University of Techonology, Sydney

Doctoral Thesis

Regression and Convex Switching System Methods for Stochastic Control Problems with Applications to Multiple-Exercise Options

Author:
Nicholas Andrew Yap Swee Guan

Supervisor:
Professor. Carl Chiarella
Co-supervisor:
Associate Professor. Juri Hinz

A thesis submitted in fulfilment of the requirements for the degree of Doctor of Philosophy

in the

Finance Discipline Group
University of Techonology, Sydney
PO Box 123, Broadway, NSW 2007, Australia

August 2015

Declaration of Authorship

I certify that the work in this thesis has not previously been submitted for a degree nor has it been submitted as part of requirements for a degree except as fully acknowledged within the text.

I also certify that the thesis has been written by me. Any help that I have received in my research work and the preparation of the thesis itself has been acknowledged. In addition, I certify that all information sources and literature used are indicated in the thesis.

Signature of Student:

_____

Date:

_____

"You must be shapeless, formless, like water. When you pour water in a cup, it becomes the cup. When you pour water in a bottle, it becomes the bottle. When you pour water in a teapot, it becomes the teapot. Water can drip and it can crash. Become like water my friend."

Bruce Lee

Abstract

In this thesis, we develop several new simulation-based algorithms for solving some important classes of optimal stochastic control problems. In particular, these methods are aimed at providing good approximate solutions to problems that involve a high-dimensional underlying processes. These algorithms are of the primal-dual kind and therefore provide a gauge of the distance to optimality of the given approximate solutions to the optimal one. These methods will be used in the pricing of the multiple-exercise option. In Chapter 1, we conduct a review of the literature that is relevant to the pricing of the multiple-exercise option and the primal and dual methods that we will be developing in this thesis.

In the next two chapters of the thesis, we focus on regression-based dual methods for optimal multiple stopping problems in probability theory. In particular, we concentrate on finding upper bounds on the price of the multiple-exercise option as it sits within this framework. In Chapter 2, we derive an additive dual for the multiple-exercise options using financial arguments, and see that this approach leads to the construction of an algorithm that has greater computational efficiency than other methods in the literature. In Chapter 3, we derive the first known dual of the multiplicative kind for the multiple-exercise option and devise a tractable algorithm to compute it.

In the penultimate chapter of the thesis, we focus on a new class of algorithms that are based on what is known as convex switching system. These algorithms provide approximate solutions to the more general class of optimal stochastic switching problems. In Chapter 4, techniques based on combinations of rigorous theory and heuristics arguments are used to improve the efficiency and applicability of the method. We then devise algorithms of the primal-dual kind to assess the accuracy of this approach. Chapter 5 concludes.

Acknowledgements

To rephrase a famous saying, I have only been able to complete this thesis because I have stood on the shoulders of giants. The giants I speak of refer to the many academics whose work have become the bedrock of modern financial economics/mathematics. I feel extremely honoured to have worked alongside some of them and privileged to have a few shape my education over the course of this PhD program. As I give thanks, I shall try my best not to leave anyone out.

First and foremost, I would like to thank my supervisor Professor Carl Chiarella for giving me the opportunity to pursue my PhD. Through Carl, I was able to receive the necessary funding and research support by tapping into his vast academic network, giving me every opportunity to succeed. He also encouraged me to pursue research in the area of swing option pricing. This entire thesis has benefitted immensely from his careful reading.

Secondly, I would like to acknowledge Professor Mark Joshi of the University of Melbourne, Department of Economics, Center of Actuarial Studies for his role in my education. Mark taught me the fundamentals of derivatives pricing out of his famous book, *The Concepts and Practice of Mathematical Finance* Joshi [2003], when I was an honours student doing actuarial studies at the University of Melbourne and sparked my desire to pursue research in this area. Chapters 2 and 3 of this thesis are largely based on the joint works Chiarella et al. [2014] and Joshi and Yap [2014] that we have collaborated on. My understanding of the area of derivatives pricing and mathematics in general has largely improved through our numerous interactions.

Thirdly, I would like to thank Professor Juri Hinz of the University of Technology, Sydney, School of Mathematical Sciences for his willingness to co-supervise me. Chapter 4 is largely based on the joint work Hinz and Yap [2014], which is an important extension of his seminal work on convex switching systems Hinz [2014]. It is in my humble opinion that this work will eventually become as important as Monte Carlo and lattice methods in solving optimal stochastic control problems and I am honoured to have been given the opportunity to contribute towards its development. On a personal level, Juri has been extremely good to me, buying me many lunches and organising a research position for me at CSIRO.

Fourthly, I would like to express my gratitude to Professor Marek Rutkowski of the University of Sydney, Department of Mathematics. Marek was kind enough to agree to be a discussant of the paper Yap [2013] which I presented at the 26th Annual PhD Conference in Economics and Business in Canberra and also encouraged me to give a two part seminar as part of his research group's weekly seminar series. Marek also provided invaluable comments on an earlier version of the Joshi and Yap [2014] paper. Through my interactions with him, I was forced to think even more deeply about the work.

I would also like to thank Professor Tony He for doing such a wonderful job co-ordinating the PhD program and Ms Stephanie Ough for helping me with all the related administrative work. Lastly, I would like to thank the many academics at UTS that have found the time to attend my half-yearly progress report presentation and critique my work. This includes Professor Eckhard Platen, Professor Erik Schogl, Professor Tony Hall, Dr Mark Craddock, Dr Hardy Hulley, Dr Christina Nikitopoulos and Ms Nadima El Hassan. I have had the absolute pleasure of taking some of the courses taught by them as part of the Master in Quantitative Finance programme at UTS. Finally, I am thankful to have been able to immerse myself in an intellectual environment on a daily basis thanks to my fellow postdoctoral and PhD friends Dr Alex Badran, Dr Jan Baldeaux, Dr Boda Kang, Dr Chi Chung Siu, Dr Kris Glover, Dr Ke Du, Dr Kai Li, Dr Yang Chang, Dr Jinan Zhai, Dr Vinay Patel and Dr Edgardo Cayon, Benjamin Cheng, Weiting Pan, Tianhao Zhi and Danny Lo.

On a personal note, I would like to thank my family for their unwavering support. Firstly, to my wife Nicole, who took a chance on me when I was a poor PhD student, thank you for the conversations over Skype every night and keeping me sane during these three and a half years. Secondly, to my maternal grandmother and late paternal grandparents, thank you for loving me unconditionally. Thirdly, my one and only brother Neil, thank you for the love and care and I believe that you will find success in whatever you do. Most importantly, to Mum and Dad who have raised me, no words can give enough thanks for all that you have done for me. The only thing I can do is dedicate this thesis to you.

Contents

List of Figures

List of Tables

Abbreviations

AB	**A**ndersen and **B**roadie
ADP	**A**pproximate **D**ynamic **P**rogramming
CSS	**C**onvex **S**witching **S**ystem
LSM	**L**east **S**quares **M**onte Carlo
MH	**M**einhausen and **H**ambly

To my parents Andrew and Beng Im. . .

Chapter 1

Introduction

When pricing financial derivatives, models that are based on a high-dimensional underlying stochastic process (or several processes) are chosen when there are many sources of risk. Multiple sources of risk may arise when pricing derivatives on multiple assets such as basket options or when trying to capture the dynamics of entire yield curves or forward curves. Monte Carlo simulation is the method of choice when pricing with such models due to its distinct advantage - which is that its convergence rate is typically independent of the dimension of the process.

Unfortunately, simulation methods in the case of early-exercisable products are not without their drawbacks, the principal one being that the optimal exercise strategy along each simulated path is unknown. While the optimal strategy may be approximated to some degree, any simulation using a suboptimal strategy leads to a lower bound on the price. The severity of such a mispricing is often ameliorated by using a more advanced simulation method to compute an upper bound on the price, giving the seller of the option indicative bounds of where the true price should be.

The combination of these two methods is known as *primal-dual* simulation where the *primal* method refers to that which gives us a lower bound and the *dual* refers to that which yields an upper one. By coupling these two approximate solutions, primal-dual simulation achieves a measure of the distance of these approximations to the optimal solution.

The quality of approximated exercise strategies are in turn heavily reliant on how well value functions are approximated at each of the time steps. In this thesis, we study

simulation methods based on two different ways of approximating value functions. The first way is the well-known regression based-method where the value function is represented by a suitably parameterized set of basis functions. The second is based on what is known as a convex switching system (CSS) which approximates value functions by a linear combination of piecewise-linear convex functions.

The main application of the methods in this thesis will be the pricing of multiple-exercise options. Therefore, in Section 1.1, we introduce multiple-exercise options and explain their raison d'etre, which is primarily to hedge against volumetric risk - the uncertainty in the desired quantity of the underlying to be bought or sold at a predetermined price.

In Section 1.2, we highlight the seminal works in the area of optimal multiple stopping and give an intuitive explanation as to why multiple-exercise options sit within this framework.

Connections between multiple stopping and primal-dual methods are drawn in Sections 1.3 and 1.4. In these sections, we provide a review of primal-dual methods for the pricing of the Bermudan option. Advances in these methods and recent attempts to extend them to the multiple-exercise case are discussed.

In Section 1.5, we turn our attention to CSS methods. We describe the philosophy behind it and how it seeks to overcome the some of the shortcomings of regression-based methods.

1.1 Multiple-exercise options

The multiple-exercise option is an early-exercisable derivative that gives the holder a fixed number of rights to buy or sell the underlying at a predetermined price at any time from its commencement up to and including its expiry. The American/Bermudan option is a special case of the multiple-exercise option where the holder has only one right.

Various methods such as trees, stochastic meshes and finite difference methods have been proposed for computing prices for both kinds of options. However, Monte Carlo simulation becomes the method of choice for computing prices for multiple-exercise options when the dimension of the underlying process is high.

While multiple-exercise options appear in interest rate markets, they are more predominant in various commodity markets in particular oil, gas and electricity. A concise summary of the different types of multiple-exercise options can be found in Carmona and Ludkovski [2010]. We will describe one such contract that will be of relevance to this thesis - the *swing option* but stress that the methods developed in this thesis are flexible enough to price other options with multiple-exercise features.

Swing options are widely used energy markets and play an important role in risk management for energy providers. To see why, consider the simple example of an electric company that requires gas on a daily basis and wishes to hedge against the uncertainty of daily prices over a given period. They may enter into an agreement with an investment bank to buy 10,000 MMBtu (million British Thermal Units) per day at a fixed price of \$2.00/MMBtu to meet their average generating needs. However, suppose there is a cold snap on one of the days resulting in an unexpected increase in demand for electricity and thus gas. The electric company therefore needs to somehow hedge against the uncertainty of daily gas prices *and* the uncertainty of demand.

A swing option addresses both issues by giving the purchaser a periodic delivery of the underlying over a fixed period of time at an initial contracted volume while also affording the right to vary (or "swing") this volume over the life of the contract subject to certain restrictions. Such restrictions may include a minimum/maximum of the daily volume purchased, a minimum/maximum of the total volume purchased and/or a limit to the maximum number of times you can "swing" this contract. Another popular restriction is the observance of a minimum waiting times between exercises which is also known as a *refraction period*. Furthermore, we note the existence of contracts with added provisions that allows for the holder to violate such restrictions at the cost of incurring penalties. For the specification of a multi-year swing contract with a large number of different provisions, we refer the reader to Chiarella et al. [2012].

1.2 Optimal multiple stopping

The relationship between the single-exercise case (the Bermudan option) and optimal stopping has been the subject of numerous papers and is already well known in the area of financial mathematics. Optimal stopping is a topic in probability theory whose

treatment deserves more than what we can provide for in this thesis. We thus refer the reader to the monograph Peskir and Shiryaev [2006]. The pricing multiple-exercise options naturally finds itself under the optimal multiple stopping umbrella.

As the later chapters are mathematically rigorous and notation heavy, we shall attempt to provide some intuition behind the classification of the multiple-exercise option as an optimal multiple stopping problem. Central to this classification is the concept of information which we will now explain. The probabilistic view of stochastic processes is that any realisation of a process has already occurred but we as the observers are denied knowledge of the path it has taken going forwards. However, time is a friend who gradually reveals this path to us, giving us more information about the process. Finding the set of times at which it is optimal in a probabilistic sense to exercise these rights is an exercise of making the best decision given the information we have so far.

Let us define the difference between the value of an option with $N-1$ rights and that of an option with N rights as the *marginal value of the N-th right*. Note that in terms of information, the marginal value will be known to us at any point in time as it is an *expectation* of the discounted future payoff of an extra right. In probability theory, a stopping time is a random variable whose value is interpreted as the time at which a given stochastic process exhibits a certain behavior of interest. This "behavior of interest" can only be determined by the information available to up until that point. The stochastic process in question here is the value of the immediate payoff of a given right and the "behavior of interest" is the time at which it exceeds its marginal value. Modelling the optimal exercise time as a stopping time enforces the subtle point that the decision to exercise at any given date has to be based only on information available up until that point, and subsequently eliminates the possibility of choosing a "better" time if one had the ability of clairvoyance! We conclude this discussion by stating that the value of the multiple-exercise option is the expected value of the cumulative sum of processes which are evaluated at the optimal set of stopping times.

Carmona and Touzi [2008] were the first to identify the connection between the pricing of multiple-exercise options and the optimal multiple stopping paradigm in probability theory. While their paper was motivated by the need to price such options, the existence of a set of optimal stopping times under the assumption of continuity of the underlying dynamics was proven. They study the option in the case of Geometric Brownian Motion

and by introducing appropriate Snell envelopes, show that the optimal multiple-stopping problem can be reduced to a sequence of ordinary optimal stopping problems that can be solved iteratively. The reduction of the multiple stopping problem to a series of single stopping problems goes back to the seminal work of Haggstrom [1967].

Perhaps the most important contribution of their paper was that it sparked a renewed interest in the area of optimal multiple stopping, which had until then received very little attention by probability theorists. What followed was an extension of the work of their to the more general case of linear diffusions in Carmona and Dayanik [2008]. Furthermore, Kobylanski et al. [2011] proved the existence result for optimal multiple stopping times in the case where the underlying process is right continuous with left limits. The development of dual approaches to optimal multiple stopping problems soon followed and will be included in the discussion on dual simulation methods for multiple-exercise options in Section 1.4.

1.3 Primal simulation methods

In terms of numerical solutions to the problem, the multiple-stopping framework naturally extends itself to Monte-Carlo simulation in the following way - if we had a way of approximating the optimal stopping times, then we can evaluate the aforementioned expected cumulative sum by taking the average of sums of simulated stopped processes. This is known as primal simulation.

Over the last two decades, we have witnessed an explosion in the literature on simulation methods for single-exercise as well as multiple-exercise options. Primal simulation methods include the popular regression-based least squares Monte Carlo (LSM) method (Carriere [1996], Longstaff and Schwartz [2001], Tsitsiklis and Van Roy [2001]) and the iterative exercise boundary method of Ibanez and Zapatero [2004]. These methods have been successfully applied to the multiple-exercise case in Meinshausen and Hambly [2004] and Ibanez [2004] respectively. Improved lower bounds for Bermudans can be obtained using the practical policy iteration method of Kolodko and Schoenmakers [2006] and is extended to the multiple-exercise case in Bender and Schoenmakers [2006]. Further improvements to regression-based methods and the policy iteration method can

be found in Beveridge et al. [2013]. The following pseudo-algorithm provides a general idea of how these methods work for the single-exercise case.

Pseudo-algorithm 1: Single-exercise primal simulation

1. Find an approximate exercise strategy.
2. Simulate a set of paths of the state variables, terminating each path once the strategy indicates exercise and compute the discounted payoff there.
3. Average the discounted payoffs over all paths.

These methods differ in their implementation of this framework. We provide a quick summary of the LSM method, which is by far the most popular and will be the method of choice in this thesis.

The LSM method produces exercise strategies by a functional approximation of the continuation value of the option at each step. The continuation value of the option at a given time is the value of the option if exercise does not occur there. Suppose $\{X_t\}_{t=0,\ldots,T} = \{X_0, \ldots, X_T\}$ is a d-dimensional discrete time Markov Chain that represents the state variables on the time interval $0, 1, \ldots, T$. At time $t = 0, 1, \ldots, T$, let $Z_t(X_t)$ denote the cash flow upon exercise and $C_t(X_t)$ be the continuation value of the option. The LSM method first requires us to make a judicious choice of basis functions $(\psi_{t,r} : \mathbb{R}^d \to \mathbb{R}^d, r = 1, \ldots, R_t)$ where the number of functions R_t may be allowed vary with time. The approximate value of the option at time t obtained via a parameterized set of basis functions is given as

$$c_t(X_t) = \sum_{r=1}^{R_t} \beta_{t,r} \psi_{t,r}(X_t)$$

where $\beta_{t,1}, \ldots, \beta_{t,R_t}$ are the coefficients that need to be determined. The following procedure describes how one uses a sequence of regressions to compute these coefficients.

Pseudo-algorithm 2: Determining basis function coefficients for 1 right

1. Simulate a sufficiently large number of paths of $\{X_t\}_{t=0,\ldots,T}$.

2. We start at the penultimate time step $T-1$ and determine the paths that are in-the-money. For each of these paths, we evaluate the cash flows occurring at T and discount it to $T-1$. This will give us the regressands. The regressor will be the basis functions evaluated at $T-1$ (i.e. $\psi_{T-1,1}(X_{T-1}), \ldots, \psi_{T-1,R_{T-1}}(X_{T-1})$) for each path. Perform a regression to find the coefficients $\beta_{T-1,1}, \ldots, \beta_{T-1,R_{T-1}}$.

3. For each t from $T-2$ to 1, the regressors are again the basis functions evaluated at of the in-the-money paths. The regressands for each path are given by finding $Z_{t+k}(X_{t+k})$ where $k = \min\{1, \ldots, T-t : Z_{t+k}(X_{t+k}) \geq c_{t+k}(X_{t+k})\}$ and discounting it to time t. Each $c_{t+k}(X_{t+k})$ can be computed since we have already computed the necessary coefficients in the previous steps. Perform a regression to find the coefficients $\beta_{t,1}, \ldots, \beta_{t,R_t}$.

We have this seen that the LSM method obtains an exercise strategy at any time for a given realisation of the state variable by comparing the approximated continuation value and the immediate payoff from exercise.

The following pseudo-algorithm provides a general idea of how these methods work for the multiple-exercise case for an option with N rights where N is a positive integer.

Pseudo-algorithm 3: Multiple-exercise primal simulation

1. Find an approximate exercise strategy for the first right.

2. For $n = 2$ To N: Use the approximate exercise strategy for the $(n-1)$-st right to approximate an exercise strategy for the (n)-th right.

3. Simulate a set of paths of the state variables.

4. On each path, find the sequence of times at which the exercise strategy for each of the remaining right indicates exercise and sum the discounted payoffs at these times.

5. Average the sums of discounted payoffs over all paths.

To compute exercise strategies and lower bounds throughout the next two chapters, we shall use a straightforward extension of the LSM method to the multiple-exercise case that is described in Section 3.1 of Schoenmakers [2012]. At times $t = 0, \ldots, T$, let

$C_t^{(n)}(X_t)$ denote the continuation value of the multiple-exercise option with n remaining rights and $c_t^{(n)}(X_t)$ be its approximation using the set of basis functions

$$(\psi_{t,r}^{(n)} : \mathbb{R}^d \to \mathbb{R}^d, r = 1, \ldots, R_t^{(n)}).$$

where

$$c_t^{(n)}(X_t) = \sum_{r=1}^{R_t^{(n)}} \beta_{t,r}^{(n)} \psi_{t,r}^{(n)}(X_t)$$

with coefficients $(\beta_{t,r}^{(n)})_{t=1,\ldots,R_t^{(n)}}$.

The coefficients for the n-th right are found in the same way as we did for the single-exercise case except that the regressands are now given by the payoff plus the continuation value and the condition for exercise at a time t is given by $Z_t(X_t) + c_t^{(n-1)}(X_t) \geq c_t^{(n)}(X_t)$.

Pseudo-algorithm 4: Determining basis function coefficients for n rights

1. Simulate a sufficiently large number of paths of $\mathbf{X}$.

2. Determine the paths that are in-the-money at $T-1$. For these paths, the regressands are the cash flows $Z_T(X_T)$ occurring at T that are discounted to $T-1$. The regressants will be the basis functions evaluated at $T-1$ along each path (i.e. $\psi_{T-1,1}^{(n)}(X_{T-1}), \ldots, \psi_{T-1,R_{T-1}}^{(n)}(X_{T-1})$) and we perform a regression to find the coefficients $\beta_{T-1,1}^{(n)}, \ldots, \beta_{T-1,R_{T-1}}^{(n)}$.

3. For each t from $T-2$ to 1, we find the regressands for all paths that are in-the-money at t by finding $Z_{t+k}(X_{t+k}) + c_{t+k}^{(n-1)}(X_{t+k})$ where $k = \min\{1, \ldots, T-t : Z_{t+k}(X_{t+k}) + c_{t+k}^{(n-1)}(X_{t+k}) \geq c_{t+k}^{(n)}(X_{t+k})\}$ along each of these paths and discounting it to t.

A key point to note is that Monte Carlo will only give us the theoretical value of the option when exercise is optimal along each path. Since the exercise strategies have been approximated, the primal method gives us a lower bound, hence the need to find an upper bound on the price.

1.4 Dual simulation methods

While the extension of primal simulation methods and their enhancements in the case of the Bermudan option to the multiple-exercise case is fairly trivial, the same cannot be said for dual methods.

The upper bound method for the Bermudan option involves a dual formulation of the original optimal stopping problem in the martingale sense (see Davis and Karatzas [1994]) and was proposed independently by Rogers [2002] and Haugh and Kogan [2004]. In theory, equality between the option's price and its dual is achieved if the martingale of the dual is the martingale part of the Doob-Meyer decomposition of the option's price process. In Joshi [2007], an interpretation of the problem is provided for whereby the dual is characterised as the seller's price, with the attainment of the optimal martingale analogous to achieving a perfect hedge. The strategy to approximate such a hedge is given by Andersen and Broadie [2004] in the form of an algorithm which constitutes a key part of their primal-dual method. To construct a perfect hedge, one requires the optimal exercise strategy, which is of course unattainable in practice. As the Anderson and Broadie (AB) hedge is constructed using a suboptimal exercise strategy and leads to hedging errors that are additive, this method is commonly known as the *additive dual*.

A competing upper bound method that utilises the multiplicative version of the Doob decomposition was developed by Jamshidian [2004]. Joshi [2007] demonstrated that by changing the hedging strategy, the "seller's price" interpretation of upper bound still holds. Furthermore, it was shown that an AB-style construction of the hedge for Jamshidian's upper bound still produces a high-biased estimate even in the presence of sub-simulations. Under this method, hedging errors are multiplicative, and so this method is aptly termed the *multiplicative dual*. While it is unclear which method gives a "better" upper bound, the multiplicative method results in a simulated random variable which has higher variance (see Chen and Glasserman [2007]).

In order to compute hedges, one needs to compute martingale increments. The martingale increments in the AB algorithm are computed via sub-simulations where at each step where the exercise strategy indicates exercise, one has to replace the continuation value of the option with its Monte Carlo counterpart, thus making it computationally expensive. Belomestny et al. [2009] avoided the use of sub-simulations by approximating

martingale increments using a regression estimate of a discretized Clark−Ocone derivative but it has been noted by practitioners that the method requires considerable care in its implementation as the coefficients are more unstable than the martingale itself (see Andersen and Piterbarg [2010]). Joshi and Tang [2014] achieved sub-simulation-free upper bounds for interest rate Bermudan options by first using regression to estimate Deltas and then constructing the martingales by substituting the product with its Delta hedge. Their method was found to give significantly tighter bounds (or smaller duality gaps) compared to the Belomestny et al. [2009] method. However, both methods yielded duality gaps that were significantly worse than those given by the AB method. For completeness, we also mention the multi-level dual approach of Belomestny et al. [2013] where the martingale dual is constructed in the spirit of Giles [2008].

Meinshausen and Hambly [2004] extended the additive dual result to the multiple-exercise case, where their method requires optimisation over a family of martingales and stopping times. Schoenmakers [2012] derived an additive dual that only required optimisation over a family of martingales and not stopping times. His method is commonly known as the "pure" martingale dual. Bender [2011] extended the work of Meinshausen and Hambly [2004] to the more realistic case of the swing option which accommodates volume constraints and in Bender et al. [2013], a further generalization of the additive dual formulation of the multiple stopping problem was given which provided a more flexible framework for capturing certain idiosyncratic features of swing options such as refraction periods and volume constraints. It is important to mention that these formulations sit within the more general framework of information relaxation duals proposed by Brown et al. [2010] and this is discussed in Chandramouli and Haugh [2012].

An efficient implementation of Schoenmaker's result can be found in Balder et al. [2013]. However, the method approximates the martingales using the same regression-based technique of Belomestny et al. [2009] and so suffers from similar drawbacks. In Bender et al. [2013], a recursive algorithm was provided for that avoided the use of such regression techniques. However, their method doesn't capitalise on one of the most desirable properties of the AB algorithm - which is the potential of avoiding having to perform sub-simulations at every step when computing conditional expectations.

This desire to incorporate this key strength provides us with the motivation for Chapter 2. By using the hedging approach of Joshi [2007], we derive an additive dual that is

expressed in terms of a single hedge. This approach not only provides a more intuitive understanding of the problem but also allows us to devise an additive dual algorithm whose efficiency rivals that of the one given in Balder et al. [2013] and yet avoids the issues of the Belomestny et al. [2009] approach that was highlighted above. Numerical results demonstrate that the method yields comparatively smaller duality gaps than those in Meinshausen and Hambly [2004].

While there has been significant progress in the area of dual representations for multiple-stopping problems in the recent years, none of these results have yet to incorporate a dual of the multiplicative kind thus providing the motivation for Chapter 3.

In a recent paper, Joshi and Tang [2014] showed that the multiplicative method for the Bermudan option can be made to be just as competitive as the additive one if a control variate is used. In Chapter 3, we derive the multiplicative dual for the multiple-exercise option in the form of an a.s. result. Unlike what is done in Chapter 2, the approach we take in this chapter will not involve hedging arguments. From this, we then derive the first known multiplicative upper bound that competes against the Schoenmakers' result in the same way the Joshi-Tang upper bound rivals the Rogers/Haugh-Kogan one. While the result requires the path-wise maximisation of a rather convoluted objective function over a set of random times, we shall see that one can decomposes the computation into a series of optimisations over one random time, thus making it more tractable. Each step of the algorithm requires us to approximate the optimal hedges and in doing so, we naturally extend the AB algorithm to the multiplicative multiple-exercise case.

1.5 The convex switching system

Optimal stopping problems are a subset of the larger class of Markov decision problems. Markov decision problems arise when a given problem involves a Markov decision process - a process whose transition laws are influenced by an action taken by a decision maker. Markov decision theory provides a mathematical framework for modeling decision making under uncertainty when decisions affect the transition law of the process.

The theoretical underpinnings of Markov decision theory are now well-understood. Mathematically rigorous treatments are available in the textbooks Bäuerle and Rieder [2011],

Bertsekas [1995], Feinberg et al. [2002], Hernández-Lerma and Lasserre [1996] and Puterman [2009]. However, practical applications remain persistently challenging despite the rich arsenal of theoretical tools available.

In this context, the area of *approximate dynamic programming* (ADP) grew from attempts at providing practical and implementable heuristics to overcome the limitations of standard stochastic programming techniques when faced with high dimensional state spaces. At the same time, ADP also provides theoretical insights as to why these heuristics perform well in practice.

In order to control a large system, a practical approach to overcome the high-dimensionality of the state space is to achieve a finite discretization of it. Alternatively, one can rely on an efficient approximation of functions on this space. In this spirit, function-based methods, a subclass of ADP methods, suggest to approximate value functions on the state space. We have previously mentioned one such example of a function-based method - the LSM method.

The issue of convergence of the LSM method in the context of an optimal stopping problem was first addressed in Clément et al. [2002]. There, they proved that the approximate solution converges to the true one given a sufficiently large dimension of the basis function space and a sufficiently large number of Monte Carlo samples to conduct the regression, which are commonly known as *training paths*. Their work was later extended to the case of more general classes of stochastic control problems in Belomestny et al. [2010].

These analysis served to highlight the two major shortfalls of the LSM method that were already well-known to practitioners. The first is the difficulty in the selection of an appropriate set of basis functions, especially for high-dimensional state spaces. The second shortfall was in determining an adequate number of training paths given that an insufficient number may lead to a non-converging oscillation of the numerical solution as the number of simulation paths used in its computation increases. In fact, Glasserman and Yu [2004] showed that the number of paths required for convergence grows exponentially in the degree of the approximating polynomials in the comparatively trivial case of Brownian motion.

In order provide an alternative methods that does not suffer from these problems, the convex switching system (CSS) Hinz [2014] was proposed. The fundamental idea behind the method is to approximate any convex function by a suitable collection of its sub-gradients at chosen points on a grid in the state space. By doing so, we move from the problem of choosing an appropriate set of basis functions to the arguably easier problem of specifying an appropriate grid.

Given such a grid, a collection of subgradients yields what is known as a *sub-gradient envelope* of the function. A subgradient envelope is the maximum of the collection of subgradients at each point in the state space. The subgradient envelope is convex and as the density of the grid increases, we obtain a sequence of convex functions. A comparison of Figure 1.1 with Figure 1.2 demonstrates the "tightening" of a function's subgradient envelope to itself for a single dimensional state space as we increase the density of the grid from 2 points to 5 on the same interval.

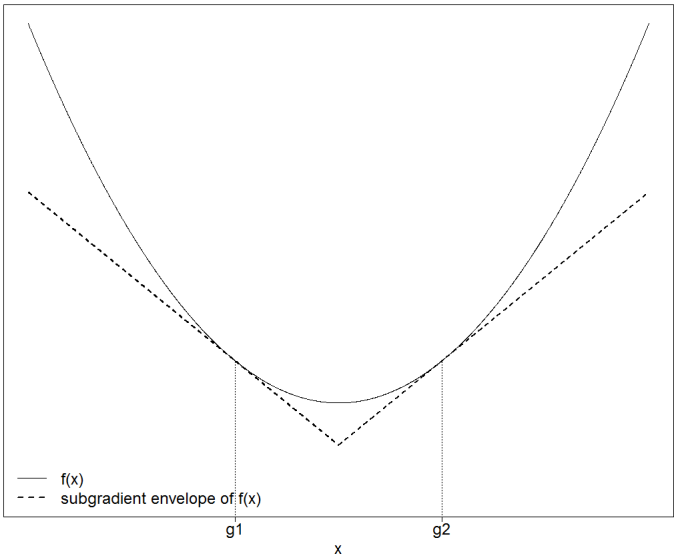

FIGURE 1.1: A function $f(x)$ and its subgradient envelope given grid points $g1$ and $g2$.

Ultimately, we observe the convergence of this sequence of subgradient envelopes to the function on any bounded set in $\mathbb{R}^d$ for a sufficiently dense grid. In fact, a direct application of Theorem 10.8 in Rockafellar [1997] proves that this convergence is in fact uniform.

In Chapter 4, we provide a recap of how this idea is used to derive an ADP method as done in Hinz [2014]. The algorithm in its raw form is computationally intensive

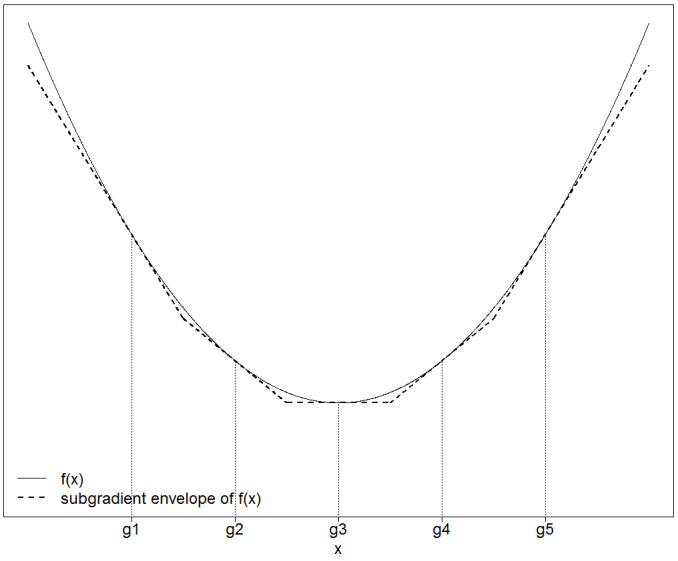

FIGURE 1.2: A function $f(x)$ and its subgradient envelope given grid points $g1$ to $g5$.

and we devise several methods to improve its efficiency. Furthermore, we look at how to overcome the restriction of convexity. Finally, we present a CSS-based primal-dual simulation algorithm to assess the distance to optimality of the approximated solutions.

Chapter 2

The additive dual for the multiple exercise option

This chapter is based on the work done in Chiarella et al. [2014] and Joshi and Yap [2014]. An earlier version of the paper, Yap [2013], was presented at the 26th Annual PhD Conference in Economics and Business. In this chapter, we derive an additive dual for the multiple-exercise derivative and its associated algorithm. In Section 2.1, we first construct a method to represent an exercise strategy for the multiple-exercise option. We then show how to construct stopping times using this construct. Using these stopping times, we express the value of an option as a multiple-stopping problem and show how one may interpret this as the buyer's price. In order to derive the dual of the problem, we first introduce the notion of a hedge in the multiple-exercise framework in Section 2.2. In Section 2.3, we derive, in the spirit of Joshi [2007], an additive dual representation of the multiple-exercise option using hedging arguments from the perspective of the seller of the option. The benefit of this approach is that the hedging strategy manifests itself in the construction of a hedge to be used in the Monte Carlo evaluation of the dual problem and by exploiting the properties of this hedge, we can obtain an efficient algorithm. This algorithm is provided for in Section 2.4 and it can be seen as the natural extension of the Andersen and Broadie (AB) algorithm Andersen and Broadie [2004] to the multiple-exercise case. By studying in the number of sub-simulations required, we show that our method requires less computational effort than the one given by Bender et al. [2013]. A comparison of our numerical results to those

computed using the Meinhausen and Hambly (MH) algorithm Meinshausen and Hambly [2004] is given in Section 2.5. Section 2.6 concludes the chapter.

2.1 Preliminaries

We work on the discrete time interval $\mathbb{T} = \{0, 1, ..., T\}$ and the underlying process we work with is a discrete time Markov chain that is defined on a filtered probability space $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ where $\mathbb{F} := \{\mathcal{F}_t : t \in \mathbb{T}\}$. We choose an asset price deflator which we call the *bond numeraire* and assume the existence of an equivalent measure $\mathbb{Q}$ under which deflated prices are martingales. In this chapter, the representation of any payoff or price will be given in deflated terms.

We consider the multiple-exercise option, $V^{(N)}$, that permits the holder to exercise an integer-valued number of rights, N, where each exercise occurs sequentially on $\mathbb{T}$. This option will be denoted by $V^{(N-n)}$ after the n-th right has been exercised. Upon exercising a right, the buyer receives a payoff of Z_t where $\{Z_t\}$ is a non-negative $\mathbb{F}$-adapted stochastic process that satisfies

$$\sum_{t=0}^{T} \mathbb{E}\big[Z_t \big| \mathcal{F}_0\big] < \infty.$$

We assume that all rights must be exercised to ensure that all stopping times are finite and to avoid having to discuss cases where not all rights are exercised before T. This is not a restriction since a right exercised out of the money simply pays zero.

A natural starting point for our analysis is to first consider the case of the Bermudan option where $N = 1$. The associated pricing problem at time t is well-known and is given by

$$V_t^{(1)} = \operatorname*{ess\,sup}_{\tau(t)} \mathbb{E}[Z_{\tau(t)} | \mathcal{F}_t] \tag{2.1}$$

where $\tau(t)$ is a stopping time taking values on $\{t, \ldots, T\}$. The financial interpretation of a stopping time is that it corresponds to the given exercise strategy for the option and the pricing problem is to find the optimal exercise strategy $\tau^*(t)$ such that

$$V_t^{(1)} = \mathbb{E}[Z_{\tau^*(t)} | \mathcal{F}_t].$$

Andersen and Broadie [2004] observe that it is necessary to find exercise strategies not just for products starting at time 0 but also for freshly created products at each exercise time. They therefore found it convenient to define exercise indicators which indicate whether a product that has not been previously exercised should be exercised at a given time, t. Thus they define an adapted indicator process $\{l_t\}_{t=0,\dots,T}$ where l_t equals 1 if exercise should occur at t and zero otherwise. To force exercise at time T, we take $l_T = 1$. Stopping times are then implied by

$$\tau_t = \inf\{s \in \{t, \dots, T\} : l_t = 1\} \tag{2.2}$$

where τ_t denotes the first instance at time t or later at which an option that is alive at time t should be exercised according to the given strategy. Such a construction of stopping times yields a lower bound $\hat{V}_t^{(1)}$ on the price where

$$\hat{V}_t^{(1)} := \mathbb{E}[Z_{\tau(t)}|\mathcal{F}_t] \leq V_t^{(1)}$$

and where equality holds if and only if $\tau(t) = \tau^*(t)$.

We need a similar construction for multi-exercisable options. One solution is simply to define a two-dimensional set $l(k,t)$ set of indicators with $l(k,t) = 1$ if and only if an option with k rights remaining should be exercised at time t. However, it is fairly straightforward to see that since payoffs are non-negative, it is always optimal to exercise an option with k rights if it is optimal to exercise one with r rights for some $r < k$. It is therefore more convenient to make a definition in terms of the least number of rights that indicates exercise where N will mean that no contracts should be exercised and 0 will specify that exercise should occur regardless of how many are left. For the remainder of this chapter, we will use n to represent any integer from 0 to $N-1$.

Definition 2.1. We define a sequence of random variables $\kappa := \{\kappa_t\}_{t=0,\dots,T}$ to be an **exercise function** for a multi-exercisable option with N rights if

(i) κ_t is a non-negative and integer-valued for $t \in \mathbb{T}$,

(ii) κ_t is $\mathcal{F}_t$-measurable for $t \in \mathbb{T}$,

(iii) $\kappa_t \geq N - t - 1$ for $t < N - 2$ and

(iv) $\kappa_{T-j} < j$.

We let $\mathcal{K}$ denote the set of all such sequences.

Property (iii) specifies the natural condition that the n-th exercise can only occur after time n. Property (iv) enforces exercise by the last possible time so that an option can never have more rights than exercise dates remaining. Clearly, this sequence implies exercise indicators for an option with k rights remaining via

$$l(k,t) = \mathbb{1}_{\{k>\kappa_t\}},$$

and thus also implies a stopping time strategy for it as above. In fact, we get a 3-dimensional family of stopping times. The dimensions are the start time of the product, t, the number of rights, k, the product has at t and the right, i, we are considering the strategy for. We will consider i to be indexed according to the original number of rights N with the lowest-indexed right exercised first. So if $k = N - n$, the first right to be exercised will be $i = N - n + 1$. We shall now state this formally in the following definition.

Definition 2.2. For a given exercise function κ, a stopping vector $\tau^{\kappa,N-n,t}$ is a collection of stopping times $(\tau_{n+1}^{\kappa,N-n,t}, \ldots, \tau_N^{\kappa,N-n,t})$ such that

$$\tau_{n+1}^{\kappa,N-n,t} = \min\{s \in \{\max(t,n), \ldots, T-N+n+1\} : N-n > \kappa_s\} \tag{2.3}$$

and

$$\tau_{n+i}^{\kappa,N-n,t} = \min\{s \in \{t, \ldots, T-N+n+i\} : s > \tau_{n+i-1}^{\kappa,N-n,t}, N-n-i+1 > \kappa_s\}, \tag{2.4}$$

for $i = 2, \ldots, N-n$.

The following properties of implied stopping times are immediate.

Proposition 2.3. *For any stopping function κ, for every $n = 0, \ldots, N-1$ and for every $t \geq n$,*

$$\tau_{N-n+1}^{\kappa,N-n,t} = \mathbb{1}_{\{N-n>\kappa_t\}} t + \mathbb{1}_{\{N-n\leq\kappa_t\}} \tau_{N-n+1}^{\kappa,N-n,t+1}.$$

Furthermore, we have that if $N - n > \kappa_t$ then for $i > N - n$, we have

$$\tau_{i+1}^{\kappa,N-n,t} = \tau_i^{\kappa,N-n-1,t}.$$

The quadruple-labelled stopping time notation can become a little cumbersome. When we are working in the context of fixed exercise function, κ, we will drop it from notation. We will also drop $N-n$ when the value of n is clear and write $\tau_{n+i}(t)$ instead. However, it is important to realize that, in general,

$$\tau_{n+2}^{\kappa,N-n,t} \neq \tau_{n+2}^{\kappa,N-n-1,t}$$

and the distinction is important. Here we have the difference between the second stopping time with $N-n$ rights remaining and the first one with $N-n-1$ left.

We restate the result of Carmona and Touzi [2008] in our language:

Theorem 2.4. *The price of the multiple-exercise option is*

$$V_t^{(N-n)} = \operatorname*{ess\,sup}_{\kappa} \mathbb{E}\left[\sum_{i=n+1}^{N} Z_{\tau_i^{\kappa,N-n,t}} \middle| \mathcal{F}_t \right], \tag{2.5}$$

where the essential supremum is taken over all exercise functions.

Note that the set of admissible strategies is more restrictive than that in Carmona and Touzi [2008], however, this is not a problem since the optimal strategy still lies in the set we consider.

The value may also be expressed as

$$V_t^{(N-n)} = \max\left(C_t^{(N-n)}, Z_t + C_t^{(N-n-1)}\right) \tag{2.6}$$

where $C_t^{(N-n)}$ is the continuation value of $V^{(N-n)}$ that is the value of a forward starting multiple-exercise option $C^{(N-n)}$ at time t. We can apply the Carmona–Touzi result and we have

$$C_t^{(N-n)} = \operatorname*{ess\,sup}_{\kappa} \mathbb{E}\left[\sum_{i=n+1}^{N} Z_{\tau_i^{\kappa,N-n,t+1}} \middle| \mathcal{F}_t \right] \tag{2.7}$$

where the first exercise time of $C^{(N-n)}$ occurs at $t+1$ if we are at t.

Fixing N, n, κ, it is immediate that the stopping times have the following properties for $i > n$, similar to the single exercisable case:

Proposition 2.5.

(i) $\max(t, i-1) \leq \tau_i(t) \leq T - N + i \quad a.s.,$

(ii) $\tau_i(t) \neq t \Rightarrow \tau_i(t) = \tau_i(t+1)$ *for* $t < T - N + i$ *and*

(iii) $\tau_{i-1}(t) < \tau_i(t)$, *for* $i < N$.

Let the optimal stopping vector be $\tau^{*,N-n,t} = (\tau^*_{n+1}(t), \ldots, \tau^*_N(t))$. We have, via the Bellman principle, that

$$\tau^*_{n+1}(t) = \min\{s \in \{t, \ldots, T - N + n + 1\} : Z_s + C_s^{(N-n-1)} \geq C_s^{(N-n)}\}$$

and for $i = n + 2, ..., N - 1$,

$$\tau^*_i(t) = \min\{s \in \{\tau_{i-1}(t), \ldots, T - N + i\} : Z_s + C_s^{(N-i)} \geq C_s^{(N-i+1)}\}.$$

We will now show that $\tau^{*,N-n,t}$ is associated to an optimal exercise function, κ^*. First, note that for any two positive integers n_1 and n_2 where $n_1 < n_2 \leq N$, it is always optimal to exercise $V^{(n_2)}$ whenever it is optimal to exercise $V^{(n_1)}$. Therefore, the optimal exercise times for $V^{(n_1)}$ will be a subset of that of $V^{(n_2)}$. With this, we prove the following.

Proposition 2.6. *There exists a* κ^* *such that for all* $n = 0, \ldots, N - 1$ *and at every* $t = n, \ldots, T - N + n$, *any stopping vector* $\tau^{\kappa,N-n,t}$ *constructed via* (3.1) *and* (2.4) *is optimal.*

Proof. Take the optimal exercise strategies for each number of rights. These define exercise indicator functions

$$l^*(k, t) = 1 \text{ if and only if } \tau^{*,k,t}_k = t.$$

If we sum these over k we get a function $f(t)$ such that options with $N, N - 1, \ldots, N - f(t) + 1$ should be exercised. Let

$$\kappa(t) = N - f(t),$$

and we are done. □

If one has a multi-exercisable derivative and an exercise function, κ, one can define an associated triggerable derivative by requiring that the exercise strategy implied by κ is used. In what follows, we will assume the existence of the products $\hat{V}^{\kappa,(N-n)}$ and

$\hat{C}^{\kappa,(N-n)}$ whose triggering are defined by the stopping times (3.1) and (2.4) that have been constructed using such a κ. The product $\hat{V}^{\kappa,(N-n)}$ can be triggered now while $\hat{C}^{\kappa,(N-n)}$ can only be triggered at the next time step. Their respective values at time t, $\hat{V}_t^{\kappa,(N-n)}$ and $\hat{C}_t^{\kappa,(N-n)}$, are given by their expectations in the $\mathbb{Q}$ measure. Since κ is not optimal their prices bound the prices of the options constructed using the optimal strategy from below

$$\hat{V}_t^{\kappa,(N-n)} := \mathbb{E}\left[\sum_{i=n+1}^{N} Z_{\tau_i^\kappa(t)}\middle|\mathcal{F}_t\right] \leq V_t^{(N-n)} \tag{2.8}$$

and

$$\hat{C}_t^{\kappa,(N-n)} := \mathbb{E}\left[\sum_{i=n+1}^{N} Z_{\tau_i^\kappa(t+1)}\middle|\mathcal{F}_t\right] \leq C_t^{(N-n)}. \tag{2.9}$$

In the next section, these products will be used to derive an upper bound for $V_0^{(N)}$.

It is important to stress at this point that our objective in this chapter is to realize an upper bound for the optimal stopping price in the chosen measure $\mathbb{Q}$. If the market is incomplete, the optimal stopping price and the upper bound will inevitably vary according to the choice of measure. In addition, the upper bound is derived via the use of a hedging martingale whose price process depends on the choice of measure. The construction of this hedge assumes the ability to buy and sell derivatives with a given pay-off and exercise strategy at the price implied by this previously fixed measure. Let us emphasize that we are not proposing nor requiring that the hedges would exist in a real market. Instead, this is a mathematical device that simplifies our derivations and renders them more intuitive.

We *define* the price processes of these contracts to be their expectation prices under the $\mathbb{Q}$-measure. Augmenting a market by allowing the ability to trade contracts at their expectation prices will not introduce arbitrage since all instruments have martingale prices processes, and it will also not change the expectation price of our original contract. In particular, we allow the trading of $\hat{V}^{\kappa,(N-n)}$ and $\hat{C}^{\kappa,(N-n)}$ for $n = 0, \ldots, N-1$ for any exercise function κ. Thus any upper bound derived in the measure $\mathbb{Q}$ for this augmented market is also an upper bound for our original market in that measure.

In order to describe the pricing problem and its dual under the buyer/seller paradigm with greater clarity, we will denote the buyer and the seller as male and female respectively. One can interpret the optimal stopping time price as the buyer's price for the

option $V^{(N-n)}$ where the $\tau_i^*(t)'s$, for $i = n+1, \ldots, N-1$ are the times at which he exercises. The dual can be interpreted as the seller's price where she constructs and maintains a hedge that is optimal in the sense that her hedging portfolio will always dominate the buyer's, regardless of how he exercises. An optimal hedge yields equality between the two prices.

The inherent difficulty in solving the dual problem from a computational perspective lies in the construction of such a hedge. Andersen and Broadie [2004] showed how one could construct the optimal hedge for a Bermudan option if the optimal exercise strategy was known, and how a "good" hedge could still be obtained using a suboptimal strategy. For a suboptimal hedge, the seller's price will be greater than that of the buyer's, thus yielding an upper bound. In the next section, we show how to construct similar hedges for the multi-exercisable option. In particular, we show that a hedge constructed using the optimal exercise strategy will be optimal in the sense that it yields equality between the buyer's and the seller's prices.

2.2 The construction of hedging martingales for multiple-exercise options

We shall first define the notion of a hedge in the multi-exercise setting. In the next section, we shall see that any such hedge will yield an upper bound on the price of the multi-exercisable product. We then describe a hedging strategy that is conditional on an exercise strategy. We see in Section 3.4 that for an optimal exercise strategy, one can construct the necessary hedges used to derive the additive dual formulation of the problem from a hedging perspective.

The hedge that we will construct will possess two properties. First, it will be dependent on the exercise times of the buyer. We shall see later that it is ideal (but not necessary) for the seller to have as part of her portfolio a position in a product with the same number of exercise rights (or triggers) remaining as the product being hedged. Therefore, reactivity of the hedge to the buyer's exercise times becomes necessary. Secondly, the hedge will also satisfy the self-financing property throughout its life.

Now consider the option $V^{(N)}$ at time 0 and suppose that the buyer does not follow the optimal exercise strategy but an arbitrary one that is represented by the stopping times

$0 \leq \tau_1 < \cdots < \tau_N \leq T$. The seller will construct her hedge at time 0, and the hedge will be a function of the buyer's exercise times. Let us now describe the mathematical notation for a hedge. At the outset the seller's hedge is denoted by $\hat{H}^{(N)}$ as she is hedging against an option with N rights. After n rights have been exercised, this hedge will be denoted by

$$\hat{H}^{(N-n),\tau_1,\ldots,\tau_n}.$$

The value of such a hedge at a time $t \in \mathbb{T}$ is given by $\hat{H}_t^{(N-n),\tau_1,\ldots,\tau_n}$.

Definition 2.7. A **hedging strategy** is a sequence of self-financing portfolios

$$\hat{H}^{(N)}, \hat{H}^{(N-1),\tau_1}, \ldots, \hat{H}^{(0),\tau_1,\ldots,\tau_N}$$

such that at each of the remaining exercise times at which the buyer exercises τ_i, $i = 1, \ldots, N$, the hedge $\hat{H}^{(N-i+1),\tau_1,\ldots,\tau_{i-1}}$ is dissolved and reconstituted as a new hedge $\hat{H}^{(N-i),\tau_1,\ldots,\tau_i}$ while maintaining the self-financing property

$$\hat{H}_{\tau_i}^{(N-i+1),\tau_1,\ldots,\tau_{i-1}} = \hat{H}_{\tau_i}^{(N-i),\tau_1,\ldots,\tau_i}.$$

Furthermore, taking $\tau_0 := 0$, we also have that

$$\mathbb{E}\left[\hat{H}_{\tau_i}^{(N-i+1),\tau_1,\ldots,\tau_{i-1}} - \hat{H}_{\tau_{i-1}}^{(N-i+1),\tau_1,\ldots,\tau_{i-1}} \middle| \mathcal{F}_{\tau_{i-1}}\right] = 0, \quad i = 1, \ldots, N-1.$$

For the remainder of the chapter, all hedges will be associated with a given hedging strategy that defined as above and we will denote the set of all such hedges by $\mathcal{H}$. In Section 3.4, it will be shown that an arbitrary hedging strategy yields an upper bound on the price.

Now consider a sequence of functions $\{f_i\}_{i=1,..,N}$ such that after each t_i, $i = 1, \ldots, N$, the seller's hedge

$$\hat{H}^{(N-i),\tau_1,\ldots,\tau_{i-1}} \text{ consists of } f_i(\hat{H}^{(N-i+1),\tau_1,\ldots,\tau_{i-1}}, Z_{\tau_i}) \text{ units of } \hat{C}^{\kappa,(N-i)}$$

and a position in numeraire bonds. By defining the functions $\{f_i\}_{i=k+1,..,N}$ to be such that the hedging portfolio always contains one unit of the triggerable product where

the remaining number of triggers is the same as the remaining number of rights of the product being hedged, we obtain the following hedging strategy.

Definition 2.8. The seller is said to follow a κ**-additive hedging strategy** if her initial hedge consists of one unit of $\hat{V}^{\kappa,(N)}$ and she adheres to the following trading rules for each of the k-th rights where $k = 1, ..., N$ at each subsequent time:

- If $\hat{V}^{\kappa,(k)}$ is not triggered and the buyer does not exercise, then seller should not change her portfolio.
- If $\hat{V}^{\kappa,(k)}$ is triggered and the buyer exercises to receive a cash payoff and $\hat{C}^{\kappa,(k-1)}$. The seller then reinvests the cash payoff in numeraire bonds.
- If $\hat{V}^{\kappa,(k)}$ is not triggered and the buyer exercises, the seller should sell $\hat{V}^{\kappa,(k)}$, buy one unit of $\hat{C}^{\kappa,(k-1)}$ and invest the net proceeds in numeraire bonds.
- If $\hat{V}^{\kappa,(k)}$ is triggered and the buyer does not exercise, the seller should sell $\hat{C}^{\kappa,(k-1)}$, buy $\hat{C}^{\kappa,(k)}$ and invest the net proceeds in numeraire bonds.

Because the seller invests any payoff from exercise (or from the product being triggered) in numeraire bonds, we shall assume that the buyer does the same. Note that this does not constitute an restriction on investor preferences, and that we are merely assuming this to construct one type of upper bound. We acknowledge that the buyer may just as easily choose to use the payoff to buy more of the option. In this instance, we would have to change the seller's hedging strategy, and it would lead to the derivation of a different upper bound. We will leave the variation in the buyer's reinvestment strategy to future work.

Now if $\kappa = \kappa^*$, then the κ^*-additive hedging strategy will always yield perfect hedges. To see this, we first observe that the seller will only trade products that are triggered at points of optimal exercise. Since the trigger is optimal and the payoff is greater than the marginal value of one right at optimal points of exercise, the seller's hedging portfolio will grow whenever the buyer does not exercise at points that are optimal. Therefore such a hedge will always dominate the buyer's portfolio. We shall call hedges that are associated with the κ^*-additive hedging strategy **additively optimal** and they will be represented by the same notation as arbitrary hedges but with the accent removed (e.g. $H^{(N-n),t_1,...,t_n}$).

2.3 The additive dual for multiple exercise

We define the *gains from trade*, $A_t^{(N-n),t_1,\ldots,t_n}$ to be the discounted value of the accumulated investment in numeraire bonds generated by any additive trading strategy that the seller adopts after the buyer has exercised the first n rights at $t_1, \ldots t_n$. Recall from the κ-additive hedging strategy that the seller's hedge is always a combination of one unit of the option and a position in numeraire bonds. The position in numeraire bonds comprises of both the accumulated payoffs and the gains from trade.

The additive dual is given by the following theorem.

Theorem 2.9. *Given the set of additively optimal hedges, the additive dual of the multiple-exercise option is given as*

$$V_0^{(N)} = H_0^{(N)} + \max_{0 \le t_1 < \ldots < t_N \le T} \left(\sum_{i=1}^{N} Z_{t_i} - H_{t_N}^{(0),t_1,\ldots,t_N} \right) \quad a.s. \tag{2.10}$$

Proof. By definition, an optimal hedge is constructed by the seller following the κ^*-additive hedging strategy. Upon the first exercise time of the buyer, τ_1, the discounted value of the seller's hedge is given as

$$H_{\tau_1}^{(N-1),\tau_1} = Z_{\tau_1} + C_{\tau_1}^{(N-1)} + A_{\tau_1}^{(N-1),\tau_1} \tag{2.11}$$

and after the n-th exercise at τ_n we have

$$H_{\tau_n}^{(N-n),\tau_1,\ldots,\tau_n} = \sum_{i=1}^{n} Z_{\tau_i} + C_{\tau_n}^{(N-n)} + A_{\tau_n}^{(N-n),\tau_1,\ldots,\tau_n}. \tag{2.12}$$

Therefore, at the final exercise date, we have that

$$H_{\tau_N}^{(0),\tau_1,\ldots,\tau_N} = \sum_{i=1}^{N} Z_{\tau_i} + A_{\tau_N}^{(0),\tau_1,\ldots,\tau_N}. \tag{2.13}$$

By following the κ^*-additive hedging strategy, any increments in the gains from trade will be positive whenever the buyer exercises at sub-optimal points and zero at other times (again we state that one can verify this by considering all four possible scenarios in Definition 2.8 using κ^*). Therefore,

$$A_{\tau_N}^{(0),\tau_1,\ldots,\tau_N} \ge 0.$$

We have thus constructed a hedge that will dominate the buyer's portfolio which is of value $\sum_{i=1}^{N} Z_{\tau_i}$ regardless of how the buyer exercises. Therefore we have that $H_0^{(N)} = V_0^{(N)}$ and the path-wise result

$$\max_{0\leq t_1<\ldots<t_N\leq T}\left(\sum_{i=1}^{N} Z_{t_i} - H_{t_N}^{(0),t_1,\ldots,t_N}\right) = 0 \quad a.s. \tag{2.14}$$

where equation (2.14) encapsulates the situation where the buyer exercises at the set of time such that the gains of trade will be zero. If we restrict ourselves to using hedges of zero initial value, we then have that

$$H_0^{(N)} = \max_{0\leq t_1<\ldots<t_N\leq T}\left(\sum_{i=1}^{N} Z_{t_i} - (H_{t_N}^{(0),t_1,\ldots,t_N} - H_0^{(N)})\right) \quad a.s.$$

and since $V_0^{(N)} = H_0$, we obtain (2.10). □

In order to construct an additively optimal hedge, we need to trade products whose triggers are dependent on the optimal exercise strategy. For any other strategy, a sub-optimal hedge is obtained. This results in an upper bound on the price of the option.

Theorem 2.10. *For an arbitrary set of hedges* $\{\hat{H}^{(N-n)}\}, n = 1, \ldots, N-1$, *we obtain the following upper bound on the price of the multiple-exercise option*

$$V_0^{(N)} \leq \hat{H}_0^{(N)} + \mathbb{E}\left[\max_{0\leq t_1<\ldots<t_N\leq T}\left(\sum_{i=1}^{N} Z_{t_i} - \hat{H}_{t_N}^{(0),t_1,\ldots t_N}\right)\middle|\mathcal{F}_0\right]. \tag{2.15}$$

Proof. We have from (2.5) that

$$V_0^{(N)} = \operatorname*{ess\,sup}_{0\leq\tau_1<\cdots<\tau_N\leq T}\mathbb{E}\left[\sum_{i=1}^{N} Z_{\tau_i}\middle|\mathcal{F}_0\right]. \tag{2.16}$$

Recall from Definition 2.7 that $H^{(N)}$ is a martingale between 0 and stopping time τ_1 and the hedges $H^{(N-k)}$ are martingales between any two stopping times τ_k and τ_{k+1}. Therefore

$$\begin{aligned} 0 &= \operatorname*{ess\,sup}_{0\leq\tau_1<\cdots<\tau_N\leq T}\mathbb{E}\left[\sum_{i=1}^{N}\mathbb{E}\left(\hat{H}_{\tau_{i-1}}^{(N-i),\tau_1,\ldots,\tau_{i-1}} - \hat{H}_{\tau_i}^{(N-i),\tau_1,\ldots,\tau_{i-1}}\middle|\mathcal{F}_{\tau_{i-1}}\right)\middle|\mathcal{F}_0\right] \\ &= \operatorname*{ess\,sup}_{0\leq\tau_1<\cdots<\tau_N\leq T}\mathbb{E}\left[\sum_{i=1}^{N}\hat{H}_{\tau_{i-1}}^{(N-i),\tau_1,\ldots,\tau_{i-1}} - \hat{H}_{\tau_i}^{(N-i),\tau_1,\ldots,\tau_{i-1}}\middle|\mathcal{F}_0\right] \end{aligned}$$

where the second equality holds true by the Tower property.

Combining equations (2.16) and (2.17),

$$\begin{aligned} V_0^{(N)} &= \operatorname*{ess\,sup}_{0\le\tau_1<\cdots<\tau_N\le T} \mathbb{E}\left[\sum_{i=1}^{N}\left(Z_{\tau_i}+\hat{H}_{\tau_{i-1}}^{(N-i),\tau_1,\ldots,\tau_{i-1}}-\hat{H}_{\tau_i}^{(N-i),\tau_1,\ldots,\tau_{i-1}}\right)\middle|\mathcal{F}_0\right] \\ &= \operatorname*{ess\,sup}_{0\le\tau_1<\cdots<\tau_N\le T} \mathbb{E}\Bigg[\sum_{i=1}^{N} Z_{\tau_i}-\hat{H}_{\tau_N}^{(1),\tau_1,\ldots\tau_{N-1}}+(\hat{H}_{\tau_{N-1}}^{(1),\tau_1,\ldots\tau_{N-1}}-\hat{H}_{\tau_{N-1}}^{(2),\tau_1,\ldots\tau_{N-2}}) \\ &\quad +\cdots+(\hat{H}_{t_1}^{(N-1),t_1}-\hat{H}_{t_1}^{(N)})+\hat{H}_0^{(N)}\Bigg|\mathcal{F}_0\Bigg] \\ &= \hat{H}_0^{(N)}+\operatorname*{ess\,sup}_{0\le\tau_1<\cdots<\tau_N\le T}\mathbb{E}\left[\sum_{i=1}^{N} Z_{\tau_i}-\hat{H}_{\tau_N}^{(0),\tau_1,\ldots\tau_N}\middle|\mathcal{F}_0\right] \\ &\le \hat{H}_0^{(N)}+\mathbb{E}\left[\max_{0\le t_1<\ldots<t_N\le T}\left(\sum_{i=1}^{N} Z_{t_i}-\hat{H}_{t_N}^{(0),t_1,\ldots t_N}\right)\middle|\mathcal{F}_0\right]. \end{aligned}$$

We have that the terms in the small brackets after the second equality sum to zero because we recall from Definition 2.7 that the hedges must satisfy the self-financing property at the times at which the buyer exercises

$$\hat{H}_{\tau_i}^{(N-i+1),\tau_1,\ldots,\tau_{i-1}}=\hat{H}_{\tau_i}^{(N-i),\tau_1,\ldots,\tau_i},\quad i=1,\ldots,N.$$

□

2.4 The collapsing tree method for the additive dual

For the remainder of this chapter we shall adopt a fixed exercise strategy κ. The seller will adopt the κ-additive hedging strategy correspondingly. Since κ is fixed, we shall drop it from the terms $\hat{V}^{\kappa,(\cdot)}$ and $\hat{C}^{\kappa,(\cdot)}$. Therefore, $\hat{V}^{(\cdot)}$ and $\hat{C}^{(\cdot)}$ will be used to denote the products that seller hedges with.

The expectation in (2.15) is evaluated via Monte Carlo simulation where each path requires the computation of the path-wise maximum of an objective function over the set of all possible exercise times. In other words, we need to compute

$$\max_{0\le t_1<\ldots<t_N\le T}\left(\sum_{i=1}^{N} Z_{t_i}-\hat{H}_{t_N}^{(0),t_1,\ldots,t_N}\right)=-\min_{0\le t_1<\cdots<t_N\le T} A_{t_N}^{(0),t_1,\ldots,t_N} \tag{2.17}$$

along each path. The equality comes from the fact that $\hat{H}_{t_N}^{(0),t_1,\ldots,t_N} = \sum_{i=1}^{N} Z_{t_i} + A_{t_N}^{(0),t_1,\ldots,t_N}$. If we proceed naively, we would be computing $\frac{T+1!}{N!(T+1-N)!}$ combinations of gains from trades per path and this yields a factorial time order of computational complexity.

In this section, we provide an algorithm which we call the *collapsing tree method* that reduces the complexity of this computation. We show that by only keeping track of the worst-performing hedges for all possible number of remaining rights at each time step, we can dramatically reduce the number of hedges that needs to be computed.

Let us first introduce some further notation that will be used from this point onwards. We will take n to be any integer from 1 to N for the remainder of this section. We will use the indicator function $\mathbb{1}_{\{\tau_n(t)=t\}}$ to denote whether the strategy indicates exercise of the n-th right at time t and use $\mathbb{1}_{\{\tau_n(t)\neq t\}}$ to denote when it does not. We shall let

$$\underline{\mathrm{A}}_t^{(n)} = \min_{0\leq t_1<\cdots<t_{N-n}<t} A_t^{(n),t_1,\ldots,t_{N-n}}$$

denote the lowest gains from trade for n remaining rights at time t given all possible combinations of exercise times chosen by the buyer prior to t. The lowest gains from trade characterizes the worst-performing hedge for the same number of remaining rights and its value at t will be denoted by $\underline{\mathrm{H}}_t^{(n)}$.

In (2.17), we try to finding the worst-performing hedge/lowest gains from trade given that all rights have been exhausted across all possible sets of exercise times. The key to the collapsing tree method involves the exploitation of the fact that the worst-performing hedge given n remaining rights at a given time step t could only have evolved from $t-1$ in a maximum of two ways - either from the worst-performing hedge given n remaining rights or from the worst-performing hedge given $n+1$ remaining rights.

First, note that the gains from trade at the last exercise time of the buyer t_N can be written as the gains from trade up to that time plus any gains arising from the product not being triggered there

$$A_{t_N}^{(0),t_1,\ldots,t_N} = A_{t_N}^{(1),t_1,\ldots,t_{N-1}} + \mathbb{1}_{\{\tau_N(t_N)\neq t_N\}}\big(\hat{V}_{t_N}^{(1)} - Z_{t_N}\big).$$

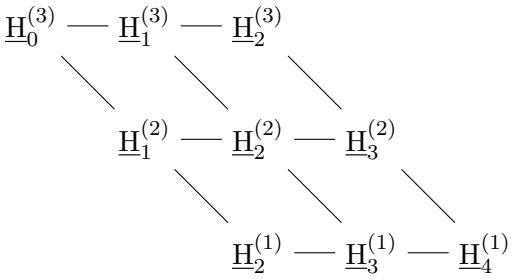

FIGURE 2.1: A tree of hedges for 3 exercise rights and 5 exercise dates

Therefore, we have that

$$\begin{aligned}\min_{0\le t_1<\cdots<t_N\le T} A_{t_N}^{(0),t_1,\ldots,t_N} &= \min_{t_N\in(N-1,\ldots,T)} \min_{0\le t_1<\cdots<t_N} A_{t_N}^{(0),t_1,\ldots,t_N} \\ &= \min_{t_N\in(N-1,\ldots,T)} \left[\underline{A}_{t_N}^{(1)} + \mathbb{1}_{\{\tau_N(t_N)\neq t_N\}}\left(\hat{V}_{t_N}^{(1)} - Z_{t_N}\right)\right].\end{aligned} \tag{2.18}$$

To compute the minimum of $A_{t_N}^{(0),t_1,\ldots,t_N}$ over all $t_1,\ldots,t_N$, we then

(i) fix the value t_N,

(ii) compute $\underline{A}_{t_N}^{(1)}$ for the fixed t_N,

(iii) use (2.18) to compute $\min_{0\le t_1<\cdots<t_N\le T} A_{t_N}^{(0),t_1,\ldots,t_N}$ for the fixed t_N, and

(iv) compute $\min_{0\le t_1<\cdots<t_N\le T} A_{t_N}^{(0),t_1,\ldots,t_N}$ over all t_N's.

However, we still need a way of finding each $\underline{A}_{t_N}^{(1)}$'s and this is where the collapsing tree approach comes in.

We shall start explaining the collapsing tree using a concrete example and generalize from there. The following example will be used to demonstrate how the computations are done and the efficiency of this method compared to the naive approach in a simple case of an option with $N = 3$ rights and 5 dates ($T = 4$). Consider the following worst-performing hedges as shown in Figure 2.1.

To compute the dual along the path, we need to compute

$$\max_{0\le t_1<t_2<t_3\le 4}(Z_{t_1} + Z_{t_2} + Z_{t_3} - \hat{H}_{t_3}^{(t_1,t_2,t_3)}) = -\min_{0\le t_1<t_2<t_3\le 4} A_{t_3}^{(0),t_1,t_2,t_3}. \tag{2.19}$$

At $t = 0$, there are no other hedges so the worst-performing hedge is $\hat{H}_0^{(3)} = \underline{\mathrm{H}}_0^{(3)}$. In the following table, we list the different hedges of the seller given the different possible exercise strategies of the buyer thus far. At this time there are only two possible scenarios.

Buyer's strategy	Hedge	Payoff	Option	Gains from trade
$t_1 = 0$	$\hat{H}_1^{(2),0}$	Z_0	$V_1^{(2)}$	$A_1^{(2),0} = \mathbb{1}_{\{\tau_1(0)\neq 0\}}(V_0^{(3)} - C_0^{(2)} - Z_0)$
$t_1 \neq 0$	$\hat{H}_1^{(3)}$	0	$V_1^{(3)}$	$A_1^{(3)} = \mathbb{1}_{\{\tau_1(0)=0\}}(Z_0 + C_0^{(2)} - C_0^{(3)})$

TABLE 2.1: Hedges at $t = 1$ for different exercise strategies of the buyer at $t = 0$

The worst-performing hedges $\underline{\mathrm{H}}_1^{(3)}$ and $\underline{\mathrm{H}}_1^{(2)}$ are thus given by $\hat{H}_1^{(3)}$ and $\hat{H}_1^{(2),0}$ respectively.

Going from $t = 1$ to $t = 2$, we then have four scenarios:

Buyer's strategy	Hedge	Payoff	Option	Gains from trade
$t_1 = 1$	$\hat{H}_2^{(2),1}$	Z_1	$\hat{V}_2^{(2)}$	$A_2^{(2),1} = A_1^{(2),1}$ $+\mathbb{1}_{\{\tau_1(1)\neq 1\}}(\hat{V}_1^{(3)} - Z_1 - \hat{C}_1^{(2)})$
$t_1 \neq 1$	$\hat{H}_2^{(3)}$	0	$\hat{V}_2^{(3)}$	$A_2^{(3)} = A_1^{(3)}$ $+\mathbb{1}_{\{\tau_2(1)=1\}}(Z_1 + \hat{C}_1^{(2)} - \hat{C}_1^{(3)})$
$t_1 = 0, t_2 \neq 1$	$\hat{H}_2^{(2),0}$	Z_0	$V_2^{(2)}$	$A_2^{(2),0} = A_1^{(2),0}$ $+\mathbb{1}_{\{\tau_2(1)=1\}}(Z_1 + \hat{C}_1^{(1)} - \hat{C}_1^{(2)})$
$t_1 = 0, t_2 = 1$	$\hat{H}_2^{(1),0,1}$	$Z_0 + Z_1$	$\hat{V}_2^{(1)}$	$A_2^{(1),0,1} = A_1^{(2),0}$ $+\mathbb{1}_{\{\tau_2(1)\neq 1\}}(\hat{V}_1^{(2)} - Z_1 - \hat{C}_1^{(1)})$

TABLE 2.2: Hedges at $t = 2$ for different exercise strategies of the buyer up to $t = 1$

We can see that $\underline{\mathrm{H}}_2^{(3)}$ and $\underline{\mathrm{H}}_2^{(1)}$ are trivially equal to $\hat{H}_2^{(3)}$ and $\hat{H}_2^{(1),0,1}$ respectively. However, $\underline{\mathrm{H}}_2^{(2)}$ could have evolved from either $\underline{\mathrm{H}}_1^{(3)}$ or $\underline{\mathrm{H}}_1^{(2)}$ and so to find $\underline{\mathrm{H}}_2^{(2)}$, we need to compute $\underline{\mathrm{A}}_2^{(2)}$, the minimum of $A_2^{(2),0}$ and $A_2^{(2),1}$.

To find $A_2^{(0),0,1,2}$ we have to see if the buyer's exercise strategies agree with the products trigger at step 2

$$A_2^{(0),0,1,2} = A_2^{(1),0,1} + \mathbb{1}_{\{\tau_3(2)=2\}}\big(\hat{V}_2^{(1)} - Z_2\big).$$

We will use the computation of the minimum of $A_{t_3}^{(0),t_1,t_2,t_3}$ when t_3 is fixed at 3 to highlight the computational gains of the method. A naive approach would have required us to take the minimum of $A_{t_3}^{(0),t_1,t_2,t_3}$ for $(t_1,t_2)=(0,1),(0,2)$ and $(1,2)$. However, $\underline{\mathrm{A}}_3^{(1)}$ is readily found by taking a minimum of two variables where

$$\underline{\mathrm{A}}_3^{(1)} = \min\left[\underline{\mathrm{A}}_2^{(1)} + \mathbb{1}_{\{\tau_3(2)=2\}}(Z_2 - \hat{C}_2^{(1)}), \underline{\mathrm{A}}_2^{(2)} + \mathbb{1}_{\{\tau_{N-n+2}(t-1)\neq t-1\}}(\hat{V}_2^{(2)} - Z_2 - \hat{C}_2^{(1)})\right].$$

We thus see that computational savings is of a polynomial order, which is important for large numbers of rights and higher values of t_N.

In general, $\underline{\mathrm{A}}^{(n)}$ may be computed along a given trajectory as follows. Suppose that at the previous time $t-1$ the buyer does not exercise an option with n remaining rights and the seller was holding hedge $\underline{\mathrm{H}}_{t-1}^{(n)}$. Then at time t, the gains from trade will be

$$\underline{\mathrm{A}}_{t-1}^{(n)} + \mathbb{1}_{\{\tau_{N-n+1}(t-1)=t-1\}}(Z_{t-1} + \hat{C}_{t-1}^{(n-1)} - \hat{C}_{t-1}^{(n)}). \tag{2.20}$$

Alternatively, let us assume at $t-1$ that the buyer had exercised the option with $n+1$ remaining rights and the seller was holding $\underline{\mathrm{H}}_{t-1}^{(n+1)}$. In this case, our gains from trade at t is

$$\underline{\mathrm{A}}_{t-1}^{(n+1)} + \mathbb{1}_{\{\tau_{N-n+2}(t-1)\neq t-1\}}(\hat{V}_{t-1}^{(n+1)} - \hat{C}_{t-1}^{(n)} - Z_{t-1}). \tag{2.21}$$

The lowest gains from trade will therefore be the lesser of (2.20) and (2.21). The following tree of hedges (Figure 2.2) tracks all the worst-performing hedges for each number of remaining exercise rights at each time.

The gains from trade will always be computed as

$$\begin{aligned}\underline{\mathrm{A}}_t^{(n)} \quad = \quad & \min\Big[\underline{\mathrm{A}}_{t-1}^{(n)} + \mathbb{1}_{\{\tau_{N-n+1}(t-1)=t-1\}}(Z_{t-1} + \hat{C}_{t-1}^{(n-1)} - \hat{C}_{t-1}^{(n)}), \\ & \underline{\mathrm{A}}_{t-1}^{(n+1)} + \mathbb{1}_{\{\tau_{N-n+2}(t-1)\neq t-1\}}(\hat{V}_{t-1}^{(n+1)} - \hat{C}_{t-1}^{(n)} - Z_{t-1})\Big]\end{aligned} \tag{2.22}$$

except in three cases. The first case is where we are still hedging an option with maximum rights. Therefore,

$$\underline{\mathrm{A}}_t^{(N)} = \underline{\mathrm{A}}_{t-1}^{(N)} + \mathbb{1}_{\{\tau_1(t-1)=t-1\}}(Z_{t-1} + \hat{C}_{t-1}^{(N-1)} - \hat{C}_{t-1}^{(N)}). \tag{2.23}$$

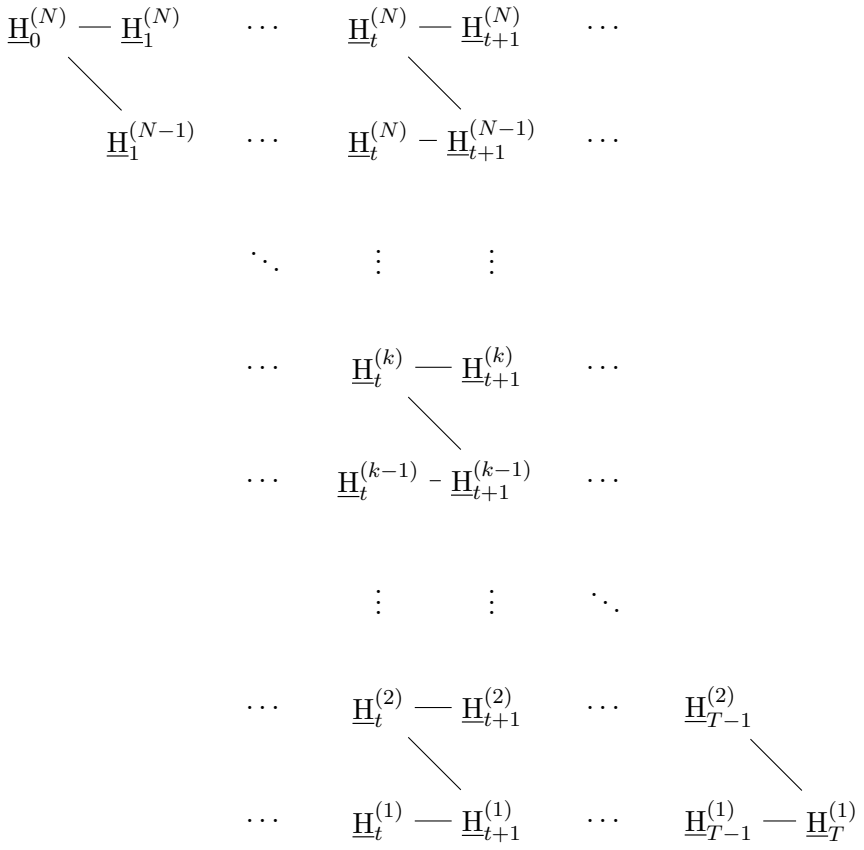

FIGURE 2.2: A tree of hedges for N exercise rights and $T+1$ exercise dates

The second case is where t is the earliest time that one can exercise the nth right. Under this case,

$$\underline{\mathrm{A}}_t^{(n)} = \underline{\mathrm{A}}_{t-1}^{(n+1)} + \mathbb{1}_{\{\tau_{N-n+2}(t-1)\neq t-1\}}(Z_{t-1} + \hat{C}_{t-1}^{(n)} - \hat{C}_{t-1}^{(n+1)}). \tag{2.24}$$

The last case is where exercise had been forced at the previous time step, $\mathbb{1}_{\{\tau_n(t-1)=t-1\}} = 1$ and so

$$\underline{\mathrm{A}}_t^{(n)} = \min\left[\underline{\mathrm{A}}_{t-1}^{(n)} + \mathbb{1}_{\{\tau_{N-n+1}(t-1)=t-1\}}(Z_{t-1} + \hat{C}_{t-1}^{(n-1)} - \hat{C}_{t-1}^{(n)}), \underline{\mathrm{A}}_{t-1}^{(n+1)}\right]. \tag{2.25}$$

2.4.1 Computational efficiency

In this sub-section we compare our method to that of Bender et al. [2013] and show the potential computational savings achieved by avoiding the computation of sub-simulations. We note that their method is more general as it encompasses multiple-exercise per period and different refraction periods for swing options. In order to provide a like-for-like

comparison with our method, let us recap their method in terms our our notation for just the single-exercise per period and unit refraction period case.

The following algorithm is applied to each path of the (outer) simulation.

- For each n from 1 to N, we initialize each θ^n_{T-n+1} as

$$\theta^n_{T-n+1} := \sum_{i=1}^{n} Z_{T-i+1}.$$

- For each t from T to 0, we compute each θ^n_t for $n = 1$ to N recursively via

$$\theta^n_t = \max\left[\theta^n_{t+1} + \hat{C}^{(n)}_{t+1} - \hat{V}^{(n)}_{t+1}, \quad Z_i + \theta^{n-1}_{t+1} + \hat{C}^{(n-1)}_{t+1} - \hat{V}^{(n-1)}_{t+1}\right].$$

The upper bound is then found by taking the average of θ^N_0 over all paths. Thus we see that unless the number of remaining rights is equal to the number of remaining exercise times (i.e. $n = T-t$), 4 sub-simulations are required for each number of remaining rights at each step. Now comparing their algorithm to ours, we observe from 2.22 that at each step and for each number of remaining rights, it is possible that no sub-simulations are required if

$$\mathbb{1}_{\{\tau_{N-n+1}(t-1)\neq t-1\}} \quad \text{and} \quad \mathbb{1}_{\{\tau_{N-n+2}(t-1)=t-1\}}.$$

However, note that in this case, at least two sub-simulations are required to compute $\underline{\mathrm{A}}^{(n+1)}_t$ and $\underline{\mathrm{A}}^{(n-1)}_t$. Conversely, if all four sub-simulations need to be computed, then each of $\underline{\mathrm{A}}^{(n+1)}_t$ and $\underline{\mathrm{A}}^{(n-1)}_t$ would require at most two sub-simulations. This means that the number of sub-simulations required is effectively halved for our method as compared to Bender et al. [2013].

2.5 Numerical results

We demonstrate our results by comparing our bounds on the price of a swing option with those given in Section 4.2 of Meinshausen and Hambly [2004]. The underlying asset S follows an exponential AR(1) process and we compute its dynamics using the following discretisation

$$\log S_t = (1-k)(\log S_{t-1} - \mu) + \mu + \sigma W_t. \tag{2.26}$$

The swing option is of maturity T, is exercisable daily and pays off $Z_t = (S_t - K)_+$ upon exercising a right. The parameters used were $S_0 = 1$, $K = 0$, $k = 0.9$, $\mu = 0$, $\sigma = 0.5$, and $T = 1000$. Interest rates are ignored. 1000 paths were used for the calibration and the evaluation of the lower bound and 20 simulations (with 50 sub-simulations) were used for the upper bound. The least-squares Monte Carlo method as described in Meinshausen and Hambly [2004] was used to find the low biased estimates and the basis function used in the regression was a polynomial of degree one. Our results are given in Table 2.3.

The simulations were performed in C++ using the Visual Studios Express 2012 Integrated Development Environment. We used an Intel Core i5-4200U 1.6Ghz processor and all runs were performed using a single thread. Computation time for the upper bound for an option with 2 rights was about 60 seconds while that of an option with 100 rights was about 5 hours. While this might seem long at first glance, we recall that the number of steps is about 1000 with approximately of 2 sub-simulations required for each right at each step, thus giving us about 200,000 sub-simulations!

It is observed that for a lower number of rights ($N < 10$), there is not much difference between our method and the MH method. However, the duality gaps as measured by the 99% confidence intervals produced by our method become significantly tighter as the number of rights increases. However, the MH method seems to yield smaller variances compared to ours as the number of rights increase too. This seems to suggest a trade-off between smaller duality gaps and certainty in the result in terms of standard deviation when deciding between the two methods. Furthermore, this decision is further complicated as the trade-off is intertwined with the number of rights.

2.6 Conclusion and future research

In this chapter we provide a new representation of the additive dual of the multiple-exercise option. In order to do so, we first introduced a new mechanism, the exercise function, which allows us to construct exercise strategies in the sense that the exercise strategy for a product with more exercise rights will always indicate exercise whenever the exercise strategy for a product with less exercise rights does.

We then define the notion of a hedge in the multiple-exercise framework. In particular, we describe a hedging strategy that allows us to perfectly hedge against the buyer of an option. This then allowed us to derive a new representation of the additive dual of the multiple-exercise option where the representation is in terms of a single hedge instead of a family of martingales.

While this approach provided an intuitive understanding of the additive dual of the multiple stopping problem from a financial perspective, the main benefit of this approach was a computational one. In particular, we showed that by monitoring the worst performing hedges, we were able to devise an efficient algorithm known as the collapsing tree approach to compute an upper bound for the multiple-exercise option.

Furthermore, we have shown that by identifying points in time where exercise is suboptimal, computational time can be saved while still obtaining an upper bound. Numerical results indicate that the method yields smaller duality gaps than ones computed using the MH method.

As we have been able to achieve an almost sure result, we have reason to believe that this approach is essentially an implementation of the "pure" martingale dual of Schoenmakers [2012]. However, we have not reconciled our approach to his result and we leave it for future work.

In Section 2.2, we noted that while we assumed that the buyer of the option invests any payoff in numeraire bonds, the buyer of the option could also choose to reinvest in the option itself, or even a combination of both. We would then need to change the hedging strategy, and this would ultimately lead to different upper bounds.

We conclude this chapter by stating that computationally, both our method and the MH method still require "good" approximations to the optimal stopping times (or an optimal exercise strategy). However, as it is not clear why our method fares better, we feel that a further analysis into how the two methods compare is warranted.

Table 2.3: Numerical results for Swing Option

						Meinshausen-Hambly		
Rights	Lower Bound	Upper Bound	Lower 99%	Upper 99%	Lower Bound	Upper Bound	Lower 99%	Upper 99%
1	4.792	4.801	4.773	4.820	4.777	4.790	4.773	4.794
2	9.099	9.141	9.048	9.192	9.024	9.085	9.016	9.091
3	12.973	13.045	12.904	13.114	12.970	13.094	12.959	13.100
4	16.810	16.878	16.765	16.923	16.786	16.899	16.773	16.906
5	20.506	20.579	20.452	20.633	20.455	20.573	20.439	20.580
10	37.370	37.485	37.300	37.555	37.334	37.531	37.305	37.540
15	52.626	52.801	52.537	52.891	52.713	52.999	52.670	53.009
20	67.123	67.204	67.097	67.230	67.105	67.513	67.050	67.525
30	93.963	94.114	93.908	94.170	93.742	94.507	93.662	94.519
40	118.619	118.810	118.548	118.882	118.457	119.611	118.353	119.625
50	141.854	142.072	141.797	142.129	141.832	143.345	141.703	143.360
60	164.433	164.714	164.352	164.795	164.112	166.020	163.960	166.037
70	185.845	186.166	185.777	186.235	185.511	187.711	185.335	187.729
80	206.669	207.042	206.602	207.109	206.045	208.682	205.844	208.702
90	226.045	226.542	225.931	226.656	225.900	228.965	225.676	228.985
100	245.529	246.052	245.451	246.130	245.157	248.630	244.910	248.651

Chapter 3

The multiplicative dual for the multiple exercise option

In this chapter, we derive the multiplicative dual for the multiple-exercise derivative and its associated algorithm. This derivation presents a different approach to the one used to derive additive dual in Chapter 2. In particular, financial arguments are avoided. In order to derive the multiplicative dual, we require a new representation of a martingale in the multiple-exercise framework and this is provided for in Section 3.2. In Section 3.3, we restate the key results in the literature for the multiplicative dual in the case of the Bermudan option in terms of our martingales. In Section 3.4, we derive the multiplicative dual for the multiple-exercise option and its associated upper bound. The algorithm to compute the multiplicative upper bound is given in Section 3.5 and the issue of the bias arising from the sub-simulations in the algorithm is naturally addressed. Section 3.6 concludes.

3.1 Preliminaries

The fundamental difficulty in solving the dual problem for the Bermudan option from a computational perspective lies in how to approximate martingales. Andersen and Broadie [2004] showed how one could construct the optimal martingale if the optimal exercise strategy was known, and how a "good" martingale could still be obtained using a suboptimal strategy. In order to derive the multiplicative dual for the multiple-exercise

option, we require a new representation of the martingale in the multiple-exercise setting. This is done in the next section.

Similar to what was done in Andersen and Broadie [2004], we will first need to approximate an exercise strategy in order to compute martingales in the multiple-exericse framework. Such as strategy will be encapsulated by a given κ. The construction of the most desirable martingale will require knowledge of $V^{(N-n)}$ and $C^{(N-n)}$. However, since we do not have them, we will use the approximations $\hat{V}^{\kappa,(N-n)}$ and $\hat{C}^{\kappa,(N-n)}$ in their place. $\hat{V}^{\kappa,(N-n)}$ and $\hat{C}^{\kappa,(N-n)}$ are multi-exercisable derivatives whose exercise are defined by the stopping times and (2.4) that have been constructed with a given κ. We shall refer to them as *triggerables* where the product is *triggered* when κ indicates exercise. The product $\hat{V}^{\kappa,(N-n)}$ can be triggered now while $\hat{C}^{\kappa,(N-n)}$ can only be triggered at the next time step. Their respective values at time t, $\hat{V}_t^{\kappa,(N-n)}$ and $\hat{C}_t^{\kappa,(N-n)}$, are given by their expectations in the $\mathbb{Q}$ measure. Note that we do not assume the existence of triggerable products in the market. We only use them as a numerical tool when performing computations. Since κ is not optimal their prices bound the prices of the options constructed using the optimal strategy from below

$$\hat{V}_t^{\kappa,(N-n)} := \mathbb{E}\left[\sum_{i=n+1}^{N} Z_{\tau_i^\kappa(t)} \middle| \mathcal{F}_t\right] \leq V_t^{(N-n)} \tag{3.1}$$

and

$$\hat{C}_t^{\kappa,(N-n)} := \mathbb{E}\left[\sum_{i=n+1}^{N} Z_{\tau_i^\kappa(t+1)} \middle| \mathcal{F}_t\right] \leq C_t^{(N-n)}. \tag{3.2}$$

3.2 Martingales in the multiple-exercise setting

Similar to the additive dual result for the multi-exercise option, the *almost sure* representation of the multiplicative dual will be in terms of a set of martingales. In this section, we will begin to construct such a set by first considering the multiplicative Doob decompositions of $V^{(N)}$ and all $C^{(N-n)}$'s on a truncated time interval $s, \ldots, T$ where s is an arbitrary time in $\mathbb{T}$.

Recall from Jamshidian [2004] that if Z_t is positive, then $V^{(1)}$ is a supermartingale that admits a multiplicative Doob decomposition

$$V_t^{(1)} = H_{s,t}^{(1)} B_{s,t}^{(1)}, \quad t = s, \ldots, T$$

where $H^{(1)}_{s,t}$ is the value of a martingale $H^{(1)}_s$ at a time t such that $H^{(1)}_{s,s} = 1$ and $B^{(1)}_s$ is a non-increasing predictable process such that $B^{(1)}_{s,s} = V^{(1)}_s$. Rescaling, we will take $H^{(1)}_{s,s} = V^{(1)}_s$ and $B^{(1)}_{s,s} = 1$.

Since $V^{(N)}$ is a supermartingale, the same decomposition holds and we now define an associated martingale.

Definition 3.1. We say that the martingale $H^{(N)}_s$ is an $(s, 0)$-martingale if there exists a predictable non-increasing process $B^{(N)}_s$ whose value at t is denote by $B^{(N)}_{s,t}$ such that $B^{(N)}_{s,s} = 1$ and

(ii)
$$V^{(N)}_t = H^{(N)}_{s,t} B^{(N)}_{s,t}, \quad t = s, \ldots, T.$$

For $n > 0$, the martingale will be associated with the multiplicative decomposition of $C^{(N-n)}$.

Definition 3.2. Let n be an integer with $1 \leq n \leq N-1$, we say that the martingale $H^{(N-n)}_s$ is an (s, n)-martingale if there exists a predictable non-increasing process $B^{(N-n)}_s$, whose value at t will be denoted by $B^{(N-n)}_{s,t}$, such that $B^{(N-n)}_{s,s} = B^{(N-n)}_{s+1,s+1} = 1$ and

$$C^{(N-n)}_s = H^{(N-n)}_{s,s} B^{(N-n)}_{s,s}, \tag{3.3}$$
$$V^{(N-n)}_t = H^{(N-n)}_{s,t} B^{(N-n)}_{s+1,t}, \quad t = s+1, \ldots, T. \tag{3.4}$$

We shall see in Section 3.4 that an *almost sure* representation of the multiplicative dual can be expressed in terms of a set of (s, n)-martingales. Given an arbitrary set of times $0 = t_0 < t_1 < \cdots < t_{N-1} < T$, we shall also represent an arbitrary set of martingales by

$$\hat{H}^{(N)}_0, \hat{H}^{(N-1)}_{t_1}, \ldots, \hat{H}^{(1)}_{t_{N-1}}$$

where the value of each $\hat{H}^{(N-n)}_{t_n}$ at time $t \geq t_n$ is given by $\hat{H}^{(N-n)}_{t_n,t}$. The set of all uniformly integrable martingales that are of positive value will be given by $\mathcal{H}$. In what follows, it will be convenient to let $\hat{H}^{(N-n)}_{t_n,t_n} = \hat{C}^{\kappa,(N-n)}_{t_n}$ (or $\hat{H}^{(N-n)}_{t_n,t_n} = \hat{V}^{\kappa,(N)}_0$ if $n = 0$) for some κ.

3.3 The multiplicative dual for single exercise

We begin by stating Jamshidian's result for the case of the Bermudan option in terms of our notation.

Theorem 3.3. *Let $N = 1$. For non-zero payoffs we have that*

$$V^{(N)} = \inf_{\hat{H}_0^{(N)} \in \mathcal{H}} \mathbb{E}\left[\max_{0 \le t \le T} \frac{Z_t}{\hat{H}_{0,t}^{(N)}} \hat{H}_{0,T}^{(1)} \middle| \mathcal{F}_0 \right] \tag{3.5}$$

where the infimum is attained if $\hat{H}_0^{(N)} = H_0^{(N)}$.

Chen and Glasserman [2007] showed that while the Monte Carlo estimate of the additive dual has zero variance given an optimal martingale, this is not the case for the multiplicative dual due to the variance of $H_{0,T}^{(1)}$. Joshi and Tang [2014] showed that one can recover a zero-variance result by using the terminal value of the martingale as a control variate. The zero-variance result is an almost sure statement. We restate their result in terms of our notation.

Theorem 3.4. *Let $N = 1$. For non-zero payoffs we have that*

$$V_0^{(N)} = \inf_{\hat{H}_0^{(N)} \in \mathcal{H}} \mathbb{E}\left[\hat{H}_{0,0}^{(N)} + \max_{0 \le t \le T} \left(\frac{Z_t}{\hat{H}_{0,t}^{(N)}} \right) \hat{H}_{0,T}^{(1)} - \hat{H}_{0,T}^{(N)} \middle| \mathcal{F}_0 \right] \tag{3.6}$$

$$= H_{0,0}^{(N)} + \max_{0 \le t \le T} \left(\frac{Z_t}{H_{0,t}^{(N)}} \right) H_{0,T}^{(N)} - H_{0,T}^{(N)} \quad a.s. \tag{3.7}$$

In Joshi [2007], it was shown that one can construct an approximation $\hat{H}_0^{(1)}$ to the optimal martingales $H_0^{(1)}$ using a multiplicative version of the Andersen-Broadie algorithm and the triggerable products $C^{\kappa,(1)}$ and $V^{\kappa,(1)}$. This is done by first finding an approximation to the optimal exercise strategy which is encapsulated by some κ and then constructing the martingales via the following recursion:

Initialisation

$$\hat{H}_{0,0}^{(1)} = B_{0,0}^{(1)} V_0^{\kappa,(1)} \quad \text{where} \quad B_{0,0}^{(1)} = 1$$

Recursion

$$\begin{aligned}\hat{H}^{(1)}_{0,t} &= B^{(1)}_{0,t} V^{\kappa,(1)}_t \quad \text{where} \\ B^{(1)}_{0,t} &= B^{(1)}_{0,t-1}\left[1 + \mathbb{1}_{\{\tau_1(t-1)=t-1\}}\left(\frac{Z_{t-1}}{C^{\kappa,(1)}_{t-1}} - 1\right)\right], \quad t = 1, \ldots, T.\end{aligned}$$

3.4 The multiplicative dual for multiple exercise

To extend the Joshi-Tang result to the case where $N > 1$, we first make the rather obvious observation that $V^{(N)}$ can be thought of as a product which can be exercised once with a payoff at t equal to $Z_t + C^{(N-1)}_t$. This payoff has the same value as $X^{(N)}_t$ units of $C^{(N-1)}_t$ where

$$X^{(N)}_t := \frac{Z_t}{C^{(N-1)}_t} + 1.$$

In general, by exercising an option with $n < N$ remaining rights, one receives a payoff of value equal to

$$X^{(n)}_t := \frac{Z_t}{C^{(n-1)}_t} + 1$$

units of the forward-starting option $C^{(n-1)}$. For the remainder of the chapter, the $X^{(\cdot)}_t$'s will be referred to as multiples. We thus obtain the following representation for the multiple-exercise option with N rights

$$V^{(N)}_0 = \inf_{\hat{H}^{(N)}_0 \in \mathcal{H}} \mathbb{E}\left[H^{(N)}_{0,0} + \max_{0 \le t_1 \le T-N+1}\left(\frac{X^{(N)}_{t_1} C^{(N-1)}_{t_1}}{H^{(N)}_{0,t_1}}\right) H^{(N)}_{0,T-N+1} - H^{(N)}_{0,T-N+1} \middle| \mathcal{F}_0\right] \tag{3.8}$$

$$= H^{(N)}_{0,0} + \max_{0 \le t_1 \le T-N+1}\left(\frac{X^{(N)}_{t_1} C^{(N-1)}_{t_1}}{H^{(N)}_{0,t_1}}\right) H^{(N)}_{0,T-N+1} - H^{(N)}_{0,T-N+1} \quad a.s. \tag{3.9}$$

The second subscript $T - N + 1$ on the martingale $H^{(N)}_{0,T-N+1}$ indicates the last time $V^{(N)}$ can be exercised.

By starting at a time t_n rather than 0, we observe that the Joshi-Tang result also holds for forward starting options

$$\begin{aligned} C_{t_n}^{(N-n)} &= \inf_{\hat{H}_{t_n}^{(N-n)} \in \mathcal{H}} \mathbb{E}\left[\hat{H}_{t_n,t_n}^{(N-n)} + \max_{t_n < t_{n+1} \leq T-N-n+1} \left(\frac{X_{t_{n+1}}^{(N-n)} C_{t_{n+1}}^{(N-n-1)}}{\hat{H}_{t_n,t_{n+1}}^{(N-n)}} \right) \hat{H}_{t_n,T-N+n+1}^{(N-n)} \right. \\ &\quad \left. - \hat{H}_{t_n,T-N+n+1}^{(N-n)} \middle| \mathcal{F}_{t_n} \right] \qquad (3.10) \\ &= H_{t_n,t_n}^{(N-n)} + \max_{t_n < t_{n+1} \leq T-N-n+1} \left(\frac{X_{t_{n+1}}^{(N-n)} C_{t_{n+1}}^{(N-n-1)}}{H_{t_n,t_{n+1}}^{(N-n)}} \right) H_{t_n,T-N+n+1}^{(N-n)} \\ &\quad - H_{t_n,T-N+n+1}^{(N-n)} \quad a.s. \qquad (3.11) \end{aligned}$$

We are now in a position to state and prove our main results.

Theorem 3.5. *Let $X_{t_N}^{(1)} = Z_{t_N}$. For positive payoffs, we have that*

$$\begin{aligned} V_0^{(N)} &= \max_{0 \leq t_1 < \ldots < t_N \leq T} \left[\sum_{j=0}^{N-1} \left(H_{t_j,t_j}^{(N-j)} - H_{t_j,T-N+j+1}^{(N-j)} \right) \prod_{i=0}^{j-1} \frac{H_{t_i,T-N+i+1}^{(N-i)}}{H_{t_i,t_{i+1}}^{(N-i)}} X_{t_{i+1}}^{(N-i)} \right. \\ &\quad \left. + \prod_{i=0}^{N-1} \frac{H_{t_i,T-N+i+1}^{(N-i)}}{H_{t_i,t_{i+1}}^{(N-i)}} X_{t_{i+1}}^{(N-i)} \right] \quad a.s. \qquad (3.12) \end{aligned}$$

Proof. Let $g_n = H_{t_n,t_n}^{(N-n)} - H_{t_n,T-N+n+1}^{(N-n)}$ and

$$h_n := \frac{H_{t_n,T-N+n+1}^{(N-n)}}{H_{t_n,t_{n+1}}^{(N-n)}} X_{t_{n+1}}^{(N-n)}$$

for $n = 0, ..., N-1$. We have from (3.9) that

$$\begin{aligned} V_0^{(N)} &= \max_{t_1} \left(H_{0,0}^{(N)} - H_{0,T-N+1}^{(N)} + \frac{X_{t_1}^{(N)} C_{t_1}^{(N-1)}}{H_{0,t_1}^{(N)}} H_{0,T-N+1}^{(N)} \right) \\ &= \max_{t_1} (g_0 + h_0 C_{t_1}^{(N-1)}) \quad a.s. \qquad (3.13) \end{aligned}$$

and from (3.11) that

$$\begin{aligned} C_{t_1}^{(N-1)} &= \max_{t_2 > t_1} \left(H_{t_1,t_1}^{(N-1)} - H_{t_1,T-N+2}^{(N-1)} + \frac{X_{t_2}^{(N-1)} C_{t_2}^{(N-2)}}{H_{t_1,t_2}^{(N)}} H_{t_1,T-N+2}^{(N-1)} \right) \quad a.s. \\ &= \max_{t_2 > t_1} (g_1 + h_1 C_{t_2}^{(N-2)}) \quad a.s. \qquad (3.14) \end{aligned}$$

Since $X^{(1)}_{t_N} = Z_{t_N}$, we have by (3.7) that

$$\begin{aligned} C^{(1)}_{t_{N-1}} &= \max_{t_N > t_{N-1}} \left(H^{(1)}_{t_{N-1},t_{N-1}} - H^{(1)}_{t_{N-1},T} + \frac{X^{(1)}_{t_N}}{H^{(1)}_{t_{N-1},t_N}} H^{(1)}_{t_{N-1},T} \right) \quad a.s. \\ &= \max_{t_N > t_{N-1}} (g_{N-1} + h_{N-1}) \quad a.s. \end{aligned} \tag{3.15}$$

Using these, we can write (3.9) as

$$\begin{aligned} V^{(N)}_0 &= \max_{t_1}[g_0 + h_0 \max_{t_2 > t_1}(g_1 + h_1 C^{(N-1)}_{t_2})] \\ &= \max_{t_1 < t_2}(g_0 + h_0 g_1 + h_0 h_1 C^{(N-1)}_{t_2}) \\ &= \max_{t_1 < t_2 < \ldots < t_N}(g_0 + h_0 g_1 + h_0 h_1 g_2 + \ldots + \prod_{i=0}^{N-2} h_i C^{(1)}_{t_{N-1}}) \\ &= \max_{t_1 < t_2 < \ldots < t_N}(g_0 + h_0 g_1 + h_0 h_1 g_2 + \ldots + \prod_{i=0}^{N-2} h_i g_{N-1} + \prod_{i=0}^{N-1} h_i) \\ &= \max_{0 \le t_1 < \ldots < t_N \le T} \Big(\sum_{j=0}^{N-1} g_j \prod_{i=0}^{j-1} h_i + \prod_{i=0}^{N-1} h_i \Big) \quad a.s. \end{aligned}$$

We recover (3.12) by substituting the relevant terms. □

As the result is an a.s. one, an optimal martingale would give us a zero-variance outcome in the Monte-Carlo simulation of $V^{(N)}_0$. For arbitrary martingales, we obtain an upper bound on the price. The upper bound is encapsulated in the following result.

Corollary 3.6. *If payoffs are positive, we have that*

$$\begin{aligned} V^{(N)}_0 = \inf_{\hat{H}^{(N)}_0, \ldots, \hat{H}^{(1)}_{t_{N-1}} \in \mathcal{H}} \mathbb{E} \Bigg\{ \max_{0 \le t_1 < \ldots < t_N \le T} \Bigg[\sum_{j=0}^{N-1} \left(\hat{H}^{(N-j)}_{t_j,t_j} - \hat{H}^{(N-j)}_{t_j,T-N+j+1} \right) \\ \times \prod_{i=0}^{j-1} \frac{\hat{H}^{(N-i)}_{t_i,T-N+i+1}}{\hat{H}^{(N-i)}_{t_i,t_{i+1}}} X^{(N-i)}_{t_{i+1}} + \prod_{i=0}^{N-1} \frac{\hat{H}^{(N-i)}_{t_i,T-N+i+1}}{\hat{H}^{(N-i)}_{t_i,t_{i+1}}} X^{(N-i)}_{t_{i+1}} \Bigg] \Bigg| \mathcal{F}_0 \Bigg\} \end{aligned} \tag{3.16}$$

where the infimum is obtained when the martingales are optimal.

The perceptive reader may question whether this result is useful from a computational perspective since each multiple $X^{(N-n)}_t$ is defined in terms of some optimal stopping price $C^{(N-n-1)}_t$. However, we shall show in Section 3.5 that by redefining it using $\hat{C}^{\kappa,(N-n-1)}_t$, we still obtain an upper bound.

Proof. Consider an arbitrary set of martingales $\{\hat{H}^{(N-n)}_{t_n}\}, n = 1, ..., N-1$. We have by (3.8) and (3.10) that

$$V_0^{(N)} \leq \mathbb{E}\left[\hat{H}^{(N)}_{0,0} + \max_{0\leq t_1\leq T-N+1}\left(\frac{X^{(N)}_{t_1}C^{(N-1)}_{t_1}}{\hat{H}^{(N)}_{0,t_1}}\right)\hat{H}^{(N)}_{0,T-N+1} - \hat{H}^{(N)}_{0,T-N+1}\middle|\mathcal{F}_0\right] \tag{3.17}$$

and

$$\begin{aligned} C^{(N-n)}_{t_n} \quad \leq \quad & \mathbb{E}\left[\hat{H}^{(N-n)}_{t_n,t_n} + \max_{t_n<t_{n+1}\leq T-N-n+1}\left(\frac{X^{(N-n)}_{t_{n+1}}C^{(N-n-1)}_{t_{n+1}}}{\hat{H}^{(N-n)}_{t_n,t_{n+1}}}\right)\hat{H}^{(N-n)}_{t_n,T-N+n+1}\right. \\ & \left. -\hat{H}^{(N-n)}_{t_n,T-N+n+1}\middle|\mathcal{F}_{t_n}\right]. \end{aligned} \tag{3.18}$$

Let $\hat{g}_n$ and $\hat{h}_n$ be defined to be the same as g_n and h_n except with arbitrary martingales. We therefore have

$$V_0^{(N)} \leq \mathbb{E}\big[\max_{t_1}(\hat{g}_0 + \hat{h}_0 C^{(N-1)}_{t_1})\big|\mathcal{F}_0\big] \tag{3.19}$$

and that for each $n = 1, ..., N-1$,

$$C^{(N-n)}_{t_n} \leq \mathbb{E}\big[\max_{t_{n+1}>t_n}(\hat{g}_n + \hat{h}_n C^{(N-n-1)}_{t_{n+1}})\big|\mathcal{F}_{t_n}\big]. \tag{3.20}$$

Substituting (3.20) for $n = 1$ into (3.19) yields

$$\begin{aligned} V_0^{(N)} \quad &\leq \quad \mathbb{E}\big\{\max_{t_1}(\hat{g}_0 + \hat{h}_0\mathbb{E}\big[\max_{t_2>t_1}(\hat{g}_1 + \hat{h}_1 C^{(N-2)}_{t_2})\big|\mathcal{F}_{t_1}\big)\big]\big|\mathcal{F}_0\big\} \\ &\leq \quad \mathbb{E}\big\{\mathbb{E}\big[\max_{t_1}(\hat{g}_0 + \hat{h}_0 \max_{t_2>t_1}(\hat{g}_1 + \hat{h}_1 C^{(N-2)}_{t_2})\big|\mathcal{F}_{t_1}\big]\big|\mathcal{F}_0\big\} \\ &= \quad \mathbb{E}\big\{\max_{t_1<t_2}(\hat{g}_0 + \hat{h}_0(\hat{g}_1 + \hat{h}_1 C^{(N-2)}_{t_2})\big|\mathcal{F}_0\big\} \end{aligned}$$

where the second inequality is an application of Jensen's inequality and the equality results from the Tower law. By repeating the previous steps for $n = 2, ..., N$, we get

$$V_0^{(N)} \leq \mathbb{E}\left[\max_{0\leq t_1<...<t_N\leq T}\Big(\sum_{j=0}^{N-1}\hat{g}_j\prod_{i=0}^{j-1}\hat{h}_i + \prod_{i=0}^{N-1}\hat{h}_i\Big)\middle|\mathcal{F}_0\right]. \tag{3.21}$$

By replacing $\hat{g}$'s and $\hat{h}$'s with their relevant terms, we obtain

$$V_0^{(N)} \leq \mathbb{E}\Bigg\{ \max_{0\leq t_1<\ldots<t_N\leq T} \Bigg[\sum_{j=0}^{N-1} \left(\hat{H}^{(N-j)}_{t_j,t_j} - \hat{H}^{(N-j)}_{t_j,T-N+j+1} \right) \prod_{i=0}^{j-1} \frac{\hat{H}^{(N-i)}_{t_i,T-N+i+1}}{\hat{H}^{(N-i)}_{t_i,t_{i+1}}} X^{(N-i)}_{t_{i+1}} + \prod_{i=0}^{N-1} \frac{\hat{H}^{(N-i)}_{t_i,T-N+i+1}}{\hat{H}^{(N-i)}_{t_i,t_{i+1}}} X^{(N-i)}_{t_{i+1}} \Bigg] \Bigg| \mathcal{F}_0 \Bigg\}.$$

Since the set of optimal martingales yields equality, we can combine this result with the previous theorem to get (3.16). □

3.5 An algorithm for the iterative construction of the multiplicative dual

In this section, we show how one attains a good approximation to the RHS of (3.16) via simulation. Several important issues are addressed in the process. For the remainder of this section, we will assume that one already has a working exercise strategy that can be represented by some κ. In order to reduce notational requirements, we shall drop the κ in $\hat{C}^{\kappa,(\cdot)}$ and $\hat{V}^{\kappa,(\cdot)}$ and take it to be that these products are priced using this strategy.

Firstly, recall that we needed to find a reasonable approximation for $X^{(N-n)}$ when computing (3.16). We thus propose the use of the approximate multiple

$$\hat{X}^{(N-n)}_t := \frac{Z_t}{\hat{C}^{(N-n-1)}_t} + 1$$

and see that it yields a valid upper bound since

$$\hat{C}^{(N-n-1)}_t \leq C^{(N-n-1)}_t \Rightarrow \hat{X}^{(N-n)}_t \geq X^{(N-n)}_t.$$

Secondly, suppose that the martingales $\hat{H}^{(\cdot)}_{\cdot,\cdot}$ and approximate multiples $\hat{X}^{(\cdot)}_{\cdot}$ have been computed. A method of computing the path-wise maximum

$$\hat{\vartheta}^{(N)}_0 := \max_{0\leq t_1<\ldots<t_N\leq T} \Bigg[\sum_{j=0}^{N-1} \left(\hat{H}^{(N-j)}_{t_j,t_j} - \hat{H}^{(N-j)}_{t_j,T-N+j+1} \right) \prod_{i=0}^{j-1} \frac{\hat{H}^{(N-i)}_{t_i,T-N+i+1}}{\hat{H}^{(N-i)}_{t_i,t_{i+1}}} \hat{X}^{(N-i)}_{t_{i+1}} + \prod_{i=0}^{N-1} \frac{\hat{H}^{(N-i)}_{t_i,T-N+i+1}}{\hat{H}^{(N-i)}_{t_i,t_{i+1}}} \hat{X}^{(N-i)}_{t_{i+1}} \Bigg] \tag{3.22}$$

along every path is then required. In subsection 3.5.1, we show how one can decompose the computation of (3.22), which is an optimisation over N random times, into a sequence of optimisations over one random time. As part of the process, we also show how each martingale is computed.

Thirdly, we will observe that any martingales and approximate multiple will be a functions of some $\hat{C}_{\cdot}^{(\cdot)}$. Since we have to estimate $\hat{C}_{\cdot}^{(\cdot)}$ via Monte Carlo (MC), it then becomes unclear whether one still obtains a high-biased estimate for $\hat{\vartheta}_0^{(N)}$ in the presence of all the MC errors. This issue is addressed in the final two subsections. In subsection 3.5.2, we start by considering how one should approximate each multiple in each step of the aforementioned sequence of optimisations. However, the effect of the MC errors of its aforementioned approximation is not fully addressed until the final subsection 3.5.3 where we will show that we are able to obtain a valid upper bound even when approximating the martingales and the multiples by their MC counterparts.

3.5.1 Finding the path-wise maximum

The computation of the martingales is just a straightforward extension of the recursion in Section 3.3. At initialisation, the martingale consists of one unit of the triggerable product and so

$$\hat{H}_{t_n,t_n}^{(N-n)} = \hat{C}_{t_n}^{(N-n)} \mathbb{1}_{\{n\neq 0\}} + \hat{V}_{t_n}^{(N)} \mathbb{1}_{\{n=0\}}. \tag{3.23}$$

We let $B_{t_n,t}^{(N-n)}$ denote the number of options in the martingale. At initialisation, $B_{t_n,t_n}^{(N-n)} = 1$ and in the case of forward-starting options, we also have that $B_{t_n,t_n+1}^{(N-n)} = 1$ for $n \geq 1$. In particular, we have that $B_{t_n,t}^{(N-n)}$ is incremented at each $t = t_n + 1, ..., T - N + n + 1$ according to

$$B_{t_n,t}^{(N-n)} = B_{t_n,t-1}^{(N-n)} \left[1 + \mathbb{1}_{\{\tau_{n+1}(t-1)=t-1\}} \left(\frac{Z_{t-1} + \hat{C}_{t-1}^{(N-n-1)}}{\hat{C}_{t-1}^{(N-n)}} - 1 \right) \right]. \tag{3.24}$$

The value of the martingale $\hat{H}_{t_n}^{(N-n)}$ at the next time t_{n+1}, requires the further computation

$$\hat{H}_{t_n,t_{n+1}}^{(N-n)} = B_{t_n,t_{n+1}}^{(N-n)} \left[\left(Z_{t_{n+1}} + \hat{C}_{t_{n+1}}^{(N-n-1)} \right) \mathbb{1}_{\{\tau_{n+1}(t_{n+1})=t_{n+1}\}} + \hat{C}_{t_{n+1}}^{(N-n)} \mathbb{1}_{\{\tau_{n+1}(t_{n+1})\neq t_{n+1}\}} \right]. \tag{3.25}$$

Clearly, we have to make do with an approximation to each of the martingales as MC approximations for each $\hat{C}_t^{(N-n)}$ and $\hat{C}_t^{(N-n-1)}$ are used in place of their true values. However, let us assume for now that we can compute the true value of the martingales rather than their approximations. Assuming that we can also compute the true value for each multiple, $\hat{\vartheta}_0^{(N)}$ can then be computed via the following procedure:

Initialisation For each $t_N = N-1, ..., T$, find

$$\hat{\vartheta}_{t_N}^{(0)} = Z_{t_N}. \tag{3.26}$$

Step 1 For each $t_{N-1} = N, ..., T-1$, calculate

$$\hat{\vartheta}_{t_{N-1}}^{(1)} = \hat{H}_{t_{N-1},t_{N-1}}^{(1)} + \max_{t_{N-1}<t_N\leq T}\left(\frac{\hat{\vartheta}_{t_N}^{(0)}}{\hat{H}_{t_{N-1},t_N}^{(1)}}\right)\hat{H}_{t_{N-1},T}^{(1)} - \hat{H}_{t_{N-1},T}^{(1)}. \tag{3.27}$$

Step $N-n$ for $n = N-2, ..., 1$ For each $t_n = n, ..., T-N+n$, calculate

$$\begin{aligned}\hat{\vartheta}_{t_n}^{(N-n)} &= \hat{H}_{t_n,t_n}^{(N-n)} + \max_{t_n<t_{n+1}\leq T-N+n+1}\left(\frac{\hat{X}_{t_{n+1}}^{(N-n)}\hat{\vartheta}_{t_{n+1}}^{(N-n-1)}}{\hat{H}_{t_n,t_{n+1}}^{(N-n)}}\right)\hat{H}_{t_n,T-N+n+1}^{(N-n)} \\ &\quad -\hat{H}_{t_{N-n},T-N+n+1}^{(N-n)}.\end{aligned} \tag{3.28}$$

Step N Finally, compute

$$\hat{\vartheta}_0^{(N)} = \hat{H}_{0,0}^{(N)} + \max_{0\leq t_1\leq T-N+1}\left(\frac{\hat{X}_{t_1}^{(N)}\hat{\vartheta}_{t_1}^{(N-1)}}{\hat{H}_{0,t_1}^{(N)}}\right)\hat{H}_{0,T-N+1}^{(N)} - \hat{H}_{0,T-N+1}^{(N)}.$$

We now move on to address the issues arising from the use of the approximations instead of the true values.

3.5.2 Approximating the multiples

In the previous subsection, we saw how one can decompose the maximisation of the given objective function over N random times into a sequence of optimisations over one random time. However, we need to ensure that each step of the procedure still yields an estimate that is biased-high when approximations for the martingales and multiples are used.

We see that Step 1 is part of the procedure for computing the dual for the Bermudan option and we know from Joshi [2007] that this will give us a high-biased estimate for $\hat{\vartheta}^{(1)}_{t_{N-1}}$, even in the presence of MC errors.

Next, at Step $N-n$ of the algorithm for each $n = N-2, ..., 1$, in order to compute each path-wise maximum $\hat{\vartheta}^{(N-n)}_{t_n}$, we first need to compute

$$\max_{t_n < t_{n+1} \leq T-N-n+1} \frac{\hat{X}^{(N-n)}_{t_{n+1}} \hat{\vartheta}^{(N-n-1)}_{t_{n+1}}}{\hat{H}^{(N-n)}_{t_n,t_{n+1}}} \tag{3.29}$$

along each path.

We now consider the computational treatment of the multiples in view of the MC errors arising from its approximation. Rather than simulating $\hat{X}^{(N-n)}_{t_{n+1}}$ and $\hat{H}^{(N-n)}_{t_n,t_{n+1}}$ individually, we can take advantage of cancelling terms by rewriting their ratio as

$$\begin{aligned}
&\frac{\hat{X}^{(N-n)}_{t_{n+1}}}{\hat{H}^{(N-n)}_{t_n,t_{n+1}}} \\
&= \frac{(Z_{t_{n+1}} + \hat{C}^{(N-n-1)}_{t_{n+1}})/\hat{C}^{(N-n-1)}_{t_{n+1}}}{B^{(N-n)}_{t_n,t_{n+1}} \left[\left(Z_{t_{n+1}} + \hat{C}^{(N-n-1)}_{t_{n+1}} \right) \mathbb{1}_{\{\tau_{n+1}(t_{n+1}) = t_{n+1}\}} + \hat{C}^{(N-n)}_{t_{n+1}} \mathbb{1}_{\{\tau_{n+1}(t_{n+1}) \neq t_{n+1}\}} \right]} \\
&= \frac{1}{B^{(N-n)}_{t_n,t_{n+1}}} \left[\frac{1}{\hat{C}^{(N-n-1)}_{t_{n+1}}} \mathbb{1}_{\{\tau_{n+1}(t_{n+1}) = t_{n+1}\}} \right. \\
&\quad \left. + \frac{1}{\hat{C}^{(N-n)}_{t_{n+1}}} \left(\frac{Z_{t_{n+1}}}{\hat{C}^{(N-n-1)}_{t_{n+1}}} + 1 \right) \mathbb{1}_{\{\tau_{n+1}(t_{n+1}) \neq t_{n+1}\}} \right].
\end{aligned} \tag{3.30}$$

We will see below that this form for $\hat{X}^{(N-n)}_{t_{n+1}} / \hat{H}^{(N-n)}_{t_n,t_{n+1}}$ is crucial to obtaining a high-biased estimate for $\hat{\vartheta}^{(N-n)}_{t_n}$.

3.5.3 Ensuring upwards bias in the presence of sub-simulation errors

Andersen and Broadie [2004] showed that using sub-simulations to compute the martingales for the additive dual still resulted in a high-biased estimate. Furthermore, Joshi [2007] showed that this approach worked when computing the martingales for the multiplicative dual for the Bermudan option. Our final proposition will show that we still get a valid upper bound for (3.16) if sub-simulations are used to compute the required variables. Its proof requires the following lemma.

Lemma 3.7. *Let ω denote an outer simulation path where $\omega \in \Omega$. Consider a collection of non-zero functions $\{g_i\}$ that is used to construct a quotient $g = \hat{g}/\check{g}$ where $\hat{g} := \prod_{i=1}^{j} g_i$ and $\check{g} := \prod_{j+1}^{k} g_i$. Let $\tilde{g}_i = g_i + \epsilon_i$ where each ϵ_i is a random variable with mean 0 and where all ϵ_i's are conditionally independent given ω. If $\mathcal{G}$ is the filtration generated by ω, we then have that $\mathbb{E}(\tilde{g}|\mathcal{G}) \geq g$.*

Proof. We have that $\mathbb{E}(\tilde{g}_i/g_i|\mathcal{G}) = 1$ and by Jensen's inequality, $\mathbb{E}(g_i/\tilde{g}_i|\mathcal{G}) \geq 1$. Therefore

$$\prod_{i=1}^{j} \mathbb{E}\Big(\frac{\tilde{g}_i}{g_i}\Big|\mathcal{G}\Big) \prod_{i=j+1}^{k} \mathbb{E}\Big(\frac{g_i}{\tilde{g}_i}\Big|\mathcal{G}\Big) \geq 1$$

which rearranges to give us $\mathbb{E}(\tilde{g}|\mathcal{G}) \geq g$. □

Proposition 3.8. *Let ω denote an outer simulation path where $\omega \in \Omega$ and let $\mathcal{G}_t$ be the information generated by ω and all sub-simulations prior to time t for $t \in \mathbb{T}$. For each $i = 0, ..., N-1$, let $\tilde{X}_{t_{i+1}}^{(N-i)}$ denote the simulated multiples and let $\tilde{H}_{t_i,t_{i+1}}^{(N-i)}$ and $\tilde{H}_{t_i,T-N+n+1}^{(N-i)}$ denote the simulated values of the arbitrary martingales $\hat{H}_{t_i,t_{i+1}}^{(N-i)}$ and $\hat{H}_{t_i,T-N+n+1}^{(N-i)}$ respectively. For each trajectory of the underlying, we have that*

$$\begin{aligned}
&\mathbb{E}\Bigg\{ \max_{0\leq t_1<\ldots<t_N\leq T} \Bigg[\sum_{j=0}^{N-1} \Big(\tilde{H}_{t_j,t_j}^{(N-j)} - \tilde{H}_{t_j,T-N+j+1}^{(N-j)} \Big) \prod_{i=0}^{j-1} \frac{\tilde{H}_{t_i,T-N+i+1}^{(N-i)}}{\tilde{H}_{t_i,t_{i+1}}^{(N-i)}} \tilde{X}_{t_{i+1}}^{(N-i)} \\
&\quad + \prod_{i=0}^{N-1} \frac{\tilde{H}_{t_i,T-N+i+1}^{(N-i)}}{\tilde{H}_{t_i,t_{i+1}}^{(N-i)}} \tilde{X}_{t_{i+1}}^{(N-i)} \Bigg] \Bigg| \mathcal{G}_0 \Bigg\} \\
&\geq \max_{0\leq t_1<\ldots<t_N\leq T} \Bigg[\sum_{j=0}^{N-1} \Big(\hat{H}_{t_j,t_j}^{(N-j)} - \hat{H}_{t_j,T-N+j+1}^{(N-j)} \Big) \prod_{i=0}^{j-1} \frac{\hat{H}_{t_i,T-N+i+1}^{(N-i)}}{\hat{H}_{t_i,t_{i+1}}^{(N-i)}} \hat{X}_{t_{i+1}}^{(N-i)} \\
&\quad + \prod_{i=0}^{N-1} \frac{\hat{H}_{t_i,T-N+i+1}^{(N-i)}}{\hat{H}_{t_i,t_{i+1}}^{(N-i)}} \hat{X}_{t_{i+1}}^{(N-i)} \Bigg].
\end{aligned} \tag{3.31}$$

Remark 3.9. We require conditional independence between each pair

$$\tilde{H}_{t_i,t_{i+1}}^{(N-i)} \quad \text{and} \quad \tilde{H}_{t_i,T-N+n+1}^{(N-i)}.$$

Therefore, they must be evaluated using separate simulations.

Proof. Let $\tilde{C}_t^{(N-n)}$ denote the simulated approximation of $\hat{C}_t^{(N-n)}$ with error $\epsilon_t^{(N-n)}$ of mean 0. From this we compute the martingales and the multiples. The maximum on

the RHS of (3.31) is attained at the collection of random times $t_1^*, ..., t_N^*$. Now

$$\mathbb{E}\left[\sum_{j=0}^{N-1}\left(\tilde{H}^{(N-j)}_{t_j^*,t_j^*}-\tilde{H}^{(N-j)}_{t_j^*,T-N+j+1}\right)\prod_{i=0}^{j-1}\frac{\tilde{H}^{(N-i)}_{t_i^*,T-N+i+1}}{\tilde{H}^{(N-i)}_{t_i^*,t_{i+1}^*}}\tilde{X}^{(N-i)}_{t_{i+1}^*}+\prod_{i=0}^{N-1}\frac{\tilde{H}^{(N-i)}_{t_i^*,T-N+i+1}}{\tilde{H}^{(N-i)}_{t_i^*,t_{i+1}^*}}\tilde{X}^{(N-i)}_{t_{i+1}^*}\right]\Bigg|\mathcal{G}_0\Bigg],$$

bounds the expectation in (3.31) from below and can be re-written as

$$\mathbb{E}\left(\tilde{H}^{(N)}_{0,0}-\tilde{H}^{(N)}_{0,T-N+1}+\frac{\tilde{H}^{(N)}_{0,T-N+1}}{\tilde{H}^{(N)}_{0,t_1^*}}\tilde{X}^{(N)}_{t_1^*}\tilde{\vartheta}^{(N-1)}_{t_1^*}\Bigg|\mathcal{G}_0\right)$$

where

$$\tilde{\vartheta}^{(N-1)}_{t_1^*}:=\mathbb{E}\left(\tilde{H}^{(N-1)}_{t_1^*,t_1^*}-\tilde{H}^{(N-1)}_{t_1^*,T-N+2}+\frac{\tilde{H}^{(N)}_{t_1^*,T-N+2}}{\tilde{H}^{(N)}_{t_1^*,t_2^*}}\tilde{X}^{(N-1)}_{t_2^*}\tilde{\vartheta}^{(N-2)}_{t_2^*}\Bigg|\mathcal{G}_{t_1^*}\right),$$

for $n = 2, ..., N-2$,

$$\tilde{\vartheta}^{(N-n)}_{t_n^*}:=\mathbb{E}\left(\tilde{H}^{(N-n)}_{t_n^*,t_n^*}-\tilde{H}^{(N-n)}_{t_n^*,T-N+n+1}+\frac{\tilde{H}^{(N-n)}_{t_1^*,T-N+n+1}}{\tilde{H}^{(N-n)}_{t_n^*,t_{n+1}^*}}\tilde{X}^{(N-n)}_{t_{n+1}^*}\tilde{\vartheta}^{(N-n-1)}_{t_{n+1}^*}\Bigg|\mathcal{G}_{t_n^*}\right)$$

and

$$\tilde{\vartheta}^{(1)}_{t_{N-1}^*}=\mathbb{E}\left(\tilde{H}^{(1)}_{t_{N-1}^*,t_{N-1}^*}-\tilde{H}^{(1)}_{t_{N-1}^*,T}+\frac{\tilde{H}^{(1)}_{t_{N-1}^*,T}}{\tilde{H}^{(1)}_{t_{N-1}^*,t_N^*}}Z_{t_N^*}\Bigg|\mathcal{G}_{t_{N-1}^*}\right).$$

One immediately identifies this decomposition as our algorithm where the maximums have been achieved at $t_1^*, ..., t_N^*$ with

$$\hat{\vartheta}^{(N-n)}_{t_n^*}=\hat{H}^{(N-n)}_{t_n^*,t_n^*}+\left(\frac{\hat{X}^{(N-n)}_{t_{n+1}^*}\hat{\vartheta}^{(N-n-1)}_{t_{n+1}^*}}{\hat{H}^{(N-n)}_{t_n^*,t_{n+1}^*}}\right)\hat{H}^{(N-n)}_{t_n^*,T-N+n+1}-\hat{H}^{(N-n)}_{t_{N-n}^*,T-N+n+1}.$$

Thus for (3.31) to hold, we require that $\tilde{\vartheta}^{(N-n)}_{t_n^*}\geq\hat{\vartheta}^{(N-n)}_{t_n^*}$ which is in turn satisfied if

$$\mathbb{E}(\tilde{H}^{(N-n)}_{t_n^*,t_n^*}|\mathcal{G}_{t_n^*}) \geq \hat{H}^{(N-n)}_{t_n^*,t_n^*}, \tag{3.32}$$

$$\mathbb{E}(\tilde{H}^{(N-n)}_{t_n^*,T-N+n+1}|\mathcal{G}_{t_n^*}) \geq \hat{H}^{(N-n)}_{t_n^*,T-N+n+1} \tag{3.33}$$

$$\text{and}\quad \mathbb{E}\left(\frac{\tilde{X}^{(N-n)}_{t_{n+1}^*}}{\tilde{H}^{(N-n)}_{t_n^*,t_{n+1}^*}}\Bigg|\mathcal{G}_{t_n^*}\right) \geq \frac{\hat{X}^{(N-n)}_{t_{n+1}^*}}{\hat{H}^{(N-n)}_{t_n^*,t_{n+1}^*}} \tag{3.34}$$

We proceed by induction, taking $\tilde{\vartheta}^{(N-n-1)}_{t^*_{n+1}} \geq \hat{\vartheta}^{(N-n-1)}_{t^*_{n+1}}$ as our induction hypothesis. (3.32) is trivially true. Since

$$\begin{aligned}
&\mathbb{E}\left(\frac{\tilde{H}^{(N-n)}_{t^*_n,T-N+n+1}}{\hat{H}^{(N-n)}_{t^*_n,T-N+n+1}}\middle|\mathcal{G}_{t^*_n}\right)\\
&\quad=\mathbb{E}\left(\frac{Z_{T-N+n+1}+\tilde{C}^{(N-n-1)}_{T-N+n+1}}{Z_{T-N+n+1}+\hat{C}^{(N-n-1)}_{T-N+n+1}}\prod_{i=t^*_n}^{T-N+n}\frac{Z_i+\tilde{C}^{(N-n-1)}_i}{Z_i+\hat{C}^{(N-n-1)}_i}\frac{\hat{C}^{(N-n)}_i}{\tilde{C}^{(N-n)}_i}\mathbb{1}_{\{\tau_{n+1}(i)=i\}}\middle|\mathcal{G}_{t^*_n}\right),
\end{aligned}$$

(3.33) is satisfied by Lemma 3.7. Furthermore, by considering (3.30), we see that

$$\begin{aligned}
&\mathbb{E}\left(\frac{\tilde{X}^{(N-n)}_{t^*_{n+1}}}{\tilde{H}^{(N-n)}_{t^*_n,t^*_{n+1}}}\Big/\frac{\hat{X}^{(N-n)}_{t^*_{n+1}}}{\hat{H}^{(N-n)}_{t^*_n,t^*_{n+1}}}\middle|\mathcal{G}_{t^*_n}\right)\\
&\quad=\mathbb{E}\left(\prod_{i=t^*_n}^{t^*_{n+1}}\frac{Z_i+\hat{C}^{(N-n-1)}_i}{Z_i+\tilde{C}^{(N-n-1)}_i}\frac{\tilde{C}^{(N-n)}_i}{\hat{C}^{(N-n)}_i}\mathbb{1}_{\{\tau_{n+1}(i)=i\}}\middle|\mathcal{G}_{t^*_n}\right)\\
&\qquad\times\left[\mathbb{1}_{\{\tau_{n+1}(t^*_{n+1})=t^*_{n+1}\}}\mathbb{E}\left(\frac{\hat{C}^{(N-n-1)}_{t^*_{n+1}}}{\tilde{C}^{(N-n-1)}_{t^*_{n+1}}}\middle|\mathcal{G}_{t^*_n}\right)\right.\\
&\qquad\left.+\mathbb{1}_{\{\tau_{n+1}(t^*_{n+1})\neq t^*_{n+1}\}}\mathbb{E}\left(\frac{\hat{C}^{(N-n)}_{t^*_{n+1}}}{\tilde{C}^{(N-n)}_{t^*_{n+1}}}\middle|\mathcal{G}_{t^*_n}\right)\mathbb{E}\left(\frac{\frac{Z_{t^*_{n+1}}}{\tilde{C}^{(N-n-1)}_{t^*_{n+1}}}+1}{\frac{Z_{t^*_{n+1}}}{\hat{C}^{(N-n-1)}_{t^*_{n+1}}}+1}\middle|\mathcal{G}_{t^*_n}\right)\right].
\end{aligned}$$

Since

$$\mathbb{E}\left(\frac{\frac{Z_{t^*_{n+1}}}{\tilde{C}^{(N-n-1)}_{t^*_{n+1}}}+1}{\frac{Z_{t^*_{n+1}}}{\hat{C}^{(N-n-1)}_{t^*_{n+1}}}+1}\middle|\mathcal{G}_{t^*_n}\right)\geq 1,$$

(3.34) is also satisfied by Lemma 3.7. □

3.6 Conclusion

In this chapter, the notion of a martingale in the the multiple-exercise framework was provided for. Using these martingales, we generalised the result of Joshi and Tang [2014] to the multiple exercise case to derive the first known multiplicative dual for the multiple-stopping problem. An algorithm that can be seen as the natural extension of the Andersen-Broadie algorithm to the multiplicative multiple-exercise case was developed

and it was shown that we still obtained an upper bound on the price even in the presence of Monte Carlo errors.

In order to produce meaningful numerical results, one requires an implementation of the algorithm on parallelized architecture due to the large number of sub-simulations required. In particular, we believe the implementation is best done in Nvidia‘s proprietary language CUDA which allows us to harness the parallel computational power of the graphics processing unit (GPU). The details of the implementation deserves an entire chapter in itself due to the intricacies of the memory management involved. This is beyond the scope of this thesis and will constitute future research.

Chapter 4

Algorithms for optimal control of stochastic switching systems

In this chapter we look at the more general framework of solving Markov decision problems with linear state dynamics. After introducing this framework in Section 4.1, Section 4.2 presents the notion of a *convex switching system* and discusses solutions to this stochastic problem class. In Section 4.3, we review and analyze the numerical scheme of Hinz [2014] that provides fast and stable solutions to convex switching problems. Section 4.4 represents a first step to relaxing the requirement of convexity. A remarkable generalization is achieved Section 4.5 where a method has been devised that allows us to by-pass any convexity requirement while yielding significant improvements in computation time. Another major contribution of the chapter is presented in Section 4.6, where we suggest an adaptation of the approach Rogers [2007] to obtain recursive schemes for upper bound estimates of an approximate solution. Section 4.7 provides two numerical examples. We conclude this chapter with Section 4.8.

4.1 Markov decision theory

We begin by reviewing the classical framework of finite-horizon Markov decision theory, where we closely follow Chapter 2 of Bäuerle and Rieder [2011] and tailor it to suit our purposes. Consider a system on the finite time horizon $0, \dots, T$ whose state varies in a measurable space $(E, \mathcal{E})$ and is affected by elements from a set A of possible actions. For

each $a \in A$, we assume that $K_t^a(x, dx')$ is a stochastic transition kernel on $(E, \mathcal{E})$. Consider a fixed sequence $(X_t)_{t=0}^T$ of random variables which can be thought of as coordinate projections acting on the product $E^{\{0,\dots,T\}}$ of copies of $(E, \mathcal{E})$. A mapping $\pi_t : E \mapsto A$ which describes the action that the controller of the system takes at time t is called a *decision rule*. A sequence of decision rules $\pi = (\pi_t)_{t=0}^{T-1}$ is called a *policy*. For each initial point $x_0 \in E$ and each policy $\pi = (\pi_t)_{t=0}^{T-1}$, there exists a probability measure $\mathbb{P}^{x_0,\pi}$ for which $\mathbb{P}^{x_0,\pi}(X_0 = x_0) = 1$ and where

$$\mathbb{P}^{x_0,\pi}(X_{t+1} \in B \,|\, X_0, \dots, X_t) = K_t^{\pi_t(X_t)}(X_t, B) \tag{4.1}$$

holds for each measurable $B \in \mathcal{E}$ and $t = 0, \dots, T-1$. That is, given that system is in state X_t at time t, the action $a = \pi_t(X_t)$ is used to pick the transition probability $K_t^{a=\pi_t(X_t)}(X_t, \cdot)$ which assigns the random evolution of the state from X_t to X_{t+1} with the distribution $K_t^{\pi_t(X_t)}(X_t, \cdot)$. For the sake of notational convenience, we use $\mathcal{K}_t^a$ to denote the one-step transition operator associated with the transition kernel K_t^a when the action $a \in A$ is chosen. In other words, for each action $a \in A$ the operator $\mathcal{K}_t^a$ acts on functions φ by

$$(\mathcal{K}_t^a \varphi)(x) = \int_E \varphi(x') K_t^a(x, dx') \qquad x \in E, \tag{4.2}$$

whenever the above integrals are well-defined.

At each time t, we are given the *t-step reward function* $r_t : E \times A \mapsto \mathbb{R}$, where $r_t(x, a)$ represents the reward for applying an action $a \in A$ when the state of the system is $x \in E$ at time t. At the end of the time horizon, at time T, it is assumed that no action can be taken. Here, if the system is in a state x, a *scrap value* $r_T(x)$, which is described by a pre-specified *scrap function* $r_T : E \to \mathbb{R}$, is collected. Given an initial point x_0, our goal is to maximize the expected finite-horizon total reward, in other words to find the argument $\pi^* = (\pi_t^*)_{t=0}^{T-1}$ such that

$$\pi^* = \text{argmax}_{\pi \in \mathcal{A}} \mathbb{E}^{x_0,\pi} \left(\sum_{t=0}^{T-1} r_t(X_t, \pi_t(X_t)) + r_T(X_T) \right), \tag{4.3}$$

where $\mathcal{A}$ is the set of all policies, and $\mathbb{E}^{x,\pi}$ denotes the expectation over the controlled Markov chain defined by (4.1). The maximization (4.3) is well-defined under diverse additional assumptions (see Bäuerle and Rieder [2011], p. 199).

The calculation of the optimal policy is addressed in the following setting. We introduce for $t = 0, \dots, T-1$ the *Bellman operator*

$$\mathcal{T}_t v(x) = \sup_{a \in A} \left(r_t(x,a) + \mathcal{K}_t^a v(x) \right), \quad x \in E \tag{4.4}$$

which acts on each measurable function $v : E \to \mathbb{R}$ where the integrals $\mathcal{K}_t^a v$ for all $a \in A$ exist. Further, consider the *Bellman recursion*

$$v_T = r_T, \quad v_t = \mathcal{T}_t v_{t+1} \qquad \text{for } t = T-1, \dots, 0. \tag{4.5}$$

Under appropriate assumptions, there exists a recursive solution $(v_t^*)_{t=0}^T$ to the Bellman recursion, which gives the so-called *value functions* and determines an optimal policy π^* via

$$\pi_t^*(x) = \operatorname{argmax}_{a \in A} \left(r_t(x,a) + \mathcal{K}_t^a v_{t+1}^*(x) \right), \quad x \in E$$

for all $t = 0, \dots, T-1$.

4.2 Convex switching systems

For the remainder of this work, we concentrate on Markov decision problems which satisfy specific additional assumptions under which the solutions to the Bellman recursion exist. This enables us to focus on finding numerical approximations.

Consider a Markov decision model whose state evolution consists of one discrete and one continuous component. To be more specific, we assume that the state space $E = P \times \mathbb{R}^d$ is the product of a finite space P and the Euclidean space $\mathbb{R}^d$. We suppose that the first component $p \in P$ is deterministically driven by a finite set A of actions in terms of a function

$$\alpha : P \times A \to A, \quad (p,a) \to \alpha(p,a),$$

where $\alpha(p,a) \in P$ is the new value of the discrete component of the state if its previous discrete component value was p and the action $a \in A$ was taken by the controller. Furthermore, we assume that the continuous state component evolves as an uncontrolled

Markov process $(Z_t)_{t=0}^T$ on $\mathbb{R}^d$ whose evolution is driven by random linear transformations

$$Z_{t+1} = W_{t+1} Z_t$$

with pre-specified independent and integrable disturbance matrices $(W_t)_{t=1}^T$. Finally, let us assume that the reward functions

$$r_t(p, z, a), \quad t = 0, \dots, T-1, \quad p \in P, \quad a \in A$$

and scrap functions

$$r_T(p, z), \quad p \in P$$

are convex and globally Lipschitz continuous in the continuous component of the state space $z \in \mathbb{R}^d$. In this setting, the transition operators are given by

$$\mathcal{K}_t^a v(p, z) = \mathbb{E}(v(\alpha(p, a), W_{t+1} z)), \qquad t = 0, \dots, T-1, \;\; a \in A \tag{4.6}$$

and the Bellman operators are

$$\mathcal{T}_t v(p, z) = \sup_{a \in A} \big(r_t(p, z, a) + \mathbb{E}(v(\alpha(p, a), W_{t+1} z)) \big) \tag{4.7}$$

for all $p \in P$, $z \in \mathbb{R}^d$ and $t = 0, \dots, T-1$. Markov decision problems satisfying these assumptions are referred to as *convex switching systems* in what follows.

4.3 Algorithmic solutions

For such systems, the backward induction described by (4.5) solves our control problem. However, by inspecting the Bellman operator

$$\mathcal{T}_t v(p, z) = \max_{a \in A} \big(r_t(p, z, a) + \mathbb{E}(v(\alpha(p, a), W_{t+1} z)) \big), \tag{4.8}$$

we see that solving the Bellman recursion results in a number of problems, the most pressing of which is that one requires a point-wise solution for each $z \in \mathbb{R}^d$. In Hinz [2014], a method was presented that targeted a solution in a "functional" form. We now provide a detailed account and justification of their approach.

First, by approximating the expectation in the Bellman operator in (4.8) by finite summation, we obtain the *modified* Bellman operator $\mathcal{T}_t^n$ that acts on a given value function according to

$$\mathcal{T}_t^n v(p,z) = \max_{a\in A}\left(r_t(p,z,a) + \sum_{k=1}^{n} \nu_{t+1}(k) v(\alpha(p,a), W_{t+1}(k)z)\right) \tag{4.9}$$

where $(W_{t+1}(k))_{k=1}^n$ represents appropriate realizations of disturbances with the corresponding probability weights $(\nu_{t+1}(k))_{k=1}^n$. By replacing the true Bellman operator (4.8) in the backward induction of (4.5) by its modified counterpart that is given by (4.9), we obtain the modified induction

$$v_T^n = r_T, \quad v_t = \mathcal{T}_t^n v_{t+1}^n \qquad \text{for } t = T-1,\dots,0. \tag{4.10}$$

Although the integration is now replaced by a finite sum, determining $(v_t^n)_{t=0}^T$ is still algorithmically intractable as the calculation must be performed at each point $z \in \mathbb{R}^d$. At this point, we turn to the important observation that since the scrap and reward functions, $r_t(p,z,a)$, $t = 0,\dots,T-1$ and $r_T(p,z)$, used in (4.9) and (4.10) are convex in the continuous component, then the resulting value functions $(v_t^n)_{t=0}^T$ must also be convex in the same component.

We now suggest an approximation of these functions $(v_t^n)_{t=0}^T$ in terms of maxima over a finite selection of their sub-gradients. Before we can begin to explain the advantage of such a piecewise linear approximation, we need to first establish a few concepts.

First, let us refer to a countable subset $G \subset \mathbb{R}^d$ as a *grid*. For a grid G, the *sub-gradient envelope* $\mathcal{S}_G f$ of a convex function f is defined to be the maximum of sub-gradients $\nabla_g f$ of f at each grid point $g \in G$ and so

$$\mathcal{S}_G f = \vee_{g\in G} \nabla_g f.$$

Given a family $\{(W_t(k))_{t=1}^T \,:\, k = 1,\dots,n\}$ of trajectories of disturbances that increases with $n \in \mathbb{N}$ and a family of grids $(G^m)_{m\in\mathbb{N}}$ whose tightness increases with $m \in \mathbb{N}$, we define for each $n, m \in \mathbb{N}$ the *double modified* Bellman operators $\mathcal{T}_t^{m,n}$ for $t = 0,\dots,T-1$

$$(\mathcal{T}_t^{m,n} v)(p,\cdot) = \mathcal{S}_{G^m} \max_{a\in A}\left(r_t(p,\cdot,a) + \sum_{k=1}^{n} \nu_{t+1}(k) v(\alpha(p,a), W_{t+1}(k)\cdot)\right)$$

Using these operators, the double modified value functions $(v_t^{m,n})_{t=0}^T$ are obtained from the backward induction which starts with

$$v_T^{m,n}(p,\cdot) = \mathcal{S}_{G^m} r_T(p,\cdot), \quad p \in P \tag{4.11}$$

and recursively determines functions

$$v_t^{m,n} = \mathcal{T}_t^{m,n} v_{t+1}^{m,n}, \ \ t = T-1, \dots 0. \tag{4.12}$$

Obviously, this approach involves two approximation parameters $n \in \mathbb{N}$ and $m \in \mathbb{N}$, which correspond to the sampled disturbances $(W_{t+1}(k))_{k=1}^n$ with their weights $(\nu_{t+1}(n))_{k=1}^n \subset \mathbb{R}_+^d$ and the grid tightening $(G^m)_{m\in\mathbb{N}}$. Under appropriate assumptions, this scheme enjoys excellent convergence properties (see Hinz [2014]). However, we shall now focus solely on its algorithmic aspect.

Since the double-modified backward induction (4.11) and (4.12) returns value functions $(v_t^{m,n})_{t=0}^T$ which are piecewise linear and convex (in the continuous component), we now address an appropriate representation of such functions in terms of matrices in order to re-write the backward induction algorithm (4.11) and (4.12) in a compact matrix form.

A matrix with d columns is called a *matrix representative* of a piecewise linear convex function $l : \mathbb{R}^d \to \mathbb{R}$ if it holds that $l(z) = \max(Lz)$ for all $z \in \mathbb{R}^d$. We shall use the expression $l \sim L$ if a piecewise linear convex function l possesses a matrix representative L. It turns out that the formation of a sub-gradient envelope can be directly described in terms of matrix representatives. Namely, if l possesses a matrix representative L then its sub-gradient envelope $\mathcal{S}_G$ on the grid $G = \{g^1, \dots, g^m\}$ possess a matrix representative $\Upsilon_G[L]$ where the row-re-arrangement operator Υ_G is defined by

$$\Upsilon_G[L]_{i,\cdot} = L_{\operatorname{argmax}(Lg^i),\cdot} \qquad \text{for all } i = 1, \dots, m.$$

In other words, when Υ_G is applied to a matrix L with d columns, the result $\Upsilon_G[L]$ of the row-rearrangement yields an $m \times d$ matrix whose i-th row is the row of L at which the maximum in Lg^i at the i-th grid point is attained. As mentioned above, the relation between the sub-gradient envelope of a function and its matrix representative is thus

given in terms of the row-rearrangement operator Υ_G:

$$l \sim L \quad \Longrightarrow \quad \mathcal{S}_G l \sim \Upsilon_G[L].$$

A similar relation holds for the summation of piecewise linear and convex functions, followed by sub-gradient envelope. Namely, it corresponds to a straight summation of their matrix representatives, after row-rearrangement:

$$l \sim L, \;\; f \sim F \quad \Longrightarrow \quad \mathcal{S}_G(f+l) \sim \Upsilon_G[L] + \Upsilon_G[F].$$

Similarly, maximization of piecewise linear and convex functions, followed by sub-gradient envelope is realized on matrix level by binding by rows of matrix representatives, followed by the row-rearrangement:

$$l \sim L, \;\; f \sim F \quad \Longrightarrow \quad \mathcal{S}_G(l \vee f) \sim \Upsilon_G[L \sqcup F].$$

Here, the binding-by-row operation $L \sqcup F$ performs a row concatenation of the two matrices L and F. Let us also introduce an equivalent but algorithmically more convenient procedure of maximization on the level of matrix representatives for later use. Given a grid $G = \{g^1, \ldots, g^m\} \subset \mathbb{R}^d$ and $m \times d$ matrices $F(a)$, $a \in A$, we introduce

$$F := \bigsqcup_{a \in A} F(a)$$

to denote a $m \times d$ matrix F whose i-th row

$$F_{i,} = F_{i,}(a(i)), \quad i = 1, \ldots, m$$

equals to the i-th row of the matrix $F(a(i))$ where the maximum at the i-th grid point g^i is reached, i.e.

$$a(i) = \operatorname{argmax}\{F_{i,}(a) \cdot g^i : i = 1, \ldots, m\}.$$

This maximization is used to obtain a sub-gradient envelope of the maximum over a family f^a, $a \in A$ of piecewise linear and convex functions in terms of the matrix representatives of their sub-gradient envelopes:

$$\mathcal{S}_G f^a \sim F(a), \;\; a \in A \quad \Rightarrow \quad \mathcal{S}_G(\bigvee_{a \in A} f^a) \sim \bigsqcup_{a \in A} F(a).$$

Finally, we emphasize that determining the sub-gradient envelope of the composition of a function with a linear mapping corresponds to a simple matrix product followed by a row-rearrangement. In other words, for each $d \times d$-matrix W, it holds that

$$l \sim L \quad \Longrightarrow \quad \mathcal{S}_G(l(W\cdot)) \sim \Upsilon_G[LW].$$

Observe that the rows of the matrix F representing a sub-gradient envelope $\mathcal{S}_G f$ of a convex piecewise linear function f can always be arranged such that $F = \Upsilon_G(F)$ holds. We say that a the sub-gradient representative F is in the *normal form* if it holds that $F = \Upsilon_G(F)$.

Since the double-modified backward induction involves maximization, summations and compositions with linear mappings applied to piecewise linear convex functions, it can be rewritten in terms of matrix operations. Let us present the resulting algorithm:

Pre-calculations: Given a grid $G^m = \{g^1, \dots, g^m\}$, implement the row-rearrangement operator $\Upsilon = \Upsilon_{G^m}$ and the row maximization operator $\sqcup_{a\in A}$. Determine a distribution sampling $(W_t(k))_{k=1}^n$ of each disturbance W_t with the corresponding weights $(\nu_t(k))_{k=1}^n$ for $t = 1, \dots, T$. Given reward functions $(r_t)_{t=0}^{T-1}$ and scrap value r_T, determine the normal form of the matrix representatives of their sub-gradient envelopes

$$\mathcal{S}_{G^m} r_t(p, \cdot, a) \sim R_t(p, a), \qquad \mathcal{S}_{G^m} r_T(p, \cdot) \sim R_T(p)$$

for $t = 0, \dots, T-1$, $p \in P$ and $a \in A$. Introduce matrix representatives $V_t(p)$ for $t = 0, \dots, T$, $p \in P$ of each value function by

$$v_t^{n,m}(p, \cdot) \sim V_t(p) \quad \text{for } t = 0, \dots, T,\ p \in P$$

which are obtained via the following matrix of the backward induction:

Initialization: Start with the matrices

$$V_T(p) = R_T(p), \qquad \text{for all } p \in P.$$

Recursion: For $t = T-1, \dots, 0$ calculate for $p \in P$

$$V_t(p) = \sqcup_{a\in A}\big(R_t(p, a) + \sum_{k=1}^{n} \nu_{t+1}(k)\Upsilon[V_{t+1}(\alpha(p, a)) \cdot W_{t+1}(k)]\big) \tag{4.13}$$

4.4 Non-convex extension

In this section, we demonstrate that for non-convex of reward and scrap functions the above algorithm can be adapted, if the functions are representable as a difference of two convex functions. More precisely, assume that for all $t = 0, \dots, T-1$, and $p \in P$, $a \in A$ it holds that

$$r_t(p, \cdot, a) = \breve{r}_t(p, \cdot, a) - \hat{r}_t(p, \cdot, a), \tag{4.14}$$

and

$$r_T(p, \cdot) = \breve{r}_T(p, \cdot) - \hat{r}_T(p, \cdot) \tag{4.15}$$

with convex functions $\breve{r}_t(p, \cdot, a)$, $\hat{r}_t(p, \cdot, a)$, $\breve{r}_T(p, \cdot)$ and $\hat{r}_T(p, \cdot)$ for $p \in P$. Given such represention, the idea is to decompose the backward induction into parallel procedures that operate on convex functions. Suppose that at the step t, the value function v_{t+1} can be represented as a difference $v_{t+1} = \breve{v}_{t+1} - \hat{v}_{t+1}$ of convex functions $\breve{v}_{t+1}(p, \cdot)$ and $\hat{v}_{t+1}(p, \cdot)$ for $p \in P$. With this, we have

$$\begin{aligned} \mathcal{T}_t v(p, z) &= \sup_{a \in A} \left(r_t(p, z, a) + \mathcal{K}_t^a v_{t+1}(p, z) \right) \\ &= \sup_{a \in A} \left([\breve{r}_t(p, z, a) + \mathcal{K}_t^a \breve{v}_{t+1}(p, z)] - [\hat{r}_t(p, z, a) + \mathcal{K}_t^a \hat{v}_{t+1}(p, z)] \right) \end{aligned}$$

showing that before maximization in $a \in A$, the result is obtained as difference of two convex functions. However, a direct application of convex function maximization (i.e. the use of the row maximization operator $\sqcup$) is not compatible with this decomposition. Therefore, we require a way to express the maximum over differences of convex functions as difference of two convex functions. The following simple observation helps here.

Consider for each $a \in A$ the difference $\breve{f}_a - \hat{f}_a$ of two convex functions $\breve{f}_a$ and $\hat{f}_a$ and let $\hat{f} := \sum_{a \in A} \hat{f}_a$. Then for each $a \in A$ the functions $\breve{f}_a - \hat{f}_a + \hat{f}$ and $\hat{f}$ are convex and yield the desired decomposition

$$\max_{a \in A} (\breve{f}_a - \hat{f}_a) = \max_{a \in A} (\breve{f}_a - \hat{f}_a + \hat{f}) - \hat{f}. \tag{4.16}$$

Having this approach in mind, we propose the following algorithm:

Pre-calculation: Decompose the reward $(r_t)_{t=0}^{T-1}$ and scrap r_T functions into a difference of convex functions as in (4.14) and (4.15) with their (normal form) matrix representatives

$$\begin{aligned} \mathcal{S}_{G^m}\breve{r}_t(p,\cdot,a) &\sim \breve{R}_t(p,a), & \mathcal{S}_{G^m}\hat{r}_t(p,\cdot,a) &\sim \hat{R}_t(p,a), \\ \mathcal{S}_{G^m}\breve{r}_T(p,\cdot) &\sim \breve{R}_T(p), & \mathcal{S}_{G^m}\hat{r}_T(p,\cdot) &\sim \hat{R}_T(p) \end{aligned} \tag{4.17}$$

for all $t = 0, \ldots, T-1$, $p \in P$ and $a \in A$. Introduce the approximate value functions $(v_t^{n,m})_{t=0}^{T}$ which possess the decomposition

$$v_t^{n,m} = \breve{v}_t^{n,m} - \hat{v}_t^{n,m} \tag{4.18}$$

where $\breve{v}_t^{n,m}(p,\cdot)$ and $\hat{v}_t^{n,m}(p,\cdot)$ are piecewise linear convex functions with matrix representatives

$$\breve{v}_t^{n,m}(p,\cdot) \sim \breve{V}_t(p) \quad \text{and} \quad \hat{v}_t^{n,m}(p,\cdot) \sim \hat{V}_t(p) \tag{4.19}$$

for $t = 0, \ldots, T$, $p \in P$.

Initialization: Start with the matrices

$$\breve{V}_T(p) = \breve{R}_T(p) \quad \text{and} \quad \hat{V}_T(p) = \hat{R}_T(p), \qquad \text{for all } p \in P.$$

Recursion: For $t = T-1, \ldots, 1$, calculate

$$\begin{aligned} \breve{\Psi}_t(p,a) &= \breve{R}_t(p,a) + \textstyle\sum_{k=1}^{n} \nu_{t+1}(k)\Upsilon[\breve{V}_{t+1}(\alpha(p,a)) \cdot W_{t+1}(k)] \\ \hat{\Psi}_t(p,a) &= \hat{R}_t(p,a) + \textstyle\sum_{k=1}^{n} \nu_{t+1}(k)\Upsilon[\hat{V}_{t+1}(\alpha(p,a)) \cdot W_{t+1}(k)] \end{aligned} \tag{4.20}$$

and determine

$$\begin{aligned} \hat{V}_t(p) &= \textstyle\sum_{a\in A} \hat{\Psi}_t(p,a) \\ \breve{V}_t(p) &= \bigsqcup_{a\in A}(\breve{\Psi}_t(p,a) - \hat{\Psi}_t(p,a) + \hat{V}_t(p)). \end{aligned} \tag{4.21}$$

for all $p \in P$.

4.5 An efficient approximation

Although numerical experiments indicate stable and reliable results, it seems that the computational performance suffers from the fact that most of the calculation time is

being spent on matrix rearrangements required by the operator Υ. We see from (4.13) that in order to calculate

$$\sum_{k=1}^{n} \nu_{t+1}(k) \Upsilon[V_{t+1}(\alpha(p,a)) \cdot W_{t+1}(k)] \tag{4.22}$$

at each step of the recursion, row-rearrangement must be performed n times, once for each disturbance matrix multiplication. This task becomes increasingly demanding for larger values of the disturbance sampling sizes n, particularly in high dimensions. Before we proceed, let us omit the time index $t+1$ in (4.22) to ease notation. We then focus on the two major sources of computational effort in evaluation of this expression, namely

$$\begin{gathered} \text{the rearrangement } \Upsilon[VW(k)] \text{ of} \\ \text{large matrices } V \cdot W(k) \end{gathered} \tag{4.23}$$

and

$$\begin{gathered} \text{the summation of matrices } \Upsilon[V \cdot W(k)] \text{ over} \\ \text{a large index range } k = 1, \ldots, n. \end{gathered} \tag{4.24}$$

The remainder of this section will be divided into two parts. In Section 4.5.1, we present a method that approximates (4.22), and addresses both problems simultaneously. The improvement in computational effort makes it feasible to obtain *approximate* solutions for large grids and distribution samples sizes. Furthermore, we will see that unlike (4.22), this approximation does not require $V = V_{t+1}$ to be convex. In Section 4.5.2, we derive a suitable first order approximation that provides a efficient way of evaluating functions without having to decompose them into convex components. By combining this approximation with the method in Section 4.5.1, we obtain an efficient algorithm where we are no longer encumbered by the requirement of convexity.

4.5.1 Estimating the conditional expectation

The crucial point is that one can approximate the procedure in (4.23) by replacing the row-rearrangement operation with an appropriate matrix multiplication. More precisely, for $k = 1, \ldots, n$ we

$$\begin{gathered} \text{construct a matrix } Y(k) \text{ such that} \\ Y(k)VW(k) \text{ approximates } \Upsilon[VW(k)]. \end{gathered} \tag{4.25}$$

Before we justify the approximation (4.25), let first us see how it can be used to address the computational problem associated with (4.24). Given (4.25), we now have the following approximation to (4.22)

$$\sum_{k=1}^{n} \nu(k) \Upsilon[VW(k)] \approx \sum_{k=1}^{n} \nu(k) Y(k) VW(k) \tag{4.26}$$

and this in turn requires an efficient calculation of sums of matrices. In practical examples, the distribution sample size n and the grid size m (row number of V) will typically be orders of magnitude of the dimension d of the disturbance matrices $W(k)$. For instance, to achieve an acceptable level of numerical convergence in typical applications, the sample size n and the grid size m must be chosen in the range of several thousands, whereas the state size dimension d is typically of several dozens. This insight shows that a significant reduction in computational effort can be achieved by an additive decomposition of the disturbance realizations. Assume that disturbance matrix W is represented as the linear combination

$$W = \bar{W} + \sum_{j=1}^{J} \epsilon_j E(j) \tag{4.27}$$

with non-random matrices $\bar{W}$ and $(E(j))_{j=1}^{J}$, and random coefficients $(\varepsilon_j)_{j=1}^{J}$. With this decomposition, each realization $W(k)$ of the disturbance matrix W is obtained as

$$W(k) = \bar{W} + \sum_{j=1}^{J} \epsilon_j(k) E(j), \qquad k = 1, \ldots, n. \tag{4.28}$$

Utilizing this, we obtain the following interchange of summations on the right-hand side of (4.26):

$$\sum_{k=1}^{n} \nu(k) Y(k) VW(k) = \left(\sum_{k=1}^{n} \nu(k) Y(k) \right) V\bar{W} + \sum_{j=1}^{J} \left(\sum_{k=1}^{n} \nu(k) \epsilon_j(k) Y(k) \right) VE(j).$$

If one pre-computes the following matrices

$$D_0 = \sum_{k=1}^{n} \nu(k) Y(k), \qquad D_j = \sum_{k=1}^{n} \nu(k) \epsilon_j(k) Y(k), \quad j = 1, \ldots, J, \tag{4.29}$$

we then obtain a significant simplification to (4.26)

$$\sum_{k=1}^{n} \nu(k) Y(k) VW(k) = D_0 V\bar{W} + \sum_{j=1}^{J} D_j VE(j) \tag{4.30}$$

which only involves a low number of matrix summations and multiplications. We shall denote this efficient calculation of the conditional expectation by $\mathcal{E}$ where

$$\mathcal{E}(V) := D_0 V \bar{W} + \sum_{j=1}^{J} D_j V E(j). \tag{4.31}$$

We now address the justification of the approximation in (4.25). Suppose that the grid $\{g^1, \ldots, g^m\}$ is represented by the matrix G where each row i contains row vector, representing the grid point g^i. Thus G will consist of m rows with $G_{i,\cdot} = g^i$ for $i = 1, \ldots, m$. Now let $\tilde{L} = \Upsilon[L]$ be the result the application of Υ to a matrix L. The matrix $\tilde{L}$ is then characterized by the following requirements:

$$\begin{gathered}\tilde{L} = \Upsilon[L] \text{ consists of } m \text{ rows which are obtained} \\ \text{from the rows of } L \text{ by a arrangement,}\end{gathered} \tag{4.32}$$

such that

$$\tilde{L}_{i,\cdot} \cdot G_{i,\cdot}^{\top} \geq L_{j,\cdot} \cdot G_{i,\cdot}^{\top} \quad \text{for all } i, j = 1, \ldots, m. \tag{4.33}$$

According to requirement (4.32), we therefore assume that

$$\begin{gathered}\Upsilon[VW(k)] \text{ consists of } m \text{ rows which are obtained} \\ \text{from the rows of } VW(k) \text{ by row-rearrangement.}\end{gathered} \tag{4.34}$$

Since any row rearrangement can be achieved by a left-multiplication with appropriate matrix, there will always exist a permutation matrix $Y_V(k)$ such that

$$Y_V(k) VW(k) = \Upsilon[VW(k)]. \tag{4.35}$$

Computing each $Y_V(k)$ requires great effort since it is not only dependent on $W(k)$, but also on each V. We suggest determining a reasonable surrogate $Y(k)$ for $Y_V(k)$ which depends only on $W(k)$ and not on V. Since $Y_V(k)$ must satisfy (4.35), we observe with (4.33) in mind that

$$(Y_V(k)V)_{i,\cdot} \cdot (W(k)G)_{i,\cdot}^{\top} \geq V_{j,\cdot} \cdot (W(k)G)_{i,\cdot}^{\top} \quad \text{for } i, j = 1, \ldots, m. \tag{4.36}$$

Now, for each $i = 1, \dots, m$ consider the row $(W(k)G)_{i,\cdot}$ and determine the closest row $G_{h_k(i),\cdot}$ in the original grid matrix by

$$h_k(i) = \operatorname{argmin}\{j = 1, \dots, m \,:\, \|(W(k)G)_{i,\cdot} - G_{j,\cdot}\|\}, \quad i = 1, \dots, m. \tag{4.37}$$

With this proximity function $h_k : \{1, \dots, m\} \to \{1, \dots, m\}$, we may consider, in place of the relation (4.36), the condition

$$(Y(k)V)_{i,\cdot} \cdot G^{\top}_{h_k(i),\cdot} \geq V_{j,\cdot} \cdot G^{\top}_{h_k(i),\cdot} \quad \text{for all } i, j = 1, \dots, m \tag{4.38}$$

with an appropriate permutation matrix $Y(k)$. While (4.38) is clearly not equivalent to (4.36), it does provide a reasonable approximation when the grid is sufficiently dense. Now define $Y(k)$ to be such that

$$Y(k)_{i,j} = \begin{cases} 1, & \text{if } j = h_k(i) \\ 0, & \text{otherwise} \end{cases} \tag{4.39}$$

and observe that with this permutation matrix $Y(k)$, the following assertion

$$V_{h_k(i),\cdot} \cdot G^{\top}_{h_k(i),\cdot} \geq V_{j,\cdot} \cdot G^{\top}_{h_k(i),\cdot} \quad \text{for all } i, j = 1, \dots, m$$

holds if V is in the normal form $\Upsilon[V] = V$. That is, the required approximation (4.25) is determined by (4.39).

The pre-calculations involved in the approximation of (4.22) (i.e. computing $D_0, \dots, D_J$) are computationally demanding. Thus, a gain in computation performance can only be realized if disturbances $(W_t)_{t=1}^T$ are identically distributed whereby the pre-calculations need only be done once. In this case, the ideas presented in this section will be encapsulated in the following algorithm.

Pre-calculations: Determine a sampling $(W(k))_{k=1}^n$ from the target distribution. For each disturbance $W(k)$, find the corresponding permutation matrix $Y(k)$ as in (4.39) using the proximity function (4.37). Use these matrices and the components of the decomposition described in (4.28) of each $W(k)$ to compute the matrices (4.29).

Continuation: Execute the algorithm (4.17) – (4.21) but replace (4.20) with

$$\begin{aligned} \breve{\Psi}_t(p,a) &= \breve{R}_t(p,a) + \mathcal{E}[\breve{V}_{t+1}(\alpha(p,a))] \quad \text{and} \\ \hat{\Psi}_t(p,a) &= \hat{R}_t(p,a) + \mathcal{E}[\hat{V}_{t+1}(\alpha(p,a))]. \end{aligned}$$

by substituting the conditional expectations with its efficient counterpart (4.31).

4.5.2 A direct approach

So far, we have worked with parallel procedures on convex functions. However, an important point to note is that in no part of the efficient conditional expectation procedure was the convexity of the target function required. With this in mind, we shall now present a further simplification to this algorithm based on a first-order approximation. Previously, we considered the convex decomposition $f = \hat{f} - \breve{f}$ of a non-convex function function f where the two convex piecewise linear functions $\hat{f}$ and $\breve{f}$ with respective matrix representatives $\hat{F}$ and $\breve{F}$. The value $f(z)$ at point z is then calculated as

$$f(z) = \max(\hat{F}z) - \max(\breve{F}z).$$

However, if only the matrix difference $\hat{F} - \breve{F}$ is known then it is possible to use a first-order approximation

$$f(z) \approx (\hat{F} - \breve{F})(h(z)).$$

where h is the so-called host function of the underlying grid G

$$h(z) = \text{argmin}\{\|z - g\| : g \in G\}$$

which returns to each argument $z \in \mathbb{R}^d$ the so-called host - the point on the grid with the smallest distance to z. The first-order approximation uses the difference $\hat{F} - \breve{F}$ directly and unlike convex decomposition, does not require a separate calculation of convex and concave parts. If one decides to use this first-order approximation to access the functions, then there is no need to trace convex and concave part separately. This gives a significant simplification and results in the following direct algorithm:

Pre-calculations: Determine the operator $\mathcal{E}$ as in (4.31), under the assumptions required therefore. Determine for $p \in P$, $a \in A$ and $t = 0, \dots, T-1$ the matrices

$$R_t(p,a) = \hat{R}_t(p,a) - \check{R}_t(p,a), \quad R_T(p) = \hat{R}_T(p) - \check{R}_T(p), \tag{4.40}$$

which are obtained as in (4.17). Introduce the approximate value functions, their convex decomposition and representatives as in (4.18) and (4.19). The matrices

$$V_t(p) = \check{V}_t(p) - \hat{V}_t(p) \quad \text{for } t = 0, \dots, T,\ p \in P$$

are obtained via the following scheme:

Initialization: Start with the matrices

$$V_T(p) = \check{R}_T(p) - \hat{R}_T(p), \qquad \text{for all } p \in P$$

Recursion: For $t = T-1, \dots, 1$ calculate for $p \in P$

$$V_t(p) = \bigsqcup_{a \in A} \left(R_t(p,a) + \mathcal{E}(V_{t+1}(p,a))\right). \tag{4.41}$$

Remark: Unlike in the convex decomposition case (4.17) – (4.21), the *direct* algorithm (4.40) – (4.41) merely returns the difference $V_t(p) = \check{V}_t(p) - \hat{V}_t(p)$. That is, the access to the approximate value functions is provided via

$$v_t^{m,n}(p,z) \approx V_t(p) \cdot h(z),$$

using the host function $z \mapsto h(z) = \operatorname{argmin}\{\|z - g\| \ : \ g \in G^m\}$ of the grid G^m. In particular, we suggest an approximation of the optimal policy $\pi^{m,n} = (\pi_t^{m,n})_{t=0}^{T-1}$ as

$$\pi_t^{m,n}(p,z) = \operatorname{argmax}_{a \in A} \left((R_t(p,a) + \mathcal{E}(V_{t+1}(p,a))) \cdot h(z)\right), \tag{4.42}$$

for $t = 0, \dots, T-1$, $z \in \mathbb{R}^d$, $p \in P$. To obtain an efficient implementation of a host function h, a tree-like structure on the grid can be used which may be established using hierarchical clustering methods.

4.6 Solution diagnostics

In Section 4.5.1, we derived a heuristic method to obtain an efficient approximation to the conditional expectation in the Bellman recursion. In Section 4.5.2, we saw that by combining this with a first order approximation, we were then able to obtain the approximation (4.42) for a given grid to the optimal policy in (4.3). In order to address the distance to optimality of this approximate solution, we first need to outline an appropriate measure for this distance.

Suppose we are given an arbitrary policy $\pi = (\pi_t)_{t=0}^{T-1}$. For such a policy one can define an associated set of *policy values* $(v_t^\pi(p,z))_{t=0}^T$ that follow the recursion

$$v_T^\pi(p,z) = r_T(p,z) \tag{4.43}$$

$$v_t^\pi(p,z) = r_t(p,z,\pi_t(p,z)) + \mathbb{E}(v_{t+1}^\pi(\alpha(p,\pi_t(p,z)), W_{t+1}z)), \tag{4.44}$$

for $t = T-1,\dots,0$. Let us consider a switching system which starts in a given initial position $p_0^\pi = p_0 \in P$ and state $Z_0 = z_0 \in \mathbb{R}^d$. At any time t, the actions and new positions are determined recursively, following policy $\pi = (\pi_t)_{t=0}^{T-1}$ as

$$a_t^\pi := \pi_t(p_t^\pi, Z_t), \quad p_{t+1}^\pi := \alpha(p_t^\pi, a_t^\pi), \quad t = 0,\dots,T-1.$$

These values define a *policy run* $\mathcal{V}_0^\pi(p_0^\pi, z_0)$ where

$$\mathcal{V}_0^\pi(p_0^\pi, Z_0) = \sum_{s=0}^{T-1} r_t(p_s^\pi, Z_s, a_s^\pi) + r_T(p_T^\pi, Z_T)$$

According to the definition, $v_0^\pi(p_0,z_0)$ is the expected value of the policy run

$$v_0^\pi(p_0,z_0) = \mathbb{E}(\mathcal{V}_0^\pi(p_0,z_0)) \quad p \in P, \quad z \in \mathbb{R}^d.$$

In practice, one can use Monte Carlo to estimate this value since given a sequence $(\omega_k)_{k\in\mathbb{N}}$ independent random draws,

$$v_0^\pi(p_0,z_0) = \mathbb{E}(\mathcal{V}_0^\pi(p_0,z_0)) = \lim_{k\to\infty} \frac{1}{K} \sum_{k=1}^{K} \mathcal{V}_0^\pi(p_0,z_0)(\omega_k), \tag{4.45}$$

holds true due to the strong law of large numbers.

Such a Monte Carlo procedure may be useful in estimating the performance of a given policy π. However, it does not clarify how far the value $v_0^{\pi}(p_0, z_0)$ is from the value $v_0^{\pi^*}(p_0, z_0)$ associated with the optimal policy π^*.

In the reminder of the section, we suggest a sound solution to this question in terms of a *diagnostic method*. Given a starting point (p_0, z_0), we explain how the gap

$$[v_0^{\pi}(p_0,\, z_0), v_0^{\pi^*}(p_0, z_0)] \tag{4.46}$$

between a given strategy π and the optimal strategy π^* can be assessed. Our methodology is based on a finite sample $\{\omega_1, \ldots, \omega_K\}$ of trajectory realizations and utilizes to a build-in variance reduction technique to derive *tight confidence bounds* for upper and lower estimates of the interval (4.46).

Let us focus on the upper bound first. Consider a sequence $\varphi = (\varphi_t)_{t=1}^T$ of random mappings

$$\varphi_t : P \times \mathbb{R}^d \times A \times \Omega \to \mathbb{R}, \qquad (p, z, a, \omega) \mapsto \varphi_t(p, z, a)(\omega), \tag{4.47}$$

which for $t = 1, \ldots, T$ satisfy

$$\mathbb{E}(\varphi_t(p, z, a)) = 0, \qquad p \in P, z \in \mathbb{R}^d, a \in A, \tag{4.48}$$

and such that the σ-algebras

$$\sigma(\varphi_t(p, z, a), W_t; a \in A, z \in \mathbb{R}^d), \qquad t = 1, \ldots, T, \tag{4.49}$$

are independent. Given these mappings $\varphi = (\varphi_t)_{t=1}^T$, we now introduce the random functions $(\bar{v}_t^{\varphi})_{t=0}^T$

$$\bar{v}_t^{\varphi} : P \times \mathbb{R}^d \times \Omega \to \mathbb{R}, \qquad t = 0, \ldots, T$$

which are recursively defined for $t = T, \ldots, 1$ via

$$\bar{v}_T^{\varphi}(p, z) = r_T(p, z) \tag{4.50}$$

$$\bar{v}_t^{\varphi}(p, z) = \max_{a \in A} \big(r_t(p, z, a) + \varphi_{t+1}(p, z, a) + \bar{v}_{t+1}^{\varphi}(\alpha(p, a), W_{t+1} z)\big). \tag{4.51}$$

Using $(\bar{v}_t^{\varphi})_{t=0}^T$, the following theorem holds:

Theorem 4.1. *(i)For each policy* $\pi = (\pi_t)_{t=0}^{T-1}$*, it holds that the policy values* $(v_t^\pi)_{t=0}^T$ *are dominated from above*

$$v_t^\pi(p,z) \le \mathbb{E}(\bar{v}_t^\varphi(p,z)), \qquad \text{for all } t=0,\dots,T,\ p\in P,\ z\in\mathbb{R}^d. \tag{4.52}$$

(ii) Given the policy values $(v_t^{\pi^*})_{t=0}^T$ *associated with the optimal policy* $\pi^* = (\pi_t^*)_{t=0}^{T-1}$*, let* $(\varphi_t^*)_{t=1}^T$ *be defined by*

$$\varphi_{t+1}^*(p,z,a) = \mathbb{E}(v_{t+1}^{\pi^*}(\alpha(p,a),W_{t+1}z)) - v_{t+1}^{\pi^*}(\alpha(p,a),W_{t+1}z) \tag{4.53}$$

for all $p\in P$*,* $z\in\mathbb{R}^d$*,* $a\in A$ *and* $t=0,\dots,T-1$*. It then holds that the mappings* $(\varphi_t^*)_{t=1}^T$ *satisfy (4.47) – (4.49) and that (4.52) holds with equality*

$$v_t^{\pi^*}(p,z) = \bar{v}_t^{\varphi^*}(p,z), \qquad \text{for all } t=0,\dots,T,\ p\in P,\ z\in\mathbb{R}^d. \tag{4.54}$$

Proof. (i) The value $(v_t^\pi)_{t=0}^T$ of the policy $\pi = (\pi_t)_{t=0}^{T-1}$ satisfies the recursion (4.44). Using this recursion and (4.48) we obtain

$$\begin{aligned} v_t^\pi(p,z) &= \mathbb{E}(r_t(p,z,\pi_t(p,z)) + \varphi_{t+1}(p,z,\pi_t(p,z))) \\ &\quad + \mathbb{E}(v_{t+1}^\pi(\alpha(p,\pi_t(p,z)),W_{t+1}z)). \end{aligned} \tag{4.55}$$

Now, let us prove the assertion (4.52) by induction. For $t=T$, the inequality (4.52) holds with equality because of the initialization

$$v_T^\pi = r_T = \bar{v}_T^\varphi \tag{4.56}$$

in (4.43) and (4.50). Given the induction assumption

$$v_{t+1}^\pi(p,z) \le \mathbb{E}(\bar{v}_{t+1}^\varphi(p,z)), \qquad \text{for all } p\in P,\ z\in\mathbb{R}^d,$$

we use (4.49) to conclude that

$$v_{t+1}^\pi(\alpha(p,\pi_t(p,z)),W_{t+1}z) \le \mathbb{E}(\bar{v}_{t+1}^\varphi(\alpha(p,\pi_t(p,z)),W_{t+1}z)\,|\,W_{t+1})$$

holds for all $p \in P$, $z \in \mathbb{R}^d$. Using this, we obtain in (4.55) an estimate

$$\begin{aligned} v_t^\pi(p,z) &\leq \mathbb{E}\left(r_t(p,z,\pi_t(p,z)) + \varphi_{t+1}(p,z,\pi_t(p,z))\right) \\ &\quad + \mathbb{E}\left(\mathbb{E}(\bar{v}_{t+1}^{\varphi}(\alpha(p,\pi_t(p,z)), W_{t+1}z) \,|\, W_{t+1})\right) \end{aligned}$$

from which the assertion follows

$$\begin{aligned} v_t^\pi(p,z) &\leq \mathbb{E}\Big(r_t(p,z,\pi_t(p,z)) + \varphi_{t+1}(p,z,\pi_t(p,z)) + \bar{v}_{t+1}^{\varphi}(\alpha(p,\pi_t(p,z)), W_{t+1}z)\Big) \\ &\leq \mathbb{E}\left(\max_{a\in A}\left[r_t(p,z,a) + \varphi_{t+1}(p,z,a) + \bar{v}_{t+1}^{\varphi}(\alpha(p,a), W_{t+1}z)\right]\right) \\ &\leq \mathbb{E}\left(\bar{v}_t^{\varphi}(p,z)\right), \end{aligned}$$

where the last step results from the recursion (4.51).

(ii) Now suppose that π^* is an optimal policy and define $\varphi^* = (\varphi_t^*)_{t=1}^T$ as in (4.53), which satisfies the assumption (4.48). Furthermore, the independence (4.49) holds since for $t = 0, \ldots, T-1$ the random component in $\varphi_{t+1}^*(p,z,a)$ is

$$v_{t+1}^{\pi^*}(\alpha(p,a), W_{t+1}z)$$

which is in turn a function of W_{t+1}. Let us now verify the assertion (4.54). By induction which is started as in (4.50) we can assume that $v_{t+1}^{\pi^*} = \bar{v}_{t+1}^{\varphi^*}$ holds. Using this, we conclude for all $p \in P$, $z \in \mathbb{R}^d$ and $a \in A$ the assertion

$$\begin{aligned} &r_t(p,z,a) + \varphi_{t+1}^*(p,z,a) + \bar{v}_{t+1}^{\varphi^*}(\alpha(p,a), W_{t+1}z) \\ &\quad = r_t(p,z,a) + \varphi_{t+1}^*(p,z,a) + v_{t+1}^{\pi^*}(\alpha(p,a), W_{t+1}z). \end{aligned} \tag{4.57}$$

On the other hand, using (4.53), we infer that for all $p \in P$, $z \in \mathbb{R}^d$ and $a \in A$ it holds that

$$\begin{aligned} &r_t(p,z,a) + \varphi_{t+1}^*(p,z,a) + v_{t+1}^{\pi^*}(\alpha(p,a), W_{t+1}z) \\ &\quad = r_t(p,z,a) + \mathbb{E}(v_{t+1}^{\pi^*}(\alpha(p,a), W_{t+1}z)) \qquad (4.58) \\ &\qquad - v_{t+1}^{\pi^*}(\alpha(p,a), W_{t+1}z) + v_{t+1}^{\pi^*}(\alpha(p,a), W_{t+1}z) \\ &\quad = r_t(p,z,a) + \mathbb{E}(v_{t+1}^{\pi^*}(\alpha(p,a), W_{t+1}z)) \qquad (4.59) \end{aligned}$$

Now, in the recursion (4.51) we replace (4.57) by (4.59) to obtain the desired result (4.54)

$$\begin{aligned} \bar{v}_t^{\varphi^*}(p,z) &= \max_{a\in A}\Big(r_t(p,z,a)+\varphi_{t+1}^*(p,z,a)+\bar{v}_{t+1}^{\varphi^*}(\alpha(p,a),W_{t+1}z)\Big) \\ &= \max_{a\in A}\Big(r_t(p,z,a)+\mathbb{E}(v_{t+1}^{\pi^*}(\alpha(p,a),W_{t+1}z))\Big) \\ &= v_t^{\pi^*}(p,z), \qquad p\in P,\ \ z\in\mathbb{R}^d,\ \ a\in A. \end{aligned}$$

□

Let us elaborate on a practical application of this technique. Suppose that we attempt to assess the distance to optimality of an approximate policy $\tilde{\pi}$, obtained by a numerical procedures described previously. According to (i) of the Theorem 4.1, an arbitrary $(\varphi_t)_{t=1}^T$ satisfying (4.48) and (4.49) yields an upper bound

$$v_0^{\tilde{\pi}}(p,z) \le v_0^{\pi^*}(p,z) \le \mathbb{E}(\bar{v}_0^{\varphi}(p,z)) \qquad p\in P,\ \ z\in\mathbb{R}^d, \tag{4.60}$$

Note that the expectation $\mathbb{E}(\bar{v}_0^{\varphi}(p,z))$ will be estimated via Monte Carlo. Thus, we obtain the following estimation procedure:

Upper bound estimation:

1) Given a switching system, implement $(\varphi_t)_{t=1}^T$ which fulfills (4.47), (4.48) and (4.49).

2) Chose a number $K\in\mathbb{N}$ of Monte Carlo trials and obtain for $k=1,\dots,K$ independent realizations $(W_t(\omega_k))_{t=1}^T$ of disturbances.

3) Starting at $z_0^k := z_0\in\mathbb{R}^d$, define for $k=1,\dots,K$ the trajectories $(z_t^k)_{t=0}^T$ recursively

$$z_{t+1}^k = W_{t+1}(\omega_k)z_t^k, \qquad t=0,\dots,T-1$$

and determine realizations

$$\varphi_{t+1}(p,z_t^k,a)(\omega_k), \qquad t=0,\dots,T-1, \quad k=1,\dots,K.$$

4) For each $k=1,\dots,K$ initialize the recursion at $t=T$ as

$$\bar{v}_T^{\varphi}(p,z_T^k) = r_T(p,z_T^k) \qquad \text{for all } p\in P$$

and continue for $t = T-1, \dots, 0$ by

$$\bar{v}_t^{\varphi}(p, z_t^k) \quad = \quad \max_{a \in A} \big(r_t(p, z_t^k, a) + \varphi_{t+1}(p, z_t^k, a)(\omega_k) + \bar{v}_{t+1}^{\varphi}(\alpha(p,a), z_{t+1}^k) \big).$$

Store the value as $\bar{v}_0^{\varphi}(p, z_0^k)$ for $k = 1, \dots, K$.

5) Determine the sample mean $\frac{1}{K} \sum_{k=1}^{K} \bar{v}_0^{\varphi}(p, z_0^k)$ and its upper confidence bound to estimate $v_0^{\pi^*}(p, z_0)$ from above.

To obtain a tight upper bound, $(\varphi_t)_{t=1}^T$ must be chosen accordingly. Thereby, the assertion (ii) of Theorem 4.1 suggests an appropriate choice. Namely, in the hypothetical case that the optimal policy value functions $(v_t^{\pi *})_{t=0}^T$ are known, the $(\varphi_t^*)_{t=1}^T$ is obtained via (4.53) will give an exact and non-random upper bound. In practice, this situation is not feasible, since an optimal strategy π^* is not known. Instead, we suggest using an approximate value function $(\tilde{\varphi}_t)_{t=0}^T$, returned by one of the algorithms described in this work. That is, following (4.53), a reasonable candidate for $t = 0, \dots, T-1$ could be given as

$$\varphi_{t+1}(p, z, a) \quad = \quad \mathbb{E}(\tilde{v}_{t+1}(\alpha(p,a), W_{t+1} z)) - \tilde{v}_{t+1}(\alpha(p,a), W_{t+1} z). \tag{4.61}$$

However, note that this choice involves an exact calculation of expectation

$$\mathbb{E}(\tilde{v}_{t+1}(\alpha(p,a), W_{t+1} z)),$$

which is not possible in practice. For this reason, we suggest a modification. We reintroduce the φ_{t+1} in (4.61), with the expectation now replaced by an arithmetic mean over a number I of independent copies $(W_{t+1}^{(i)})_{i=1}^I$ of W_{t+1}. That is, given independent random variables W_{t+1} and $W_{t+1}^{(i)}$ for $i = 1, \dots, I$ and $t = 0, \dots, T-1$ such that the distribution of $W_{t+1}^{(i)}$ equals to that of W_{t+1}, we define for all $t = 0, \dots, T-1$, $a \in A$, $p \in P$, and $z \in \mathbb{R}^d$

$$\varphi_{t+1}(p, z, a) = \frac{1}{I} \sum_{i=1}^{I} \tilde{v}_{t+1}(\alpha(p,a), W_{t+1}^{(i)} z) - \tilde{v}_{t+1}(\alpha(p,a), W_{t+1} z). \tag{4.62}$$

With this definition, $(\varphi_t)_{t=1}^T$ satisfies conditions (4.48) and (4.49), and we thus obtain a valid and computable upper bound.

Let us turn now to the estimation of the lower boundary of the interval (4.46). Given a strategy $\pi = (\pi_t)_{t=0}^{T-1}$, the value $v_0^\pi(p_0, z_0)$ can in principle be approached as in (4.45) from test runs of the strategy in a series of independent back-testing experiments. However, it turns out that a slight adaptation of the upper bound technique provides far better results, due to a built-in variance reduction technique. Similar to part (ii) of the previous theorem, which indicates that the variance of the Monte Carlo trials reduces if approximate solution is close to the optimal one, we establish a recursive procedure with a control variate built-in. The idea is simple: Given a nearly-optimal policy $\pi = (\pi_t)_{t=0}^{T-1}$ we alter the recursion (4.50), (4.51) replacing the maximization by an exact choice of the action according to the policy $\pi = (\pi_t)_{t=0}^{T-1}$.

Given a sequence $\varphi = (\varphi_t)_{t=1}^T$ satisfying (4.48) and (4.49) we introduce the random functions $(\underline{v}_t^{\pi,\varphi})_{t=0}^T$

$$\underline{v}_t^{\pi,\varphi} : P \times \mathbb{R}^d \times \Omega \to \mathbb{R}, \qquad t = 0, \dots, T$$

which are recursively defined for $t = T, \dots, 1$ via

$$\begin{aligned} \underline{v}_T^{\pi,\varphi}(p,z) &= r_T(p,z) && (4.63)\\ \underline{v}_t^{\pi,\varphi}(p,z) &= r_t(p,z,\pi_t(p,z)) + \varphi_{t+1}(p,z,\pi_t(p,z)) \\ &\quad + \underline{v}_{t+1}^{\pi,\varphi}(\alpha(p,\pi_t(p,z)), W_{t+1}z). && (4.64) \end{aligned}$$

The following theorem holds for $(\underline{v}_t^{\pi,\varphi})_{t=0}^T$.

Theorem 4.2. *(i) Given $\varphi = (\varphi_t)_{t=1}^T$ as in (4.47) satisfying (4.49) and a policy $\pi = (\pi_t)_{t=0}^{T-1}$, introduce $(\underline{v}_t^{\pi,\varphi})_{t=0}^T$ by (4.63) and (4.64). It holds that*

$$v_t^\pi(p,z) = \mathbb{E}(\underline{v}_t^{\pi,\varphi}(p,z)), \qquad \text{for all } t = 0, \dots, T,\ p \in P,\ z \in \mathbb{R}^d. \tag{4.65}$$

(ii) Given the value $(v_t^{\pi^*})_{t=0}^T$ of the optimal policy $\pi^* = (\pi_t^*)_{t=0}^{T-1}$, define $(\varphi_t^*)_{t=1}^T$ by

$$\varphi_{t+1}^*(p,z,a) = \mathbb{E}(v_{t+1}^{\pi^*}(\alpha(p,a), W_{t+1}z)) - v_{t+1}^{\pi^*}(\alpha(p,a), W_{t+1}z) \tag{4.66}$$

for all $p \in P$, $z \in \mathbb{R}^d$, $a \in A$ and $t = 0, \dots, T-1$. Then the mappings $(\varphi_t^*)_{t=1}^T$ satisfy (4.47) – (4.49) such that (4.65) holds with equality:

$$v_t^{\pi^*}(p,z) = \underline{v}_t^{\pi^*,\varphi^*}(p,z), \qquad \text{for all } t = 0, \dots, T,\ p \in P,\ z \in \mathbb{R}^d. \tag{4.67}$$

Proof. (i) The value $(v_t^\pi)_{t=0}^T$ of the policy $\pi = (\pi_t)_{t=0}^{T-1}$ satisfies the recursion (4.44). Using this recursion and (4.48) we obtain

$$\begin{aligned} v_t^\pi(p,z) &= \mathbb{E}(r_t(p,z,\pi_t(p,z)) + \varphi_{t+1}(p,z,\pi_t(p,z))) \\ &\quad + \mathbb{E}(v_{t+1}^\pi(\alpha(p,\pi_t(p,z)), W_{t+1}z). \end{aligned} \tag{4.68}$$

Now, let us prove the assertion (4.65) by induction. For $t = T$ the inequality (4.65) holds with equality because of the initialization

$$v_T^\pi = r_T = \underline{v}_T^{\pi,\varphi} \tag{4.69}$$

in (4.43) and (4.63). Given the induction assumption

$$v_{t+1}^\pi(p,z) = \mathbb{E}(\underline{v}_{t+1}^{\pi,\varphi}(p,z)), \qquad \text{for all } p \in P,\ z \in \mathbb{R}^d,$$

we use (4.49) to conclude that

$$v_{t+1}^\pi(\alpha(p,\pi_t(p,z)), W_{t+1}z) = \mathbb{E}(\underline{v}_{t+1}^{\pi,\varphi}(\alpha(p,\pi_t(p,z))) \,|\, W_{t+1})$$

holds for all $p \in P$, $z \in \mathbb{R}^d$. Using this, we obtain in (4.68) the equality

$$\begin{aligned} v_t^\pi(p,z) &= \mathbb{E}\left(r_t(p,z,\pi_t(p,z)) + \varphi_{t+1}(p,z,\pi_t(p,z))\right) \\ &\quad + \mathbb{E}\left(\mathbb{E}(\underline{v}_{t+1}^{\pi,\varphi}(\alpha(p,\pi_t(p,z))) \,|\, W_{t+1})\right) \\ &= \mathbb{E}\big(r_t(p,z,\pi_t(p,z)) + \varphi_{t+1}(p,z,\pi_t(p,z)) + \underline{v}_{t+1}^{\pi,\varphi}(\alpha(p,\pi_t(p,z)))\big). \end{aligned}$$

By using the recursion (4.64), the assertion (4.65) follows.

(ii) Let us now verify the assertion (4.67). By induction which is started as in (4.63) we can assume that $v_{t+1}^{\pi^*} = \underline{v}_{t+1}^{\pi^*,\varphi^*}$ holds. Using this, we conclude for all $p \in P$, $z \in \mathbb{R}^d$ and $a \in A$ the assertion

$$\begin{aligned} &r_t(p,z,a) + \varphi_{t+1}^*(p,z,a) + \underline{v}_{t+1}^{\pi^*,\varphi^*}(\alpha(p,a), W_{t+1}z) \\ &\quad = r_t(p,z,a) + \varphi_{t+1}^*(p,z,a) + v_{t+1}^{\pi^*}(\alpha(p,a), W_{t+1}z). \end{aligned} \tag{4.70}$$

On the other hand, using (4.53), we infer that for all $p \in P$, $z \in \mathbb{R}^d$ and $a \in A$ it holds that

$$\begin{aligned} & r_t(p,z,a) + \varphi_{t+1}^*(p,z,a) + v_{t+1}^{\pi^*}(\alpha(p,a), W_{t+1}z) \\ &= r_t(p,z,a) + \mathbb{E}(v_{t+1}^{\pi^*}(\alpha(p,a), W_{t+1}z)) \\ &\quad - v_{t+1}^{\pi^*}(\alpha(p,a), W_{t+1}z) + v_{t+1}^{\pi^*}(\alpha(p,a), W_{t+1}z) \\ &= r_t(p,z,a) + \mathbb{E}(v_{t+1}^{\pi^*}(\alpha(p,a), W_{t+1}z)). \end{aligned} \tag{4.71}$$

Now, in the recursion (4.64) we replace (4.70) by (4.71) to obtain the desired claim (4.54):

$$\begin{aligned} \underline{v}_t^{\pi^*,\varphi^*}(p,z) &= r_t(p,z,\pi_t^*(p,z)) + \varphi_{t+1}^*(p,z,\pi_t^*(p,z)) \\ &\quad + \underline{v}_{t+1}^{\pi^*,\varphi^*}(\alpha(p,\pi_t^*(p,z)), W_{t+1}z) \\ &= r_t(p,z,\pi_t^*(p,z)) + \mathbb{E}(v_{t+1}^{\pi^*}(\alpha(p,\pi_t^*(p,z)), W_{t+1}z) \\ &= v_t^{\pi^*}(p,z), \qquad p \in P,\ z \in \mathbb{R}^d,\ a \in A. \end{aligned}$$

□

The practical implementation of the lower bound estimation is based on the same realization of $(\varphi_t)_{t=1}^T$ as in (4.62), using independent copies of disturbances. Let us summarize this procedure as follows:

Lower bound estimation:

1) Given approximate value functions $(\tilde{v}_t)_{t=0}^T$ and a corresponding strategy $\tilde{\pi} = (\tilde{\pi}_t)_{t=0}^{T-1}$, chose $\varphi = (\varphi_t)_{t=0}^{T-1}$ as in (4.62).

2) Given $K \in \mathbb{N}$ Monte Carlo trials, obtain for $k = 1, \dots, K$ independent realizations $(W_t(\omega_k))_{t=1}^T$ of disturbances.

3) Starting at $z_0^k := z_0 \in \mathbb{R}^d$, define for $k = 1, \dots, K$ trajectories $(z_t^k)_{t=0}^T$ recursively

$$z_{t+1}^k = W_{t+1}(\omega_k) z_t^k, \qquad t = 0, \dots, T-1$$

and determine realizations

$$\varphi_{t+1}(p, z_t^k, a)(\omega_k), \qquad t = 0, \ldots, T-1, \quad k = 1, \ldots, K.$$

4) For each $k = 1, \ldots, K$ initialize the recursion at $t = T$ as

$$\underline{v}_T^{\tilde{\pi},\varphi}(p, z_T^k) = r_T(p, z_T^k) \qquad \text{for all } p \in P$$

and continue for $t = T-1, \ldots, 0$ and for all $p \in P$ by

$$\begin{aligned} \underline{v}_t^{\tilde{\pi},\varphi}(p, z_t^k) &= r_t(p, z_t^k, \tilde{\pi}_t(p, z_t^k)) + \varphi_{t+1}(p, z_t^k, \tilde{\pi}_t(p, z_t^k))(\omega_k) \\ &\quad + \underline{v}_{t+1}^{\tilde{\pi},\varphi}(\alpha(p, \tilde{\pi}_t(p, z_t^k)), z_{t+1}^k). \end{aligned} \tag{4.72}$$

Store the value as $\underline{v}_0^{\tilde{\pi},\varphi}(p, z_0^k)$ for $k = 1, \ldots, K$, $p \in P$.

5) Calculate the sample mean $\frac{1}{K}\sum_{k=1}^{K} \underline{v}_0^{\tilde{\pi},\varphi}(p, z_0^k)$ and use its lower confidence bounds to estimate $v_0^{\pi^*}(p, z_0)$ for each $p \in P$ from below.

4.7 Examples

In this section, we provide numerical results for the cases of the Bermudan option and the swing option and compare it to those given in Longstaff and Schwartz [2001] and Meinshausen and Hambly [2004] where numerical results were obtained using regression-based methods.

We will now perform value function approximations using the method outlined in Section 4.5 and the associated diagnostics established in Section 4.6 on two examples of Markov decision problems - the optimal stopping problem and the optimal multiple stopping problem. Such problems constitute two important subclasses of Markov decision problems (see Chapters 10 and 11 of Bäuerle and Rieder [2011]) and have studied upper bound estimation using duality approaches them in the previous two chapters. To illustrate our approach, we obtain in Section 4.7.1 bounds on the price of the *Bermudan* put option and in Section 4.7.2, we use these methods to obtain bounds on the price of the swing option.

For both applications, we will consider the evolution of the discounted asset price $(S_t)_{t=0}^T$ in discrete time, with respect to a risk-neutral measure. The dynamics $(S_t)_{t=0}^T$ of the discounted price depends on the asset type. For the Bermudan put, the discounted price process $(S_t)_{t=0}^T$ is modelled as a martingale in the risk-neutral measure. For the swing option, we suppose that the price process $(S_t)_{t=0}^T$ is modeled by the exponential of an Ornstein-Uhlenbeck process to explain the mean-reverting price property naturally expected for commodity prices.

The logarithm $(\tilde{Z}_t)_{t=0}^T$ of the price forms the continuous component of our state dynamics. In practice, a further transformation of the state space is usually required before linear state dynamics can be achieved. In most cases, an augmentation with 1 via

$$Z_t = \begin{bmatrix} \tilde{Z}_t \\ 1 \end{bmatrix}, \qquad t = 0, \ldots, T.$$

is needed to represent the evolution of the continuous state component. In this representation, the system state follows a multiplicative dynamic

$$Z_{t+1} = W_{t+1} Z_t, \qquad t = 0, \ldots, T-1$$

with independent and identically distributed matrix-valued random variables $(W_t)_{t=1}^T$. The entries of these disturbance matrices reflect the underlying process model.

The grid choice is a key ingredient in the algorithm. For multi-variate state processes, a convenient way of grid construction is by simulation of appropriate trajectories. Thus, we create a grid of a desired size by simulating and storing a sufficient number of paths of $(Z_t)_{t=0}^{k_p T}$ of an appropriate length $k_p T \in \mathbb{N}$. In our examples, we have used a number of steps that is twice of the time horizon ($k_p = 2$). The distribution of disturbances is approximated by a discrete distribution. For this, a sample of $(W(k))_{k=1}^n$ of independent realizations was generated and stored. All required steps from Section 4.5 and the Monte Carlo simulation for diagnostics refer to this discrete distribution approximation. For bound computations, we use confidence intervals based on K simulated trajectories. More precisely, we quote the intervals as

$$\left[\underline{\mu} - \Phi^{-1}(1 - \frac{x}{2}) \frac{\underline{\sigma}}{\sqrt{K}}, \quad \overline{\mu} + \Phi^{-1}(1 - \frac{x}{2}) \frac{\overline{\sigma}}{\sqrt{K}} \right] \tag{4.73}$$

where $1-x$ denotes the confidence level and $(\underline{\mu}, \underline{\sigma})$ and $(\overline{\mu}, \overline{\sigma})$ denote the sample mean and sample standard deviation of $(\underline{v}_0^{\tilde{\pi},\varphi}(p, z_0^k))_{k=1}^K$ and $(\bar{v}_0^{\varphi}(p, z_0^k))_{k=1}^K$ respectively.

4.7.1 The Bermudan put option

This option allows the holder to sell the underlying asset at a pre-specified strike price on a discrete set of exercise dates up to and including the expiry date of the option. The fair price of a Bermudian put is given by the supremum

$$\sup_{\tau} \mathbb{E}[(Ke^{-\rho\tau} - S_\tau)^+]$$

where τ runs through all $\{0, \ldots, T\}$-valued stopping times. First let us express this control problem as a switching system. We use the position set $P = \{1, 2\}$ to indicate whether the option has been exercised ($p = 1$) or not ($p = 2$). The action set $A = \{1, 2\}$ represents the choice between exercising ($a = 1$) or not exercising ($a = 2$). The control α of the discrete component of the state space

$$(\alpha(p,a))_{p,a=1}^2 \sim \begin{bmatrix} \alpha(1,1) & \alpha(1,2) \\ \alpha(2,1) & \alpha(2,2) \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}$$

ensures that $p = 1$ is absorbing. The continuous state space component follows

$$\tilde{Z}_{t+1} = \tilde{Z}_t + \gamma + \beta\epsilon_{t+1}, \quad \tilde{Z}_0 = \ln(S_0) \tag{4.74}$$

where $(\varepsilon)_{t=1}^T$ are independent standard normally distributed random variables. We set the parameters as $\gamma = -\frac{1}{2}\sigma^2\Delta$ and $\beta = \sigma\sqrt{\Delta}$ where $\Delta > 0$ is the time duration (in years) from time point t to $t+1$ and $\sigma > 0$ represents the volatility of the process that is measured on yearly scale. Given this price process, the disturbances are given as

$$W_{t+1} = \begin{bmatrix} 1 & \gamma + \beta\epsilon_{t+1} \\ 0 & 1 \end{bmatrix}, \quad t = 0, ..., T-1.$$

For the two-dimensional space $\mathbb{R}^2$ with the evolution of the continuous component as above, let us now determine all reward and the scrap functions. Consider a realization of the continuous component $(z^{(1)}, z^{(2)}) \in \mathbb{R}^2$ at the current time $t = 0, \ldots T-1$, then,

given a Bermudian put option, the action $a \in \{1, 2\}$ leads to the reward

$$r_t(p, (z^{(1)}, z^{(2)}), a) = (Ke^{-\rho t} - e^{z^{(1)}})^+ \big(p - \alpha(p, a)\big) \tag{4.75}$$

for all $p \in P$ and $a \in A$. At final time T, we suppose that the put option is exercised automatically, which gives the scrap value

$$r_T(p, (z^{(1)}, z^{(2)})) = (Ke^{-\rho T} - e^{z^{(1)}})^+ \big(p - \alpha(p, 1)\big) \tag{4.76}$$

for all $p \in P$. Note that the reward and scrap functions are not convex in the continuous

TABLE 4.1: Bermudan put option numerical results

S_0	σ	maturity	confidence interval	LSM mean	LSM se
36	0.2	1	[4.4763, 4.4768]	4.472	.0100
36	0.2	2	[4.8296, 4.8312]	4.821	.0120
36	0.4	1	[7.0989, 7.0992]	7.091	.0200
36	0.4	2	[8.4965, 8.4968]	8.488	.0240
38	0.2	1	[3.2481, 3.2489]	3.244	.0090
38	0.2	2	[3.7355, 3.7370]	3.735	.0110
38	0.4	1	[6.1451, 6.1452]	6.139	.0190
38	0.4	2	[7.6580, 7.6583]	7.669	.0220
40	0.2	1	[2.3119, 2.3129]	2.313	.0090
40	0.2	2	[2.8765, 2.8776]	2.879	.0100
40	0.4	1	[5.3093, 5.3094]	5.308	.0180
40	0.4	2	[6.9075, 6.9077]	6.921	.0220
42	0.2	1	[1.6150, 1.6158]	1.617	.0070
42	0.2	2	[2.2053, 2.2060]	2.206	.0100
42	0.4	1	[4.5797, 4.5798]	4.588	.0170
42	0.4	2	[6.2351, 6.2354]	6.243	.0210
44	0.2	1	[1.1081, 1.1087]	1.118	.0070
44	0.2	2	[1.6836, 1.6843]	1.675	.0090
44	0.4	1	[3.9449, 3.9450]	3.957	.0170
44	0.4	2	[5.6324, 5.6326]	5.622	.0210

These results were produced using a grid size of $m = 1024$ and disturbances of size $n = 4096$. Diagnostics is based on $K = 1024$ sample paths and 99% confidence bounds are calculated by setting $x = 0.01$ in (4.73). For comparison, the means and standard errors obtained by least squares Monte Carlo are given in the last two columns *LSM mean* and *LSM se* respectively, they are cited from Longstaff and Schwartz [2001], where numbers were quoted with three decimal points.

component $z = (z^{(1)}, z^{(2)})$. Hence we decompose them into the difference of two convex

functions

$$r_t(p,\cdot,a) = \breve{r}_t(p,\cdot,a) - \hat{r}_t(p,\cdot,a) \tag{4.77}$$

$$r_T(p,\cdot) = \breve{r}_T(p,\cdot) - \hat{r}_T(p,\cdot) \tag{4.78}$$

given by

$$\begin{aligned}
\breve{r}_t(p,(z^{(1)},z^{(2)}),a) &= (e^{z^{(1)}} - Ke^{-\rho t})^+\big(p-\alpha(p,a)\big)\\
\hat{r}_t(p,(z^{(1)},z^{(2)}),a) &= (e^{z^{(1)}} - Ke^{-\rho t})\big(p-\alpha(p,a)\big)\\
\breve{r}_T(p,(z^{(1)},z^{(2)})) &= (e^{z^{(1)}} - Ke^{-\rho T})^+\big(p-\alpha(p,1)\big)\\
\hat{r}_T(p,(z^{(1)},z^{(2)})) &= (e^{z^{(1)}} - Ke^{-\rho T})\big(p-\alpha(p,1)\big)
\end{aligned}$$

for all $p \in P, a \in A$ and $(z^{(1)}, z^{(2)}) \in \mathbb{R}^2$.

We compare our results with the low-biased estimates given in the literature for the Bermudan put where the risk-free rate is 0.06 and the strike is set at 40. The results are given in Table 4.1 for different combinations of initial prices, volatilities and maturities.

4.7.2 The swing option

We consider a specific case of the swing option, referred to as a *unit-time refraction period* condition. This condition limits the holder to exercise one right at any time. Given the discounted asset price $(S_t)_{t=0}^T$, the price of a swing option with N rights is given by the supremum

$$\sup_{0\le\tau_1<\cdots<\tau_N\le T} \mathbb{E}\Big[\sum_{k=1}^{N}(S_{\tau_k} - Ke^{-\rho\tau_k})^+\Big]$$

over all stopping times $\tau_1,\ldots,\tau_N$ with values in $\{0,\ldots,T\}$. In order to represent this control problem as a switching system, we use the position set $P = \{1,\ldots,N+1\}$ to represent the number of rights remaining. That is $p \in P$ stands for the situation when there are $p-1$ rights remaining to be exercised. The action set $A = \{1,2\}$ is the same

as in the case of the Bermudan Put with control matrix α given by

$$(\alpha(p,a))_{p,a} \sim \begin{bmatrix} \alpha(1,1) & \alpha(1,2) \\ \alpha(2,1) & \alpha(2,2) \\ \dots & \dots \\ \alpha(N+1,1) & \alpha(N+1,2) \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & 2 \\ \dots & \dots \\ N & N+1 \end{bmatrix}.$$

Having modeled the discounted stock price process as an exponential mean-reverting process with reversion parameter $\kappa \in [0,1[$, long run mean $\mu > 0$ and volatility $\sigma > 0$, we obtain the logarithm of the discounted price process as

$$\tilde{Z}_{t+1} = (1-\kappa)(\tilde{Z}_t - \mu) + \mu + \sigma\epsilon_{t+1}, \quad \tilde{Z}_0 = \ln(S_0). \tag{4.79}$$

with the disturbance matrix

$$W_{t+1} = \begin{bmatrix} (1-\kappa) & \kappa\mu + \sigma\epsilon_{t+1} \\ 0 & 1 \end{bmatrix}, \quad t = 0, ..., T-1.$$

The reward and scrap values are given by

$$r_t(p,(z^{(1)},z^{(2)}),a) = (e^{z^{(1)}} - Ke^{-\rho t})^+\big(p - \alpha(p,a)\big) \quad t = 0,\dots,T-1 \tag{4.80}$$

and

$$r_T(p,(z^{(1)},z^{(2)})) = (e^{z^{(1)}} - Ke^{-\rho T})^+\big(p - \alpha(p,1)\big) \tag{4.81}$$

respectively for all $p \in P$ and $a \in A$.

In Table 4.2, we compare our results to those given in Meinshausen and Hambly [2004] with bounds on the swing option price where the underlying process is assumed to follow the dynamics (4.79) with parameters

$$\rho = 0, \quad \kappa = 0.9, \quad \mu = 0, \quad \sigma = 0.5, \quad S_0 = 1, \quad K = 0 \quad \text{and} \quad T = 1000.$$

TABLE 4.2: Swing option numerical results

	CSS	MH
Position (Rights + 1)	confidence interval	confidence interval
2	[4.737, 4.761]	[4.773, 4.794]
3	[9.005, 9.031]	[9.016, 9.091]
4	[13.001, 13.026]	[12.959, 13.100]
5	[16.805, 16.830]	[16.773, 16.906]
6	[20.465, 20.491]	[20.439, 20.580]
11	[37.339, 37.363]	[37.305, 37.540]
16	[52.694, 52.718]	[52.670, 53.009]
21	[67.070, 67.095]	[67.050, 67.525]
31	[93.811, 93.835]	[93.662, 94.519]
41	[118.639, 118.663]	[118.353, 119.625]
51	[142.059, 142.084]	[141.703, 143.360]
61	[164.368, 164.392]	[163.960, 166.037]
71	[185.757, 185.781]	[185.335, 187.729]
81	[206.362, 206.386]	[205.844, 208.702]
91	[226.284, 226.308]	[225.676, 228.985]
101	[245.601, 245.625]	[244.910, 248.651]

Results were produced using a grid size of $m = 1024$ and disturbances of size $n = 1024$. Diagnostics is based on $K = 1024$ sample paths and 99% confidence bounds are calculated by setting $x = 0.01$. The columns under *MH* denote the results from Meinshausen and Hambly [2004].

4.8 Conclusion

In this work we present a novel class of algorithms to solve stochastic switching problems whose processes follow linear state space dynamics. Our methodology is directly applicable to high-dimensional problems and shows remarkable numerical efficiency and excellent precision. More importantly, we adapt the primal-dual approach to estimate the distance to optimality of approximate solutions using Monte Carlo techniques. With this, we establish a sound and reliable diagnostics and quality assessment tool for a posterior justification of the numerical approximation. The authors believe that such combination of efficient numerical schemes with a subsequent diagnostic check can be very useful in practical applications. This approach may help in development and justification of further approximate methods. In this context, natural extensions of the present scheme (say, from linear to piecewise linear dynamics) can be examined in detail. We address this promising direction in further research.

Chapter 5

Conclusion

In this thesis, we devised new computational methods for some specific classes of optimal stochastic control problems based on high-dimensional underlying processes. The second and third chapter of this thesis was dedicated to solving the multiple-stopping problem by considering its dual formulation while the fourth chapter was dedicated to solving the more general class of optimal switching problems. This thesis was significantly motivated by the need to price the multiple-exercisable derivative, a mathematical problem which can be formulated under either framework. We devised primal-dual algorithms that allowed us to construct bounds on its price of such a derivative. With these bounds, traders of such derivatives can have confidence that their computed prices do not deviate too far from the theoretical ones.

We introduced the multiple-exercise option in Chapter 1. The seminal papers in the area of optimal multiple stopping were briefly reviewed and we provided an intuitive explanation as to why multiple-exercise options sit within this framework. We then reviewed some of primal-dual methods for solving optimal stopping problems and recent extensions of theses methods to the multiple stopping case. These methods require the use of regression-based approximations to the optimal policy and we explained some of the shortcomings of this approach. To address some of these problems, we introduced an alternative method - the convex switching system which would form the basis for the penultimate chapter of the thesis.

In Chapter 2, we derived an additive dual for the multiple-exercise derivative and an efficient algorithm to compute it. In order to do so, we first expressed the value of

an option as a multiple-stopping problem and show how one may interpret this as the buyer's price. As part of the formulation of the primal problem, we provided a new way to construct stopping times in the multiple-exercise framework. This gave us a way of expressing exercise strategies that were consistent across options with different numbers of rights. In order to derive the additive dual, we adopted a hedging argument approach similar to that of Joshi [2007] where the dual is associated with seller's price of the option. This approach required us to introduce the notion of a hedge in the multiple-exercise framework. We showed that by exploiting the properties of this hedge, we can obtain an efficient algorithm to compute the dual. In doing so, we naturally extended the Andersen and Broadie algorithm to the multiple-exercise case. A comparison of the numerical results to those computed using another known method was provided for. It was shown that this method yielded smaller duality gaps for larger numbers of exercise rights.

The multiplicative dual for the multiple-exercise option and its accompanying algorithm is derived in Chapter 3. The main result Theorem 3.5 is a generalisation of the multiplicative dual result for the single-exercise case given in Joshi and Tang [2014] to the multiple-exercise case. It also represents the first known multiplicative dual representation of the multiple stopping problem. An algorithm that can be seen as the natural extension of the AB algorithm to the multiplicative multiple-exercise case was developed and it was shown that a valid upper bound was still obtained even in the presence of Monte Carlo errors.

In the penultimate chapter, we moved from optimal multiple-stopping problems to the more general framework of stochastic switching problems. We reviewed the convex switching system approach that was first proposed in Hinz [2014] as a viable alternative to regression-based methods to solving problems of this kind. This approach involves the approximation of convex value functions using their sub-gradient envelopes which are computed recursively in accordance with the Bellman equations. Unfortunately, the original procedure only works for value functions that are convex and so convexity of all scrap and reward functions is required.

In Chapter 4, we took a first step to overcoming this restriction by providing a parallel procedure that allowed us to compute value functions in the special case where a non-convex function can be decomposed into the difference of two convex ones. A remarkable

generalization was then achieved where a method, which we term the *direct approach*, was devised that allowed us to bypass any convexity requirement while yielding significant improvements in computation time. In order to gauge its accuracy, we proposed a diagnostic method that measures the accuracy of the direct approach by first embedding it in a primal-dual Monte Carlo method which is then used to construct confidence intervals. The confidence intervals would then be a measure of the distance to optimality of an approximated value functions computed using the direct approach. We then showed how to compute bounds on the prices of Bermudan puts and swing options using this method and compared them to other results found in the literature.

With this, we conclude the summary of this thesis. Possible future research directions based the contributions in this thesis have been highlighted at the end of each of the individual chapters. We hope future researchers in the area will find it to be as interesting as we have.

Bibliography

Andersen, L. and Broadie, M. [2004], 'Primal-dual simulation algorithm for pricing multidimensional American options', *Management Science* **50**(9), 1222–1234.

Andersen, L. and Piterbarg, V. [2010], *Interest Rate Modeling. Volume 3: Products and Risk Management*, Atlantic Financial Press.

Balder, S., Mahayni, A. and Schoenmakers, J. [2013], 'Primal–dual linear Monte Carlo algorithm for multiple stopping-an application to flexible caps', *Quantitative Finance* **13**(7), 1003–1013.

Bäuerle, N. and Rieder, U. [2011], *Markov Decision Processes with Applications to Finance: Markov Decision Processes with Applications to Finance*, Springer.

Belomestny, D., Bender, C. and Schoenmakers, J. [2009], 'True upper bounds for Bermudan products via non-nested Monte Carlo', *Mathematical Finance* **19**(1), 53–71.

Belomestny, D., Kolodko, A. and Schoenmakers, J. [2010], 'Regression methods for stochastic control problems and their convergence analysis', *SIAM Journal on Control and Optimization* **48**(5), 3562–3588.

Belomestny, D., Schoenmakers, J. and Dickmann, F. [2013], 'Multilevel dual approach for pricing American style derivatives', *Finance and Stochastics* **17**(4), 717–742.

Bender, C. [2011], 'Dual pricing of multi-exercise options under volume constraints', *Finance and Stochastics* **15**, 1–26.

Bender, C. and Schoenmakers, J. [2006], 'An iterative method for multiple stopping: convergence and stability', *Advances in Applied Probability* **38**(3), pp. 729–749.

Bender, C., Schoenmakers, J. and Zhang, J. [2013], 'Dual representations for general multiple stopping problems', *Mathematical Finance* .

Bertsekas, D. P. [1995], *Dynamic Programming and Optimal Control*, Vol. 1, Athena Scientific Belmont, MA.

Beveridge, C., Joshi, M. and Tang, R. [2013], 'Practical policy iteration: Generic methods for obtaining rapid and tight bounds for Bermudan exotic derivatives using Monte Carlo simulation', *Journal of Economic Dynamics and Control* **37**(7), 1342–1361.

Brown, D. B., Smith, J. E. and Sun, P. [2010], 'Information relaxations and duality in stochastic dynamic programs', *Operations research* **58**(4-part-1), 785–801.

Carmona, R. and Dayanik, S. [2008], 'Optimal multiple stopping of linear diffusions', *Mathematics of Operations Research* **33**(2), pp. 446–460.

Carmona, R. and Ludkovski, M. [2010], 'Swing options', *Encyclopedia of Quantitative Finance* .

Carmona, R. and Touzi, N. [2008], 'Optimal multiple stopping and valuation of swing options', *Mathematical Finance* **18**(2), 239–268.

Carriere, J. F. [1996], 'Valuation of the early-exercise price for options using simulations and nonparametric regression', *Insurance: Mathematics and Economics* **19**(1), 19–30.

Chandramouli, S. S. and Haugh, M. B. [2012], 'A unified approach to multiple stopping and duality', *Operations Research Letters* **40**(4), 258–264.

Chen, N. and Glasserman, P. [2007], 'Additive and multiplicative duals for American option pricing', *Finance and Stochastics* **11**, 153–179.

Chiarella, C., Clewlow, L. and Kang, B. [2012], 'The evaluation of gas swing contracts with regime switching', *Topics in Numerical Methods for Finance* **19**, 155–176.

Chiarella, C., Joshi, M. S. and Yap, N. [2014], 'Computing the additive dual for multiple-exercise options using a collapsing tree approach', *Preprint* .

Clément, E., Lamberton, D. and Protter, P. [2002], 'An analysis of a least squares regression method for American option pricing', *Finance and Stochastics* **6**(4), 449–471.

Davis, M. H. A. and Karatzas, I. [1994], 'A deterministic approach to optimal stopping', *Probability, Statistics and Optimization: A Tribute to Peter Whittle* pp. 455–466.

Feinberg, E. A., Shwartz, A. and Altman, E. [2002], *Handbook of Markov Decision Processes: Methods and Applications*, Kluwer Academic Publishers Boston, MA.

Giles, M. B. [2008], 'Multilevel Monte Carlo path simulation', *Operations Research* **56**(3), 607–617.

Glasserman, P. and Yu, B. [2004], 'Number of paths versus number of basis functions in American option pricing', *The Annals of Applied Probability* **14**(4), 2090–2119.

Haggstrom, G. W. [1967], 'Optimal sequential procedures when more than one stop is required', *The Annals of Mathematical Statistics* pp. 1618–1626.

Haugh, M. B. and Kogan, L. [2004], 'Pricing American options: A duality approach', *Operations Research* **52**(2), 258–270.

Hernández-Lerma, O. and Lasserre, J. B. [1996], *Discrete-time Markov control processes*, Springer.

Hinz, J. [2014], 'Optimal stochastic switching under convexity assumptions', *SIAM Journal on Control and Optimization* **52**(1), 164–188.

Hinz, J. and Yap, N. [2014], 'Algorithms of optimal control of stochastic switching systems', *Preprint* .

Ibanez, A. [2004], 'Valuation by simulation of contingent claims with multiple early exercise opportunities', *Mathematical Finance* **14**(2), 223–248.

Ibanez, A. and Zapatero, F. [2004], 'Monte Carlo valuation of American options through computation of the optimal exercise frontier', *Journal of Financial and Quantitative Analysis* **39**(2), 253–275.

Jamshidian, F. [2004], 'Numeraire-invariant option pricing & American, Bermudan, and trigger stream rollover'. Version 1.6.

Joshi, M. S. [2003], *The Concepts and Practice of Mathematical Finance*, Cambridge University Press.

Joshi, M. S. [2007], 'A simple derivation of and improvements to Jamshidian's and Rogers' upper bound methods for Bermudan options', *Applied Mathematical Finance* **14**(3), 197–205.

Joshi, M. S. and Tang, R. [2014], 'Effective sub-simulation-free upper bounds for the Monte Carlo pricing of callable derivatives and various improvements to existing methodologies', *Journal of Economic Dynamics and Control* **40**, 25–45.

Joshi, M. S. and Yap, N. [2014], 'The multiplicative dual for multiple-exercise options', *http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2430558* .

Kobylanski, M., Quenez, M.-C., Rouy-Mironescu, E. et al. [2011], 'Optimal multiple stopping time problem', *The Annals of Applied Probability* **21**(4), 1365–1399.

Kolodko, A. and Schoenmakers, J. [2006], 'Iterative construction of the optimal Bermudan stopping time', *Finance and Stochastics* **10**(1), 27–49.

Longstaff, F. A. and Schwartz, E. [2001], 'Valuing American options by simulation: a simple least-squares approach', *Review of Financial Studies* **14**(1), 113–147.

Meinshausen, N. and Hambly, B. M. [2004], 'Monte Carlo methods for the valuation of multiple-exercise options', *Mathematical Finance* **14**(4), 557–583.

Peskir, G. and Shiryaev, A. [2006], *Optimal Stopping and Free-boundary Problems*, Springer.

Puterman, M. L. [2009], *Markov Decision Processes: Discrete Stochastic Dynamic Programming*, Vol. 414, John Wiley & Sons.

Rockafellar, R. T. [1997], *Convex Analysis*, Princeton University Press.

Rogers, L. [2007], 'Pathwise stochastic optimal control', *SIAM Journal on Control and Optimization* **46**(3), 1116–1132.

Rogers, L. C. G. [2002], 'Monte Carlo valuation of American options', *Mathematical Finance* **12**(3), 271–286.

Schoenmakers, J. [2012], 'A pure martingale dual for multiple stopping', *Finance and Stochastics* **16**, 319–334.

Tsitsiklis, J. and Van Roy, B. [2001], 'Regression methods for pricing complex American-style options', *Neural Networks, IEEE Transactions on* **12**(4), 694 –703.

Yap, N. A. [2013], Additive and multiplicative upper bounds for multiple-exercise options, *in* 'Proceedings of the 26th Annual PhD Conference in Economics and Business', The Australia National University, Canberra, Australia.