu(t, x) \). Now, for any \((t, x) \in [0, T] \times \mathbb{R}, \)

\[
|b(t, x)| = | - \frac{1}{T} [x - T\alpha] | \leq \frac{1}{T} (|x| + |T\alpha|) \\
\leq C_1 (1 + |x|) \quad (A.14)
\]

with \( C = \max \{ \frac{|T\alpha|}{T}, \frac{1}{T}, \sqrt{2\varepsilon} \} \). Thus \( b \) satisfies a linear growth condition in \( x \). From definition (A.13), the drift \( b \) also satisfies a Lipschitz condition since \( \forall x_1, x_2 \in \mathbb{R}: \)

\[
|b(t, x_1) - b(t, x_2)| = | - \frac{1}{T} [x_1 - T\alpha] - [ - \frac{1}{T} (x_2 - T\alpha)] | \\
= | - \frac{1}{T} (x_1 - x_2)| \\
\leq C |x_1 - x_2| \quad (A.15)
\]

The diffusion coefficient \( \sigma \) also satisfies the linear growth and Lipschitz property as \( b \), which means the Itô conditions have been verified. Finally, \( \Theta^\nu \) is Borel-measurable since it is a linear function of \( u \), which is Borel-measurable by definition. Moreover, \( \Theta^\nu \) is bounded since \( u \) takes values in the compact set \( U \). The conditions of Theorems V.10.2 and V.10.3 of [57] are therefore satisfied.

A.2 Approximate Dynamic Programming

A discrete-time dynamic programming procedure was used to approximate the solution to the deterministic control problem of Chapter 3.1. The implementation described below follows the computational scheme outlined by [13, pp. 16–19] and exemplified in [13, pp. 327–331]. Similar computational procedures are described in [15, pp. 280–287] and Appendix A of [7].

According to [7, p. 17], the method relies on the use of approximate value functions \( V_h \) for step-size \( h > 0 \) in the time-like variable \( t \), which satisfy point-wise in \( Q = (0, T) \times \mathbb{R}, \) a “finite increment approximate version” of the HJB equation (3.36). The corresponding optimal controls are suboptimal for the original problem. Convergence results of the approximate value functions and controls as \( h \downarrow 0 \) are presented in Chapter I, Section 7 and Appendix A of [7] for the infinite time horizon case.

Let \( x_{\min}, x_{\max} \in \mathbb{R} \) with \( x_{\max} > x_{\min} \). The \((t, x)\) plane is represented using a uniform grid (or mesh) of discrete values \( t = t_k \in [0, T] \) in the time-like variable, and \( x = x^j \in [x_{\min}, x_{\max}] \) in the spatial variable (indoor temperature). Given two positive
integers \( M \) and \( N \), the nodes on the grid are given by \((t_k, x_j)\) where \( t_k = k \cdot h \) for \( k = 0, 1, \ldots, M \) and \( x_j = x_{\text{min}} + j \Delta x \) for \( j = 0, 1, \ldots, N \). The spacing parameters \( h > 0 \) and \( \Delta x > 0 \) are constant and chosen so that \( T = M \cdot h \) and \( x_{\text{max}} = x_{\text{min}} + N \cdot \Delta x \).

At each time \( t_k = k \cdot h \) with \( k < M \), there is a “path” that the state variable can follow from one node \((t_k, x_{j_1})\) to another \((t_{k+1}, x_{j_2})\). These paths correspond to the evolution of the indoor temperature, and are dependent on the value of the control variable. Figure A.1 below illustrates the idea.

Figure A.1: Dynamic Programming Grid

Let \( \mathcal{U}_d^h(0, y) \) be the subclass of controls \( u(\cdot) \in \mathcal{U}_d(0, y) \) that are piecewise constant:

\[
u(t) = \sum_{k=0}^{M-1} u_k \mathbf{1}_{[t_k, t_{k+1})}(t), \quad t \in [0, T], \quad u_k \in U.
\]

Basic results on first order linear ordinary differential equations (for example, [1, p. 47]) show that the solution \( x(\cdot) \) to (3.1)–(3.2) corresponding to a given \( u(\cdot) \in \mathcal{U}_d^h(0, y) \) satisfies:

\[
x(t_{k+1}) = e^{-\frac{h}{\tau}} x(t_k) + \left(1 - e^{-\frac{h}{\tau}}\right) \left(T^a + T^g u_k\right), \quad k = 0, \ldots, M - 1.
\] (A.16)

Recall that the dynamic programming principle (Theorem 3.6.1) says that the value function for the continuous-time optimal control satisfies

\[
V(t_k, x_k) = \inf_{u \in \mathcal{U}_d(t_k, x_k)} \left\{ \int_{t_k}^{t_{k+1}} u(s) \tilde{S}(s) + \kappa_1 B(x(s)) ds + V(t_{k+1}, x_{k+1}) \right\}.
\] (A.17)

The integral in equation (A.17) can be approximated linearly as

\[
\int_{t_k}^{t_{k+1}} u(s) \tilde{S}(s) + \kappa_1 B(x(s)) ds \approx [u_k \tilde{S}_k + \kappa_1 B(x_k)] h
\] (A.18)

For continuous integrands this follows from a first-order Taylor expansion at the point \( t_k \). Note that this approximation is exact for the discretised control problem. For
notational ease, let \( x_k = x(t_k) \), \( u_k = u(t_k) \) and \( \tilde{S}_k = \tilde{S}(t_k) \) for \( k = 0, \ldots, M \). Define \( J_h \) as the following approximation to the performance index (3.4)

\[
J_h(0, y, u) = h \sum_{k=0}^{M-1} \left[ u_k \tilde{S}_k + \kappa_1 B(x_k) \right] + \psi(x_M)
\]

and define the corresponding value function \( V_h \) by

\[
V_h(0, y) = \inf_{U_h^0(0,y)} J_h(0, y, u).
\]

Using (A.17) above, the value function \( V_h \) satisfies the following recursive relationship for each \( k = 0, \ldots, M - 1 \):

\[
V_h(t_k, x_k) = \inf_{u \in U_h^k(t_k, x_k)} \left\{ [u_k \tilde{S}_k + \kappa_1 B(x_k)] h + V_h(t_{k+1}, x_{k+1}) \right\}.
\]

Suppose \( x^* \) is the discrete trajectory corresponding to an optimal control \( u^{h,*} \) for the discretised problem. For each \( k = 0, \ldots, M - 1 \), \( u_k^{h,*} := u^{h,*}(t_k) \) achieves the minimum on the right-hand side of the relation (A.21) over all fixed values \( u_k \in U \).

Suppose \( u^{h,*} \) can be represented by \( u^{h,*}(t_k) = u^{h,*}(t_k, x_k^*) \), where \( u^{h,*} : \{t_0, \ldots, t_M\} \times \mathbb{R} \) is a feedback control map which satisfies for each \( t_k \) and \( z \in \mathbb{R} \):

\[
\mathbf{u}_k^{h,*}(t_k, z) \in \arg\min_{v \in U} \left\{ [v \tilde{S}_k + \kappa_1 B(z)] h + V(t_{k+1}, z) \right\}.
\]

According to Section 3.6.3, an optimal control only takes values in the boundary \( \{0, 1\} \) of the control set. The continuous-time optimal control problem (3.5) can be solved approximately by determining the control law from (A.22) and the corresponding trajectory using (A.16).

Starting at time \( t_M = T \), the value function can be initialised to \( V(t_M, x) = V_h(t_M, x) = \psi(x) \) for all \( x \in \mathbb{R} \). At time \( t_{M-1} \), values \( V_h(t_{M-1}, x_{M-1}) \) can be computed for \( x_{M-1} = x^j, j = 0, \ldots, N \) using equation (A.21). The values for \( x_M \) which are needed for this computation can be calculated using equation (A.16). This process is repeated backward in time for \( k = M - 2, \ldots, 0 \) and an interpolation / extrapolation method is used to determine the value of \( V_h(t_{k+1}, x_{k+1}) \) if it had not already been computed. The values \( \mathbf{u}_k^{h,*}(t_k, x^j), V_h(t_k, x^j) \) for \( k = 0, \ldots, M - 1 \) and \( j = 0, \ldots, N \) are stored in a look-up table and an approximately optimal trajectory \( x^* \) starting from \( x_0^* = y \) can be found by using (A.16) and the look-up table forward in time.
Appendix B

Appendix to Chapter 4

B.1 Supplementary proofs

Lemma B.1.1. For each $i \in I$, let $U^i \in \mathcal{S}^2$ and $Y^i \in \mathcal{S}^2$ be defined as in equations (4.12) and (4.13) respectively. Let $\tau_n \in \mathcal{T}$ and $\iota_n: \Omega \to I$ be $\mathcal{F}_{\tau_n}$-measurable. Then,

\[ Y^i_{\tau_n} = \operatorname{ess sup}_{\tau \in \mathcal{T}} E \left[ \sum_{s=1}^{\tau-1} \Psi_i(s) + U^i_{\tau} \bigg| \mathcal{F}_{\tau} \right] \text{ on } [\tau_n, T]. \]  

(B.1)

Proof. For notational simplicity define $(\tilde{U}^i_t)_{t \in T}$ by $\tilde{U}^i_t = \sum_{s=0}^{t-1} \Psi_i(s) + U^i_t$. For any $i \in I$ and any time $s \leq t$, $\Psi_i(s)$ is $\mathcal{F}_t$-measurable and using this in equation (4.13) shows,

\[ Y^i_t = \operatorname{ess sup}_{\tau \in \mathcal{T}} E \left[ \sum_{s=1}^{\tau-1} \Psi_i(s) + U^i_{\tau} \bigg| \mathcal{F}_{\tau} \right] \]

(B.2)

Since $U^i, \Psi_i \in \mathcal{S}^2$, the Snell envelope of the process $(\sum_{s=0}^{t-1} \Psi_i(s) + U^i_t)_{t \in T}$ exists (cf. Proposition 4.3.1) and is denoted by $\tilde{Y}^i$. Furthermore, using equation (B.2), $\tilde{Y}^i$ satisfies

\[ \tilde{Y}^i_t = \operatorname{ess sup}_{\tau \in \mathcal{T}} E \left[ \tilde{U}^i_{\tau} \bigg| \mathcal{F}_{\tau} \right] = Y^i_t + \sum_{s=0}^{t-1} \Psi_i(s) \]  

(B.3)

In particular $\tilde{Y}^i$ is the smallest supermartingale which dominates $\tilde{U}^i$. Note that as $Y^i, \Psi_i \in \mathcal{S}^2$, the supermartingale property carries over to stopping times by Doob’s Optional Sampling Theorem (Theorem II.59.1 of [117]).
APPENDIX B. APPENDIX TO CHAPTER 4

Consider the process $\sum_{i \in I} 1_{\{\tau_n=i\}} \bar{Y}^i$ on $[\tau_n, T]$ and remember that the sum over $I$ is finite. Let $r, t \in \mathbb{T}$ be arbitrary times satisfying $r \leq t$. Note that the indicator function $1_{\{\tau_n=i\}}$ is non-negative, and each $1_{\{\tau_n=i\}}$ is $\mathcal{F}_{\tau_n}$-measurable and therefore $\mathcal{F}_r$-measurable on $\{\tau_n \leq r\}$. Using these observations together with the supermartingale property yields: almost surely,

$$E \left[ \sum_{i \in I} 1_{\{\tau_n=i\}} \bar{Y}_t^i \Bigg| F_r \right] 1_{\{\tau_n \leq r\}} = \sum_{i \in I} 1_{\{\tau_n=i\}} E \left[ \bar{Y}_t^i \Bigg| F_r \right] 1_{\{\tau_n \leq r\}} \leq \sum_{i \in I} 1_{\{\tau_n=i\}} 1_{\{\tau_n \leq r\}} \bar{Y}_r^i$$

This shows $\sum_{i \in I} 1_{\{\tau_n=i\}} \bar{Y}_t^i$ is a supermartingale on $[\tau_n, T]$. For each $i \in I$, the dominating property of the Snell envelope and non-negativity of $1_{\{\tau_n=i\}}$ leads to:

$$1_{\{\tau_n \leq t\}} 1_{\{\tau_n=i\}} \bar{Y}_t^i \geq 1_{\{\tau_n \leq t\}} 1_{\{\tau_n=i\}} \bar{U}_t^i$$

and summing over $i \in I$ then gives,

$$\bar{Y}_t^{\tau_n} := \sum_{i \in I} 1_{\{\tau_n=i\}} \bar{Y}_t^i \geq \sum_{i \in I} 1_{\{\tau_n=i\}} \bar{U}_t^i =: \bar{U}_t^{\tau_n} \text{ on } \{\tau_n \leq t\}.$$

The process $\bar{Y}_t^{\tau_n}$ is therefore a supermartingale dominating $\bar{U}_t^{\tau_n}$ on $[\tau_n, T]$. Similar arguments as above can be used to show that $\bar{Y}^{\tau_n}$ is the smallest supermartingale with this property, and is therefore the Snell envelope of $\bar{U}_t^{\tau_n}$. Proposition 4.3.1 leads to a representation for $\bar{Y}^{\tau_n}$ similar to (B.3), and the $\mathcal{F}_r$-measurability of the summation term leads to equation (B.1).

Lemma B.1.2. Let $\alpha^* = (\tau_n^*, \iota_n^*)_{n \geq 0}$ be the sequence defined in (4.14). Suppose that Assumption 4 holds for the switching costs. Then $\alpha^* \in \mathcal{A}_{t,i}$.

Proof. The times $\{\tau_n^*\}_{n \geq 0}$ are non-decreasing by definition, $\tau_0^* = t$ and each $\tau_n^* \in \mathcal{T}_i$ since $U^i$ and $Y^i$ are adapted for every $i \in I$. Corollary II-1-4 of [104] states that for any adapted process $Z$ and stopping time $\tau$, $Z_\tau$ is $\mathcal{F}_\tau$-measurable. The sets $A_{j_n}^{i_n-1}$ in equation (4.14) are therefore $\mathcal{F}_{\tau_n^*}$-measurable sets which means the modes $\{\iota_n^*\}_{n \geq 0}$ are also $\mathcal{F}_{\tau_n^*}$-measurable. Furthermore, $\iota_n^* \neq \iota_n^* + 1$ almost surely for $n \geq 0$.

The last thing to verify is $P \left( \{\tau_n^* < T\} \cap \{\tau_n^* = \tau_{n+1}^*\} \right) = 0$ for $n \geq 1$. Assume contrarily that for some $n \geq 1$, the event $\{\tau_n^* < T\} \cap \{\tau_n^* = \tau_{n+1}^*\}$ has positive probability (recall $\tau_{n+1}^* \geq \tau_n^*$). Using the definition for $\tau_n^*$ and $\tau_{n+1}^*$, $P$-almost surely:

$$Y_{\tau_n^*}^{\tau_{n+1}^*} = U_{\tau_n^*}^{\tau_{n+1}^*}, \quad Y_{\tau_n^*}^{\tau_{n+1}^*} = U_{\tau_n^*}^{\tau_{n+1}^*}.$$
APPENDIX B. APPENDIX TO CHAPTER 4

By definition of $i^*$ and $i^*_{n+1}$, it is also true that:

\[
Y_{\tau_n^{i-1}} = -\gamma_{i^*_{n-1},i^*_n} (\tau_n^*) + Y_{\tau_n^{i^*_n}}^{i^*_n} \quad \text{on } \{\tau_n^* < T\} \cap \{\tau_n^* = \tau_{n+1}^*\}
\]

(B.4)

Let $H$ be the event defined by

\[
H = \{i^*_{n-1} = i, i^*_n = j, i^*_{n+1} = k\} \cap \{\tau_n^* < T\} \cap \{\tau_n^* = \tau_{n+1}^*\}
\]

where $i, j, k \in \mathbb{I}$ are three modes satisfying $i \neq j$ and $j \neq k$. Note that $i^*_{n+1}$ is $\mathcal{F}_{\tau_n^*}$-measurable on the event $\{\tau_n^* = \tau_{n+1}^*\}$. Substituting for $Y_{\tau_n^{i^*_n}}^{i^*_n}$ in (B.4) and using condition (4.8) for the switching costs gives,

\[
Y_{\tau_n^{i^*_n}} = -\gamma_{i,j} (\tau_n^*) - \gamma_{j,k} (\tau_n^*) + Y_{\tau_n^{k}}^{k} < -\gamma_{i,k} (\tau_n^*) + Y_{\tau_n^{k}}^{k} \quad \text{on } H
\]

The previous arguments have just shown

\[-\gamma_{i,k} (\tau_n^*) + Y_{\tau_n^{k}}^{k} > -\gamma_{i,j} (\tau_n^*) + Y_{\tau_n^{j}}^{j} = \max_{l \neq i} \{-\gamma_{i,l} (\tau_n^*) + Y_{\tau_n^{l}}^{l}\} \quad \text{on } H
\]

which is a contradiction for every $k \in \mathbb{I}$. Since $i \neq j$ and $j \neq k$ were arbitrary modes, for $n \geq 1$ it is true that $P(\{\tau_n^* < T\} \cap \{\tau_n^* = \tau_{n+1}^*\}) = 0$. \(\square\)

### B.2 Approximate dynamic programming for optimal switching problems

Approximate dynamic programming with least squares Monte Carlo regression (LSMC) is a popular and powerful probabilistic numerical method for solving optimal switching problems. It was introduced in [20] as an extension of other algorithms for solving optimal stopping problems such as [88, 130, 131]. The main aspects of this algorithm are summarised in this section.

#### B.2.1 Probabilistic Setup

Let $0 < T < \infty$ be a given integer and define a sequence of times $T := \{0, \ldots, T\}$. Let $X = \{X_t\}_{t=0}^T$ be a sequence of $\mathbb{R}^d$-valued random variables with respect to a given measurable space $(\Omega, \mathcal{F})$. Let $\mathbb{F} = \{\mathcal{F}_t\}_{t=0}^T$ be the filtration on this space generated by $X$. Suppose there exists a family of probability measures $\{P_{sx}: s \in \mathbb{T}, x \in \mathbb{R}^d\}$ such that $X$ is a Markov chain with respect to $(\Omega, \mathcal{F}, \mathbb{F}, P_{sx})$ and $P_{sx}(\{X_s = x\}) = 1$. When $s = 0$ the notation $P_x \equiv P_{0x}$ is used.
B.2.2 Discrete-time optimal switching

Let \( \mathbb{I} = \{1, 2, \ldots, q\} \) denote the finite set of modes. Admissible strategies \( \alpha = (\tau_n, i_n)_{n \geq 0} \) for the optimal switching problem should satisfy the properties outlined in Definition 4.2.1. Recall that \( \mathcal{A}_i \) denotes the class of admissible strategies which start in mode \( i \in \mathbb{I} \) at time 0. The performance index for the optimal switching problem is given by (cf. (4.5)):

\[
J(\alpha; 0, i) := \mathbb{E}_x \left[ \sum_{t=0}^{T-1} \Psi_{i_t}(t, X_t) + \Gamma_{i_T}(X_T) - \sum_{n \geq 1} \gamma_{i_{n-1}, i_n}(\tau_n, X_{\tau_n}) 1_{[\tau_n < T]} \right], \quad \alpha \in \mathcal{A}_i \quad (B.5)
\]

where \( \Psi_i: \mathbb{T} \times \mathbb{R}^d \to \mathbb{R}, \Gamma_i: \mathbb{R}^d \to \mathbb{R} \) and \( \gamma_{i,j}: \mathbb{T} \times \mathbb{R}^d \to \mathbb{R} \) are given deterministic functions for \( i, j \in \mathbb{I} \). The assumptions of Section 4.2.3 are assumed to hold when these functions are evaluated along the path of \( X \).

The value function \( V \) for the optimal switching problem is given by

\[
V(0, i) = \sup_{\alpha \in \mathcal{A}_i} J(\alpha; 0, i).
\]

Under appropriate assumptions, it has the following representation (cf. Section 4.5):

\[
V(T, i) = \Gamma_i(X_T) \quad (B.6)
\]

\[
V(t, i) = \max_{j \in \mathbb{I}} \left\{ -\gamma_{i,j}(t, X_t) + \mathbb{E}_x \left[ \Psi_j(t, X_t) + V(t+1, j) \mid \mathcal{F}_t \right] \right\}, \quad t = T - 1, \ldots, 0. \quad (B.7)
\]

An optimal strategy can then be obtained in feedback form along the path of \( X \) as follows: for each \( i \in \mathbb{I} \) and \( t = T - 1, \ldots, 0 \):

\[
u^*(t, i) = \arg \max_{j \in \mathbb{I}} \left\{ -\gamma_{i,j}(t, X_t) + \mathbb{E}_x \left[ \Psi_j(t, X_t) + V(t+1, j) \mid \mathcal{F}_t \right] \right\}. \quad (B.8)
\]

B.2.3 The least squares Monte Carlo regression procedure

Regression-based Monte Carlo algorithms are often used to obtain approximate solutions to optimal stopping problems. The approach has become quite popular for multidimensional optimal stopping problems and well known publications such as [22, 88, 130, 131] demonstrate its utility for valuing American-style derivative securities.
on multiple assets. An introduction to the numerical method with further references can be found in Chapter 8 of [62].

At the start of the algorithm, $N_S$ sample paths of the Markov chain $X$ are simulated with initial value $X_0 = x$. The algorithm initialise the value function according to equation (B.6), $V^{(l)}(T, i) = \Gamma_i(X_T^{(l)})$, for each simulated value of $X_T^{(l)}$ for $l = 1, \ldots, N_S$. It then proceed to the next iteration at time $t = T - 1$.

Suppose an approximation to $V(t + 1, i)$ has been computed at $t = T - 1, \ldots, 0$ for every $i \in I$ along the simulated path $l = 1, \ldots, N_S$. Denote this approximation by \( \hat{V}_{l}^{(t+1, i)} \), noting that it is exact at $t = T - 1$. The conditional expectations appearing in equation (B.7) are computed in the following way. Observe that the conditional expectation in (B.7) is an element of the separable Hilbert space $L^2(F_t)$. It therefore has an orthonormal expansion in terms of appropriately chosen basis functions on $L^2(F_t)$, \( \{B_k\}_{k=1}^{\infty} \) – see Chapter 24 of [120], for example. On the other hand, by the Markov property of $X$ the following holds \( P_x\text{-a.s.:} \)

$$E_x[\Psi_j(t, X_t) + V(t + 1, j) | X_t = y] = \sum_{k=1}^{\infty} \alpha_k B_k(y; t, j),$$  \hspace{1cm} (B.10)

where the $\alpha_k$ are real-valued coefficients. An approximation to the conditional expectation can be obtained by truncating the infinite sum on the right-hand side of (B.10) to a finite number of terms $N_b$:

$$E_x[\Psi_j(t, X_t) + V(t + 1, j) | X_t = y] \approx \sum_{k=1}^{N_b} \alpha_k B_k(y; t, j)$$

$$=: \hat{E}(y; t, j).$$  \hspace{1cm} (B.11)

Utilising the precomputed sample paths, the basis functions \( \{B_k(y^{(l)}; t, j)\}_{k=1}^{N_b} \) are evaluated at $X_t^{(l)} = y^{(l)}$, $l = 1, \ldots, N_S$ for a fixed $j \in I$. Values for a single set of coefficients $\{\alpha_k\}_{k=0}^{N_b}$ are obtained by an empirical regression of the values $\Psi_j(t, y^{(l)}) + V^{(l)}(t + 1, j)$ against $\{B_k(y^{(l)}; t, j)\}_{k=0}^{N_b}$. The $N_S$ fitted values $\hat{E}(y^{(l)}; t, j)$ from the regression are obtained following the linear combination in equation (B.11). The optimal
mode along each path \( l = 1, \ldots, N_S \) can then be found by:

\[
    u^{(l),*}(t, i) = j^{(l),*} := \arg \max_{j \in \mathbb{I}} \left\{ -\gamma_{i,j}(t, y^{(l)}) + \hat{E}(y^{(l)}; t, j) \right\}
\]  

(B.12)

There are two popular update rules which can be used to set the value for \( \hat{V}^{(l)}(t, i) \). The Tsitsiklis-Van Roy update rule (TvR), based on the paper \([130]\), uses the fitted values from the regression to update the approximate value function:

\[
    \hat{V}^{(l)}(t, i) = -\gamma_{i,j^{(l),*}}(t, y^{(l)}) + \hat{E}(y^{(l)}; t, j^{(l),*}) .
\]  

(B.13)

Alternatively, the Longstaff-Schwartz rule (LS), based on the paper \([88]\), only uses the regression to determine the optimal mode. The value function’s approximation is determined using its value from the previous iteration:

\[
    \hat{V}^{(l)}(t, i) = -\gamma_{i,j^{(l),*}}(t, y^{(l)}) + \hat{V}^{(l)}(t+1, j^{(l),*}) .
\]  

(B.14)

This procedure for evaluating the approximate value functions is continued from \( t = T - 1 \) until \( t = 0 \). The value function of the optimal switching problem is then estimated by the average of the path-wise approximate value functions:

\[
    V(0, i) \approx \frac{1}{N_S} \sum_{l=1}^{N_S} \hat{V}^{(l)}(0, i) .
\]  

(B.15)

Remark B.2.1. The update rule used for the approximate value function can have a significant impact on the results of the algorithm. It was suggested in \([40, p. 1404]\) that in comparison to TvR (cf. equation (B.13)), LS (cf. equation (B.14)) may give a better approximation of the optimal stopping policy—particularly when the process is non-stationary and either the time horizon is finite or there is slow convergence to a stationary distribution.

Remark B.2.2. The number and type of basis functions used to estimate the continuation value (conditional expectation), as well as the number of Monte Carlo sample paths are also important considerations of the regression-based algorithm. Empirical evidence in papers such as \([21, 88, 122]\) suggest that the number of basis functions \( N_b \) can be as low as 2 and still give good results.

Orthogonal families such as the Legendre, Laguerre and Hermite polynomials are popular choices for the basis functions. However, an empirical analysis carried out
in [122, p. 153] showed that, despite the increased computational cost, there is little
to no improvement in the precision of the results obtained using an orthogonal set of
basis functions when compared to the simpler monomials.

In [63] it was shown for American options that the number of sample paths required
to achieve worst-case convergence of the continuation value estimators grows at least
exponentially with respect to the number of basis functions when using schemes such
as Tvr or LS. A similar conclusion is reached by [40, p. 1411], who also admits that
Tvr generally requires less sample paths for convergence but at the expense of a less
accurate estimator of the optimal stopping rule.
Appendix C

Appendix to Chapter 5

C.1 Admissibility of the candidate optimal strategy

This section verifies $\alpha^* \in \mathcal{A}_{t,i}$ (cf. Definition 5.2.1). One readily verifies (by right-continuity) that $\{\tau_n^*\}_{n \geq 0} \subset \mathcal{T}$ is non-decreasing with $\tau_0^* = t$, and each $\iota_n^*$ is an $\mathcal{F}_{\tau_n^*}$-measurable $\mathbb{I}$-valued random variable with $\iota_0^* = i$ and $\iota_n^* \neq \iota_{n+1}^*$ for $n \geq 0$. The remaining properties are established in a number of steps, beginning with the following lemma on the switching times.

Lemma C.1.1. Let $\{\tau_n^*\}_{n \geq 0}$ be the switching times defined in equation (5.26) of Theorem 5.5.2. Then these times satisfy the finiteness condition $P(\{\tau_n^* < T, \forall n \geq 0\}) = 0$.

Proof. First, note that by definition the sequence $\{\tau_n^*\}_{n \geq 0}$ is non-decreasing. Furthermore, the sequence is strictly increasing in the following sense: $P(\{\tau_n^* < T \cap \{\tau_n^* = \tau_{n+1}^*\}) = 0$ for $n \geq 1$. To see this, notice that by definition of the stopping times $\tau_n^*$ and $\tau_{n+1}^*$ (and right continuity of $Y$, $\gamma_{i,j}$), P-a.s.:

$$Y_{\tau_{n-1}^*}^{\iota_{n-1}^*} = \max_{k \neq \iota_{n-1}^*} \left\{-\gamma_{\iota_{n-1}^*,k} (\tau_n^*) + Y_{\tau_n^*}^k \right\} \text{ on } \{\tau_n^* < T \cap \{\tau_n^* = \tau_{n+1}^*\}$$

$$Y_{\tau_{n+1}^*}^{\iota_{n+1}^*} = \max_{l \neq \iota_{n+1}^*} \left\{-\gamma_{\iota_{n+1}^*,l} (\tau_{n+1}^*) + Y_{\tau_{n+1}^*}^l \right\} \text{ on } \{\tau_n^* < T \cap \{\tau_n^* = \tau_{n+1}^*\}$$

Use the definition of $\iota_n^*$ and $\iota_{n+1}^*$ to get,

$$Y_{\tau_n^*}^{\iota_n^*} = Y_{\tau_n^*}^{\iota_n^*} - \gamma_{\iota_{n-1}^*,\iota_n^*}^n (\tau_n^*) \text{ on } \{\tau_n^* < T \cap \{\tau_n^* = \tau_{n+1}^*\}$$

$$Y_{\tau_{n+1}^*}^{\iota_{n+1}^*} = Y_{\tau_{n+1}^*}^{\iota_{n+1}^*} - \gamma_{\iota_{n}^*,\iota_{n+1}^*}^n (\tau_n^*) \text{ on } \{\tau_n^* < T \cap \{\tau_n^* = \tau_{n+1}^*\}$$

240
then substitute for \( Y_{i_n^*}^{n+1} \) and use Assumption 6 to get the following contradiction

\[
Y_{i_n^*}^{n+1} - \gamma_{i_{n-1}^* i_n^*} (\tau_n^*) < Y_{i_n^*}^{n+1} - \gamma_{i_{n-1}^* i_{n+1}^*} (\tau_n^*), \quad P \text{ a.s. on } \{ \tau_n^* < T \} \cap \{ \tau_n^* = \tau_{n+1}^* \}
\]

Therefore \( P(\{ \tau_n^* < T \} \cap \{ \tau_n^* = \tau_{n+1}^* \}) = 0 \) for \( n \geq 1 \).

Following the argument in [70, pp. 192–193], assume contrarily that \( P(\{ \tau_n^* < T, \forall n \geq 0 \}) > 0 \). By definition of \( \{ \tau_n^* \}_{n \geq 0} \), this would imply that

\[
\left\{ Y_{i_{n+1}^*}^{n+1} - \gamma_{i_{n+1}^*} (\tau_{n+1}^*), i_n^* \neq i_{n+1}^*, \forall n \geq 0 \right\} > 0.
\] (C.1)

Since \( \mathbb{I} \) is finite, this implies the existence of a sequence of elements \( i_0, \ldots, i_k, i_0 \in \mathbb{I} \) satisfying \( i_0 \neq i_1 \) and a subsequence \( (n_q)_{q \geq 0} \) such that:

\[
P\left( \left\{ Y_{i_q^*}^{i_{n+1}^*} = Y_{i_{n+1}^*}^{i_{n+1}^*} - \gamma_{i_{n+1}^*} (\tau_{n+1}^*), l = 0, \ldots, k; \ (i_{k+1} = i_0), \forall q \geq 0 \right\} > 0 \right).
\]

Now, as the times are bounded above by \( T \), the subsequence of times \( \{ \tau_{n_q}^* \}_{q \geq 0} \) is convergent and \( \lim_{q} \tau_{n_q}^* \rightarrow \tau^* \leq T \). Taking the limit with respect to \( q \) in equation (C.1) and using the quasi-left-continuity of the switching costs then gives

\[
P\left( \left\{ Y_{i_q^*}^{i_{n+1}^*} = Y_{i_{n+1}^*}^{i_{n+1}^*} - \gamma_{i_{n+1}^*} (\tau_{n+1}^*), l = 0, \ldots, k; \ (i_{k+1} = i_0) \right\} > 0 \right).
\]

Repeated substitution of \( Y_{i_q^*}^{i_{n+1}^*} \) for \( l = 0, \ldots, k + 1 \) and recalling \( i_{k+1} = i_0 \) yields

\[
P(\{ \gamma_{i_{0} i_1} (\tau^*) + \ldots + \gamma_{i_k i_0} (\tau^*) = 0 \}) > 0,
\]

which is a contradiction due to Assumption 6 on the switching costs, and shows \( P(\{ \tau_n^* < T, \forall n \geq 0 \}) = 0 \).

\[\square\]

### C.1.1 An alternative representation for the cumulative switching cost

The rest of this section is devoted to verifying condition (5.3) for the strategy \( \alpha^* \).

Recall that the cumulative cost of switching \( n \geq 1 \) times is given by,

\[
C_n^{\alpha^*} = \sum_{k=1}^{n} \gamma_{i_{k-1}^* i_k^*} (\tau_k^*) I_{\{ \tau_k^* < T \}}
\]
Since the switching costs satisfy $\gamma_{i,j} \in S^2$ for every $i, j$ in the finite set $\mathbb{I}$, $C_n^\alpha \in L^2$ for every $n \geq 1$. We define a sequence

$$N_n^* := \sum_{k=1}^n 1_{\{\tau_k^* < T\}}, \quad n = 1, 2, \ldots$$

which we use to rewrite the expression for $C_n^\alpha$ as follows:

$$C_n^\alpha = \sum_{k=1}^{N_n^*} \gamma_{i_{k-1},i_k}(\tau_k^*). \quad (C.2)$$

The following proposition gives an alternative representation of $C_n^\alpha$ in terms of the processes $Y^1, \ldots, Y^m$ and their Meyer decomposition with random superscripts (cf. Lemma 5.5.1).

**Proposition C.1.2.** Let $\alpha^* = (\tau^*_n, i^*_n)_{n \geq 0} \in \mathcal{A}_{t,i}$ be the switching control strategy defined in equation (5.26) of Theorem 5.5.2 and let $u^*$ be the associated mode indicator function. Then $C_n^\alpha$, the cumulative cost of switching $n \geq 1$ times under $\alpha^*$, satisfies

$$C_n^\alpha = Y^*_{i_{N_n^*}} - Y^*_{i_0} + \int_{\tau_0^*}^{\tau_{N_n^*}} \psi_{i^*_t}(s) ds - \sum_{k=1}^{N_n^*} (M^*_{i_k^* - 1} - M^*_{i_{k-1}^* - 1}) \quad \text{P-a.s.} \quad (C.3)$$

where $M^*_{i_k}$, $k \geq 0$, is the martingale component of the Meyer decomposition (5.22) in Lemma 5.5.1.

**Proof.** By definition of the strategy $\alpha^*$ (cf. (5.26)), optimality of the time $\tau_n^*$ and the definition of $i_n^*$, for $n \geq 1$ the cost of switching at $\tau_n^*$ is,

$$\gamma_{i_{n-1},i_n^*}(\tau_n^*) 1_{\{\tau_n^* < T\}} = \left(Y^*_{i_n^*} - Y^*_{i_{n-1}^*}\right) 1_{\{\tau_n^* < T\}} \quad \text{P-a.s.} \quad (C.4)$$

Therefore, from equation (C.2) and (C.4) the cost of the first $n$ switches can be rewritten as,

$$C_n^\alpha = \sum_{k=1}^{N_n^*} \left(Y^*_{i_k^*} - Y^*_{i_{k-1}^*}\right) \quad \text{P-a.s.} \quad (C.5)$$

Now, Lemma 5.5.1 proved that the following Meyer decomposition holds for $k \geq 0$ (cf. equation (5.22)):

$$Y^*_{i_k^*} + \int_0^t \psi_{i_k^*}(s) ds = M^*_{i_k^*} - A^*_{i_k^*}, \quad \text{P-a.s.} \quad \forall \tau_k^* \leq t \leq T. \quad (C.6)$$
where, on \([\tau_k^*, T]\), \(M^{i^k}\) is a uniformly integrable càdlàg martingale and \(A^{i^k}\) is a predictable, continuous and increasing process. The Meyer decomposition is used to rewrite equation (C.5) for the cumulative switching costs as follows: P-a.s.,

\[
C_n^\tau = \sum_{k=1}^{N_k^*} (M^{i^k}_{\tau_k^*} - M^{i^k}_{\tau_k^* - 1}) - \sum_{k=1}^{N_k^*} (A^{i^k}_{\tau_k^*} - A^{i^k}_{\tau_k^* - 1}) \\
- \sum_{k=1}^{N_k^*} \left( \int_{0}^{\tau_k^*} \psi_{i^k}(s) ds - \int_{0}^{\tau_k^*} \psi_{i^k-1}(s) ds \right).
\]

(C.7)

The first summation term in equation (C.7) can be rewritten as:

\[
\sum_{k=1}^{N_k^*} (M^{i^k}_{\tau_k^*} - M^{i^k}_{\tau_k^* - 1}) = M^{i^k}_{N_k^*} - M^{i^k}_{0} - \sum_{k=1}^{N_k^*} (M^{i^k}_{\tau_k^* - 1} - M^{i^k}_{\tau_k^* - 1})
\]

(C.8)

For every \(k \geq 0\), by the definition of \(\tau_{k+1}^*\) and property 4 of Proposition 5.3.2, the process \(\left( Y^{i^k}_t + \int_{0}^{t} \psi_{i^k}(s) ds \right) \) is a martingale P-a.s. for every \(\tau_k^* \leq t \leq \tau_{k+1}^*\). By using the Meyer decomposition (C.6), one observes that \(\forall k \geq 0, A^{i^k}_{\tau_k^*}\) is constant P-a.s. \(\forall \tau_k^* \leq t \leq \tau_{k+1}^*\). In particular,

\[
\forall k \geq 1, \quad A^{i^k}_{\tau_k^* - 1} = A^{i^k}_{\tau_k^*} \quad \text{P-a.s.}
\]

and the summation term in (C.7) with respect to \(A^{i^k}_{\tau_k^* - 1}\) is simplified as follows,

\[
\sum_{k=1}^{N_k^*} (A^{i^k}_{\tau_k^*} - A^{i^k}_{\tau_k^* - 1}) = \sum_{k=1}^{N_k^*} (A^{i^k}_{\tau_k^*} - A^{i^k}_{\tau_k^* - 1}) = A^{i^k}_{N_k^*} - A^{i^k}_{0} \quad \text{P-a.s.}
\]

(C.9)

By writing out the terms and using the definition of the mode indicator function \(u^*\), the third summation term in (C.7) is simplified as follows: P-a.s.,

\[
- \sum_{k=1}^{N_k^*} \left( \int_{0}^{\tau_k^*} \psi_{u^k}(s) ds - \int_{0}^{\tau_k^*} \psi_{u^k-1}(s) ds \right) \\
= \int_{0}^{\tau_1^*} \psi_{u^0}(s) ds + \sum_{k=1}^{N_k^* - 1} \int_{\tau_k^*}^{\tau_{k+1}^*} \psi_{u^k}(s) ds - \int_{0}^{\tau_k^*} \psi_{u^k}(s) ds \\
= \int_{0}^{\tau_1^*} \psi_{u^0}(s) ds + \sum_{k=1}^{N_k^* - 1} \int_{\tau_k^*}^{\tau_{k+1}^*} \psi_{u^k}(s) ds - \int_{0}^{\tau_k^*} \psi_{u^k}(s) ds \\
= \int_{0}^{\tau_1^*} \psi_{u^0}(s) ds + \int_{\tau_1^*}^{\tau_{N_k^*}} \psi_{u^k}(s) ds - \int_{0}^{\tau_k^*} \psi_{u^k}(s) ds
\]

(C.10)
Substituting equations (C.8), (C.9), and (C.10) into equation (C.7) for the cumulative switching cost yields,

\[
C^{n*} = M_{\tau_{N_n}}^{N_n} - M_{\tau_0}^{N_n} - \sum_{k=1}^{N_n} \left( M_{\tau_k}^{k-1} - M_{\tau_k}^{k-1} \right) - A_{\tau_{N_n}}^{N_n} + A_{\tau_0}^{N_n} \\
+ \int_{\tau_0}^{\tau_1} \psi\iota_0(s)ds + \int_{\tau_1}^{\tau_{N_n}} \psi\iota_1(s)ds - \int_{\tau_1}^{\tau_{N_n}} \psi\iota_1(s)ds \\
= Y_{\tau_{N_n}}^{N_n} - \left( Y_{\tau_0}^{0} + \int_{\tau_0}^{\tau_1} \psi\iota_0(s)ds \right) + \int_{\tau_1}^{\tau_{N_n}} \psi\iota_1(s)ds + \int_{\tau_1}^{\tau_{N_n}} \psi\iota_1(s)ds \\
- \sum_{k=1}^{N_n} \left( M_{\tau_k}^{k-1} - M_{\tau_k}^{k-1} \right)
\]

\[
= Y_{\tau_{N_n}}^{N_n} - Y_{\tau_0}^{0} + \int_{\tau_0}^{\tau_{N_n}} \psi\iota_1(s)ds - \sum_{k=1}^{N_n} \left( M_{\tau_k}^{k-1} - M_{\tau_k}^{k-1} \right) \quad \mathbb{P} - \text{a.s.}
\]

The second and third equations were obtained by using the Meyer decomposition (C.6) and the definition of \( u^* \) respectively.

Remark C.1.1. There is a close resemblance between equation (C.3) and the reflected backward stochastic differential equation (RBSDE) (5.13) of [70], which is not surprising given the connection between RBSDEs and Snell envelopes. The RBSDE in [70] has a stochastic integral formed by integrating a predictable process in \( \mathcal{M}^2 \) with respect to a standard Brownian motion, whereas (C.3) above has a martingale summation term. If \( \mathcal{F} \) is the completed natural filtration of a standard Brownian motion, as assumed in [70], then it satisfies a predictable representation property for local martingales (and therefore uniformly integrable martingales). It is then possible to obtain a stochastic integral in equation (C.3) similar to the one in [70]. This integral representation can be extended to more general filtrations, as explained in the survey [30] and book [76].

C.1.2 Convergence of the family of cumulative switching costs

A discrete-parameter martingale.

Admissibility of the strategy \( \alpha^* \) depends on the asymptotic behaviour of the martingale summation term \( \sum_{k=1}^{N_n} \left( M_{\tau_k}^{k-1} - M_{\tau_k}^{k-1} \right) \) as \( n \to \infty \). For \( k \geq 0 \), define an
\(\mathcal{F}_{\tau_k}\)-measurable random variable \(\xi_k\) by,

\[
\xi_k := \begin{cases} 
M_{\tau_k}^{k-1} - M_{\tau_{k-1}}^{k-1} & \text{on } k \geq 1 \text{ and } \{\tau_k^* < T\}, \\
0 & \text{otherwise}.
\end{cases}
\]  

(C.11)

Note that the limit \(\xi_\infty\) is a well-defined \(\mathcal{F}_T\)-measurable random variable which satisfies

\[
\xi_\infty := \lim_k \xi_k = \begin{cases} 
0, & \text{on } \{N(\alpha^*) < \infty\}, \\
0, & \text{a.s. on } \{N(\alpha^*) = \infty\}.
\end{cases}
\]

where the second line holds since \(M^i, i \in \mathbb{I}\), is quasi-left-continuous, and the switching times \(\{\tau_k^*\}_{k \geq 1}\) are almost surely (strictly) increasing toward \(T\) on \(\{N(\alpha^*) = \infty\}\) (cf. proof of Lemma C.1.1). In this case set \(\tau_\infty^* := u^*_T\).

Some properties of the sequence \(\{\xi_k\}_{k \geq 0}\) defined in (C.11) are described in the next few lines. First, notice that by Proposition 5.3.3, \(M^i \in \mathcal{S}^2\) for \(i \in \mathbb{I}\) so that, since the set \(\mathbb{I}\) is finite, the sequence \(\{\xi_k\}_{k \geq 0}\) is in \(L^2\). The properties of these square-integrable martingales and conditional expectations also show that for \(n \geq 1\):

\[
E \left[ \sum_{k=1}^n (\xi_k)^2 \right] = E \left[ \sum_{k=1}^n (M_{\tau_k}^i - M_{\tau_{k-1}}^i)^2 1_{\{\tau_k^* < T\}} \right] \\
\leq \sum_{i=1}^m \sum_{k=1}^n E \left[ (M_{\tau_k}^i - M_{\tau_{k-1}}^i)^2 \right] \\
= \sum_{i=1}^m \sum_{k=1}^n E \left[ (M_{\tau_k}^i)^2 - 2 \cdot M_{\tau_{k-1}}^i \cdot E[M_{\tau_k}^i \mid \mathcal{F}_{\tau_{k-1}}^*] + (M_{\tau_{k-1}}^i)^2 \right] \\
= \sum_{i=1}^m \sum_{k=1}^n E \left[ (M_{\tau_k}^i)^2 - (M_{\tau_{k-1}}^i)^2 \right] \\
\leq \sum_{i=1}^m E \left[ \sup_{0 \leq s \leq T} |M_s^i|^2 \right] \leq 4 \cdot m \cdot \max_{i \in \mathbb{I}} E \left[ (M_T^i)^2 \right].
\]  

(C.12)

Finally, almost surely for \(1 \leq k \leq N_\infty^*\),

\[
E[\xi_k \mid \mathcal{F}_{\tau_{k-1}}^*] = E[M_{\tau_k}^i - M_{\tau_{k-1}}^i \mid \mathcal{F}_{\tau_{k-1}}^*] = \sum_{i \in \mathbb{I}} 1_{\{\tau_{k-1}^* = i\}} E[M_{\tau_k}^i - M_{\tau_{k-1}}^i \mid \mathcal{F}_{\tau_{k-1}}^*] = 0
\]

and letting \(n \to \infty\) shows that \(E[\xi_k \mid \mathcal{F}_{\tau_{k-1}}^*] = 0\) for \(k \geq 1\). Now define an increasing family of sub-\(\sigma\)-algebras of \(\mathcal{F}\), \(\mathcal{G} = (\mathcal{G}_n)_{n \geq 0}\), by \(\mathcal{G}_n := \mathcal{F}_{\tau_n}^*\). By applying Lemma C.1.1 and Proposition 5.3.1 (since \(\mathcal{F}\) is quasi-left-continuous by assumption), \(\mathcal{G}\) satisfies

\[
\mathcal{G}_\infty := \bigvee_n \mathcal{G}_n = \bigvee_n \mathcal{F}_{\tau_n}^* = \mathcal{F}_T.
\]
The previous discussion showed that \((\xi_n, G_n)_{n \geq 0}\) forms an \(L^2\) martingale difference sequence, and \((X_n, G_n)_{n \geq 0}\), with \(X_n\) defined by

\[
X_n := \sum_{k=0}^{n} \xi_k
\]  

(C.13)

is a discrete-parameter martingale in \(L^2\). The probability space \((\Omega, \mathcal{F}, \mathbb{P})\) with filtration \(\mathcal{G} = (G_n)_{n \geq 0}\) will be used to discuss convergence and integrability properties of \((X_n)_{n \geq 0}\).

**Convergence of the discrete-parameter martingale.**

As discussed previously, the \(\mathcal{G}\)-martingale \((X_n)_{n \geq 0}\) is in \(L^2\). It is not hard to verify, for example by the conditional Jensen inequality, that the sequence \((X^2_n)_{n \geq 0}\) is a positive \(\mathcal{G}\)-submartingale. By Doob’s Decomposition (Proposition VII-1-2 of [104]), \((X^2_n)_{n \geq 0}\) can be decomposed uniquely as

\[
X^2_n = Q_n + R_n
\]  

(C.14)

where \((Q_n)_{n \geq 0}\) is an integrable \(\mathcal{G}\)-martingale and \((R_n)_{n \geq 0}\) is an increasing process (starting from 0) with respect to \(\mathcal{G}\). Convergence of \((X_n)_{n \geq 0}\) in \(L^2\) depends on the properties of the compensator \((R_n)_{n \geq 0}\), and this is made more precise by the following proposition.

**Proposition C.1.3** ([104], Proposition VII-2-3). If \((X_n)_{n \geq 0}\) is a square-integrable \(\mathcal{G}\)-martingale such that (without loss of generality) \(X_0 = 0\), and \((R_n)_{n \geq 0}\) denotes the increasing process associated with the \(\mathcal{G}\)-submartingale \((X^2_n)_{n \geq 0}\) by the Doob decomposition (C.14), then:

1. if \(E[R_\infty] < \infty\), the martingale \((X_n)_{n \geq 0}\) converges in \(L^2\); further, \(E[(\sup_{n \geq 0} |X_n|)^2] \leq 4E[R_\infty]\);

2. in every case the martingale \((X_n)_{n \geq 0}\) converges almost surely to a finite limit on the event \(\{R_\infty < \infty\}\).

These results and following lemma are the main tools needed to prove admissibility of \(\alpha^*\) and \(L^2\) convergence of \(\{C_{n^*}^\alpha : n \geq 1\}\).
Lemma C.1.4. Let \((X_n)_{n \geq 0}\) be the \(\mathbb{G}\)-martingale defined in (C.13) and \((R_n)_{n \geq 0}\) denote the increasing process associated with the \(\mathbb{G}\)-submartingale \((X_n^2)_{n \geq 0}\) by the Doob decomposition (C.14). Then \(\mathbb{E}[R_\infty] < \infty\).

Proof. Note that the monotone convergence theorem ensures \(\mathbb{E}[R_\infty]\) is well-defined (but may be infinite), and Fatou’s Lemma gives

\[
\mathbb{E}[R_\infty] \leq \liminf_{n \to \infty} \mathbb{E}[R_n].
\] (C.15)

For \(n \geq 1\), the random variable \(R_n\) can be decomposed as follows [104, p. 148]:

\[
R_n = \sum_{k=0}^{n-1} R_{k+1} - R_k = \sum_{k=0}^{n-1} \mathbb{E}[(X_{k+1} - X_k)^2 \mid \mathcal{G}_k] = \sum_{k=0}^{n-1} \mathbb{E}[(\xi_{k+1})^2 \mid \mathcal{G}_k].
\] (C.16)

Using equation (C.16) in (C.15) and applying the tower property of conditional expectations leads to

\[
\mathbb{E}[R_\infty] \leq \liminf_{n \to \infty} \mathbb{E} \left[ \sum_{k=0}^{n-1} (\xi_{k+1})^2 \right].
\] (C.17)

On the other hand, the inequalities leading up to (C.12) above show that the term on the right-hand side of (C.17) is finite, which proves \(\mathbb{E}[R_\infty] < \infty\).

\[\square\]

Theorem C.1.5 (Square-integrable cumulative switching costs). The family \(\{C^\alpha_n : n \geq 1\}\) converges in \(L^2\) and also satisfies \(\mathbb{E}[\sup_{n \geq 0} |C^\alpha_n|^2] < \infty\).

Proof. Proposition C.1.2 gave the following representation for \(C^\alpha_n\), \(n \geq 1\),

\[
C^\alpha_n = Y^\alpha_{\tau^*_N} - Y^\alpha_{\tau^*_0} + \int_{\tau^*_0}^{\tau^*_N} \psi_{s^\alpha} (s) ds - \sum_{k=1}^{N^\alpha_n} (M^\alpha_{\tau^*_k} - M^\alpha_{\tau^*_k-1})
\]
\[
= Y^\alpha_{\tau^*_N} - Y^\alpha_{\tau^*_0} + \int_{\tau^*_0}^{\tau^*_N} \psi_{s^\alpha} (s) ds - X_n \wedge N(\alpha^*) \quad \mathbb{P} \text{-a.s.}
\] (C.18)

Since \(N(\alpha^*) < \infty\) almost surely, \(\tau^*_N\) and \(\tau^*_n\) converge almost surely to \(\tau^*_N(\alpha^*) \leq T\) and \(\tau^*_N(\alpha^*) = u^*_T\), respectively. Next, note that \(Y^\alpha \in \mathcal{S}\) and \(\psi_i \in \mathcal{M}\) for every \(i \in \mathbb{I}\), and \((X_n)_{n \geq 0}\) converges in \(L^2\) and \(\mathbb{E}[(\sup_{n \geq 0} |X_n|)^2] < \infty\) by Proposition C.1.3 and Lemma C.1.4. From these observations and equation (C.18), it follows that \(\{C^\alpha_n : n \geq 1\}\) converges in \(L^2\) and also satisfies \(\mathbb{E}[\sup_{n \geq 0} |C^\alpha_n|^2] < \infty\). \[\square\]
C.2 Supplementary proofs.

Lemma C.2.1. Assume condition (C2), then ∀α = (τ_n, ι_n)_{n≥0} ∈ A:

\[ ∀N ≥ 1, \quad -\sum_{n=1}^{N} γ_{τ_{n-1},ι_n} ≤ \max_{j∈I} [-γ_{ι_0,j}] \quad P - a.s. \]  \hspace{1cm} (C.19)

Proof. First, define \( \tilde{A} \) to be the class of double sequences \((τ_n, ι_n)_{n≥0}\) satisfying all the conditions of Definition 5.2.1 for \( A \), with the exception \( P(\bigcup_{n≥1}\{ι_{n-1} = ι_n\}) ≥ 0 \). Note that \( A ⊂ \tilde{A} \) by this definition. The proof now follows by an induction argument courtesy of [92, p. 399]. For any \( \tilde{α} ∈ \tilde{A} \), (C.19) is clearly true for \( N = 1 \). Now, suppose that for all \( \tilde{α} ∈ \tilde{A} \), (C.19) is satisfied for \( N ≥ 1 \). Let \( α = (τ_n, ι_n)_{n≥0} ∈ \tilde{A} \) be arbitrary and notice that \(-γ_{ι_{N-1},ι_N} - γ_{ι_{N},ι_{N+1}} < -γ_{ι_{N-1},ι_{N+1}} a.s. \) by (5.16) so

\[ -\sum_{n=1}^{N+1} γ_{τ_{n-1},ι_n} ≤ -\sum_{n=1}^{N-1} γ_{τ_{n-1},ι_n} \quad P - a.s. \]

Define a new double sequence \( \tilde{α} = (\tilde{τ}_n, \tilde{ι}_n)_{n≥0} \) by \((\tilde{τ}_n, \tilde{ι}_n) = (τ_n, ι_n)\) for \( n = 1, \ldots, N-1 \) and \((\tilde{τ}_n, \tilde{ι}_n) = (τ_{n+1}, ι_{n+1})\) for \( n ≥ N \). Note that \( \tilde{α} ∈ \tilde{A} \) and by the induction hypothesis one gets

\[ -\sum_{n=1}^{N+1} γ_{τ_{n-1},ι_n} ≤ -\sum_{n=1}^{N} γ_{τ_{n-1},ι_n} ≤ \max_{j∈I} [-γ_{ι_0,j}] \quad P - a.s. \]

The claim holds since \( α = (τ_n, ι_n)_{n≥0} ∈ \tilde{A} \) was arbitrary and \( A ⊂ \tilde{A} \). \( \square \)
Appendix D

Appendix to Chapter 6

D.1 The relationship between upper and lower values

This section begins with an elementary lemma on the relationship between upper and lower values for the Dynkin game.

**Lemma D.1.1.** Let $W^+$ and $W^-$ denote the respective upper and lower value of a Dynkin game with payoff $D(\sigma, \tau)$:

\[
W^+ = \inf_{\sigma \in \mathcal{T}} \sup_{\tau \in \mathcal{T}} E[D(\sigma, \tau)] \\
W^- = \sup_{\tau \in \mathcal{T}} \inf_{\sigma \in \mathcal{T}} E[D(\sigma, \tau)]
\]  

(D.1)

Then, $W^- \leq W^+$.

**Proof.** Let $\hat{\sigma}$ and $\hat{\tau}$ be arbitrary stopping times for the game. Notice that by definition of the infimum,

\[
\inf_{\sigma \in \mathcal{T}} E[D(\sigma, \hat{\tau})] \leq E[D(\hat{\sigma}, \hat{\tau})].
\]

In particular, it is true that,

\[
\inf_{\sigma \in \mathcal{T}} E[D(\sigma, \hat{\tau})] \leq \sup_{\tau \in \mathcal{T}} E[D(\hat{\sigma}, \tau)].
\]  

(D.2)

Since the term on the right-hand side of equation (D.2) is an upper bound for the left hand side over all choices of $\hat{\tau}$, and the supremum is the smallest upper bound, this leads to

\[
\sup_{\tau \in \mathcal{T}} \inf_{\sigma \in \mathcal{T}} E[D(\sigma, \tau)] \leq \sup_{\tau \in \mathcal{T}} E[D(\hat{\sigma}, \tau)].
\]  

(D.3)
Similarly, the left-hand side of equation (D.3) is a lower bound for the right-hand side over all choices of $\tilde{\sigma}$. Since the infimum is the greatest of these lower bounds, one concludes with
\[
\sup_{\tau \in \mathcal{T}} \inf_{\sigma \in \mathcal{T}} \mathbb{E} [D(\sigma, \tau)] \leq \inf_{\sigma \in \mathcal{T}} \sup_{\tau \in \mathcal{T}} \mathbb{E} [D(\sigma, \tau)].
\]
\[\square\]

D.2 Proof of existence of the game value

The rest of this appendix is devoted to the proof of Theorem 6.5.1. Let $\varepsilon > 0$ be given and let $\alpha^1 \in \mathcal{A}_1$, $\alpha^1 = (\tau^{(1)}_n, \iota^{(1)}_n)_{n \geq 0}$, satisfy:
\[
V(1) \leq J(\alpha^1; 1) + \varepsilon.
\]

The switching control $\alpha^1$ is therefore $\varepsilon$-optimal for the problem starting in mode $i = 1$. Let $u^1$ be the mode indicator function (cf. (6.14)) associated with $\alpha^1$ and define $\hat{\tau}$ to be the first switching time under $\alpha^1$:
\[
\hat{\tau} = \tau^{(1)}_1 = \inf \{ s \geq 0 : u^1_s = 0 \} \land T. \tag{D.4}
\]

Given an arbitrary stopping time $\sigma \in \mathcal{T}$, define the function $u^0$ as follows:
\[
u^0 := \begin{cases} 
  u^0_s = 0, & s \in [0, \hat{\tau} \land \sigma] \\
  u^0_s = u^1_s, & s \in (\hat{\tau} \land \sigma, T]
\end{cases} \tag{D.5}
\]

By construction, $u^0$ is also a valid mode indicator and it is associated with a switching control $\alpha^0$. Furthermore, one verifies that $\alpha^0 \in \mathcal{A}_0$ since $\alpha^1 \in \mathcal{A}_1$. The strategies $\alpha^0$ and $\alpha^1$, their associated mode indicator functions, and the stopping times $\hat{\tau}$ and $\sigma$ are referenced in the following steps which ultimately prove Theorem 6.5.1.

**Proposition D.2.1.** For $i \in \{0, 1\}$, let $V(i)$ be the value function for the optimal switching problem with performance index (6.20). Let $\sigma, \tau \in \mathcal{T}$ and $D(\sigma, \tau)$ be defined as in equation (6.4). Then
\[
V(1) - V(0) \leq \sup_{\tau \in \mathcal{T}} \inf_{\sigma \in \mathcal{T}} \mathbb{E} [D(\sigma, \tau)]. \tag{D.6}
\]
Proof. The proof is guided by the method used in [65, p. 436]. Notice that the claim (D.6) is true if and only if, for every \( \varepsilon > 0 \), there exists \( \hat{\tau} \in \mathcal{T} \) such that for all \( \sigma \in \mathcal{T} \):

\[
V(1) - V(0) \leq \mathbb{E}[D(\sigma, \hat{\tau})] + \varepsilon.
\]

Let \( \varepsilon > 0 \) be given and \( \hat{\tau}, \alpha^0 \) and \( \alpha^1 \) be as defined in Section D.2 so that

\[
V(1) - V(0) \leq V(1) - J(\alpha^0; 0) \leq J(\alpha^1; 1) - J(\alpha^0; 0) + \varepsilon.
\]

Writing out the terms of the performance index under the different controls yields:

\[
J(\alpha^1; 1) - J(\alpha^0; 0) + \varepsilon
= \mathbb{E} \left[ \int_0^T \psi(s) u_1^s ds + \Gamma u_T^1 - \sum_{n \geq 1} \gamma(\tau_n^{(1)}, \iota_n^{(1)}) 1_{\{\tau_n^{(1)} < T\}} \right]
\]

\[
- \mathbb{E} \left[ \int_0^T \psi(s) u_0^s ds + \Gamma u_T^0 - \sum_{n \geq 1} \gamma(\tau_n^{(0)}, \iota_n^{(0)}) 1_{\{\tau_n^{(0)} < T\}} \right] + \varepsilon
= \mathbb{E} \left[ \int_0^{\hat{\tau} \land \sigma} \psi(s) \left[u_1^s - u_0^s\right] ds \right] + \mathbb{E} \left[ \Gamma \left[u_T^1 - u_T^0\right] \right]
\]

\[
+ \mathbb{E} \left[ \sum_{n \geq 1} \gamma(\tau_n^{(0)}, \iota_n^{(0)}) 1_{\{\tau_n^{(0)} < T\}} - \sum_{n \geq 1} \gamma(\tau_n^{(1)}, \iota_n^{(1)}) 1_{\{\tau_n^{(1)} < T\}} \right] + \varepsilon \quad (D.7)
\]

The three terms involving the expectation operator in equation (D.7) are looked at separately:

**Term A:** \( \mathbb{E} \left[ \int_0^T \psi(s) \left[u_1^s - u_0^s\right] ds \right] \)

**Term B:** \( \mathbb{E} \left[ \Gamma \left[u_T^1 - u_T^0\right] \right] \)

**Term C:** \( \mathbb{E} \left[ \sum_{n \geq 1} \gamma(\tau_n^{(0)}, \iota_n^{(0)}) 1_{\{\tau_n^{(0)} < T\}} - \sum_{n \geq 1} \gamma(\tau_n^{(1)}, \iota_n^{(1)}) 1_{\{\tau_n^{(1)} < T\}} \right] \)

**Term A.** Since \( u_1^s = 1 \) and \( u_0^s = 0 \) for \( s \in [0, \hat{\tau} \land \sigma] \), and \( u_0^s = u_1^s \) for \( s \in (\hat{\tau} \land \sigma, T] \), one gets

\[
\mathbb{E} \left[ \int_0^T \psi(s) \left[u_1^s - u_0^s\right] ds \right] = \mathbb{E} \left[ \int_0^{\hat{\tau} \land \sigma} \psi(s) ds \right]. \quad (D.8)
\]

**Term B.** Recall that \( \hat{\tau} \leq T \) and \( \sigma \leq T \). The two events \( \{\hat{\tau} \land \sigma < T\} \) and \( \{\hat{\tau} = \sigma = T\} \) are disjoint and the probability of their union is one. Use these facts to rewrite the second expectation in equation (D.7) as follows:

\[
\mathbb{E} \left[ \Gamma \left[u_T^1 - u_T^0\right] \right] = \mathbb{E} \left[ \Gamma \left[u_T^1 - u_T^0\right] \begin{cases} 1_{\{\sigma = \hat{\tau} = T\}} + 1_{\{\hat{\tau} \land \sigma < T\}} \end{cases} \right].
\]
Using (D.5) on the event \{\hat{\tau} \land \sigma < T\} gives \(u_T^1 = u_T^0\). On the event \{\sigma = \hat{\tau} = T\}, there is no change in the mode indicator function on \([0, T]\) for either policy, which means \(u_T^1 = 1\) and \(u_T^0 = 0\). These observations give

\[
E \left[ \Gamma (u_T^1 - u_T^0) \right] = E \left[ \Gamma 1_{\{\sigma = \tau = T\}} \right].
\] (D.9)

**Term C** By Lemma D.2.2 below, the following inequality for the third expectation holds:

\[
E \left[ \sum_{n \geq 1} \gamma(\tau_n^{(0)}, \iota_n^{(0)}) 1_{\{\tau_n^{(0)} < T\}} - \sum_{n \geq 1} \gamma(\tau_n^{(1)}, \iota_n^{(1)}) 1_{\{\tau_n^{(1)} < T\}} \right] \\
\leq E \left[ \gamma_-(\sigma) 1_{\{\sigma \leq \hat{\tau}\}} 1_{\{\sigma < T\}} - \gamma_+ (\hat{\tau}) 1_{\{\sigma > \hat{\tau}\}} \right]
\] (D.10)

**Conclusion** Using equations (D.8), (D.9) and (D.10), the claim holds since:

\[
V(1) - V(0) \\
\leq E \left[ \int_0^{\hat{\tau} \land \sigma} \psi(s)ds + \Gamma 1_{\{\sigma = \tau = T\}} + \gamma_-(\sigma) 1_{\{\sigma \leq \hat{\tau}\}} 1_{\{\sigma < T\}} - \gamma_+ (\hat{\tau}) 1_{\{\sigma > \hat{\tau}\}} \right] + \varepsilon \\
= E \left[ D(\sigma, \hat{\tau}) \right] + \varepsilon
\]

□

**Lemma D.2.2.** Let \(\sigma \in \mathcal{T}\) be arbitrary and \(\hat{\tau}, \alpha^0, \alpha^1\) be as defined in Section D.2. Then,

\[
E \left[ \sum_{n \geq 1} \gamma(\tau_n^{(0)}, \iota_n^{(0)}) 1_{\{\tau_n^{(0)} < T\}} - \sum_{n \geq 1} \gamma(\tau_n^{(1)}, \iota_n^{(1)}) 1_{\{\tau_n^{(1)} < T\}} \right] \\
\leq E \left[ \gamma_-(\sigma) 1_{\{\sigma \leq \hat{\tau}\}} 1_{\{\sigma < T\}} - \gamma_+ (\hat{\tau}) 1_{\{\sigma > \hat{\tau}\}} \right]
\] (D.11)

**Proof.** The events \(\{\sigma > \hat{\tau}\}\) and \(\{\sigma \leq \hat{\tau}\}\) are disjoint and the probability of their union occurring is one. This leads to the following:

\[
E \left[ \sum_{n \geq 1} \gamma(\tau_n^{(0)}, \iota_n^{(0)}) 1_{\{\tau_n^{(0)} < T\}} - \sum_{n \geq 1} \gamma(\tau_n^{(1)}, \iota_n^{(1)}) 1_{\{\tau_n^{(1)} < T\}} \right] \\
= E \left[ \left( \sum_{n \geq 1} \gamma(\tau_n^{(0)}, \iota_n^{(0)}) 1_{\{\tau_n^{(0)} < T\}} - \sum_{n \geq 1} \gamma(\tau_n^{(1)}, \iota_n^{(1)}) 1_{\{\tau_n^{(1)} < T\}} \right) 1_{\{\sigma > \hat{\tau}\}} \right] \\
+ E \left[ \left( \sum_{n \geq 1} \gamma(\tau_n^{(0)}, \iota_n^{(0)}) 1_{\{\tau_n^{(0)} < T\}} - \sum_{n \geq 1} \gamma(\tau_n^{(1)}, \iota_n^{(1)}) 1_{\{\tau_n^{(1)} < T\}} \right) 1_{\{\sigma \leq \hat{\tau}\}} \right].
\]

The two terms in the last line shall be analysed separately:
Term A’: \[ E \left( \sum_{n \geq 1} \gamma(\tau_n^{(0)}, \ell_n^{(0)}) \mathbf{1}_{\{\tau_n^{(0)} < T\}} - \sum_{n \geq 1} \gamma(\tau_n^{(1)}, \ell_n^{(1)}) \mathbf{1}_{\{\tau_n^{(1)} < T\}} \right) \mathbf{1}_{\{\sigma > \tau\}} \]

Term B’: \[ E \left( \sum_{n \geq 1} \gamma(\tau_n^{(0)}, \ell_n^{(0)}) \mathbf{1}_{\{\tau_n^{(0)} < T\}} - \sum_{n \geq 1} \gamma(\tau_n^{(1)}, \ell_n^{(1)}) \mathbf{1}_{\{\tau_n^{(1)} < T\}} \right) \mathbf{1}_{\{\sigma \leq \tau\}} \]

Term A’ Remember from its definition in (D.4), that \( \hat{\tau} \) is the first switching time under \( \alpha^1 \). Hence by definition of \( u^1 \) and equation (6.14) of the mode indicator:

\[ u^1_s = 1 \text{ for } s \in [0, \hat{\tau}] \quad \text{and} \quad u^1_s = 0 \text{ for } s \in (\hat{\tau}, \tau_2^{(1)}]. \]

There is a cost of \( \gamma(\hat{\tau}, 0) \) incurred for switching to mode 0 at \( \hat{\tau} \) and, using equation (6.18), this means \( \gamma(\hat{\tau}, \ell^{(1)} = 0) = \gamma_+ \hat{\tau} \). Since \( \hat{\tau} \land \sigma = \hat{\tau} \) on the event \( \{\sigma > \hat{\tau}\} \), by definition of \( \alpha^0 \) it follows that

\[ u^0_s = 0 \text{ for } s \in [0, \hat{\tau}] \quad \text{and} \quad u^0_s = u^1_s \text{ for } s \in (\hat{\tau}, \tau_2^{(1)}]. \]

Since \( u^1_s = 0 \) for \( s \in (\hat{\tau}, \tau_2^{(1)}], \) this means there is no switch under policy \( \alpha^0 \) until \( \tau_2^{(1)} \), and the switching costs under \( \alpha^0 \) and \( \alpha^1 \) satisfy \( \gamma(\tau_n^{(1)}, \ell^{(1)}_n) = \gamma(\tau_n^{(0)}, \ell^{(0)}_n) \) for \( n \geq 2 \). Therefore Term A’ reduces to:

\[
E \left[ \left( \sum_{n \geq 1} \gamma(\tau_n^{(0)}, \ell_n^{(0)}) \mathbf{1}_{\{\tau_n^{(0)} < T\}} - \sum_{n \geq 1} \gamma(\tau_n^{(1)}, \ell_n^{(1)}) \mathbf{1}_{\{\tau_n^{(1)} < T\}} \right) \mathbf{1}_{\{\sigma > \tau\}} \right] \\
= E \left[ \left( \sum_{n \geq 1} \gamma(\tau_n^{(0)}, \ell_n^{(0)}) \mathbf{1}_{\{\tau_n^{(0)} < T\}} - \sum_{n \geq 1} \gamma(\tau_n^{(1)}, \ell_n^{(1)}) \mathbf{1}_{\{\tau_n^{(1)} < T\}} \right) - \gamma(\tau_1^{(1)}, \ell_1^{(1)}) \mathbf{1}_{\{\sigma > \tau\}} \right] \\
= E \left[ -\gamma_+ \hat{\tau} \mathbf{1}_{\{\sigma > \tau\}} \right] \tag{D.12}
\]

Term B’ In order to deal with Term B’, first recall that \( \sigma \leq T \) so that:

\[
E \left[ \left( \sum_{n \geq 1} \gamma(\tau_n^{(0)}, \ell_n^{(0)}) \mathbf{1}_{\{\tau_n^{(0)} < T\}} - \sum_{n \geq 1} \gamma(\tau_n^{(1)}, \ell_n^{(1)}) \mathbf{1}_{\{\tau_n^{(1)} < T\}} \right) \mathbf{1}_{\{\sigma \leq \tau\}} \right] \\
= E \left[ \left( \sum_{n \geq 1} \gamma(\tau_n^{(0)}, \ell_n^{(0)}) \mathbf{1}_{\{\tau_n^{(0)} < T\}} - \sum_{n \geq 1} \gamma(\tau_n^{(1)}, \ell_n^{(1)}) \mathbf{1}_{\{\tau_n^{(1)} < T\}} \right) \mathbf{1}_{\{\sigma \leq \tau\}} \mathbf{1}_{\{\sigma \leq T\}} \right] \\
= E \left[ \left( \sum_{n \geq 1} \gamma(\tau_n^{(0)}, \ell_n^{(0)}) \mathbf{1}_{\{\tau_n^{(0)} < T\}} - \sum_{n \geq 1} \gamma(\tau_n^{(1)}, \ell_n^{(1)}) \mathbf{1}_{\{\tau_n^{(1)} < T\}} \right) \mathbf{1}_{\{\sigma \leq \tau\}} \mathbf{1}_{\{\sigma < T\}} \right] \\
+ E \left[ \left( \sum_{n \geq 1} \gamma(\tau_n^{(0)}, \ell_n^{(0)}) \mathbf{1}_{\{\tau_n^{(0)} < T\}} - \sum_{n \geq 1} \gamma(\tau_n^{(1)}, \ell_n^{(1)}) \mathbf{1}_{\{\tau_n^{(1)} < T\}} \right) \mathbf{1}_{\{\sigma \leq \tau\}} \mathbf{1}_{\{\sigma = T\}} \right] \tag{D.13}
\]

The two terms in equation (D.13) are dealt with below:
Term B’-1:

\[
E \left[ \left( \sum_{n \geq 1} \gamma(\tau_n^{(0)}, t_n^{(0)}) \mathbf{1}_{\{\tau_n^{(0)} < T\}} - \sum_{n \geq 1} \gamma(\tau_n^{(1)}, t_n^{(1)}) \mathbf{1}_{\{\tau_n^{(1)} < T\}} \right) \mathbf{1}_{\{\sigma \leq \hat{\tau}\}} \mathbf{1}_{\{\sigma < T\}} \right]
\]

Term B’-2:

\[
E \left[ \left( \sum_{n \geq 1} \gamma(\tau_n^{(0)}, t_n^{(0)}) \mathbf{1}_{\{\tau_n^{(0)} < T\}} - \sum_{n \geq 1} \gamma(\tau_n^{(1)}, t_n^{(1)}) \mathbf{1}_{\{\tau_n^{(1)} < T\}} \right) \mathbf{1}_{\{\sigma \leq \hat{\tau}\}} \mathbf{1}_{\{\sigma = T\}} \right]
\]

Term B’-2. The next few lines establish that the Term B’-2 is zero. Note that \( \tau \leq T \), which gives:

\[
E \left[ \left( \sum_{n \geq 1} \gamma(\tau_n^{(0)}, t_n^{(0)}) \mathbf{1}_{\{\tau_n^{(0)} < T\}} - \sum_{n \geq 1} \gamma(\tau_n^{(1)}, t_n^{(1)}) \mathbf{1}_{\{\tau_n^{(1)} < T\}} \right) \mathbf{1}_{\{\sigma \leq \hat{\tau}\}} \mathbf{1}_{\{\sigma = T\}} \right] = E \left[ \left( \sum_{n \geq 1} \gamma(\tau_n^{(0)}, t_n^{(0)}) \mathbf{1}_{\{\tau_n^{(0)} < T\}} - \sum_{n \geq 1} \gamma(\tau_n^{(1)}, t_n^{(1)}) \mathbf{1}_{\{\tau_n^{(1)} < T\}} \right) \mathbf{1}_{\{\sigma = \hat{\tau} = T\}} \right] = 0.
\]

Term B’-1. Turn now to Term B’-1, and rewrite it as follows:

\[
E \left[ \left( \sum_{n \geq 1} \gamma(\tau_n^{(0)}, t_n^{(0)}) \mathbf{1}_{\{\tau_n^{(0)} < T\}} - \sum_{n \geq 1} \gamma(\tau_n^{(1)}, t_n^{(1)}) \mathbf{1}_{\{\tau_n^{(1)} < T\}} \right) \mathbf{1}_{\{\sigma \leq \hat{\tau}\}} \mathbf{1}_{\{\sigma < T\}} \right] = E \left[ \left( \sum_{n \geq 1} \gamma(\tau_n^{(0)}, t_n^{(0)}) \mathbf{1}_{\{\tau_n^{(0)} < T\}} - \sum_{n \geq 1} \gamma(\tau_n^{(1)}, t_n^{(1)}) \mathbf{1}_{\{\tau_n^{(1)} < T\}} \right) \mathbf{1}_{\{\sigma < \hat{\tau}\}} \right] 
\]

\[
\times \mathbf{1}_{\{\sigma < T\}}.
\]

Under policy \( \alpha^0 \), \( u_s^0 = 0 \) for all times \( s \in [0, \hat{\tau} \wedge \sigma] \), and \( u_s^0 = u_s^1 \) for \( s \in (\hat{\tau} \wedge \sigma, T] \). On the event \( \{\sigma < \hat{\tau}\} \), it is true that \( u_s^0 = 1 \) for \( s \in (\sigma, \hat{\tau}] \), so \( \sigma \) becomes the first switching time for policy \( \alpha^0: \sigma = \tau_i^{(0)} \) on \( \{\sigma < \hat{\tau}\} \). The cost for switching to mode 1 is then:

\[
\gamma(\sigma, \ell^{(0)} = 1) = \gamma_-(\sigma).
\]
Therefore, on the event \( \{\sigma < \hat{\tau}\} \cap \{\sigma < T\} \), the switching times and new modes under \( \alpha^0 \) and \( \alpha^1 \) satisfy \( \tau_n^{(0)} = \tau_{n-1}^{(0)} \) and \( t_n^{(0)} = t_{n-1}^{(0)} \) for \( n \geq 2 \), which means:

\[
\mathbb{E}\left[ \left( \sum_{n \geq 1} \gamma(\tau_n^{(0)}, t_n^{(0)}) 1_{\{\tau_n^{(0)} < T\}} - \sum_{n \geq 1} \gamma(\tau_n^{(1)}, t_n^{(1)}) 1_{\{\tau_n^{(1)} < T\}} \right) 1_{\{\sigma < \hat{\tau}\}} \times 1_{\{\sigma < T\}} \right] \\
= \mathbb{E}\left[ \gamma(\sigma) 1_{\{\sigma < \hat{\tau}\}} 1_{\{\sigma < T\}} \right].
\] (D.14)

On the event \( \{\sigma = \hat{\tau}\} \cap \{\sigma < T\} \), the stopping time \( \sigma \) is the first switching time under policy \( \alpha^1 \). Then, similar to the derivation of equation (D.12), one arrives at:

\[
\mathbb{E}\left[ \left( \sum_{n \geq 1} \gamma(\tau_n^{(0)}, t_n^{(0)}) 1_{\{\tau_n^{(0)} < T\}} - \sum_{n \geq 1} \gamma(\tau_n^{(1)}, t_n^{(1)}) 1_{\{\tau_n^{(1)} < T\}} \right) 1_{\{\sigma = \hat{\tau}\}} 1_{\{\sigma < T\}} \right] \\
= \mathbb{E}\left[ -\gamma(\sigma) 1_{\{\sigma = \hat{\tau}\}} 1_{\{\sigma < T\}} \right] \\
\leq \mathbb{E}\left[ \gamma(\sigma) 1_{\{\sigma = \hat{\tau}\}} 1_{\{\sigma < T\}} \right].
\] (D.15)

The inequality is as a result of the “no arbitrage” condition (6.10) in Assumption 8.

**Conclusion**  Use (D.14) and (D.15) in equation (D.13) to obtain:

\[
\mathbb{E}\left[ \left( \sum_{n \geq 1} \gamma(\tau_n^{(0)}, t_n^{(0)}) 1_{\{\tau_n^{(0)} < T\}} - \sum_{n \geq 1} \gamma(\tau_n^{(1)}, t_n^{(1)}) 1_{\{\tau_n^{(1)} < T\}} \right) 1_{\{\sigma \leq \hat{\tau}\}} \right] \\
\leq \mathbb{E}\left[ \gamma(\sigma) 1_{\{\sigma \leq \hat{\tau}\}} 1_{\{\sigma < T\}} \right]
\] (D.16)

Finally, use (D.12) and (D.16) to conclude

\[
\mathbb{E}\left[ \left( \sum_{n \geq 1} \gamma(\tau_n^{(0)}, t_n^{(0)}) 1_{\{\tau_n^{(0)} < T\}} - \sum_{n \geq 1} \gamma(\tau_n^{(1)}, t_n^{(1)}) 1_{\{\tau_n^{(1)} < T\}} \right) 1_{\{\sigma \leq \hat{\tau}\}} \right] \\
\leq \mathbb{E}\left[ \gamma(\sigma) 1_{\{\sigma \leq \hat{\tau}\}} 1_{\{\sigma < T\}} - \gamma(\hat{\tau}) 1_{\{\sigma > \hat{\tau}\}} \right]
\]

\( \square \)

**Proposition D.2.3.** For \( i \in \{0, 1\} \), let \( V(i) \) be the value function for the optimal switching problem with performance index (6.20). Let \( D(\sigma, \tau) \) be the payoff for the Dynkin game as in equation (6.4). Then,

\[
V(1) - V(0) \geq \inf_{\sigma \in T} \sup_{\tau \in T} \mathbb{E}\left[ D(\sigma, \tau) \right].
\] (D.17)
Proof. The proof is similar to the one given for Proposition D.2.1. The claim (D.17) is true if and only if, for every \( \varepsilon > 0 \), there exists \( \hat{\sigma} \in \mathcal{T} \) such that, for all \( \tau \in \mathcal{T} \):

\[
V(1) - V(0) \geq \mathbb{E} [D(\hat{\sigma}, \tau)] - \varepsilon,
\]

\[
\iff V(0) - V(1) \leq -\mathbb{E} [D(\hat{\sigma}, \tau)] + \varepsilon.
\]

(D.18)

Let \( \varepsilon > 0 \) be given and \( \alpha^0 = (\tau_n^{(0)}, \iota_n^{(0)})_{n \geq 0} \) be an admissible switching control with initial mode \( \iota_0^{(0)} = 0 \) and:

\[
V(0) \leq J(\alpha^0; 0) + \varepsilon.
\]

Let \( u^0 \) be mode indicator associated with \( \alpha^0 \). Define \( \hat{\sigma} \) to be the first switching time under \( \alpha^0 \),

\[
\hat{\sigma} = \tau_1^{(0)} = \inf \{ s \geq 0 : u^0_s = 1 \} \land T. \tag{D.19}
\]

Let \( \tau \in \mathcal{T} \) be arbitrary and define \( u^1 \) by:

\[
u^1 := \begin{cases} 
    u^1_s = 1, & s \in [0, \hat{\sigma} \land \tau] \\
    u^1_s = u^0_s, & s \in (\hat{\sigma} \land \tau, T]
\end{cases} \tag{D.20}
\]

Then \( u^1 \) is a valid mode indicator to which one associates an admissible switching policy \( \alpha^1 \in \mathcal{A}_1 \). The rest of the proof follows in the same way as Proposition D.2.1 with the roles of \( \alpha^1 \) and \( \alpha^0 \) reversed.

\[ \Box \]

Proof of Theorem 6.5.1. Propositions D.2.1 and Proposition D.2.3 above have shown that:

\[
\inf_{\sigma \in \mathcal{T}} \sup_{\tau \in \mathcal{T}} \mathbb{E} [D(\sigma, \tau)] \leq V(1) - V(0) \leq \sup_{\tau \in \mathcal{T}} \inf_{\sigma \in \mathcal{T}} \mathbb{E} [D(\sigma, \tau)].
\]

However, by applying Lemma D.1.1 one gets

\[
\sup_{\tau \in \mathcal{T}} \inf_{\sigma \in \mathcal{T}} \mathbb{E} [D(\sigma, \tau)] \leq \inf_{\sigma \in \mathcal{T}} \sup_{\tau \in \mathcal{T}} \mathbb{E} [D(\sigma, \tau)] \leq V(1) - V(0) \leq \sup_{\tau \in \mathcal{T}} \inf_{\sigma \in \mathcal{T}} \mathbb{E} [D(\sigma, \tau)]
\]

and proves the claim. \[ \Box \]
Appendix E

Additional Tables

<table>
<thead>
<tr>
<th>month</th>
<th>season</th>
<th>day type</th>
<th>I</th>
<th>II</th>
<th>III</th>
<th>IV</th>
<th>V</th>
</tr>
</thead>
<tbody>
<tr>
<td>January</td>
<td>winter</td>
<td>weekday</td>
<td>29%</td>
<td>52%</td>
<td>10%</td>
<td>4%</td>
<td>5%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>weekend</td>
<td>30%</td>
<td>57%</td>
<td>7%</td>
<td>3%</td>
<td>2%</td>
</tr>
<tr>
<td>February</td>
<td>winter</td>
<td>weekday</td>
<td>31%</td>
<td>51%</td>
<td>14%</td>
<td>3%</td>
<td>1%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>weekend</td>
<td>15%</td>
<td>75%</td>
<td>8%</td>
<td>2%</td>
<td>0%</td>
</tr>
<tr>
<td>March</td>
<td>spring</td>
<td>weekday</td>
<td>32%</td>
<td>52%</td>
<td>10%</td>
<td>2%</td>
<td>3%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>weekend</td>
<td>25%</td>
<td>60%</td>
<td>13%</td>
<td>0%</td>
<td>2%</td>
</tr>
<tr>
<td>April</td>
<td>spring</td>
<td>weekday</td>
<td>57%</td>
<td>36%</td>
<td>4%</td>
<td>1%</td>
<td>1%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>weekend</td>
<td>58%</td>
<td>38%</td>
<td>3%</td>
<td>0%</td>
<td>0%</td>
</tr>
<tr>
<td>May</td>
<td>spring</td>
<td>weekday</td>
<td>59%</td>
<td>31%</td>
<td>5%</td>
<td>3%</td>
<td>2%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>weekend</td>
<td>72%</td>
<td>26%</td>
<td>2%</td>
<td>0%</td>
<td>0%</td>
</tr>
<tr>
<td>June</td>
<td>summer</td>
<td>weekday</td>
<td>48%</td>
<td>30%</td>
<td>14%</td>
<td>4%</td>
<td>4%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>weekend</td>
<td>67%</td>
<td>26%</td>
<td>5%</td>
<td>1%</td>
<td>1%</td>
</tr>
<tr>
<td>July</td>
<td>summer</td>
<td>weekday</td>
<td>48%</td>
<td>33%</td>
<td>9%</td>
<td>3%</td>
<td>7%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>weekend</td>
<td>65%</td>
<td>29%</td>
<td>4%</td>
<td>0%</td>
<td>1%</td>
</tr>
<tr>
<td>August</td>
<td>summer</td>
<td>weekday</td>
<td>55%</td>
<td>32%</td>
<td>5%</td>
<td>4%</td>
<td>4%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>weekend</td>
<td>66%</td>
<td>25%</td>
<td>6%</td>
<td>0%</td>
<td>3%</td>
</tr>
<tr>
<td>September</td>
<td>autumn</td>
<td>weekday</td>
<td>35%</td>
<td>50%</td>
<td>6%</td>
<td>2%</td>
<td>7%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>weekend</td>
<td>40%</td>
<td>48%</td>
<td>2%</td>
<td>5%</td>
<td>6%</td>
</tr>
<tr>
<td>October</td>
<td>autumn</td>
<td>weekday</td>
<td>18%</td>
<td>53%</td>
<td>14%</td>
<td>7%</td>
<td>8%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>weekend</td>
<td>22%</td>
<td>65%</td>
<td>10%</td>
<td>2%</td>
<td>1%</td>
</tr>
<tr>
<td>November</td>
<td>autumn</td>
<td>weekday</td>
<td>11%</td>
<td>50%</td>
<td>18%</td>
<td>5%</td>
<td>15%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>weekend</td>
<td>21%</td>
<td>49%</td>
<td>22%</td>
<td>6%</td>
<td>2%</td>
</tr>
<tr>
<td>December</td>
<td>winter</td>
<td>weekday</td>
<td>20%</td>
<td>51%</td>
<td>11%</td>
<td>10%</td>
<td>8%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>weekend</td>
<td>21%</td>
<td>61%</td>
<td>10%</td>
<td>4%</td>
<td>3%</td>
</tr>
</tbody>
</table>

Table E.1: Occurrence of different categories of price variability from 2005 to 2014 based on the month and type of day. The categories are defined in Table 3.1.