On the valuation of barrier and American options in local volatility models with jumps

A thesis presented for the degree of

Doctor of Philosophy of Imperial College London

by

Bjorn Eriksson

Department of Mathematics

Imperial College London

London SW7 2AZ

United Kingdom

OCTOBER 31, 2013

Abstract

In this thesis two novel approaches to pricing of barrier and American options are developed in the setting of local volatility models with jumps: the moments method and the Markov chain method.

The moments method is a valuation approach for barrier options that is based on a characterisation of the exit location measure and the expected occupation measure of the price process of the underlying in terms of the corresponding moments. It is shown how the value of barrier-type derivatives can be expressed using these moments, which are in turn shown to be characterised by an infinite-dimensional linear system. By solving finite-dimensional linear programming problems, which are obtained by restricting to moments of a finite degree, upper and lower-bounds are found for the values of the options in question.

The Markov chain method for the valuation of American options is based on an approximation of the underlying price process by a continuous-time Markov chain. The value-function of the American option driven by the approximating chain is identified by solving the associated optimal stopping problem. In particular, a novel explicit characterisation of the optimal exercise boundary is derived in terms of the generator of the Markov chain. Using this characterisation it is shown that the optimal exercise boundary and the corresponding value-function can be evaluated efficiently.

For both of the presented methods convergence results are established. The methods are implemented for a range of local volatility models with jumps, and a number of numerical examples are discussed in detail to illustrate the scope of the methods.

I certify that this thesis, and the research to which it refers, are the product of my own work, and that any ideas or quotations from the work of other people, published or otherwise, are fully acknowledged in accordance with the standard referencing practices of the discipline.

Signed: Bjorn Eriksson

Copyright

Acknowledgements

First and foremost I would like to thank my supervisor Martijn Pistorius. His never ending enthusiasm and support was invaluable throughout the work on my PhD.

I would also like to thank staff and PhD students at the mathematical finance groups at King's college and Imperial college London for making this a pleasurable time. Special thanks go out to my fellow PhD friends Duncan Steele, Alexander Kabanov and Haluk Yener for help with proof reading my thesis.

The research presented in this thesis was supported by EPSRC grant EP/D039053.

Last but not least, I thank my family for their support.

Bjorn Eriksson

Table of contents

List of Figures

List of Tables

Chapter 1

Introduction

This thesis is devoted to the study of numerical methods for the valuation of American options and barrier contracts. Those securities are commonly traded in the financial markets. The majority of the traded options belong to the class of vanilla options, which can be either European or American. With the exception of indices, which are often represented by European options, most stock and equity vanilla options are in fact of American type. We also note that a significant part of the class of exotic derivative securities contain barrier features. For example, in the foreign exchange markets barrier options such as double-no-touch and knock-in and knock-out put and call options are commonly traded derivatives. It is thus of considerable interest to financial market participants to have efficient valuation methods for this class of derivative securities.

While the valuation of European options is reasonably well understood, there are still many open questions concerning the valuation of American options and barrier contracts. The valuation of an American option requires knowledge of the law of the entire underlying price process since its value is given by an optimal stopping problem. Furthermore, the value of a barrier option depends on the laws of path-functionals of the underlying price process such as the first-passage time to the barrier and the overshoot over the barrier (the latter in the case of rebate payments under a price process that is discontinuous). Determining these laws is more complex than calculating the law of the marginal distribution that is required for the valuation of the European option. Due to the complexity of these problems, a good deal of methods have been developed for specific underlyings. They often rely on the characteristics of the specific process making them difficult to generalise. The aim of this work has been to develop methods to value American and barrier options in a class of processes that contain a wide range of different underlyings, specifically aiming towards processes allowing jumps. The class of processes in each method depends on the characteristics of the method developed. Both classes are sufficiently general to contain a large number of the standard processes used in mathematical finance today, and in particular contain local volatility models with jumps. The use of this class is motivated by recent studies of financial markets which show that stock-price data typically contains features such as fat tails, asymmetry, excess kurtosis and jumps. These are all features that can be captured by processes contained in the class of local volatility models with jumps.

In this thesis two solution methods for the posed problems are presented: Firstly, a method for valuing a general class of double barrier options under a piecewise polynomial jump diffusion is introduced. This method expresses the value of the option using moments of measures. Those moments are characterised using linear programming. Secondly, a method for valuing the American option in a Markov chain setting is introduced. An explicit characterisation of the optimal stopping boundary of the American option under Markov chains is developed. Subsequently a method to value the American option using this boundary is developed. In addition, a dynamic programming method to value American option under Markov chains is presented. It is shown that the two described methods can be used to value American options under price processes given by local volatility models with jumps by combining these methods with a Markov chain approximation of the underlying price process. Both methods are numerically implemented for a range of models and convergence proofs are given.

The remainder of this thesis is structured as follows. In Chapter 2 the relevant elements of derivative pricing theory are discussed in relation to barrier and American options, and the existing methods are reviewed. In Chapter 3 the moments method for valuing double barrier contracts is developed. Chapter 4 is devoted to the method for valuing the American option under Markov chains. The results in Chapter 3 have been published in Eriksson and Pistorius [34].

Chapter 2

Financial derivative contracts

In this chapter elements of the theory of arbitrage-free valuation of financial derivative contracts are discussed. First, a brief overview of the arbitrage pricing theory of general financial derivative contracts is given. This is the valuation framework that is deployed throughout the thesis. Subsequently, the application of arbitrage pricing theory to the valuation of barrier contracts and American options is outlined, and a review of the existing numerical methods for valuation of these derivative securities is presented.

2.1 No-arbitrage pricing of financial contracts

The general theory of no-arbitrage pricing of financial contracts is covered in any text book on mathematical finance. This section takes its inspiration from Björk [12] and Cont and Tankov [24, Ch:9].

Consider a market consisting of a risk free bank account $(S_0(t))$ with constant interest rate (r) and a risky asset $(S_1(t))$ defined on a filtered probability space $(\Omega, \mathcal{F}, \mathbf{F}, \mathbb{P}_0)$, where $\mathbf{F} = \{\mathcal{F}_t\}_{t\in[0,T]}$ denotes the standard filtration generated by S_1 and T denotes the time horizon. Assume that the risky asset pays dividends according to a continuous yield $d \geq 0$. In general it is possible to allow for N risky assets but for the purpose of this thesis $N = 1$ is sufficient. Let $S^v(t) = (S_0(t), S_1(t))$ be the prices of the assets. Let $h_i(t)$ indicate ownership of $h_i(t)$ units of asset i at time t, then $h = (h_0(t), h_1(t))$ is called a portfolio on the assets $S^v(t)$. The value of a portfolio is given by $V(t) = h(t)S^v(t)$, where the multiplication of the two vectors is the scalar product. A portfolio is called admissible if there exist a nonnegative α such that for all $t \in [0, T]$

$$\int_0^t h(u)\mathrm{d}S^v(u) \geq -\alpha.$$

An admissible portfolio h is called self-financing if for all $t \in (0, T]$

$$\mathrm{d}V(t) = h(t)\mathrm{d}S^v(t)$$

holds. The admissibility requirement is to remove the so called "doubling strategies" that, when allowing infinite debt, will be guaranteed to be always profitable. In a self-financing portfolio there are no withdrawals of funds from or deposits into the bank account, so that any change of the portfolio is a redistribution between the given assets. An admissible self-financing portfolio is called an arbitrage opportunity if there is a possibility of a sure gain without the potential of loss. More precisely, the definition of arbitrage opportunity is given as follows:

Definition 2.1.1 (Arbitrage)**.** Let h be a self-financing portfolio and $V^h(t)$ its value at time t. Then h is an arbitrage possibility if there exist a time $t \in (0, T]$ such that

$$\begin{aligned} V^h(0) &= 0, \\ \mathbb{P}_0(V^h(t) \geq 0) &= 1, \\ \mathbb{P}_0(V^h(t) > 0) &> 0. \end{aligned}$$

In the sequel, it is assumed that the market is efficient and, in particular, free of arbitrage opportunities. The condition to guarantee the absence of arbitrage opportunities is that the risky asset process S_1, when appropriately discounted, is a *martingale.*

Definition 2.1.2 (Martingale)**.** A stochastic process $M = (M_t)_{t\in[0,T]}$ defined on the probability space $(\Omega, \mathcal{F}, \mathbf{F}, \mathbb{P}_0)$ is a martingale if M_t is $\mathcal{F}_t$ measurable, $\mathbf{E}[|M_t||\mathcal{F}_t] < \infty$ and

$$\mathbf{E}[M_t|\mathcal{F}_s] = M_s \quad \text{for all } t, s \in [0, T].$$

Two probability measures $\mathbb{P}_0$ and $\mathbb{P}$ on the measurable space $(\Omega, \mathcal{F})$ are called equivalent probability measures, denoted $\mathbb{P}_0 \sim \mathbb{P}$, if

$$\forall A \in \mathcal{F} \quad \mathbb{P}_0(A) = 0 \Leftrightarrow \mathbb{P}(A) = 0.$$

The precise characterisation of the absence of arbitrage is as follows:

Theorem 2.1.1 (Fundamental theorem of asset pricing)**.** *Let the market, as above, consist of a risk free bank account ($S_0(t)$) with constant interest rate (r) and a risky asset ($S_1(t)$) defined on a filtered probability space* $(\Omega, \mathcal{F}, \mathbf{F}, \mathbb{P}_0)$*, where* $\mathbf{F} = \{\mathcal{F}_t\}_{t\in[0,T]}$ *denotes the standard filtration generated by S_1. Let the risky asset pay dividends according to a continuous yield $d \geq 0$. The market model is arbitrage free if and only if there exists a probability measure $\mathbb{P}$ on $(\Omega, \mathcal{F})$, with $\mathbb{P} \sim \mathbb{P}_0$ such that the discounted risky asset price process $e^{-(r-d)t}S_t$ is a martingale with respect to $\mathbb{P}$.*

The measure $\mathbb{P}$ is called the risk neutral measure or the pricing measure. This measure is not necessarily unique and in general there are several ways to select a risk neutral probability measure. One possibility is to select the measure $\mathbb{P}$ by calibration of model prices of traded financial contracts to market quotes. In the remainder, it is assumed that a pricing measure $\mathbb{P}$ has been selected, and that all stochastic processes are specified under this probability measure. Expectations are always taken with respect to this pricing measure $\mathbb{P}$ unless specified otherwise.

The arbitrage free value of a financial contract is given by the expected value under the risk neutral measure of future cash flows. For example, the value V_t at time t of a contingent claim with payoff at maturity T given by $\phi(S_T)$ for some function ϕ of the price S_T of the risky asset at time T is given by

$$V_t = \mathbf{E}[e^{-r(T-t)}\phi(S_T)|\mathcal{F}_t],$$

where $\mathbf{E}$ denotes the expectation under $\mathbb{P}$.

2.2 Barrier options

Barrier options and barrier-type contracts are derivatives whose payoffs are activated or cancelled depending on whether some reference rate or price quote cross a predetermined level. Such contracts exist in foreign exchange, fixed income, commodities and equity markets. They are among the most widely traded exotic derivatives, which makes their valuation an important topic. An example of contracts that are frequently traded in foreign exchange markets are the double-no-touch contracts, which pay a fixed amount if the foreign exchange rate has not left a finite interval at maturity and zero otherwise. Some barrier contracts are said to belong to the so-called first-generation exotics.

Two fundamental types of barrier options are knock-in and knock-out options. A knock-in option is a derivative where the payoff is only activated once a barrier is crossed. Conversely a knock-out option is a derivative that become worthless once the barrier is crossed. Note that the value of the sum of a knock-in and a knock-out option with the same payoff is equal to the value of a European contract with said payoff, which is a well-known parity result. As a consequence, assuming it is known how to value the European contract, it is sufficient to know how to value either knock-in or knock-out options. In this thesis the focus will be on knock-out options.

Consider a Markov process S defined on some filtered probability space $(\Omega, \mathcal{F}, \mathbf{F}, \mathbb{P})$, where $\mathbf{F} = \{\mathcal{F}_t\}_{t\in[0,T]}$ denotes the standard filtration generated by S. Assume that S pays dividends according to a continuous yield $d \geq 0$. In general, the payoff of a knock-out barrier-type contract consists of a payment $h(S_T)$ at maturity T if the underlying has not entered the knock-out set B before maturity, a flow of payments $g(s, S_s)$ until the moment τ_B that S_t enters B and a rebate $R(\tau_B, S_{\tau_B})$ paid at τ_B. Denoting the rate of discounting by r, the expected discounted value V of this double barrier knock-out option is given by

$$\begin{aligned} V(0, x_0) = \mathbf{E}_{0,x_0}\left[e^{-rT}h(S_T)\mathbf{I}(T < \tau_B) + \int_0^{\tau_B \wedge T} e^{-rs}g(s, S_s)\mathrm{d}s\right] \\ + \mathbf{E}_{0,x_0}\left[e^{-r\tau_B}R(\tau_B, S_{\tau_B})\mathbf{I}(\tau_B \le T)\right], \end{aligned} \tag{2.1}$$

where $\mathbf{E}_{0,x_0}[\cdot] = \mathbf{E}[\cdot|S_0 = x_0]$, $\tau_B = \inf\{t \ge 0 : S_t \in B\}$, $\mathbf{I}$ is the indicator function and $a \wedge b = \min(a, b)$.

If the function g in Equation (2.1) is zero, it can be shown, by an application of the strong Markov property of S, that the process $(e^{-(r-d)(t\wedge\tau_B)}V(t \wedge \tau_B, S_{t\wedge\tau_B}))$ is a martingale during the lifespan of the option.[1]

In particular, for $B = \mathbb{R}\backslash[l, u]$, Dynkin's lemma implies that if $V(t, x)$ is a sufficiently regular solution to the system

$$\begin{aligned} \mathcal{L}_t V(t, x) &= 0 \quad t \in [0, T), x \in (l, u) \\ V(T, x) &= h \quad x \in (l, u) \\ V(t, x) &= R(t, x) \quad t \in [0, T], x \in \mathbb{R}\backslash(l, u), \end{aligned}$$

where $\mathcal{L}_t f = \partial f/\partial t + \mathcal{L}f - rf$ denotes the "discounted" infinitesimal generator of the time-space process (t, S_t), it is the value function of the derivative in Equation (2.1). For a model with continuous sample paths $\mathcal{L}_t V(t, x) = 0$ is a partial differential equation (PDE) while, if the model contains jumps the operator $\mathcal{L}_t$ contain an integral part whereby $\mathcal{L}_t V(t, x) = 0$ is a partial integro-differential equation (PIDE). This equation is the starting point in a number of option pricing methods.

[1] If g is non-zero the discounted value function can still be shown to be a martingale however the payments already made have to be taken into account.

The valuation of a barrier option is more involved than the valuation of a standard European type option, since the expectation in Equation (2.1) depends on the entire path of S_t. The valuation of barrier options has attracted a good deal of attention and there currently exist a body of literature dealing with different aspects of their pricing. In particular, for double barrier options, Geman and Yor [36] and Pelsser [64] developed a Laplace transform approach in the geometric Brownian motion (GBM) setting. Sepp [71] derived semi-analytical expressions in a jump-diffusion setting with exponential jumps, also using a transform approach. Carr and Crosby [20] considered double no touch contracts in a setting with exponential jumps, allowing the process dynamics to change after a barrier is breached. Davydov and Linetsky [30] used eigenfunction expansions to price double barrier options in a CEV setting. The mentioned papers exploit specific features of the model under consideration and can therefore not readily be extended to the setting of processes that exhibit jumps and non-constant local volatility.

In Chapter 3 a moments method is presented for the valuation of barrier options under a general class of polynomial jump-diffusions.[2]

2.3 American options

An American option is an option that can be exercised at any time prior to maturity. In the case that the maturity is infinite (that is to say, there is no maturity) the contract is known as a perpetual American derivative. Consider a time-homogeneous Markov processes S_t, with state space $\mathbb{E} = [0, \infty)$ defined on some filtered probability space $(\Omega, \mathcal{F}, \mathbf{F}, \mathbb{P})$, where $\mathbf{F} = \{\mathcal{F}_t\}_{t \in [0,T]}$ denotes the standard filtration generated by S and $\mathbb{P}$ is a risk neutral probability measure. Let the interest rate r be constant and denote by $d \geq 0$ the continuous dividend yield. Assume, as usual, that the discounted

[2]We mention that, recently, after the research project, the results of which are presented in Chapter 3, had been completed, an alternative general method for the valuation of barrier options has been proposed in Mijatović and Pistorius [62] based on a Markov chain approximation of the underlying. This method is also applicable to a wide class of Markov processes.

price process $\{e^{-(r-d)t}S_t\}_{t\in[0,T]}$ is a martingale.

Note that the American option with payoff $\phi^+(x) = \max(0, \phi(x))$ is equivalent to the American option with payoff $\phi(x)$, since if $\phi(x) < 0$ the holder will not exercise. In the following, unless specifically stated, it is assumed that $\phi(x) \geq 0$.

Assume that the payoff function $\phi : \mathbb{E} \to \mathbb{R}^+$ is non-negative and satisfies the integrability condition

$$\mathbf{E}_{0,x_0}\left[\sup_{t\in[0,T]} \phi(S_t)\right] < \infty, \qquad x \in \mathbb{E}.$$

At any time t prior to maturity T the holder of an American option has to decide whether or not to exercise the option by comparing the immediate exercise value $\phi(S_t)$ with the expected discounted future cash flows. The value V_t^* of the American option at time $t \in [0,T]$ with payoff function ϕ is given by

$$V_t^* = \operatorname*{ess.\,sup}_{\tau\in\mathcal{T}_{t,T}} \mathbf{E}[e^{-r\tau}\phi(S_\tau)|\mathcal{F}_t],$$

where $\mathcal{T}_{t,T}$ denotes the set of $\mathbf{F}$-stopping times taking values between t and T. The process $V^* = \{V_t^*, t \in [0,T]\}$ is called the *Snell-envelope* of the collection of discounted payoffs $\Pi = \{e^{-rt}\phi(S_t), t \in [0,T]\}$: it is the smallest $\mathbf{F}$-supermartingale that is bounded below by Π. The Markov property of S implies $V_t^* = V(t, S_t)$ where the value function of the American option $V = \{V(t,x), t \in [0,T], x \in \mathbb{E}\}$ is given by

$$\begin{aligned} V(t,x_0) &= \sup_{\tau\in\mathcal{T}_{t,T}} \mathbf{E}_{t,x_0}\left[e^{-r(\tau-t)}\phi(S_\tau)\right] \\ &= \sup_{\tau\in\mathcal{T}_{0,T-t}} \mathbf{E}_{0,x_0}\left[e^{-r\tau}\phi(S_\tau)\right], \qquad (t,x_0) \in [0,T]\times\mathbb{E}, \end{aligned} \tag{2.2}$$

where the second line is a consequence of the time-homogeneity of the Markov process S.

According to the general theory of optimal stopping (see Peskir and Shiryaev [65], Shiryaev [72] or Boyarchenko and Levendorskii [14]), if S_t is Markovian the space $[0,T] \times \mathbb{R}^+$ can be partitioned into a stopping region $\mathbb{S}$ and a continuation region $\mathbb{C}$

given by

$$\mathbb{C} = \{(t,x) \in [0,T] \times \mathbb{R}^+ : V(t,x) > \phi(x)\}$$
$$\mathbb{S} = \{(t,x) \in [0,T] \times \mathbb{R}^+ : V(t,x) = \phi(x)\}.$$

By construction, while (t, S_t) is in $\mathbb{C}$ it is optimal to hold on to the option, and while (t, S_t) is in $\mathbb{S}$ it is optimal to immediately exercise the option. The shape of $\mathbb{C}$ and $\mathbb{S}$ depend on the payoff ϕ and the process S_t. Under weak continuity assumptions (see Peskir and Shiryaev [65] for details) on V and ϕ, the stopping time

$$\tau_{\mathbb{S}}(t) = \inf\{s \geq t : S_s \in \mathbb{S}\}.$$

is an optimal stopping time for the optimisation problem in Equation (2.2). Note that

$$\mathbb{P}(\tau_{\mathbb{S}}(t) \leq T) = 1$$

holds, since the non-negativity of $\phi(x)$ implies the equality $V(T,x) = \phi(x)$ for all $x \in \mathbb{R}^+$ so that $(T, \mathbb{R}^+)$ is contained in the stopping region $\mathbb{S}$.

From the theory of Snell envelopes it is known that (see for example Elliot and Kopp [33, p:192] who in turn cite El Karoui [32] for the proof of this statement)

- $e^{-rs}V(s, S_s)_{s \in [t,T]}$ is a super martingale
- $e^{-r(s \wedge \tau_{\mathbb{S}}(t))}V(s \wedge \tau_{\mathbb{S}}(t), S_{s \wedge \tau_{\mathbb{S}}(t)})_{s \in [t,T]}$ is a martingale.

For specific forms of the payoff ϕ and the underlying S more is known about the form of the continuation and stopping regions and about the value function. For example, American call and put options with strike K, which have payoffs $\phi(x) = (x-K)^+$ and $\phi(x) = (K-x)^+$, have been studied in detail in various settings - we refer to Peskir and Shiryaev [65] and Lamberton and Mikou [52] for further details.

Under the GBM model, when S satisfies the SDE

$$\mathrm{d}S_t = (r-d)S_t\mathrm{d}t + \sigma S_t \mathrm{d}W_t, \quad t \in (0,T], \quad S_0 = x_0 > 0,$$

where W is a Wiener process, the value-function of the American put option is $C^{1,2}$ on the continuation region $\mathbb{C}$. A function is $C^{1,2}$ if it is continuously differentiable in its first argument and twice continuously differentiable in its second argument. The value function of the American option under the GBM model satisfies the free-boundary problem given by

$$\begin{aligned}
&\mathcal{L}_t V \leq 0 \\
&\mathcal{L}_t V = 0 \text{ in } \mathbb{C} \\
&V(t,x) = \phi(x) \text{ in } \mathbb{S} \\
&V(t,x) > \phi(x) \text{ in } \mathbb{C} \\
&V(T,x) = \phi(x) \\
&\text{Fit condition (see below)}
\end{aligned}$$

where, for any $C^{1,2}$ function $f : [0,T] \times \mathbb{R}^+ \to \mathbb{R}$, $\mathcal{L}_t$ given by

$$\mathcal{L}_t f = \frac{\partial f}{\partial t} + \mathcal{L} f - rf$$

is the discounted time-space generator of (t, S_t) where

$$\mathcal{L} f = (r-d)x\frac{\partial f}{\partial x} + \frac{\sigma^2 x^2}{2}\frac{\partial^2 f}{\partial x^2}$$

is the infinitesimal generator of S_t. The fit condition is given by

$$\begin{aligned}
&V_x(t,x) = \phi'(x) \text{ for } x = b(t) \text{ (smooth fit)} \\
&V(t,x) = \phi(x) \text{ for } x = b(t).
\end{aligned}$$

In the case of various other sufficiently regular Markov process models for the underlying price process the above free-boundary characterisation essentially remains valid but it may be needed to replace the smooth fit condition by a continuous fit condition. Heuristics suggest that if the boundary is regular (irregular) the smooth fit (continuous fit) condition hold. See for example Peskir and Shiryaev [65, Ch:9] and Lamberton and Mikou [53] for precise statements.

Pricing American options has a long history and there is a large body of existing methods. The main approaches to valuing American options are (i) tree methods, (ii) methods based on solution of the associated free-boundary problem and (iii) Monte Carlo methods. Although the Monte Carlo methods generally surpass the methods from classes (i,ii) as far as generality of the underlying price process and the payoff is concerned, it is generally recognised that methods from the classes (i,ii) are more efficient for those models for which these have been developed. While in the literature methods from the classes (i,ii) have been studied mostly for diffusion models, and also Lévy models, the general case of local-volatility models with jumps appears to have received far less attention.

In Chapter 4 an efficient method is presented for the valuation of American options under a large class of Markov processes, which includes local volatility models with jumps. The method is based on approximation of the underlying price process by a Markov chain, in the spirit of Kushner [49] and Kushner and Dupuis [50].

We next present a brief overview of the various existing methods in the literature. The earliest study of the free-boundary problem associated to an American put option with finite maturity under the GBM model can be traced back at least as far a McKean [60]. The problem of valuing the corresponding perpetual American put option was solved analytically in Merton [61]. In this case the free-boundary is constant, and the corresponding problem admits an explicit solution.

For the GBM model a number of increasingly accurate analytical approximations have been developed by Johnson [42], Geske and Johnson [37], Barone-Adesi and Whaley [9] and Jua and Zhong [43]. These methods are based on manipulating the PDE and at some point approximating some terms in such a way that the PDE is simplified.

Another approach is to numerically solve the associated free-boundary problem using finite difference methods to approximate the PDE or PIDE operator. In the GBM setting an early application of such a numerical approach is presented in Brennan and Schwartz [17]. We next mention some examples of more recent contribution in this direction. In Toivanen [75] the American option under Kou's double exponential jump diffusion model is valued by using a finite difference scheme and a mixture of implicit and explicit time stepping. Two methods of handling the optimal boundary are tried: a penalty method and an operator splitting method. d'Halluin et al. [31] study jump diffusion's with a discretised PDE, the optimal boundary is handled by a constraint in a penalty method. In Hirsa and Madan [39] the option under a Variance Gamma model is valued by employing an implicit-explicit finite difference scheme. The infinite activity is handled by splitting the integral term and handling the area around the singularity by expanding the integrand. Wang et al. [76] price the American option in the CGMY model by using an implicit finite difference method. The integral terms are dealt with using quadrature and the infinite activity issues are addressed by using a Taylor expansion. The linear system is solved using a preconditioned iterative solver. Almendral and Oosterlee [5] value the American option under variance gamma model by using a linear complementarity problem and an implicit-explicit finite difference method which is solved iteratively and speeded up by a fast Fourier transform. In Almendral and Oosterlee [4] a method for the CGMY model is developed using a finite-difference method for integro-differential equations. The space discretisation is done by a collocation method and the time discretisation by an explicit backward differentiation formula. The singularity of the Lévy measure is dealt with using an integration by parts technique. These are just some examples of PDE/PIDE method - many others exist.

A related approach is the tree method. In fact the binomial method which was introduced in Cox et al. [26], was one of the earliest valuation methods for the finite maturity option. In the following years many refinements and extensions have been developed, Lamberton [51] for example produce error bounds for the GBM. One of the natural extensions is to tri- or mult-nomial trees, Ahn and Song [2] shows convergence for the trinomial tree in the GBM case. Maller et al. [59] study a multinomial tree for exponential Lévy models. In Szimayer and Maller [74] a grid based approximation for pure jump Lévy processes is suggested.

The final main approach is to use Monte Carlo methods. American Monte Carlo methods are more involved than the standard Monte Carlo method that can be used to value European or barrier options. The problem is that the value at a given time and level (t, x) depends both on the value of exercising and on the expected value of exercising in the future. In theory, to get the value at some point on a path the optimal stopping problem with this point as a starting point must be solved. In principle at every time point along every path a new simulation is needed, leading to an explosion in computational complexity. A number of ideas to overcome this problem has been suggested.

One approach is the regression method of Broadie and Glasserman [18], Carriere [23] and Longstaff and Schwartz [55] where the expectation of the value function is approximated using regression.

Another approach to using Monte Carlo method to price American options was presented in Milstein et al. [63]. Their method is developed for a local volatility setting. The main idea is to first calculate the optimal barrier and then use the optimal barrier in constructing the Monte Carlo method and therefore escaping the need to calculate the value of the option in each point along each sample path. The optimal boundary can either be found using some other method or using a Monte Carlo method suggested in Milstein et al. [63, Sec:4].

The dual approach is another Monte Carlo based method to pricing American options. Introduced in Rogers [66], the method uses a convex duality to rewrite the value of the option as $v = \inf_{M \in H_0^1} \mathbf{E}\left[\sup_{0 \le t \le T}(Z_t - M_t)\right]$ where H_0^1 is a class of martingales and Z is the payoff process. Given a class of martingales the expectation is calculated using a Monte Carlo method. Building on this method a number of papers have followed, some popular examples are the primal-dual methods containing, the nested Andersen and Broadie [6] and the non-nested Belomestny et al. [10]. A further extension of this idea is a class of methods that, rather than optimising over a set of martingales, iteratively construct a minimizing martingale; these were developed in Rogers [67] and Schoenmakers et al. [69].

Parallel to these the quantisation method was developed in Bally et al. [7] and following papers. The approach is to use a Monte Carlo simulation to build an optimal grid in each time step: the generation of the grid produces some factors that are of use when calculating conditional expectations. The valuation of the option is then done using dynamic programming on this grid, where thanks to these factors the conditional expectations are easily calculated. This method has been developed for local volatility models and is of most use in multidimensional problems.

A general method not contained in the main approaches given above is to use quadrature as in Lord et al. [56]. They use quadrature together with the fast Fourier transform and convolution. Due to the convolution the American option have to be approximated by a Bermudan option.

Another approach to valuing the American options, which has been developed in the setting of the GBM model in Kim [44], Jacka [40], Carr et al. [21], Barone-Adesi and Elliott [8] and Allegretto et al. [3], is based on a characterisation or approximation of the early exercise boundary.

In addition to the general methods that were reviewed above, we mention that in the literature there exist also a number of specialised approaches that exploit specific features of the underlying. As examples we mention Kou and Wang [47] where an analytic approximation using Laplace expansions is developed, and Boyarchenko and Levendorskii [15] where the value of an American option in a regime-switching Levy model is computed using finite differences, Carr's randomisation and Wiener-Hopf factorisation.

Chapter 3

Valuing barrier contracts using moments of measures[1]

This chapter is devoted to a method of moments approach to price double barrier-type contracts in the general setting of a piecewise polynomial type jump-diffusion. The method allows the contract to contain a number of interesting features including rebates and payments at a continuous rate. In this approach the rate of discounting is to be a piecewise polynomial function of time.

In this approach, the first step is to express the value of the option as an integral with respect to two measures: the (discounted) expected exit measure and the (discounted) expected occupation measure. The former describes the law of the underlying at expiration or at crossing the barrier whichever is earlier, while the latter describes the law of the process until this moment. By restricting to payoffs that are piecewise polynomial functions of the underlying, the value of the option can be expressed as a linear combination of moments of these two measures. The moments of those two measures are subsequently shown to satisfy an infinite dimensional linear system. The price

[1]This chapter is an adaptation of the article Eriksson and Pistorius [34] published as International Journal of Theoretical and Applied Finance 14(7), 1139–1158 (2011). DOI:10.1142/S0219024911006644 ©World Scientific Publishing Company http://www.worldscientific.com/loi/ijtaf

can thus be associated to the two linear programming problems of minimisation and maximisation of a linear criterion over the spaces of measures. By adding conditions on the moments, which guarantee that a given sequence is equal to the moments of a measure, one is led to an infinite dimensional linear programming problem. Finally, by restricting to a finite number of moments the problem is expressed as a finite dimensional linear programming problem.

The method is numerically illustrated by valuing an American corridor, a double-no-touch and a double knock-out call option under different models. In all cases tight bounds are found with short execution times. A convergence proof is provided to show that the values of the linear programming problems converge monotonically to the value of the option if the number of moments employed is increased.

The method takes its inspiration from a series of earlier papers. Starting with Kurtz and Stockbridge [48] and Bhatt and Borkar [11], these papers independently show how stochastic optimal control problems can be formulated as linear programming problems over measures. Building on these two works, Helmes et al. [38] developed a method of moments algorithm to calculate the moments of the first exit time distribution of a diffusion. Lasserre et al. [54] developed an semi-definite programming (SDP) relaxation approach to price a class of exotic options in a general diffusion setting. By characterising the moments of the underlying diffusion price process Lasserre et al. [54] derived upper and lower bounds for the values of Asian, European and barrier options in terms of semi-definite programs and theoretical and numerical convergence results were provided.

The remainder of the chapter is organised as follows. In Section 3.1 the model and the problem setting is specified. Section 3.2 is devoted to the method of moments, the algorithm is described and a convergence proof is provided. Section 3.3 provides the implementation and numerical examples.

3.1 Problem setting

By following a similar argument as that in Chapter 2, the value of a double barrier knock-out contract, with payoff function $h(S_T)$, rebate R and a flow of payments at a continuous rate $g(s, S_s)$ until the earlier of maturity and the moment of knock-out is given by

$$v = \mathbf{E}_{0,x_0}\left[e^{-\alpha_T} h(S_T)\mathbf{I}(T < \tau_B) + \int_0^{\tau_B \wedge T} e^{-\alpha_s} g(s, S_s)\mathrm{d}s + R\, e^{-\alpha_{\tau_B}}\mathbf{I}(\tau_B \leq T)\right],$$

where S denotes the stochastic price process of the underlying, $\mathbb{E}\backslash B$ is the knock-out set, $\tau_B = \inf\{t \geq 0 : S_t \notin B\}$, $\mathbf{I}$ is the indicator function and α_t is the cumulative discounting

$$\alpha_t = \int_0^t r(s)\mathrm{d}s,$$

where $r : [0, T] \to \mathbb{R}^+$ denotes the risk-free rate of discounting.

3.1.1 Model: piecewise polynomial jump-diffusion

In this chapter the stochastic process S modelling the price process of the underlying will be restricted to the class of piecewise polynomial jump-diffusions. More specifically, the underlying S is assumed to be defined on some filtered probability space $(\Omega, \mathcal{F}, \mathbf{F}, \mathbb{P})$ and to evolve according to the stochastic differential equation

$$\begin{aligned} \mathrm{d}S_t &= b(t, S_t)\mathrm{d}t + \sigma(t, S_t)\mathrm{d}W_t + \lambda(t, S_t)\mathrm{d}J_t, \quad t > 0, \\ S_0 &= x, \end{aligned} \tag{3.1}$$

where x is a real constant, W is a Wiener process and J is a pure jump Lévy process that is a martingale and independent of W. Here b, σ and λ are functions that will be specified below. The state space $\mathbb{E}$ of S is assumed to be equal to $\mathbb{R}_+$ or $\mathbb{R}$.

Recall the notion of piecewise polynomial as it will play an important role in the model description:

Definition 3.1.1. A function $f : [0,T] \times \mathbb{R} \to \mathbb{R}$ is called a *piecewise polynomial* if there exists a finite partition $\{C_i^f\}_{i=1}^k$ of $[0,T] \times \mathbb{E}$ for some integer $k > 0$ such that

$$f(t,x) = \sum_{i=1}^{k} f_i(t,x)\mathbf{I}((t,x) \in C_i^f), \tag{3.2}$$

where f_i are polynomials in (t,x). A finite partition $\{C_i\}_{i=1}^k$ of $[0,T] \times \mathbb{S}$ is a collection of disjoint sets of the form $C_i = C_i^* \times C_i^{**}$, where C_i^*, C_i^{**} are intervals, such that $\cup_{i=1}^k C_i = [0,T] \times \mathbb{S}$.

The following class of stochastic processes will be considered:

Definition 3.1.2. A stochastic process S is called a *piecewise polynomial jump-diffusion* if it is the strong solution of (3.1) where b, σ^2 and λ are *continuous* piecewise polynomials that have linear growth.

Under the conditions stated in the definition, the SDE (3.1) admits a unique strong solution: this follows as a consequence of classical existence and uniqueness results (see for example Jacod and Shiryaev [41, Ch. III.2]), as the continuous piecewise polynomials b, σ and λ are Lipschitz continuous and are assumed to have linear growth. This class of processes contains the class of so-called m-polynomial processes, that was studied in Cuchiero et al. [29].

The precise description of the setting considered in this paper is as follows:

Assumption 3.1.1. *(i) The process S is a polynomial jump-diffusion, and r is a piecewise polynomial.*

(ii) The functions g and h are piecewise polynomials, and R is a non-negative constant.

Let $C^{1,2}([0,T]\times\mathbb{E})$ be the set of functions $f : [0,T]\times\mathbb{E} \to \mathbb{R}$ that are once continuously differentiable in time t and twice continuously differentiable in space x. To the process S is associated the operator $\mathcal{L} : C^{1,2}([0,T]\times\mathbb{E}) \to C([0,T]\times\mathbb{R})$ that acts on functions

$f \in C^{1,2}([0,T] \times \mathbb{E})$ with linear growth as

$$\mathcal{L}f = \frac{\partial f}{\partial t} + b\,\frac{\partial f}{\partial x} + \frac{\sigma^2}{2}\,\frac{\partial^2 f}{\partial x^2} + \mathcal{B}f, \tag{3.3}$$

where $\mathcal{B}f$ is an integro-differential operator given by

$$\mathcal{B}f(t,x) = \int_{\mathbb{R}} \left[f(x + \lambda(t,x)y) - f(t,x) - \lambda(t,x)\frac{\partial f}{\partial x}(t,x)y \right] \eta(\mathrm{d}y),$$

where η denotes the Lévy measure of J. Note that $\int_{|y|>1} |y|\eta(\mathrm{d}y) < \infty$ as J is a martingale. The operator $\mathcal{L}$ is closely related to the infinitesimal generator of the time-space process (t, S_t). In fact, for functions $f \in C_c^{1,2}([0,T] \times \mathbb{E})$ it holds that

$$f(t,S_t) - f(0,S_0) - \int_0^t \mathcal{L}f(s,S_s)\mathrm{d}s$$

is a martingale, where $C_c^{1,2}([0,T] \times \mathbb{E})$ is the set of functions $f : [0,T] \times \mathbb{E} \to \mathbb{R}$ that have compact support and are once continuously differentiable in time t and twice continuously differentiable in space x. If the time-space process (t, S_t) is a Feller process, then any such f lies in the domain of the infinitesimal generator of the semigroup of (t, S_t), and $\mathcal{L}f$ is the infinitesimal generator of (t, S_t). A Markov process S is a Feller process if, for any function $f \in C_0(E)$, the following hold true:

$x \mapsto \mathbf{E}_{0,x}[f(S_t)] \in C_0(E)$ for any $t > 0$.

$\lim_{t\to 0} \mathbf{E}_{0,x}[f(S_t)] = f(x)$ for any $x \in \mathbb{E}$,

where $C_0(E)$ the set of continuous functions that tend to zero at infinity. See Ethier and Kurtz [35] for background on Markov processes and their infinitesimal generators.

Note that the operator $\mathcal{L}$ maps polynomials to piecewise polynomials, which is an essential property needed in the moment approach, as shown in section 3.2 below.

3.1.2 Examples

Next some examples of polynomial jump-diffusion models are presented.

(i) In the landmark paper Black and Scholes [13] the classical geometric Brownian motion (GBM) was suggested as a model of the stock price. The GBM satisfies the SDE

$$\mathrm{d}S_t = bS_t\mathrm{d}t + \sigma S_t \mathrm{d}W_t, \qquad t > 0, \qquad S_0 = x_0 > 0, \tag{3.4}$$

where W denotes a standard one-dimensional Brownian motion. The infinitesimal generator acts on $f \in C_c^2(\mathbb{R}_+)$ as

$$\mathcal{L}f(s) = \frac{s^2\sigma^2}{2}\frac{\mathrm{d}^2 f}{\mathrm{d}s^2}(s) + bs\frac{\mathrm{d}f}{\mathrm{d}s}(s). \tag{3.5}$$

(ii) Lévy models (for an overview see for example Schoutens [70] or Cont and Tankov [24]). A Lévy process is a Markov process with stationary and independent increments; the infinitesimal generator $\mathcal{L}$ of the semi-group of S acts on $f \in C_c^2(\mathbb{R})$ as

$$\mathcal{L}f(x) = \frac{\sigma^2}{2}\frac{\mathrm{d}^2 f}{\mathrm{d}x^2}(x) + b\frac{\mathrm{d}f}{\mathrm{d}x}(x) + \int_{\mathbb{R}} [f(x+y) - f(x) - yf'(x)\mathbf{I}(|y| < 1)]\eta(\mathrm{d}y),$$

where (b, σ^2, η) is the Lévy triplet with $b, \sigma \in \mathbb{R}$ and η the Lévy measure which satisfies the integrability assumption $\int_{\mathbb{R}}[1 \wedge y^2]\eta(\mathrm{d}y) < \infty$. An example of a Lévy process is the variance gamma process (Madan et al. [58]) which is a Lévy process without diffusion component ($\sigma^2 = 0$) with Lévy measure given by

$$\eta(\mathrm{d}x) = \frac{C}{x}e^{-Mx}\mathbf{I}(x > 0)\mathrm{d}x + \frac{C}{|x|}e^{-G|x|}\mathbf{I}(x < 0)\mathrm{d}x, \tag{3.6}$$

with C, G, M positive constants. In this case the infinitesimal generator $\mathcal{L}$ takes the simpler form

$$\mathcal{L}f(x) = df'(x) + \int_{\mathbb{R}} [f(x+y) - f(x)]\, \eta(\mathrm{d}y),$$

where $d \in \mathbb{R}$ is the drift of the variance gamma process.

(iii) Geometric Lévy process with time-dependent interest rate. In such a model the asset price $\{S_t\}_{t\in[0,T]}$ is specified by the SDE

$$\frac{\mathrm{d}S_t}{S_{t-}} = (r(t) - d(t) - c)\mathrm{d}t + \mathrm{d}X_t, \qquad t > 0,$$

where $S_0 = x_0 \in (0, \infty)$, $r, d : [0, T] \to \mathbb{R}_+$ are polynomials representing short rate and dividend yield and X is a Lévy process with characteristic triplet (c, σ^2, η), where the Lévy measure η has support in $(-1, \infty)$ and

$$\int_{(-1,\infty)} z\eta(\mathrm{d}z) < \infty.$$

The former condition on η guarantees that $S_t > 0$ for all $t > 0$, the latter that $\mathbf{E}_{0,x}[|X_t|] < \infty$ for all $t > 0$. The discounted process $\{e^{-\int_0^t [r(s)-d(s)]\mathrm{d}s} S_t\}_{t\geq 0}$ is a martingale, and the infinitesimal generator $\mathcal{L}$ of (t, S_t) acts on $f \in C_c^{1,2}([0, T] \times \mathbb{R}_+)$ as

$$\mathcal{L}f(t, x) = \frac{\partial f}{\partial t}(t, x) + \frac{\sigma^2 x^2}{2}\frac{\partial^2 f}{\partial x^2}(t, x) + [r(t) - d(t)]x\frac{\partial f}{\partial x}(t, x) + \mathcal{B}f(t, x),$$

where

$$\mathcal{B}f(t, x) = \int_{(-1,\infty)} [f(t, x(1 + z)) - f(t, x) - xf'(t, x)z]\eta(\mathrm{d}z).$$

(iv) Affine processes (see for example Cuchiero et al. [28] for applications of affine models in finance). An example of an affine diffusion introduced in Cox et al. [27] is the Cox Ingersoll Ross (CIR) model, which is a mean-reverting diffusion satisfying the SDE

$$\mathrm{d}S_t = a(b - S_t)\mathrm{d}t + \sigma\sqrt{S_t}\mathrm{d}W_t, \qquad a, b > 0, \tag{3.7}$$

with the infinitesimal generator acting on $f \in C_c^2(\mathbb{R}_+)$ as

$$\mathcal{L}f(x) = \frac{\sigma^2 x}{2}\frac{\mathrm{d}^2 f}{\mathrm{d}x^2}(x) + a(b - x)\frac{\mathrm{d}f}{\mathrm{d}x}(x). \tag{3.8}$$

3.2 Method of moments

In the sequel an important role will be played by the sub-probability measures ν and μ given by

$$\nu(A) = \mathbf{E}_{0,x_0}[e^{-\alpha_\tau}\mathbf{I}((\tau, S_\tau) \in A)], \qquad \mu(A) = \mathbf{E}_{0,x_0}\left[\int_{T_0}^{\tau} e^{-\alpha_s}\mathbf{I}((s, S_s) \in A)\mathrm{d}s\right],$$

for Borel sets $A \in \mathcal{B}([T_0, T] \times \mathbb{E})$, where $T_0 < T$ and

$$\tau_B = \inf\{t \geq T_0 : S_t \notin B\}, \qquad \tau = \tau_B \wedge T,$$

for some finite interval $B = (b_-, b_+)$ where $\inf \emptyset = +\infty$. The measures ν and μ are called the *discounted exit location measure* and the *discounted occupation measure* of the time-space process (t, S_t) with respect to $[T_0, T] \times B$. The expected discounted time that the process (t, S_t) spends in a Borel set $A \subset B$ before it exits B for the first time is given by $\mu(A)$, whereas $\nu(A')$ is equal to the discounted probability that (τ, S_τ) takes a value in $A' \subset [T_0, T] \times \mathbb{E}$. Note that while the measures ν and μ are defined on $[T_0, T] \times \mathbb{E}$, from the definition it is clear that $\nu(A) = 0$ for any $A \subset [T_0, T) \times B$ and $\mu(A) = 0$ for any A such that $A \cap [T_0, T) \times B = \emptyset$. The value v of the contract can be expressed in terms of μ and ν as

$$\begin{aligned} v &= \mathbf{E}_{0,x_0}\left[e^{-\alpha_T}h(S_T)\mathbf{I}(T \leq \tau) + \int_{T_0}^{\tau} e^{-\alpha_s}g(s, S_s)\mathrm{d}s + R\, e^{-\alpha_\tau}\mathbf{I}(\tau < T)\right] \\ &= \int_{[T_0,T]\times\mathbb{R}} h(x)\nu(\mathrm{d}t, \mathrm{d}x) + \int_{[T_0,T]\times B} g(t, x)\mu(\mathrm{d}t, \mathrm{d}x), \end{aligned}$$

where $h(x) = R$ if $x \in \mathbb{E}\backslash B$. In fact, in view of the form (3.2) of g and h, v can be expressed in terms of the moments of μ and ν, as follows:

$$v = \sum_{i,j}\sum_{m} d_{i,j}(m)\nu_{i,j}^{(m)} + \sum_{i,j}\sum_{m} b_{i,j}(m)\mu_{i,j}^{(m)},$$

where $d_{i,j}(m)$ and $b_{i,j}(m)$ are some constants, and the notation $\nu^{(m)}(A) = \nu(A \cap C_m)$ and $\mu^{(m)}(A) = \mu(A \cap \tilde{C}_m)$ has been used for the restriction of ν and μ on some partitions ($\{C_m\}_{m=1}^{k}$ and $\{\tilde{C}_m\}_{m=1}^{\tilde{k}}$) determined by the forms of r, σ, b, λ, g and h. The notation $\rho_{i,j} = \int_{[T_0,T]\times\mathbb{E}} t^i x^j \rho(\mathrm{d}t, \mathrm{d}x)$ for the ijth moment of a measure ρ on $[T_0, T] \times \mathbb{E}$ has also been introduced.

3.2.1 The adjoint equation

The measures μ and ν are closely related to each other and to the generator of the underlying process S. Informally, for suitably regular f and all bounded stopping times $\tau \geq T_0$, Dynkin's lemma yields that

$$\mathbf{E}_{0,x_0}[e^{-\alpha_\tau} f(\tau, S_\tau)] - \mathbf{E}_{0,x_0}[e^{-\alpha_{T_0}} f(T_0, S_{T_0})] = \mathbf{E}_{0,x_0}\left[\int_{T_0}^{\tau} e^{-\alpha_t}(\mathcal{L}f - rf)(t, S_t)\mathrm{d}t\right],$$

where $\mathcal{L}f$ is given in (3.3), which can be expressed in terms of the measures ν, μ and the distribution $\ell(\mathrm{d}x) = \mathbf{E}_{0,x_0}\left[e^{-\alpha_{T_0}}\mathbf{I}(S_{T_0} \in \mathrm{d}x)\right]$ of S at T_0 as

$$\int_{[T_0,T]\times\mathbb{E}} f(t,x)\nu(\mathrm{d}t, \mathrm{d}x) = \int_{\mathbb{E}} f(T_0, x)\ell(\mathrm{d}x) + \int_{[T_0,T]\times B} (\mathcal{L}f - rf)(t,x)\mu(\mathrm{d}t, \mathrm{d}x). \quad (3.9)$$

The identity (3.9) is called the *basic adjoint equation* (see for example Helmes et al. [38]). In view of the form of operator $\mathcal{L}$, under appropriate integrability conditions on the Lévy measure η, a formal application of $\mathcal{L}$ to a monomial $f_{ij}(t,x) = t^i x^j$ yields a piecewise polynomial. More specifically, for some coefficients $c_{k,\ell}^{(m)}(i,j)$ that are determined by the form of $\mathcal{L}$, it holds that

$$(\mathcal{L}f_{ij} - rf_{ij})(t,x) = \sum_{k,\ell}\sum_{m} c_{k,\ell}^{(m)}(i,j) t^\ell x^k \mathbf{I}\left((t,x) \in \widetilde{C}_m\right).$$

Hence, by applying (3.9) to a monomial $f_{ij}(t,x) = t^i x^j$ the following infinite system of equations, linking the moments of μ, ν and ℓ, is obtained:

$$\begin{aligned}\int_{[T_0,T]\times\mathbb{E}} & t^i x^j \nu(\mathrm{d}t,\mathrm{d}x) \\ &= T_0^i \int_{\mathbb{E}} x^j \ell(\mathrm{d}x) + \sum_{k,\ell}\sum_m c_{k,\ell}^{(m)}(i,j) \int_{[T_0,T]\times B} t^k x^\ell \mu^{(m)}(\mathrm{d}t,\mathrm{d}x), \qquad i,j = 0,1,\dots\end{aligned}$$

Equivalently, in compact notation,

$$\nu_{i,j} = T_0^i \ell_j + \sum_{k,\ell}\sum_m c_{k,\ell}^{(m)}(i,j)\mu_{k,\ell}^{(m)} \qquad i,j = 0,1,\dots, \tag{3.10}$$

where ℓ_j denotes the j-th moment of ℓ. The following result provides sufficient conditions to justify this informal analysis:

Proposition 3.2.1. *Suppose that the Lévy measure η of J satisfies the following integrability condition:*

$$\int_{|y|>1} |y|^k \eta(\mathrm{d}y) < \infty \qquad \text{for all } k = 0,1,2,\dots. \tag{3.11}$$

Then Equation (3.10) *holds.*

Proof. Fix $i,j \geq 0$ arbitrary and let $f = f_{ij}$ since f is a $C^{1,2}$ function, it can be verified by an application of Ito's lemma to $e^{-\alpha_t} f(t,S_t)$ that

$$M_t = e^{-\alpha_t} f(t,S_t) - e^{-\alpha_{T_0}} f(T_0,S_{T_0}) - \int_{T_0}^t e^{-\alpha_s}(\mathcal{L}f - rf)(s,S_s)\mathrm{d}s \tag{3.12}$$

is a local martingale. As $\tau = \tau_B$ is a bounded stopping time, $M^\tau = \{M_{t\wedge\tau}\}$ is also a local martingale. Here

$$M_t = \int_0^t e^{-\alpha_t}\sigma(s,S_s)\frac{\partial f}{\partial x}(s,S_s)\mathrm{d}W_s + \int_{[0,t]\times\mathbb{R}} e^{-\alpha_s} g(s,S_s,y)\phi(\mathrm{d}y,\mathrm{d}s),$$

where ϕ denotes the compensated jump measure associated to J (with compensator

$\eta(dy)dt)$ and

$$g(t,x,y) = f(t, x + y\lambda(t,x)) - f(t,x).$$

Note that $g(t,x,y)$ is a piecewise polynomial function in y, as $f = f_{ij}$ is a monomial. Further note that, under the integrability conditions (3.11) the process

$$Z_k(t) = \int_{[0,t]\times\mathbb{R}} |y|^k \phi(\mathrm{d}y, \mathrm{d}t)$$

is a square integrable martingale. As a consequence, $\mathbf{E}_{0,x_0}[M_\tau^2] < \infty$ and M^τ is a square integrable martingale. Taking expectations in (3.12) yields that

$$\mathbf{E}_{0,x_0}[e^{-\alpha_\tau} f(\tau, S_\tau)] = \mathbf{E}_{0,x_0}\left[e^{-\alpha_{T_0}} f(T_0, S_{T_0})\right] + \mathbf{E}_{0,x_0}\left[\int_{T_0}^{\tau} e^{-\alpha_s}(\mathcal{L}f - rf)(s, S_s)\mathrm{d}s\right].$$

Since i and j were arbitrary, the validity of (3.10) now follows. □

Remark 3.2.1. Given the entrance measure ℓ, the system of adjoint equations (3.9) for a sufficiently large class of functions f uniquely determines the pair of measures (ν, μ), under a mild regularity condition on the process (t, S_t). If the stopped process $(t \wedge \tau, S_{t\wedge\tau})$ is a Feller process then an extension of Echeverria's Theorem (see Ethier and Kurtz [35, Prop:4.9.19]) implies that if a pair $(\tilde{\mu}, \tilde{\nu})$ of measures defined on the measurable spaces $(\mathcal{U} := (T_0, T) \times B, \mathcal{B}(\mathcal{U}))$ and $(\mathbb{E}\backslash\mathcal{U}, \mathcal{B}(\mathbb{E}\backslash\mathcal{U}))$, satisfies Equation (3.9) for all functions in the domain of the infinitesimal generator, then it holds that $(\tilde{\mu}, \tilde{\nu}) = (\mu, \nu)$.

Truncation

If the minimum $\underline{\lambda}$ of $\lambda(t,x)$ over $[0,T]\times[b_-, b_+]$ is strictly positive, a modification of the adjoint equations can be derived that is valid even when the integrability conditions (3.11) are not satisfied. To that end, note that a double knock-out option becomes worthless at the first time that a jump occurs of size larger than $L_+ := (b_+ - b_-)/\underline{\lambda}$ or smaller than $L_- := (b_- - b_+)/\underline{\lambda}$, since any such jump of the process J in (3.1) will take S out of the interval $[b_-, b_+]$. Hence, the value of the knock-out option does not change if S is replace by a process $\tilde{S}$ that is driven by a Lévy process $\tilde{J}$ with the truncated Lévy measure

$$\tilde{\eta}(\mathrm{d}x) = \mathbf{I}(U_- < x < U_+)\eta(\mathrm{d}x) + \eta_+\delta_{U_+}(\mathrm{d}x) + \eta_-\delta_{U_-}(\mathrm{d}x),$$

for some $U_+ \geq L_+$ and $U_- \leq L_-$, where $\eta_- = \eta((-\infty, U_-])$ and $\eta_+ = \eta([U_+, \infty))$ and δ_a denotes the delta measure at a.

In summary, denoting by $\tilde{\nu}_{ij}^{(m)}$, $\tilde{\mu}_{ij}^{(m)}$ the ij-th moments of the exit and occupation measures $\tilde{\nu}^{(m)}$, $\tilde{\mu}^{(m)}$ of the process $\tilde{S}$, the following result holds:

Corollary 3.2.1. *If $\underline{\lambda} > 0$, then v is given by*

$$v = \sum_{i,j} \sum_{m} d_{i,j}(m)\tilde{\nu}_{i,j}^{(m)} + \sum_{i,j} \sum_{m} b_{i,j}(m)\tilde{\mu}_{i,j}^{(m)},$$

where $\tilde{\nu}_{ij}^{(m)}$ and $\tilde{\mu}_{ij}^{(m)}$ solve the system of equations

$$\sum_{m} \tilde{\nu}_{ij}^{(m)} - T_0^i \lambda_j = \sum_{k,\ell} \sum_{m} \tilde{c}_{k,\ell}^{(m)}(i,j)\tilde{\mu}_{k,\ell}^{(m)} \qquad i,j = 0,1,\ldots, \tag{3.13}$$

where the coefficients $\tilde{c}_{k,\ell}^{(m)}(i,j)$ are defined by

$$\tilde{\mathcal{L}}f_{ij} - rf_{ij} = \sum_{k,\ell} \sum_{m} \tilde{c}_{k,\ell}^{(m)}(i,j)t^k x^l,$$

with $\tilde{\mathcal{L}}f = \frac{\partial f}{\partial t} + b\frac{\partial f}{\partial x} + \frac{\sigma^2}{2}\frac{\partial^2 f}{\partial x^2} + \mathcal{B}f$, with $\mathcal{B}f$ given by

$$\mathcal{B}f(t,x) = \int_{\mathbb{R}} \left[f(x + \lambda(t,x)y) - f(t,x) - \lambda(t,x)\frac{\partial f}{\partial x}(t,x)y \right] \tilde{\eta}(\mathrm{d}y).$$

3.2.2 Linear programs

By optimising over the pair of measures that satisfies the adjoint equations, the value v can be bounded in terms of linear programs over sets of measures, as follows:

$$\inf_{\nu,\mu} L(\nu,\mu) \leq v \leq \sup_{\nu,\mu} L(\nu,\mu), \tag{3.14}$$

where L is the linear functional of the moments of ν and μ given by

$$L(\nu,\mu) := \sum_{i,j}\sum_{m} d_{i,j}(m)\nu_{i,j}^{(m)} + \sum_{i,j}\sum_{m} b_{i,j}(m)\mu_{i,j}^{(m)},$$

and where the infimum and the supremum are taken over the pairs of measures (ν,μ) supported on $((T_0,T]\times\mathbb{E}\backslash\mathcal{S},\mathcal{S})$ where $\mathcal{S}=(T_0,T)\times B$ that satisfy the linear adjoint equations derived before. In general there is no guarantee that a given solution of the system (3.10) or of (3.13) is the moment sequence of some measure. Therefore to formulate these optimisation problems completely in terms of moment sequences it is needed to express the condition that μ and ν be measures in terms of their moments. The problem to determine whether a given sequence is the moment sequence of some measure and, if so, whether this measure is uniquely determined (in which case the measure is called *moment-determinate*) is classical and has been extensively studied. It is well known that the Cramér condition

$$\int_{\mathbb{R}} e^{c|x|} m(\mathrm{d}x) < +\infty, \text{ for some } c>0$$

is a sufficient condition for a measure m to be moment determinate, since under the Cramér condition the moment-generating function of m exists in a neighbourhood of zero. In particular, any measure with compact support is moment-determinate. Further, the following Hausdorff conditions are necessary and sufficient for a given sequence m_i to correspond to the moments of a measure m with support on the interval $[a, b]$ (see for example Shohat and Tamarkin [73, Thm:1.5]):

$$\sum_{j=0}^{n} \binom{n}{j} (-1)^j \tilde{m}_{j+k} \geq 0 \quad \forall n, k = 0, 1, 2, ... \tag{3.15}$$

where the $\tilde{m}_i$ are linear combinations of the m_j, as follows:

$$\tilde{m}_l = (b-a)^{-l} \sum_{i=0}^{l} \binom{l}{i} (-a)^{l-i} m_i.$$

In fact, the $\tilde{m}_i$ are themselves the moments of a measure $\tilde{m}$ that is the affine transformation of m supported on $[0, 1]$. That these conditions are necessary immediately follows by observing that $\int_0^1 y^k (1-y)^n \tilde{m}(\mathrm{d}y)$ is non-negative and by expressing $\tilde{m}_i$ in terms of m_j. More generally, given an array $(m_{ij}, i, j = 0, 1, \ldots)$, the two dimensional Hausdorff-conditions (see for example Shohat and Tamarkin [73, Thm:1.6])

$$\sum_{i=0}^{m} \sum_{j=0}^{n} \binom{m}{i} \binom{n}{j} (-1)^{i+j} \tilde{m}_{i+l, j+k} \geq 0. \quad \forall n, m, k, l = 0, 1, 2, ..., \tag{3.16}$$

where the $\tilde{m}_{i,j}$ are related to the $m_{i,j}$ by

$$\tilde{m}_{k,l} = (b-a)^{-k} (d-c)^{-l} \sum_{i=0}^{k} \sum_{j=0}^{l} \binom{k}{i} \binom{l}{j} (-a)^{k-i} (-c)^{l-j} m_{i,j},$$

are necessary and sufficient conditions to guarantee that there exists a measure m supported on $[a, b] \times [c, d]$ such that $m_{ij} = \int_a^b \int_c^d x^i y^j m(\mathrm{d}x, \mathrm{d}y)$. See Shohat and Tamarkin [73] for proofs and further background on moment problems.

In the sequel it will be assumed that the state space has been truncated (as in Section 3.2.1), so that the occupation and exit measures have compact support and thus are moment-determinate.

If a single barrier problem is considered, as in Remark 3.2.4, depending on the process considered the measures might not be moment-determinate.

3.2.3 Approximations and convergence

To be able to calculate lower and upper bounds for the value v, the optimisation problems in (3.14) is approximated by restricting the total number of moments used to N. If B is a finite interval, employing the moment conditions (3.15) and (3.16) results in the following (finite) linear programming problems:

$$v_{\pm}^{(N)} := \frac{\max}{\min} \left\{ \begin{array}{l} \sum_{i,j} \sum_{m} d_{i,j}(m) \nu_{i,j}^{(m)} + \sum_{i,j} \sum_{m} b_{i,j}(m) \mu_{i,j}^{(m)} \\ \\ \text{subject to} \\ \\ \nu_{i,j} - t_0^i \lambda_j = \sum_{k,\ell} \sum_{m} c_{k,\ell}^{(m)}(i,j) \mu_{k,\ell}^{(m)}, \qquad i+j \leq N,\ k+\ell \leq N \\ \quad \text{with} \quad \nu = \sum_{m} \nu^{(m)}, \quad \mu = \sum_{m} \mu^{(m)} \\ \\ \text{conditions (3.15)–(3.16) for } \nu_{i,j}^{(m)},\ \mu_{i,j}^{(m)}, \quad i+j \leq N \end{array} \right\}.$$

It can be shown that the values $v_{-}^{(N)}$, $v_{+}^{(N)}$ of the linear programs converge to the value v of the option as N tends to infinity:

Proposition 3.2.2. *Suppose that the system* (3.13) *has a unique solution and that B is a finite interval. Then*

$$v_{-}^{(N)} \uparrow v \quad \textit{and} \quad v_{+}^{(N)} \downarrow v,$$

as $N \to \infty$.

Proof. Since the number of equations grows with N, it follows that $v_{-}^{(N)}$ is monotone increasing, since the minimum taken over a smaller set of elements is larger.

Note that v is a finite linear combination of moments. Because of the fact that the support of the different measures is compact, if follows that each moment is bounded, so that v is bounded and $v_-^{(N)}$ is the minimisation of a linear function over a bounded set. Thus, the minimum $v_-^{(N)}$ is finite and attained at a vector $q^N = (q_i^N)_i$ that satisfies the corresponding linear system of equations. Thus, for each fixed i there exists a q_i^* such that, for N along a subsequence, $q_i^N \to q_i^*$. In fact, by a diagonal argument it follows that there exists a subsequence $\tilde{N}$ such that, as $\tilde{N} \to \infty$,

$$q_i^{\tilde{N}} \to q_i^* \quad \text{for all } i.$$

Clearly, q^* satisfies the infinite system and thus under the assumption that there exists a unique sequence that solves the infinite system, it follows that q^* must be equal to $(\mu_i, \nu_i)_i$. Moreover, μ and ν are the unique measures corresponding to these moments, as they are both moment-determinate. Thus, $v = L(q^*)$ and $v_-^{(N)} \uparrow L(q^*)$

The proof of the convergence of the sequence $(v_+^{(N)})$ is similar and omitted. □

Remark 3.2.2. A sufficient condition for uniqueness is that $(t \wedge \tau, S_{t\wedge\tau})$ is a Feller process. Indeed, if q^* is a solution of the system (3.13), then the moment-conditions in conjunction with the fact that measures with compact support are moment-determinate yield that q^* is equal to a sequence of moments of a pair $(\tilde{\mu}, \tilde{\nu})$ of measures. An extension of Echeverria's theorem then implies that $(\tilde{\mu}, \tilde{\nu}) = (\mu, \nu)$ (see the Remark 3.2.1).

Remark 3.2.3. The presented approach can in principle be extended to a multi-dimensional jump-diffusion S with polynomial coefficients. For example, if B is a hyper-cube, the adjoint equations and the moment conditions take analogous forms. The limitation in practice will be the capacity of the LP solvers to deal with large size programs.

Remark 3.2.4. In the case of single barrier options, when the set B is a half-line, the measures in question will have unbounded support. Sufficient conditions to characterise the moments of measures with unbounded support can be phrased in terms of the semi-definiteness of certain moment matrices. The resulting optimisation problems are then semi-definite programming (SDP) problems. In a diffusion setting Lasserre et al. [54] provided convergence results for single barrier option prices computed employing this SDP approach.

3.3 Numerical examples

In this section the method is illustrated by valuing three different derivative contracts under different models. For the numerical examples the freeware LP solver `lp_solve`[2] was run on an Intel Core Duo T2500 2GHz (with 2GB RAM). The numerical outcomes were compared with Monte Carlo results obtained employing a standard Euler scheme.

3.3.1 Double knock-out call option in the GBM model

In this benchmark example a double knock-out call option in the geometric Brownian motion (GBM) model is considered. By standard arbitrage principles the arbitrage-free value v of such an option is given by

$$\begin{aligned} v &= e^{-rT}\mathbf{E}_{0,x_0}[(S_T - K)^+\mathbf{I}(\tau \geq T)] \quad \text{with} \\ \tau &= \inf\{t \geq T_0 : S_t \notin [B_d, B_u]\}, \end{aligned}$$

where $S = \{S_t\}_{t\geq 0}$ is given by (3.4) with $b = r - d$ with r denoting the short rate and d the dividend yield, and $T_0 \in [0, T)$ denotes the moment that the continuous monitoring commences. The case $T_0 > 0$ corresponds to a forward starting double knock-out call option.

For the ease of notation the discounting factor e^{-rT} is dropped in the following section. As a geometric Brownian motion has continuous paths, the time-space process (t, S_t) will exit $[T_0, T] \times [B_d, B_u]$ either if S hits one of the barriers B_u or B_d or maturity T is reached, so that the support for the exit location measure ν is given by

$$\Omega = \{[T_0, T] \times \{B_d\}\} \cup \{\{T\} \times [B_d, B_u]\} \cup \{[T_0, T] \times \{B_u\}\}.$$

[2]This package can be downloaded free of charge from `http://lpsolve.sourceforge.net/` (together with its reference guide). It runs under in a number of different packages, such as MATLAB, SCILAB, R and OCTAVE.

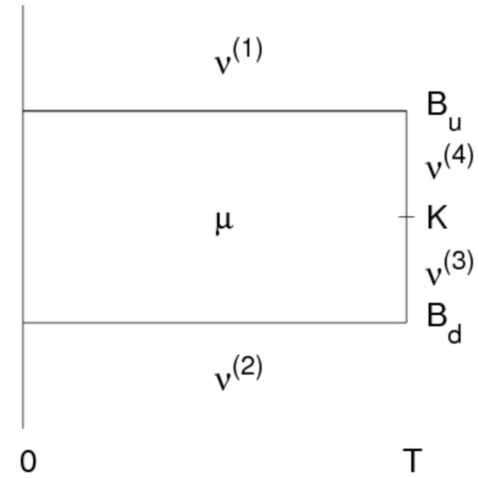

Figure 3.1: Domain for the measures ν and μ in the case of a double knock-out call option in the GBM model.

The set Ω is partitioned into four parts with the restricted measures

$\nu^{(1)}$ and $\nu^{(2)}$ with support on $[T_0, T] \times \{B_d\}$, $[T_0, T] \times \{B_u\}$

$\nu^{(3)}$ and $\nu^{(4)}$ with support on $\{T\} \times [B_d, K]$ and $\{T\} \times [K, B_u]$ respectively.

The expected occupation measure μ is supported on the domain $[T_0, T] \times [B_d, B_u]$ – See also Figure 3.1 for an illustration.

The value v can then be expressed as follows:

$$v = \int_\Omega (x - K)^+ \mathbf{I}(t = T)\nu(\mathrm{d}t, \mathrm{d}x) = \nu_1^{(4)} - K\nu_0^{(4)}.$$

Using the form (3.5) of the infinitesimal generator of the Geometric Brownian motion, the resulting system of basic adjoint equations for this problem can be seen to be given by

$$\begin{aligned} &B_u^m \nu_n^{(1)} + B_d^m \nu_n^{(2)} + T^n \nu_m^{(3)} + T^n \nu_m^{(4)} - \\ &\quad n\mu_{n-1,m} - \left(bm + \frac{\sigma^2}{2} m(m-1) \right) \mu_{n,m} = T_0^n l_m, \end{aligned}$$

Case 1	$b = 0.1$	$\sigma = 0.1$		
Degree of moment	9	10	11	12
Upper Bound	0.9250	0.9211	0.9182	0.9161
Lower Bound	0.9096	0.9100	0.9102	0.9103
Relative Error	0.76%	0.57%	0.42%	0.32%
Exact solution	0.9103			
Case 2	$b = 0.2$	$\sigma = 0.2$		
Degree of moment	8	9	10	11
Upper Bound	1.1656	1.1611	1.1569	1.1534
Lower Bound	1.1064	1.1163	1.1256	1.1293
Relative Error	0.54%	0.30%	0.08%	0.07%
Exact solution	1.1421			
Case 3	$b = 0.1$	$\sigma = 0.3$		
Degree of moment	7	8	9	10
Upper Bound	0.9675	0.9598	0.9536	0.9482
Lower Bound	0.8257	0.8437	0.8525	0.8551
Relative Error	0.75%	0.18%	0.03%	0.19%
Exact solution	0.9034			

Table 3.1: Numerical results for a double knock-out call option in the GBM with spot $x_0 = 2$, strike $K = 1.3$, starting point of monitoring period $T_0 = 0$ and maturity $T = 1$, and upper and lower barrier levels $B_u = 5$ and $B_d = 1$. The relative error reported is the error of the average of the lower and upper bounds with respect to the exact value. Calculation times of a single bound varied from 0.5 seconds (7 moments) to 5 seconds (12 moments).

for all n, m such that $n + m \leq N$, when using all moments up to degree N. Note that if $T_0 = 0$ it holds $l_m = x_0^m$. To complete the setup of the problem, add the LP moment conditions for the measures $\nu^{(i)}$ and μ with supports as given above.

The numerical results for three sets of parameter values are given in Table 3.1. The exact value was calculated employing the analytical formula in Pelsser [64, Section 3.3].

3.3.2 The American corridor option in the CIR model

An American corridor option is a derivative contract traded in the Foreign exchange markets that entitles the holder to a flow of payments until either the underlying rate leaves the corridor or maturity is reached, whichever comes earlier (see for example

Weithers [77] or Wystup [79] for background). The value of this contract is given by

$$v = \mathbf{E}_{0,x_0}\left[\int_{T_0}^{\tau} e^{-rt}\mathrm{d}t\right] = \int_{\Theta} \mu(\mathrm{d}t, \mathrm{d}s) = \mu_{0,0} \quad \text{with}$$
$$\tau = T \wedge \inf\{t \geq T_0 : S_t \notin [B_d, B_u]\},$$

where $\Theta = \{[0,T] \times [B_d, B_u]\}$ and T denotes again the maturity. The underlying will be modelled by a Cox Ingersoll Ross (CIR) process, evolving according to the SDE (3.7). Since the CIR process is continuous, the supports of the different measures are given as follows:

$$\nu^{(1)} \text{ and } \nu^{(2)} \text{ supported on } [0,T] \times \{B_d\} \text{ and } [0,T] \times \{B_u\},$$
$$\nu^{(3)} \text{ supported on } \{T\} \times [B_d, B_u],$$
$$\mu \text{ is supported on } [0,T] \times [B_d, B_u].$$

In view of the form of the infinitesimal generator for the CIR process (3.8), the basic adjoint equation for this problem can be assembled as

$$B_u^m \nu_n^{(1)} + B_d^m \nu_n^{(2)} + T^n \nu_m^{(3)} - n\mu_{n-1,m} +$$
$$(am + r)\mu_{n,m} - \left(abm + \frac{\sigma^2}{2} m(m-1)\right)\mu_{n,m-1} = T_0^n l_m$$

for all m, n such that $m + n \leq N$, and add the appropriate LP moment conditions, as before.

The price of an American corridor with maturity $T = 1$ under the CIR model with time-dependent parameters that are constant from $T_0 = 0$ to $t_1 = 0.5$ and from $t_1 = 0.5$ to $T = 1$ is computed, by applying the algorithm twice. The results are reported in Table 3.2. Observe that with respect to the Monte Carlo results, the relative error of the average of the upper and lower bounds is not more than 1% using moments up to degree 11.

$a_1 = 0.5$	$b_1 = 0.1$	$a_2 = 0.75$	$b_2 = 0.05$
$\sigma_2 = 0.3$	$\sigma_1 = 0.3$	$\sigma_1 = 0.2$	$\sigma_1 = 0.1$
Upper Bound	0.8520	0.8863	0.9048
Lower Bound	0.7932	0.8307	0.8518
Relative Error	0.76%	0.04%	0.26%
$\sigma_2 = 0.2$	$\sigma_1 = 0.3$	$\sigma_1 = 0.2$	$\sigma_1 = 0.1$
Upper Bound	0.8797	0.9288	0.9432
Lower Bound	0.8103	0.8521	0.8883
Relative Error	1.10%	0.10%	0.46%
$a_1 = 0.25$	$b_1 = 0.1$	$a_2 = 0.75$	$b_2 = 0.05$
$\sigma_2 = 0.3$	$\sigma_1 = 0.3$	$\sigma_1 = 0.2$	$\sigma_1 = 0.1$
Upper Bound	0.9009	0.9266	0.9389
Lower Bound	0.8407	0.8780	0.8940
Relative Error	0.13%	0.07%	0.02%
$\sigma_2 = 0.2$	$\sigma_1 = 0.3$	$\sigma_1 = 0.2$	$\sigma_1 = 0.1$
Upper Bound	0.9289	0.9474	0.9516
Lower Bound	0.8565	0.8951	0.9105
Relative Error	0.40%	0.31%	0.76%

Table 3.2: Values of the American corridor option with lower and upper levels $B_d = 0.5$ and $B_u = 1.5$, short rate $r = 0.1$ and maturity $T = 1$, in a Cox Ingersoll Ross model with spot $x_0 = 1$, where the parameters are given by (a_1, b_1, σ_1) on $[0, t_1)$ and change at time $t_1 = 0.5$ to (a_2, b_2, σ_2), and can be found in the table. The relative error reported is the error of the average of the lower and upper bounds with respect to the value calculated by a Monte Carlo simulation (with standard error smaller than 0.0001). All results were obtained by applying the algorithm twice, on $[0, t_1]$ and on $[t_1, T]$, using moments up to degree 11. Calculation times of a single bound were 6-7 seconds.

3.3.3 Double-no-touch option in the VG model with non-constant interest rate

A double no touch option pays one unit at maturity T if, at maturity, the underlying S_t has not crossed either of the barriers B_d or B_u. Its value can be expressed as

$$v = \mathbf{E}_{0,x_0}[e^{-\alpha_T}\mathbf{I}(\tau \geq T)] \quad \text{where}$$
$$\tau = \inf\{t \geq 0 : S_t \notin [B_d, B_u]\}.$$

Let the price process of the underlying $\{S_t\}_{t\in[0,T]}$ evolve as the exponential process

$$S_t = x_0 e^{\int_0^t [r(s)-d(s)]\mathrm{d}s} \frac{e^{X_t}}{\mathbf{E}_{0,X_0}[e^{X_t}]}, \tag{3.17}$$

where X_t is a variance gamma (VG) process with $\mathbf{E}_{0,X_0}[e^{X_t}] < \infty$, and the functions $r, d : [0, T] \to \mathbb{R}_+$ are the risk-free interest rate and the dividend yield. Note that, as a consequence of the independence of increments, of X the discounted process $e^{-\int_0^t [r(s)-d(s)]\mathrm{d}s} S_t$ is a martingale.

The stopping time τ can equivalently be represented as

$$\tau = \inf\{t \geq 0 : Y_t \notin [b_d, b_u]\}, \qquad \text{with } b_d = \log(B_d) \text{ and } b_u = \log(B_u)$$

and $Y_t = \log S_t = \beta(t) + X_t$ for some function β. In view of Equation (3.17), $\beta : [0, T] \to \mathbb{R}$ is equal to

$$\beta(t) = r(t) - d(t) - \int_{\mathbb{R}} (e^x - 1)\eta(\mathrm{d}x),$$

with η denoting the Lévy measure (3.6) of X. Thus, it follows that the infinitesimal generator $\mathcal{L}$ of Y acts on $f \in C_c^2(\mathbb{E})$ as

$$\mathcal{L}f(t, x) = \frac{\partial f}{\partial t}(t, x) + (r(t) - d(t) - c)\frac{\partial f}{\partial x}(t, x) + \int_{\mathbb{R}} [f(t, x + y) - f(t, x)]\eta(\mathrm{d}y),$$

where, in view of the form (3.6) of η,

$$c := \int_{-\infty}^{\infty} (e^x - 1)\eta(\mathrm{d}x) = C\left(\log\left(\frac{G}{1 + G}\right) + \log\left(\frac{M}{1 - M}\right)\right).$$

The dividend yield d is set to zero and the interest rate r is chosen to be

$$r(t) = r_b + r_s t^2.$$

Since a variance gamma process is a finite variation jump process, Y may jump across the barriers b_d or b_u when exiting (b_d, b_u). As a consequence the exit location measure ν of Y is supported on

$$\Omega = \{[0, T] \times [b_u, \infty)\} \cup \{\{T\} \times [b_d, b_u]\} \cup \{[0, T] \times (-\infty, b_d]\}.$$

In view of the form of the payoff of a double-no-touch, it suffices to split the domain of the exit-location measure ν into three parts $K_1 = [0,T] \times (-\infty, b_d]$, $K_2 = [0,T] \times [b_u, \infty)$ and $K_3 = \{T\} \times [b_d, b_u]$, and the value v can be expressed in terms of moments as follows

$$v = \mathbf{E}_{0,x_0}[e^{-\alpha_T} I(\tau \geq T)] = \int_{K_3} \nu^{(3)}(\mathrm{d}t, \mathrm{d}y) = \nu_0^{(3)},$$

where as before $\nu^{(i)} = \nu(\cdot \cap K_i)$. In order to be able to calculate the value v of the option using the LP moment conditions, the Lévy measure η is truncated as described in Section 3.2.1 to achieve bounded support. Since any jump with absolute size larger than $L \geq b_u - b_d$ will trigger an immediate knock-out, the value v is the same as the value of a double-no-touch in the model (3.17) where X is a Lévy process with Lévy measure $\tilde{\eta}$ given by

$$\tilde{\eta}(\mathrm{d}x) = \mathbf{I}(|x| < L)\eta(\mathrm{d}y) + \lambda_- \delta_{-L}(\mathrm{d}y) + \lambda_+ \delta_L(\mathrm{d}y),$$

where δ_a denotes the delta measure at a. The support of the corresponding exit measure is then given by

$$\tilde{\Omega} = \{[0,T] \times [b_u, b_u + L]\} \cup \{\{T\} \times [b_d, b_u]\} \cup \{[0,T] \times [b_d - L, b_d]\}$$

and the three restrictions of ν are $\tilde{\nu}^{(3)} = \nu^{(3)}$ and

$$\begin{aligned} &\tilde{\nu}^{(1)} \quad \text{supported on } [0,T] \times [b_u, b_u + L], \\ &\tilde{\nu}^{(2)} \quad \text{supported on } [0,T] \times [b_d - L, b_d]. \end{aligned}$$

As the occupation measures μ and $\tilde{\mu}$ of the original and the truncated processes are equal the basic adjoint equation is then given by

$$\begin{aligned} \tilde{\nu}^{(1)}_{n,m} + \tilde{\nu}^{(2)}_{n,m} + T^n \nu^{(3)}_m = {} & n\mu_{n-1,m} - r_b \mu_{n,m} - r_s \mu_{n+2,m} + (r_b - c) m \mu_{n,m-1} \\ & + r_s m \mu_{n+2,m-1} + \sum_{k=1}^{m} \binom{m}{k} \tilde{c}(k) \mu_{n,m-k} + T_0^n l_m, \end{aligned}$$

Case 1	$r_b = 0.05$	$r_s = 0.05$		
Degree of moment	6	7	8	9
Upper Bound	0.9356	0.9355	0.9355	0.9355
Lower Bound	0.8453	0.8757	0.9042	0.9143
Relative Error	4.79%	3.17%	1.64%	1.10%
Monte Carlo	0.9352	Std Error	0.0002	
Case 2	$r_b = 0.05$	$r_s = 0.1$		
Degree of moment	6	7	8	9
Upper Bound	0.9203	0.9201	0.9200	0.9200
Lower Bound	0.8196	0.8533	0.8836	0.8957
Relative Error	5.38%	3.56%	1.91%	1.26%
Monte Carlo	0.9194	Std Error	0.0002	
Case 3	$r_b = 0.1$	$r_s = 0.1$		
Degree of moment	7	8	9	10
Upper Bound	0.8752	0.8752	0.8752	0.8752
Lower Bound	0.7980	0.8319	0.8449	0.8565
Relative Error	6.68%	4.73%	2.84%	2.09%
Monte Carlo	0.8746	Std Error	0.0002	

Table 3.3: Numerical results for the double-no-touch option with barriers at $B_u = 2$ and $B_d = 0.5$ and maturity $T = 1$, driven by an exponential VG process with parameters $C = 0.5$, $G = 8$ and $M = 12$, and with $x_0 = 1$. Note that the upper bounds are accurate already for 6-7 moments, with relative errors less than 0.1%. Calculation times varied from 0.4 (6 moments) to 2 seconds (9 moments).

where $\tilde{c}(k)$ denotes the k-th moment of the truncated Lévy measure $\tilde{\eta}$,

$$\tilde{c}(k) = \int_{\mathbb{R}} y^k \, \tilde{\eta}(\mathrm{d}y) = \int_{\mathbb{R}} [(-L) \vee y \wedge L]^k \eta(\mathrm{d}y).$$

with $x \vee y = \max\{x, y\}$ and $x \wedge y = \min\{x, y\}$.

The numerical results are presented in Table 3.3. In all cases tight bounds that agree with the Monte Carlo simulation result are found. Also note that the upper bound is close to the true value already for 6–7 moments.

3.4 Conclusion

A method of moments approach for pricing of double barrier-type contracts driven by polynomial jump-diffusions has been presented, allowing for a non-constant interest rate. An infinite-dimensional linear program was derived, which was then approximated by a sequence of LP problems. Theoretical convergence results were provided, and the method was illustrated with numerical examples. Accurate results with tight upper and lower bounds were obtained with a small number of moments, with execution times that were significantly faster than Monte Carlo simulation. When attempting to use the algorithm with higher order moments (say, beyond 14-15th moment), the solver `lp_solve` experienced numerical instability. To extend the range of practical applicability of this moment-matching algorithm, for example to the case of two-dimensional homogeneous Markov processes, development of a dedicated LP solver would be needed that can handle a larger number of moments. This is a task left for future investigation.

Chapter 4

The valuation of American options under Markov chains

In this chapter the optimal stopping problem associated to an American option in the setting of a continuous-time Markov chain with discrete state space is considered. Stochastic processes from this class have served as models for the evolution of random quantities that take values in lattices. Models from this class, which contains the classical birth-death processes, have recently also been deployed to model the state of the order book or the limit price—see for example Abergel and Jedidi [1] and references therein. Furthermore, Markov chains have been deployed as models on a discrete state space that closely approximate continuous space diffusions, jump-diffusions and general Feller processes. In a continuous-time Markov chain setting the optimal stopping problem associated to the valuation of an American option with a payoff that is a function of the Markov chain is solved. While it follows from the general theory of optimal stopping that the optimal stopping time is given by the first passage time into a certain set (see Peskir and Shiryaev [65]), the characterisation of the value function as the solution of a corresponding free boundary problem and the identification of the optimal boundary involve non-standard arguments. Taking advantage of the explicit form of the semi-group, it is demonstrated that the value function of such an American option is the unique solution in distributional sense of an associated free boundary problem, and it is deduced that the value function is in fact a classical solution by showing that it is continuously differentiable as function of time (see Theorem 4.2.3 below). In cases when the payoff and Markov process are sufficiently regular, fit principles have been used to identify and characterise the optimal boundary (See Section 2.3 for details). However, in the case of a discrete state space with finite transition rates the fit principles no longer apply due to the lack of smoothness that is a result of the fact that the state space is discrete. In the absence of any fit principles, an explicit characterisation of the optimal stopping boundary, directly in terms of the infinitesimal generator of the Markov chain, is derived in the case that the optimal stopping boundary is monotone (see Theorem 4.2.5 below).

Deploying this characterisation, an algorithm is designed for the computation of the value function of an American option under a continuous-time Markov chain model. By an application of the dynamic programming principle, an algorithm is also obtained for the computation of the value function of a Bermudan option under a continuous-time Markov chain model. Note that this algorithm can also be used to approximate the value of the American option, for example by taking a sufficiently large set of equidistant exercise date, as in that case the values of the Bermudan and the American options will be close. By constructing the Markov chain such that it closely follows the evolution of a given Feller process (for example, by using the construction from Mijatović and Pistorius [62]), these algorithms, with the constructed Markov chain as input, provide methods for the valuation of American options under the Feller process in question. An advantage of the Markov chain model is its computational tractability: It is demonstrated in this chapter that the described algorithms provide efficient and accurate methods for the valuation of American options, and the computation of the optimal boundary, using the powerful tools of matrix-based computations. The idea of valuation using Markov chain approximation goes back at least as far as Kushner [49] in the case of diffusions, and was further developed in for example Mijatović and Pistorius [62]. To illustrate its effectiveness, the algorithm is implemented for a number of models, and results (such as estimates of the errors) are reported in Section 4.4. Proof of convergence of the approximation methods are also given.

The remainder of the chapter is organised as follows. Section 4.1 contains preliminaries and notation that is used throughout the chapter. Section 4.2 is devoted to the valuation of an American option driven by a continuous-time Markov chain and contains a characterisation of the optimal boundary, an algorithm for solving the free boundary problem and an algorithm for valuing the Bermudan option when the underlying is a continuous-time Markov chain. Convergence of the algorithm is established in Section 4.3, and a number of numerical examples are analysed in Section 4.4. Appendix 4.A contain the algorithm to build a continuous-time Markov chain approximating a Feller process and the method for the valuation of knock-out barrier options driven by a continuous-time Markov chain introduced in Mijatović and Pistorius [62].

4.1 Problem setting

In this section the problem setting is specified. In Section 4.1.1 the continuous-time Markov chain is described and in Section 4.1.2 the American options considered are presented. Section 4.1.3 contains a version of Dynkin's lemma for continuous-time Markov chains.

4.1.1 Markov chains

Next the notation is set that will be used throughout the chapter. Let M be a continuous-time time-homogeneous Markov chain with discrete state space $\mathbb{G} = \{x_i, i \in \mathbb{N}\}$, defined on some filtered probability space $(\Omega, \mathcal{G}, \mathbf{G}, \mathbb{P})$ where $\mathbf{G} = \{\mathcal{G}_t\}_{t\in[0,T]}$ denotes the filtration generated by M. Assume that M is a Feller process, this implies the existence of an infinitesimal generator of M, denoted by Λ. To avoid explosion of the chain M in finite time the set of Markov chains is restricted to those with an infinitesimal generator Λ that has uniformly bounded elements:

Assumption 4.1.1. *The infinitesimal generator Λ of M satisfies the condition*

$$\sup_{x\in\mathbb{G}} |\Lambda(x,x)| < \infty.$$

Denoting by $l^\infty(\mathbb{G})$ the collection of bounded real-valued functions with domain $\mathbb{G}$, the semi-group of M is equal to the collection $(P_t, t \geq 0)$ of maps $P_t : l^\infty(\mathbb{G}) \to l^\infty(\mathbb{G})$ that is expressed in terms of the infinitesimal generator $\Lambda : l^\infty(\mathbb{G}) \to l^\infty(\mathbb{G})$ of M by

$$\begin{aligned}
&(P_t f)(x) = \sum_{y\in\mathbb{G}} P_t(x,y)f(y) \qquad t \geq 0, x \in \mathbb{G}, f \in l^\infty(\mathbb{G}),\\
&P_t(x,y) = \mathbb{P}(M_t = y | M_0 = x) =: \mathbb{P}_x(M_t = y), \qquad x,y \in \mathbb{G},\\
&\text{with } \ P_t = \exp(t\Lambda), \qquad t \in \mathbb{R}_+,\\
&\exp(t\Lambda) = \sum_{n=0}^{\infty} \frac{t^n}{n!}\Lambda^n
\end{aligned}$$

with $\Lambda^n = \Lambda^{n-1} \circ \Lambda$, that is, $\Lambda^n f = \Lambda^{n-1}(\Lambda f)$ for any $f \in l^\infty(\mathbb{G})$. The infinitesimal generator Λ is given by

$$\begin{aligned}
&(\Lambda f)(x) = \sum_{y\in\mathbb{G}} \Lambda(x,y)f(y), \qquad \Lambda(x,y) = (\Lambda\delta_y)(x), \qquad x \in \mathbb{G}, f \in l^\infty(\mathbb{G}),\\
&\text{with } (1-\delta_y(x)) \cdot \Lambda(x,y) \geq 0 \text{ and } \sum_{z\in\mathbb{G}} \Lambda(x,z) = 0, \quad x,y \in \mathbb{G},
\end{aligned}$$

where δ_y is the Kronecker delta, which is the map on $\mathbb{G}$ that is equal to 1 if x and y are equal and zero otherwise. In particular, it follows that the expected value of the payoff $\phi(M_T)$ at time T, where ϕ is an arbitrary map from the set $l^\infty(\mathbb{G})$, is given by

$$\mathbf{E}_{t,x}[\phi(M_T)] = \mathbf{E}_{0,x}[\phi(M_{T-t})] = (\exp((T-t)\Lambda)\phi)(x), \qquad x \in \mathbb{G}, t \in [0,T]. \quad (4.1)$$

For a bounded function $f : [0,T] \times \mathbb{G} \to \mathbb{R}$ the notation

$$(P_u f)(t,x) = (P_u f_t)(x), \qquad t,u \in [0,T],$$

is used, where f_t is the map $f_t : \mathbb{G} \to \mathbb{R}$ given by $f_t(x) = f(t, x)$. Discounting at rate $r \geq 0$ can be incorporated by replacing the infinitesimal generator Λ by the sub-generator matrix $\Lambda^{(r)}$ given by

$$\Lambda^{(r)} = \Lambda - r\mathbb{I},$$

where $\mathbb{I} : l^\infty(\mathbb{G}) \to l^\infty(\mathbb{G})$ is the identity map, so that Equation (4.1) generalizes to

$$\mathbf{E}_{t,x}[e^{-r(T-t)}\phi(M_T)] = (\exp((T-t)\Lambda^{(r)})\phi)(x), \qquad x \in \mathbb{G}, t \in [0, T]. \tag{4.2}$$

Remark 4.1.1. The Markov property of the chain M together with the identity in Equation (4.2) imply that the discounted process $\{e^{-rt}M_t, t \in \mathbb{R}_+\}$ is a martingale precisely if

$$\mathbf{E}_{0,x}[e^{-rt}M_t] = x \quad \text{for all } x \in \mathbb{G}.$$

4.1.2 American options

An American option with payoff function given by ϕ and with maturity $T > 0$, on an underlying with price process given by a Markov chain $M = \{M_t, t \in [0, T]\}$, is a derivative security that entitles its holder to receive the payoff $\phi(M_t)$ at any time t prior to the maturity T that she wishes to exercise the contract. The most common type of American options are the American call option with strike K, which has payoff $\phi(s) = (s - K)^+$ (with $x^+ = \max\{x, 0\}$ for $x \in \mathbb{R}$), and the American put option with strike K, which has payoff given by $\phi(s) = (K - s)^+$. Assume that the payoff function $\phi : \mathbb{G} \to \mathbb{R}_+$ is non-negative and satisfies the integrability condition

$$\mathbf{E}_{0,x}\left[\sup_{t\in[0,T]} \phi(M_t)\right] < \infty, \qquad x \in \mathbb{G}. \tag{4.3}$$

The value V_t^* of the American option at time $t \in [0,T]$ with payoff function ϕ is given by

$$V_t^* = \operatorname*{ess.\,sup}_{\tau \in \mathcal{T}_{t,T}} \mathbf{E}[e^{-r\tau}\phi(M_\tau)|\mathcal{G}_t],$$

where $\mathcal{T}_{t,T}$ denotes the set of $\mathbf{G}$-stopping times taking values between t and T. The process $V^* = \{V_t^*, t \in [0,T]\}$ is called the *Snell-envelope* of the collection of discounted payoffs $\Pi = \{e^{-rt}\phi(M_t), t \in [0,T]\}$: it is the smallest $\mathbf{G}$-supermartingale that is bounded below by Π. The Markov property of M implies $V_t^* = V(t, M_t)$ where the value function of the American option $V = \{V(t,x), t \in [0,T], x \in \mathbb{G}\}$ is given by

$$V(t,x) = \sup_{\tau \in \mathcal{T}_{t,T}} \mathbf{E}_{t,x}\left[e^{-r(\tau-t)}\phi(M_\tau)\right] \tag{4.4}$$

$$= \sup_{\tau \in \mathcal{T}_{0,T-t}} \mathbf{E}_{0,x}\left[e^{-r\tau}\phi(M_\tau)\right], \qquad (t,x) \in [0,T] \times \mathbb{G}, \tag{4.5}$$

where the second line is a consequence of the homogeneity of the Markov process M.

4.1.3 Dynkin's Lemma

In the sequel the following version of Dynkin's lemma will be frequently deployed in the analysis.

Lemma 4.1.1 (Dynkin)**.** *Assume that the function $F : [0,T] \times \mathbb{G} \to \mathbb{R}$ is bounded and that, for any $x \in \mathbb{G}$, the map $t \to F(t,x)$ is absolutely continuous with density $f(t,x)$ that is non-negative for almost every $t \in [0,T]$. Then it holds that for any $t \in [0,T]$ and any $\mathbf{G}$-stopping time τ taking values in $[t,T]$*

$$\mathbf{E}_{t,x}\left[e^{-r(\tau-t)}F(\tau, M_\tau)\right] = F(t,x) + \mathbf{E}_{t,x}\left[\int_t^\tau e^{-r(s-t)}(\overline{\Lambda}F)(s, M_s)\mathrm{d}s\right] \tag{4.6}$$

with the map $\overline{\Lambda}F : [0,T] \times \mathbb{G} \to \mathbb{R}$ defined by

$$(\overline{\Lambda}F)(t,x) = f(t,x) + (\Lambda^{(r)}F)(t,x), \qquad t \in [0,T], x \in \mathbb{G} \tag{4.7}$$

Proof. Assume first in addition that $F(t,x)$ is continuously differentiable in time for every $x \in \mathbb{G}$. An application of Itô's lemma to the semi-martingale $\{e^{-rt}F(t,M_t)\}_{t\in[0,T]}$ shows that the process $\{N_t\}_{t\in[0,T]}$ with

$$N_t = e^{-rt}F(t,M_t) - F(0,M_0) - \int_0^t e^{-rs}\left[\frac{\partial F}{\partial t} + (\Lambda F - rF)\right](s,M_s)\mathrm{d}s$$

is a local martingale. In view of the assumptions on F and Λ it follows that N is in fact a uniformly integrable martingale. An application of Doob's Optional Stopping Theorem implies that for every stopping time $\tau \in \mathcal{T}_{t,T}$ the equality $\mathbf{E}_{t,x}[N_\tau] = 0$ holds so that Equation (4.6) holds.

Assume next that F is as stated in the Lemma, with density f. Since the set $\mathcal{H}$ of functions $H : [0,T] \times \mathbb{G} \to \mathbb{R}$ that is continuously differentiable at $t \in [0,T]$ for every $x \in \mathbb{G}$ is dense in the set of continuous real-valued functions with domain $[0,T] \times \mathbb{G}$, there exists a sequence of functions $(H_n)_n$ in $\mathcal{H}$ that almost everywhere converges to F. An application of Lebesgue's Dominated Convergence Theorem, which is justified by the facts that F is bounded and Λ has uniformly bounded diagonal, shows that Equation (4.6) holds under the stated assumptions. □

4.2 American Options under the Markov chain model

This section develops conditions on the value function of the American option when the underlying is a Markov chain. These are then used to present two methods to value American options: One based on an explicit characterisation of the optimal boundary, and the other on a direct application of the Dynamic Programming Principle.

According to the general theory of optimal stopping (see Peskir and Shiryaev [65]), the solution of the optimal stopping problem in Equation (4.4) is expressed in terms of a stopping region $\mathbb{S}$ and a continuation region $\mathbb{C}$ given by

$$\mathbb{S} = \{(s,x) \in [0,T] \times \mathbb{G} : V(s,x) = \phi(x)\},$$
$$\mathbb{C} = \{(s,x) \in [0,T] \times \mathbb{G} : V(s,x) > \phi(x)\}.$$

In particular, $\tau_{\mathbb{S}}(t)$ given by

$$\tau_{\mathbb{S}}(t) = \inf\{s \in [t,T] : M_s \in \mathbb{S}\}$$

is a **G**-stopping time in the set $\mathcal{T}_{t,T}$ that achieves the supremum in Equation (4.4). By combining with the strong Markov property of M it follows

$$\{e^{-r(t\wedge\tau)}V(t\wedge\tau, M_{t\wedge\tau}), t \in [0,T]\} \qquad \text{is a martingale for } \tau = \tau_{\mathbb{S}}(0). \tag{4.8}$$

$\mathbb{S}$ can be decomposed as follows

$$\mathbb{S} = \bigcup_{x\in\mathbb{G}} \mathbb{S}(x) \times \{x\}, \qquad \mathbb{S}(x) = \{s \in [0,T] : V(s,x) = \phi(x)\}.$$

The following lemma provides a useful property that the generator matrix of a continuous-time Markov chain satisfies.

Lemma 4.2.1. *Let f be a non-negative function, $\forall x \in \mathbb{G}$, $f(x) \geq 0$. Then for any $g_i \in \mathbb{G}$ such that*

$$f(g_i) = 0$$

it holds that

$$(\Lambda^{(r)} f)(g_i) \geq 0$$

where $\Lambda^{(r)}$ is a (sub-) generator matrix.

Proof. Since $\Lambda^{(r)}$ is a (sub-) generator matrix

$$\Lambda(g_i, g_j) \geq 0 \quad \forall i \neq j$$
$$\Lambda(g_i, g_i) \leq 0 \quad \forall i$$

holds. Therefore the only negative value in row $\Lambda^{(r)}(i,:)$ of the row-column vector multiplication $(\Lambda^{(r)} f)(g_i)$ is cancelled by the corresponding value in f being zero. Therefore the lemma holds. □

In the following result two properties of the value function and its generator are recorded that will be used later:

Proposition 4.2.2. *The following hold for the value function V:*

(i) For each $x \in \mathbb{G}$, the map $t \mapsto V(t,x)$ is decreasing and continuous.

(ii) For each $x \in \mathbb{G}$, the map $\Lambda^{(r)} : [0,T] \to \mathbb{R}$ given by $t \mapsto [\Lambda^{(r)} f_t](x)$ with $f_t(x) = V(t,x)$ is continuous and is decreasing when restricted to $\mathbb{S}(x)$.

Proof. (i) Since for any $s, t \in [0,T]$ with $t < s$ the set of stopping times $\mathcal{T}_{0,T-t}$ and $\mathcal{T}_{0,T-s}$ satisfy $\mathcal{T}_{0,T-t} \supseteq \mathcal{T}_{0,T-s}$, it follows from the representation in Equation (4.5) that $V(t,x) \geq V(s,x)$ holds for each $x \in \mathbb{G}$. Lebesgue's Dominated Convergence Theorem, the fact that ϕ satisfies the integrability condition in Equation (4.3) and the triangle inequality imply that $V(t,x)$ is continuous as a function of t, for any fixed $x \in \mathbb{G}$.

For any $t_1, t_2 \in [0,T]$, $t_2 \geq t_1$, and $g \in \mathbb{G}$ such that (t_1, g) and (t_2, g) are elements of $\mathbb{S}$, the function $f : \mathbb{G} \to \mathbb{R}$ given by $f(x) = V(t_1, x) - V(t_2, x)$ satisfies the conditions of Lemma 4.2.1, by virtue of the facts that $t \mapsto V(t,x)$ is decreasing (by part (i)) and that $V(t_1, g) = V(t_2, g) = \phi(g)$ holds (by the definition of $\mathbb{S}$). Hence $(\Lambda^{(r)} f)(g)$ is nonnegative, which shows the stated monotonicity.

Since $(\Lambda^{(r)} V)(t,h) = \sum_{g \in \mathbb{G}} \Lambda^{(r)}(h,g) V(t,g)$ holds for each $h \in \mathbb{G}$, it follows from part (i) that $t \mapsto (\Lambda^{(r)} V)(t,h)$ is a finite sum of continuous functions and is therefore also continuous. $\square$

The monotonicity of $t \mapsto V(t,x)$ stated in Proposition 4.2.2(i) implies that if a point (t,x) lies in $\mathbb{S}$ then also any point of the form (s,x) for $s > t$ lies in $\mathbb{S}$. Thus, since $t \mapsto V(t,x)$ is continuous, the set $\mathbb{S}(x)$ is closed and is of the form

$$\mathbb{S}(x) = [\tau(x), T] \qquad \text{for some } \tau(x) \in [0,T].$$

Let the set $\{t_i\}$ be the increasing sequence of distinct times in the set $\{\tau(x)\}_{x\in\mathbb{G}} \cup \{T\}$. Denoting the first element of the set $\{t_i\}$ by t_1, it follows that $t_1 = 0$. Furthermore, it holds $\max\{t_i\} = T$. The set $\{t_i\}$ can be used to partition the interval $[0,T]$ as

$$\begin{aligned} [0,T) &= \bigcup_i T_i \\ T_i &= [t_i, t_{i+1}). \end{aligned}$$

By construction $\mathbb{S}$ is constant on each set T_i, let

$$\mathbb{S}_{T_i} = \{x \in \mathbb{G} : t \in T_i \Rightarrow (t,x) \in \mathbb{S}\}$$

be the optimal stopping set on the interval T_i. The stopping times

$$\tau_{T_i}(t) = \inf\{s \geq t : M_s \in \mathbb{S}_{T_i}\}$$

are the first hitting times of these sets. Note that for any $t \in T_i$

$$\tau_{\mathbb{S}}(t) \in T_i \Rightarrow \tau_{\mathbb{S}}(t) = \tau_{T_i}(t). \tag{4.9}$$

As the stopped discounted price process is a martingale, for any $t \in T_i$ Equation (4.9) implies

$$V(t,x) = \mathbf{E}_{t,x}\left[e^{-r(\tau_{T_i}(t)\wedge t_{i+1}-t)} V\left(\tau_{T_i}(t) \wedge t_{i+1}, M_{\tau_{T_i}(t)\wedge t_{i+1}}\right)\right].$$

This shows that within each T_i the contract can be seen as a knock-out option with maturity t_{i+1}, knock-out set $\mathbb{S}_{T_i}$, rebate ϕ and payoff given by the value function at t_{i+1}, which is itself also given by a knock-out option unless $t_{i+1} = T$ in which case $V(T,x) = \phi(x)$.

Associated to the value function of the American option is the system of variational inequalities given by

$$(\Lambda_t V)(t,x) \leq 0 \qquad \text{for } (t,x) \in [0,T] \times \mathbb{G}, \tag{4.10}$$

$$(\Lambda_t V)(t,x) = 0 \qquad \text{for } (t,x) \in \mathbb{C}, \tag{4.11}$$

$$V(t,x) = \phi(x) \qquad \text{for } (t,x) \in \mathbb{S}, \tag{4.12}$$

$$V(t,x) > \phi(x) \qquad \text{for } (t,x) \in \mathbb{C}, \tag{4.13}$$

where Λ_t denotes the infinitesimal generator of the time-space process (t, M_t), which acts on functions F in the set $C^1([0,T] \times \mathbb{G})$ [the set of functions $F : [0,T] \times \mathbb{G} \to \mathbb{R}$ that are continuously differentiable as function of the first argument], as follows:

$$\Lambda_t F = \frac{\partial F}{\partial t} + \Lambda^{(r)} F.$$

Since a priori the value function V is only known to be continuous and decreasing as function of t, V may not be a classical solution of the system in Equations (4.10)—(4.13) of variational inequalities. A function $V : [0,T] \times \mathbb{G} \to \mathbb{R}$ is called a solution in *distributional sense* of the system in Equations (4.10)—(4.13) if V satisfies Equations (4.10)—(4.13) with the map $\Lambda_t V$ replaced by the map $\overline{\Lambda} V$ that was defined in Equation (4.7).

The following existence and uniqueness result can now be shown:

Theorem 4.2.3. *The function V defined in* (4.4) *is the unique continuous decreasing function that solves the system of variational inequalities in Equations* (4.10)—(4.13) *in distributional sense.*

Furthermore

$$(\Lambda^{(r)}V)(\tau(x), x) = 0 \qquad \textit{for any } x \in \mathbb{G} \textit{ satisfying } \tau(x) < T, \tag{4.14}$$

$$(\Lambda^{(r)}V)(t, x) \leq 0 \qquad \textit{for any } x \in \mathbb{G} \textit{ and } t \in [0, T] \textit{ with } t > \tau(x) \tag{4.15}$$

holds. In particular, the value function V is a classical solution of the system in Equations (4.10)—(4.13)*.*

Proof. (Existence) That V is decreasing and continuous follows from Proposition 4.2.2. Note that Equations (4.12) and (4.13) hold true by definition of the stopping and continuation regions $\mathbb{S}$ and $\mathbb{C}$. Next Equation (4.10) is shown to hold true. Since $t \mapsto V(t, x)$ is decreasing and continuous, $V(\cdot, x)$ admits a density that is almost everywhere non-positive. For any $x \in \mathbb{G}$ and any $t \in [0, T]$ and any stopping time $\tau \in \mathcal{T}_{t,T}$, by Dynkin's lemma, Lemma 4.1.1,

$$\mathbf{E}_{t,x}\left[e^{-r(\tau-t)}V(\tau, M_\tau)\right] = V(t, x) + \mathbf{E}_{t,x}\left[\int_t^\tau e^{-r(s-t)}(\overline{\Lambda}_t V)(s, M_s)\mathrm{d}s\right] \tag{4.16}$$

holds, where $\overline{\Lambda}_t V$ is defined in Equation (4.7). As the discounted value-process $e^{-rt}V(t, M_t)$ is a supermartingale, for any pair $t_1, t_2 \in [0, T]$ with $t_1 < t_2$ and any $x \in \mathbb{G}$ the inequality $\mathbf{E}_{t_1,x}\left[e^{-rt_2}V(t_2, M_{t_2})\right] \leq e^{-rt_1}V(t_1, x)$ holds. This inequality yields in view of Equation (4.16) the relation

$$B(t_1, t_2, x) := \mathbf{E}_{t_1,x}\left[\int_{t_1}^{t_2} e^{-r(s-t_1)}(\overline{\Lambda}_t V)(s, M_s)\mathrm{d}s\right] \leq 0. \tag{4.17}$$

To see that Equation 4.17 implies that Equation (4.10) is satisfied (in distributional sense), note that the left-hand side of Equation (4.17) is equal to

$$B(t_1, t_2, x) = \sum_{y \in \mathbb{G}} \int_{t_1}^{t_2} e^{-r(s-t_1)}(\overline{\Lambda}_t V)(s, y)\mathbb{P}_{x,t_1}(M_s = y)\mathrm{d}s.$$

Since $\mathbb{P}_{x,t_1}(M_s \neq x) = -\Lambda(x, x)(s - t_1) + o(t_2 - t_1)$ $(t_2 \searrow t_1)$ holds for all $s \leq t_2$ (as M is a continuous-time Markov chain), it follows that $(\overline{\Lambda}_t V)(s, y)$ is non-positive for almost every $t \in [0,T]$ and for all $y \in \mathbb{G}$. Thus, the claim follows from Equation (4.17).

Finally, it is checked that Equation (4.11) is satisfied. Since the stopped process $e^{-r(t\wedge\tau_{\mathbb{S}})}V(t\wedge\tau_{\mathbb{S}}, M_{t\wedge\tau_{\mathbb{S}}})$ is a $\mathbb{P}_{t,x}$-martingale for any $(t,x) \in \mathbb{C}$ (see Equation (4.8)), it follows that for any $t_1, t_2 \in [0,T]$ with $t_1 < t_2$

$$\mathbf{E}_{t_1,x}\left[e^{-r(t_2\wedge\tau_{\mathbb{S}})}V(t_2\wedge\tau_{\mathbb{S}}, M_{t_2\wedge\tau_{\mathbb{S}}})\right] = \mathbf{E}_{t_1,x}\left[e^{-r(t_1\wedge\tau_{\mathbb{S}})}V(t_1\wedge\tau_{\mathbb{S}}, M_{t_1\wedge\tau_{\mathbb{S}}})\right] = e^{-rt_1}V(t_1,x)$$

holds. So that, in view of the equality in Equation (4.16), the equality

$$\mathbf{E}_{t_1,x}\left[\int_{t_1}^{t_2\wedge\tau_{\mathbb{S}}} e^{-r(s-t_1)}(\overline{\Lambda}_t V)(s, M_s)\mathrm{d}s\right] = 0$$

follows. A line of reasoning that is similar to the one used in the previous paragraph shows $(\overline{\Lambda}_t V)(t,x) = 0$ for almost every $t \in [0,T]$ and every $x \in \mathbb{G}$ with $(t,x) \in \mathbb{C}$, so that Equation (4.11) holds (in distributional sense).

(Uniqueness) Assume that the function $\widetilde{V}$ is a continuous decreasing function that solves in the distributional sense (4.10)—(4.13). An application of Equation (4.13) implies

$$\mathbf{E}_{t,x}\left[e^{-r(\tau-t)}\phi(M_\tau)\right] \leq \mathbf{E}_{t,x}\left[e^{-r(\tau-t)}\widetilde{V}(\tau, M_\tau)\right] \tag{4.18}$$

for any $\tau \leq \tau_{\mathbb{S}}(t)$. By applying Dynkin's lemma to the right hand side of Equation (4.18) and using Equation (4.11)

$$\mathbf{E}_{t,x}\left[e^{-r(\tau-t)}\widetilde{V}(\tau, M_\tau)\right] = \widetilde{V}(t,x).$$

Hence, choosing $\tau = \tau_{\mathbb{S}}(t)$ in Equation (4.18) and using (4.12)

$$V(t,x) = \mathbf{E}_{t,x}\left[e^{-r(\tau_{\mathbb{S}}(t)-t)}\phi(M_{\tau_{\mathbb{S}}(t)})\right] = \mathbf{E}_{t,x}\left[e^{-r(\tau_{\mathbb{S}}(t)-t)}\widetilde{V}(\tau_{\mathbb{S}}(t), M_{\tau_{\mathbb{S}}(t)})\right] = \widetilde{V}(t,x)$$

holds. Which shows that the solution of the system of Equations (4.10)—(4.13) is unique.

(Equations (4.14) and (4.15)) Since $t \mapsto V(t,x)$ is decreasing (Proposition 4.2.2), it follows that $V(\cdot,x)$ admits a density that is non-positive for almost every $t \in [0,T]$ and any $x \in \mathbb{G}$ with $(t,x) \in \mathbb{C}$. Hence, in combination with the equality in Equation (4.11) and the continuity of $t \mapsto (\Lambda^{(r)}V)(t,x)$

$$(\Lambda^{(r)}V)(t,x) \geq 0 \qquad \text{for all } (t,x) \in \mathbb{C}.$$

holds. Since the map $t \mapsto V(t,x)$ restricted to the interval $\mathbb{S}(x) = [\tau(x), T]$ is constant equal to $\phi(x)$, the density of $V(\cdot,x)$ is equal to zero for almost every $t \in [\tau(x),T]$ and $x \in \mathbb{G}$ for which $\tau(x)$ is strictly smaller than T. Thus, in view of the relation in Equation (4.10) and the continuity of the map $t \mapsto (\Lambda^{(r)}V)(t,x)$

$$0 \geq (\Lambda^{(r)}V)(t,x) \qquad \text{for any } t \in [\tau(x),T] \text{ and } x \in \mathbb{G} \text{ with } \tau(x) < T.$$

holds. Since the map $t \mapsto (\Lambda^{(r)}V)(t,x)$ is continuous, non-negative for $t < \tau(x)$ and non-positive for $t > \tau(x)$, the intermediate value theorem implies $(\Lambda^{(r)}V)(\tau(x),x)$ is equal to zero, and the proof of Equations (4.14) and (4.15) is complete.

(Classical solution) To show that V is a classical solution it suffices to show that at every t in $[0,T]$ and x in $\mathbb{G}$ the map $t \mapsto V(t,x)$ is continuously differentiable. Noting that the function V restricted to the interval (t_i, t_{i+1}) is a knock-out barrier contract with knock-out set $\mathbb{S}_{T_i}$, it follows that V is continuously differentiable at every t in (t_i, t_{i+1}) with derivative given by

$$\frac{\partial V}{\partial t}(t,x) = -(\widetilde{\Lambda}_r^{(i)}V)(t,x), \qquad t \in (t_i, t_{i+1})$$

where $\widetilde{\Lambda}_r^{(i)}$ is defined in Section 4.A.2 on valuing barrier options under Markov chains as the generator of the chain stopped at entering $\mathbb{S}_{T_i}$. Note that

$$(\widetilde{\Lambda}_r^{(i)}V)(t,x) = (\Lambda^{(r)}V)(t,x), \quad t \in T_i \text{ and } (t,x) \in \mathbb{C}. \tag{4.19}$$

To complete the proof of the continuous differentiability of V consider the case $t = t_i$.

Due to the construction of $\widetilde{\Lambda}_r^{(i)}$, for any $x \notin \mathbb{S}_{T_i} \backslash \mathbb{S}_{T_{i-1}}$

$$(\widetilde{\Lambda}_r^{(i-1)} V)(t,x) = (\widetilde{\Lambda}_r^{(i)} V)(t,x)$$

holds, where $\mathbb{S}_{T_i} \backslash \mathbb{S}_{T_{i-1}}$ is any element in $\mathbb{S}_{T_i}$ that is not an element of $\mathbb{S}_{T_{i-1}}$. For any $x \in \mathbb{S}_{T_i} \backslash \mathbb{S}_{T_{i-1}}$ it follows that $\tau(x) = t_i$. Note that on $\mathbb{S}$, as the value function V is constant in time, it holds $\frac{\partial V}{\partial t}(t,x) = 0$. Therefore the right limit of $\frac{\partial V}{\partial t}(t_i,x)$ is zero. The left limit of $\frac{\partial V}{\partial t}(t_i,x)$ is, by Equation (4.19), equal to $(\Lambda^{(r)} V)(t_i,x)$ which, in view of Equation (4.14), is also equal to zero. Thus, the function V is continuously differentiable at all $t \in [0,T]$ and $x \in \mathbb{G}$ and the proof is complete. □

4.2.1 Free boundary method

In this section a characterisation of the stopping region $\mathbb{S}$ is presented. To simplify the presentation the following assumption will be made throughout this section:

Assumption 4.2.1. *The stopping region is of the form*

$$\mathbb{S} = \{(t,x) \in [0,T] \times \mathbb{G} : x \leq B(t)\},$$

where the optimal boundary $t \mapsto B(t)$ is increasing as a function of time t with $B(T)$ taking a finite value.

Observe that if the sequences $\mathcal{X} = \{x_1, x_2, \ldots\}$ and $\{\tau(x_1), \tau(x_2), \ldots\}$ are non-decreasing, then the optimal boundary is given by $B(t) = \sup\{x_i \in \mathcal{X} : t \in [\tau(x_i), T]\}$. This form of the optimal boundary is for example encountered in the case of an American put option under a continuous-time Markov chain model that is spatially homogeneous (see Figure 4.1).

Denote by

$$\mathbb{B} = \{B(\tau(x)), x \in \mathbb{G}\} = \{b_i\}_i, \qquad b_i < b_{i+1},$$

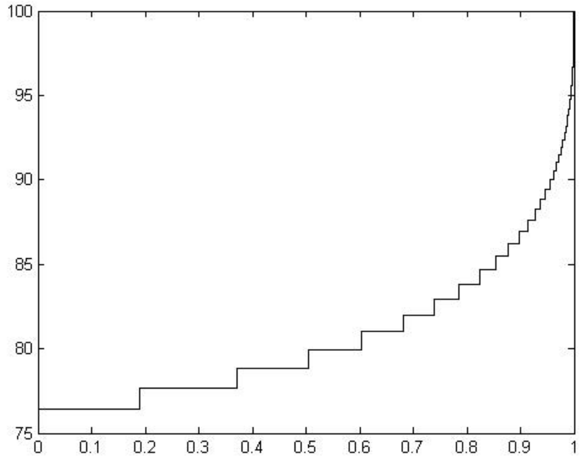

Figure 4.1: The optimal boundary corresponding to an at-the-money American put option with strike $S_0 = K = 100$ and maturity $T = 1$ when interest rate and dividend yield are given by $r = 0.1$ and $\delta = 0$ and the underlying is given by a Markov chain that closely approximates a geometric Brownian motion with volatility $\sigma = 0.3$. The chain has a state space of size 200 and was constructed by matching the instantaneous moments of the Markov chain with those of the Brownian motion, using the procedure described in Mijatović and Pistorius [62].

the set of distinct elements in $\{B(\tau(x)), x \in \mathbb{G}\}$ that the optimal boundary takes (in increasing order). In this setting the knock-out sets are of the form

$$\mathbb{S}_{T_i} = \{x \in \mathbb{G} : x \leq b_i\}.$$

In line with this the stopping times $\tau_{T_i}(t)$ are given by

$$\tau_{T_i}(t) = \tau_{b_i}(t) = \inf\{s \geq t : M_s \leq b_i\}.$$

At this point it is good to take note of the facts that (i) the sequence $\{t_i\}_i$ is increasing and (ii) the boundary B is constant in between the epochs t_i and has a discontinuity at the epochs t_i. The connection between the times t_i, the barrier levels b_i, time intervals T_i and the stopping sets $\mathbb{S}_{T_i}$ are is illustrated in Figure 4.2. Given the times t_i and the optimal barrier levels b_i the American option can be valued recursively: The value function V of the American option is equal to the value function of a barrier

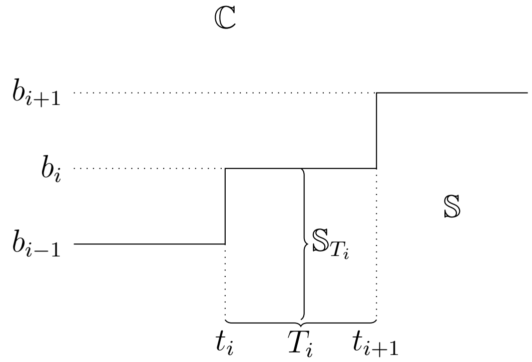

Figure 4.2: Illustration of the connection between the terms b_i, t_i, T_i and $\mathbb{S}_{T_i}$.

option contract with time-dependent barrier B that entitles the holder to a rebate payment $\phi(M_{\tau_B})$ if the epoch $\tau_B = \inf\{t \geq 0 : M_t \leq B(t)\}$ is strictly smaller than T and to a payment $\phi(M_T)$ in the case that the epoch τ_B is larger or equal to T. From the Markov property of M applied at the epochs t_i that the barrier B has jumps it follows that the function V is equal to the final value V_1 of the following recursion:

$$\begin{aligned} V_i(t,x) &= \mathbf{E}_{t,x}\Big[e^{-r(\tau_{b_i}(t)-t)}\phi(M_{\tau_{b_i}(t)})\mathbf{I}(\tau_{b_i}(t) < t_{i+1}) \\ &\qquad + e^{-r(t_{i+1}-t)}V_{i+1}(t_{i+1}, M_{t_{i+1}})\mathbf{I}(\tau_{b_i}(t) > t_{i+1})\Big] \\ &= \mathbf{E}_{t,x}\Big[e^{-r(\tau_{b_i}(t)\wedge t_{i+1}-t)}V_{i+1}(t_{i+1}, M_{\tau_{b_i}(t)\wedge t_{i+1}})\Big] \quad t \in [0, t_{i+1}] \qquad (4.20) \\ V_{E+1}(t,x) &= \phi(x) \quad t \in [0,T] \end{aligned}$$

for all $i = 1, \ldots, E$ and $x \in \mathbb{G}$. Note that

$$\begin{aligned} &V(t,x) = V_i(t,x) = \phi(x) \text{ for any pair } (x,t) \text{ with } x \in \mathbb{S} \text{ and } t \leq t_{i+1} \\ &V(t,x) = V_i(t,x) \qquad \text{for any } t \in T_i \text{ and } x \in \mathbb{G}. \end{aligned}$$

Thus, the optimal value function V is equal to V_i on the time interval T_i. (Also note in passing that the value of $V_i(t,x)$ for $t < t_i$ with (t,x) in the continuation region $\mathbb{C}$ is in general *not* equal to $V(t,x)$.)

Next the collection of epochs $\{t_i\}_i$ will be characterise in terms of the value of the time-space generator Λ_t applied to the functions V_i.

This require the following auxiliary result:

Lemma 4.2.4. *For any $i \in \mathbb{N}$ with $b_i \in \mathbb{B}$ and any $x \in \mathbb{G}$, the function $V_i(\cdot, x) : [0, t_{i+1}] \to \mathbb{R}$ given by $t \mapsto V_i(t, x)$ is decreasing and continuous. As a consequence, the function $(\Lambda^{(r)} V_i)(\cdot, b_i) : [0, t_{i+1}] \to \mathbb{R}$ given by $t \mapsto (\Lambda^{(r)} V_i)(t, b_i)$ is continuous and decreasing on $[0, t_{i+1}]$.*

Proof. Let $x \in \mathbb{G}$ and i with $b_i \in \mathbb{B}$ be arbitrary and given. The function $t \mapsto V_i(t, x)$ restricted to the interval T_i is equal to the function $t \mapsto V(t, x)$, which was shown to be decreasing in Proposition 4.2.2. Next turn to the case $t \leq t_i$. Note that for any $t \in [0, t_{i+1}]$

$$\begin{aligned} V_i(t, x) &= \mathbf{E}_{t,x} \left[e^{-r(\tau_{b_i}(t) \wedge (t_{i+1} - t))} V_{i+1} \left(t_{i+1}, M_{\tau_{b_i}(t) \wedge (t_{i+1} - t)} \right) \right] \\ &= \left(e^{(t_{i+1} - t)\tilde{\Lambda}_r^{(i)}} V_{i+1} \right) (t_{i+1}, x) \end{aligned} \tag{4.21}$$

holds. Where $\widetilde{\Lambda}_r^{(i)}$ is the (sub-)generator matrix of M stopped upon first entrance into the set $\mathbb{S}_{T_i}$, which is specified in Equation (4.49) in the appendix.

Thus, for any $t, s \in [0, t_{i+1}]$ with $t < s$

$$V_i(t, x) - V_i(s, x) = \left(e^{(t_{i+1} - s)\tilde{\Lambda}_r^{(i)}} \left(e^{(s-t)\tilde{\Lambda}_r^{(i)}} - \mathbf{I} \right) V_{i+1} \right) (t_{i+1}, x) \tag{4.22}$$

holds. Since $t \mapsto V_i(t, x)$ is decreasing for $t \in T_i$ deduce from Equations (4.21) and (4.22)

$$V_i(t, x) - V_i(t_{i+1}, x) = \left[\left(\exp[(t_{i+1} - t)\widetilde{\Lambda}_r^{(i)}] - I \right) V_{i+1} \right] (t_{i+1}, x) \geq 0 \tag{4.23}$$

for any $t \in T_i$ and $x \in \mathbb{G}$. In view of Equations (4.22) and (4.23) it follows

$$V_i(t, x) - V_i(s, x) \geq 0 \text{ for any } s, t \in [0, t_{i+1}] \text{ with } s - t \in [0, t_{i+1} - t_i]. \tag{4.24}$$

As the difference $t_{i+1} - t_i$ is strictly positive, the statement in Equation (4.24) implies $V_i(t, x) - V_i(s, x) \geq 0$ for any $s, t \in [0, t_{i+1}]$ with $t \leq s$. The proof of the monotonicity of $V_i(\cdot, x)$ is complete.

The continuity of $t \mapsto V_i(t,x)$ for any $x \in \mathbb{G}$ follows from the form of the value function of a barrier option (see Theorem 4.A.1). The continuity of $t \mapsto (\Lambda^{(r)}V_i)(t, b_i)$ follows since $(\Lambda^{(r)}V_i)(t, b_i)$ is equal to a linear combination of $V_i(t,x)$ over different $x \in \mathbb{G}$.

By applying Lemma 4.2.1 to $f(x) = V_i(t,x) - V_i(s,x)$ (noting that $V_i(t, b_i) = \phi(b_i)$ for any $t \in [0, t_{i+1}]$), it follows that the monotonicity of $V_i(\cdot, x)$ implies the monotonicity of $t \mapsto (\Lambda^{(r)}V_i)(t, b_i)$ on the interval $[0, t_{i+1}]$. $\square$

Theorem 4.2.5. *Let V_i be defined by* (4.20)*. For any $i \in \mathbb{N}$ with $b_i \in \mathbb{B}$ and $t_i < T$, it holds*

$$(\Lambda_t V_i)(t,x) = 0 \quad \textit{for } x > b_i,\ t \in [0, t_{i+1}), \tag{4.25}$$

$$(\Lambda_t V_i)(t,x) = (\Lambda^{(r)}V_i)(t,x) = 0 \quad \textit{for } x = b_i,\ t = t_i, \tag{4.26}$$

$$(\Lambda_t V_i)(t,x) = (\Lambda^{(r)}V_i)(t,x) \le 0 \quad \textit{for } x \le b_i,\ t \in T_i. \tag{4.27}$$

$$(\Lambda_t V_i)(t,x) = (\Lambda^{(r)}V_i)(t,x) > 0 \quad \textit{for } x = b_i,\ t < t_i, \tag{4.28}$$

Proof. Since $V_i(t,x) = V(t,x)$ for $t \in T_i$, Equations (4.26) and (4.27) hold in view of Theorem 4.2.3.

The function V_i is the value function of a down-and-out barrier option with maturity t_{i+1} rebate $\phi(x)$ and terminal payoff function $V_{i+1}(t_{i+1}, x)$. Since the process $e^{-r(t\wedge t_{i+1}\wedge \tau_{b_i}(t))}V_i(t \wedge t_{i+1} \wedge \tau_{b_i}(t), M_{t\wedge t_{i+1}\wedge \tau_{b_i}(t)})$ is a martingale, it follows by an analogous reasoning as the one that was used in the proof of Theorem 4.2.3 that $(\Lambda_t V_i)(t,x) = 0$ for $x > b_i$ and $t < t_{i+1}$. Hence, Equation (4.25) holds true.

Finally, turn to the proof of Equation (4.28). Start by observing that $(\Lambda^{(r)}V_i)(t, b_i)$ is non-negative on the interval $t \in [0, t_i]$ in view of Lemma 4.2.4 and Equation (4.26). Next that $(\Lambda^{(r)}V_i)(t, b_i)$ is in fact strictly positive on the interval $[0, t_i)$ is shown.

By an application of Dynkin's lemma, Lemma 4.1.1,

$$\begin{aligned} V_i(t,x) - V_{i-1}(t,x) = \mathbf{E}_{t,x}\left[e^{-r(\tau-t)}\left\{V_i(\tau, M_\tau) - V_{i-1}(\tau, M_\tau)\right\}\right] \\ - \mathbf{E}_{t,x}\left[\int_t^\tau e^{-r(s-t)}\{(\Lambda_t V_i)(s,M_s) - (\Lambda_t V_{i-1})(s,M_s)\}\mathrm{d}s\right] \end{aligned} \tag{4.29}$$

for all $x \in \mathbb{G}$, $t \leq t_i$ and $\tau \in \mathcal{T}_{t,t_i}$. Since by Equation (4.25)

$$(\Lambda_t V_i)(s,x) = 0 \qquad \text{for any } x > b_i,\ s \in [0, t_{i+1}) \text{ and any } i \in \mathbb{N} \text{ with } b_i \in \mathbb{B},$$

and the collection $\{b_i\}_i$ is increasing, choosing in Equation (4.29) τ to be equal to

$$\tau_i = \min\{\tau_{b_{i-1}}(t), t_i\}$$

shows that the right-most expectation in Equation (4.29) is equal to

$$\begin{aligned} &\mathbf{E}_{t,x}\left[\int_t^{\tau_i} e^{-r(s-t)}\{(\Lambda_t V_i)(s,M_s) - (\Lambda_t V_{i-1})(s,M_s)\}\mathrm{d}s\right] \\ &\quad = \mathbf{E}_{t,x}\left[\int_t^{\tau_i} e^{-r(s-t)}(\Lambda_t V_i)(s,M_s)\mathbf{I}_{\{M_s=b_i\}}\mathrm{d}s\right]. \end{aligned} \tag{4.30}$$

Furthermore, $V_i(\tau_i, M_{\tau_i}) = V_{i-1}(\tau_i, M_{\tau_i})$ holds for the following two reasons: (a) it holds $V_{i-1}(t_i, M_{t_i}) = V_i(t_i, M_{t_i})$ by definition of V_{i-1} and (b) on the set $\{\tau_{b_{i-1}}(t) < t_i\}$

$$V_i\left(\tau_{b_{i-1}}(t), M_{\tau_{b_{i-1}}(t)}\right) = V_{i-1}\left(\tau_{b_{i-1}}(t), M_{\tau_{b_{i-1}}(t)}\right) = \phi\left(M_{\tau_{b_{i-1}}(t)}\right)$$

follow as $M_{\tau_{b_{i-1}}(t)} \leq b_{i-1} < b_i$ holds by the definition of $\tau_{b_{i-1}}(t)$ and the fact that b_i is increasing as a function of i. Hence identity

$$\mathbf{E}_{t,x}\left[e^{-r(\tau_i-t)}V_i(\tau_i, M_{\tau_i})\right] = \mathbf{E}_{t,x}\left[e^{-r(\tau_i-t)}V_{i-1}(\tau_i, M_{\tau_i})\right] \tag{4.31}$$

is deduced. Combining Equations (4.29), (4.30) and (4.31) shows

$$V_i(t,x) - V_{i-1}(t,x) = \mathbf{E}_{t,x}\left[\int_t^{\tau_i} e^{-r(s-t)}(\Lambda_t V_i)(s,M_s)\mathbf{I}_{\{M_s=b_i\}}\mathrm{d}s\right]. \tag{4.32}$$

On the one hand, the construction of the value functions $\{V_i\}$ and the definition of the collection $\{b_i\}$ imply

$$V_{i-1}(t, b_i) > \phi(b_i) = V_i(t, b_i), \qquad t \in [0, t_i), \tag{4.33}$$

while, on the other hand, the equality $V_i(t, b_i) = \phi(b_i)$ for all $t \in [0, t_{i+1}]$ implies $\partial V_i(t, b_i)/\partial t = 0$ for $t \in (0, t_{i+1})$ so that

$$(\Lambda_t V_i)(t, b_i) = (\Lambda^{(r)} V_i)(t, b_i) \qquad t \in (0, t_{i+1}) \tag{4.34}$$

holds. Thus, from Equations (4.32), (4.33) and (4.34),

$$\mathbf{E}_{t,b_i}\left[\int_t^{\tau_i} e^{-r(s-t)}(\Lambda^{(r)} V_i)(s, b_i)\mathbf{I}_{\{M_s=b_i\}}\mathrm{d}s\right] > 0, \qquad \text{for any } t \in [0, t_i] \tag{4.35}$$

is deduced. Since the map $t \mapsto (\Lambda^{(r)} V_i)(t, b_i)$ is continuous, decreasing and non-negative on the interval $[0, t_i]$, a contradiction argument shows that $(\Lambda^{(r)} V_i)(t, b_i) > 0$ for any $t \in [0, t_i)$. $\square$

The characterisation of the free boundary given in Theorem 4.2.5 can be deployed to compute the optimal boundary and the corresponding value of an American option under the Markov chain model. For the presentation of a valuation algorithm restrict the set of Markov chains to those with a finite state space (of size N, say).

To identify the epochs $\{t_i\}$ a numerical method has to be deployed since the equations

$$(\Lambda_t V_i)(t, b_i) = 0$$

are highly non-linear in t. Except in degenerate cases, one may expect the map $s \mapsto (\Lambda_t V_i)(s, b_i)$ to be strictly decreasing, in which case the equation $(\Lambda_t V_i)(t, b_i) = 0$ admits a unique solution and it is efficient to use a solver such as the Newton-Raphson method (which is the method that was used in the examples in Section 4.4). A procedure for computation of the value function of an American option under a Markov

Algorithm 4.1: Markov chain free boundary algorithm

find the index i of the largest grid point $x_i \in \mathbb{G}$ such that $(\Lambda^{(r)}\phi)(x_i) < 0$
set $t^* \leftarrow T$
while $t^* > 0$
    find $s < t^*$ such that $\left(\overline{\Lambda}_r^{(i)}\left[\exp\left((t^*-s)\widetilde{\Lambda}_r^{(i)}\right)\phi\right]\right)(x_i) = 0$;
    if $s > 0$
        set $\phi \leftarrow \exp\left((t^*-s)\widetilde{\Lambda}_r^{(i)}\right)\phi$;
    else if $s \leq 0$
        set $\phi \leftarrow \exp\left(t^*\widetilde{\Lambda}_r^{(i)}\right)\phi$;
    set $i \leftarrow i-1$; set $t^* \leftarrow s$;
end
return ϕ

chain model based on a solution of the corresponding free boundary problem that was outlined in the previous paragraph is described in Algorithm 4.1. In Algorithm 4.1 the notation $\overline{\Lambda}_r^{(i)}$ is used for the generator that satisfy

$$\widetilde{\Lambda}_r^{(i)} + \overline{\Lambda}_r^{(i)} = \Lambda^{(r)}$$

where $\widetilde{\Lambda}_r^{(i)}$ is the generator of the chain stopped on entering the set $\mathbb{S}_{T_i}$, as described in Appendix 4.A.2.

Remark 4.2.1. Algorithm 4.1 is based on the observation that, in view of the definition of the matrix $\overline{\Lambda}_r^{(i)}$ and the relation $\frac{\mathrm{d}}{\mathrm{d}t}\exp(tA) = A\exp(tA)$ that holds for any square matrix A the following equality holds:

$$\left(\Lambda_t \exp\left((t^*-t)\widetilde{\Lambda}_r^{(i)}\right)\right)\Big|_{t=s} = O \Leftrightarrow \Lambda^{(r)}\exp\left((t^*-s)\widetilde{\Lambda}_r^{(i)}\right) = \widetilde{\Lambda}_r^{(i)}\exp\left((t^*-s)\widetilde{\Lambda}_r^{(i)}\right),$$

where O denotes a zero matrix of appropriate size.

Remark 4.2.2. The above algorithm concerns the case of an American option with monotone increasing exercise boundary. By obvious modifications a similar algorithm can be phrased for the case of an American option with monotone decreasing exercise boundary. Furthermore, in view of the local nature of the characterisation it is not unreasonable to expect that the algorithm may be extended to the case of American options with more complex forms of the stopping regions. Detailed investigations of those extensions are left for future research.

4.2.2 Dynamic programming method

A Bermudan option, with payoff function ϕ and finite set of admissible exercise times $\mathbb{T} \subset [0, T]$, is a derivative security that may be exercised at any time $\tau \in \mathbb{T}$ yielding payoff $\phi(M_\tau)$. For the ease of presentation the restriction to the case of an equidistant grid given by

$$\mathbb{T} = \{i\Delta : i = 0, \dots, M\} \quad \text{with} \qquad \Delta = T/M,$$

is made. The value $V(t, x)$ of the Bermudan option at time $t \in \mathbb{T}$ in case of $\{M_t = x\}$ is given by

$$V(t, x) = \max_{\tau \in \mathcal{T}_{t,T}^M} \mathbf{E}_{t,x}[e^{-r(\tau - t)}\phi(M_\tau)],$$

for $t \in \mathbb{T}$, and $x \in \mathbb{G}$, where $\mathcal{T}_{t,T}^M$ is the set of $\mathbf{G}$-stopping times τ taking values in $[t, T] \cap \mathbb{T}$, where $\mathbf{G} = \{\mathcal{G}_t, t \in [0, T]\}$ denotes the filtration generated by the Markov chain M. At any time $t \in \mathbb{T}$, the holder of the Bermudan option has the choice between immediately exercising or continuing to wait. The former results in a payoff of $\phi(M_t)$, while in the latter case, the expected reward of postponing exercise, assuming that the holder continues to follow an optimal strategy from time t to maturity, is $\mathbf{E}_{t,M_t}[e^{-r\Delta}V(t + \Delta, M_\Delta)]$. Thus, for any $t \in \mathbb{T}$, the value $V(t, x)$ is at least equal to the larger of $\phi(x)$ and $\mathbf{E}_{t,x}[e^{-r\Delta}V_{t+\Delta}(M_\Delta)]$. The Dynamic Programming principle states that in fact equality holds: with $V^{(i)}(x) = V(i\Delta, x)$,

$$V^{(i)}(x) = \max\left(\phi(x), \mathbf{E}_{i\Delta,x}\left[e^{-r\Delta}V^{(i+1)}(M_{(i+1)\Delta})\right]\right)$$

Algorithm 4.2: Procedure to compute the value of a Bermudan option

```
set  $\triangle \leftarrow \frac{T}{M}$
set  $V \leftarrow O \in \mathbb{R}^{N\times(M+1)}$
set  $V(\,:\,,\ M+1) \leftarrow \phi(\,:\,)$
evaluate  $A = \exp(\triangle\Lambda^{(r)})$
for  $i = M$  to  1
        $V(\,:\,, i) \leftarrow A[V(\,:\,,\ i+1)]$;
        $V(\,:\,, i) \leftarrow \max(\phi(\,:\,), V(\,:\,,\ i))$;
        $i \leftarrow i-1$;
end
return  $V$
```

holds, for $i = 0, \ldots, M-1$, and $x \in \mathbb{G}$. Noting that in view of the form of the semigroup in Equation (4.1)

$$\mathbf{E}_{t,x}\left[e^{-r\Delta}V^{(i+1)}(M_\Delta)\right] = \left[\exp\left(\Delta\Lambda^{(r)}\right)V^{(i+1)}\right](x).$$

holds. By deploying the Dynamic Programming Principle, the recursive procedure to compute the values of $V^{(i)}(x)$ ranging over all initial values $x \in \mathbb{G}$ and all time-steps $i = 0, \ldots, M$ described in Algorithm 4.2 is obtained.

Remark 4.2.3. (i) The algorithm returns the matrix $(V^{(i)}(x), (i\Delta, x) \in \mathbb{T}\times\mathbb{G})$ of values of the Bermudan option on the time-space grid $\mathbb{T}\times\mathbb{G}$, where $V(\,:\,, i)$ denotes the ith column of the matrix V and contains the values $V^{(i+1)}(x)$ for $x \in \mathbb{G}$.

(ii) Note that when, as assumed above, the time-grid $\mathbb{T}$ is equidistant, the exponentiation of the matrix $\Delta\Lambda$ only needs to be computed once. If the time-grid $\mathbb{T}$ is chosen non-equidistant, the above algorithm will computationally be a good deal more expensive, since a costly exponentiation would need to be carried out at every iteration of the recursive procedure.

4.3 Convergence

In this section it is shown that the convergence of a sequence of Markov chains converging to a limit process S carries over to convergence of the corresponding sequence of values of American options to the value of the American option under the limiting model. The price process S is assumed to be a Markov process with state space $\mathbb{R}_+$ that is defined on some filtered probability space $(\Omega, \mathcal{F}, \mathbf{F}, \mathbb{P})$, where $\mathbf{F} = \{\mathcal{F}_t\}_{t\geq 0}$ denotes the standard filtration generated by S and Ω denotes the Skorokhod space of right-continuous functions with left-hand limits that map $\mathbb{R}_+$ to $\mathbb{R}$. The interest rate and dividend yield are taken to be constant equal to r and d, and it is assumed as before that the discounted price process $\{e^{-\gamma t}S_t\}_{t\geq 0}$ with $\gamma = r - d$ is a martingale. Assume in addition that S is a Feller process that solves the stochastic differential equation given by

$$\frac{\mathrm{d}S_t}{S_{t-}} = \gamma \mathrm{d}t + \sigma(S_{t-})\mathrm{d}W_t + p(\mathrm{d}t \times \mathrm{d}x), \qquad t > 0,$$

with $S_0 = x > 0$, where W denotes a Wiener process and p denotes a compensated random measure with compensator given by the random measure $\nu(S_{t-}, \mathrm{d}z)\mathrm{d}t$, where, for every $x \in \mathbb{R}_+$, $\nu(x, \mathrm{d}y)$ is a measure with support in $(-1, \infty)$ satisfying the integrability condition

$$\int_{(-1,\infty)} y^2 \nu(x, \mathrm{d}y) < \infty.$$

The value function $v : [0, T] \times \mathbb{R}_+ \to \mathbb{R}_+$ of the American option with payoff $\phi : \mathbb{R}_+ \to \mathbb{R}_+$ on the underlying process S is denoted by

$$v(t, x) = \sup_{\tau \in \mathcal{T}_{t,T}(\mathbf{F})} \mathbf{E}_{t,x}\left[e^{-r(\tau - t)}\phi(S_\tau)\right], \quad (t, x) \in [0, T] \times \mathbb{R}_+,$$

with the set $\mathcal{T}_{t,T}(\mathbf{F})$ equal to the collection of $\mathbf{F}$-stopping times taking values in between t and T. The Bermudan option, which is an American-type option for which the epoch of exercise is restricted to take values in the grid $\mathbb{T}$ is a closely related derivative security, with value function $v^M : [0,T] \times \mathbb{R}_+ \to \mathbb{R}_+$ given by

$$v^M(t,x) = \sup_{\tau \in \mathcal{T}^M_{t,T}(\mathbf{F})} \mathbf{E}_{t,x}\left[e^{-r(\tau-t)}\phi(S_\tau)\right], \quad (t,x) \in [0,T] \times \mathbb{R}_+,$$

where $\mathcal{T}^M_{t,T}(\mathbf{F})$ denotes the collection of $\mathbf{F}$-stopping times taking values in the grid $\mathbb{T}$ intersected with the interval $[t,T]$.

Let $M^{(n)}$ denote a sequence of Markov chains, defined on the measurable space $(\Omega, \mathcal{F})$, that weakly converges to the Feller process S, where the weak convergence is in the Skorokhod J_1 topology (see for example Jacod and Shiryaev [41]). Let $V^{(n),M}$ and $V^{(n)}$ denote the value functions of a Bermudan option with M equidistant exercise times and an American option on the Markov chain $M^{(n)}$, respectively. Below it is shown that as n and M tend to infinity then both $V^{(n)}(0,x)$ and $V^{(n),M}(0,x)$ tend to the value $v(x)$ of the American option when the spot S_0 is equal to x. More precisely, assume that the subsequent grids $(\mathbb{G}^{(n)})_{n\in\mathbb{N}}$ are all nested (that is, $\mathbb{G}^{(n)}$ is contained in $\mathbb{G}^{(n+1)}$ for any positive integer n) and that the union $\cup_{n\in\mathbb{N}}\mathbb{G}^{(n)}$ is dense in $\mathbb{R}$, and consider the following convergence of a sequence of functions $f^{(n)} : \mathbb{G}^{(n)} \to \mathbb{R}$ to a function $f : \mathbb{R} \to \mathbb{R}$:

$$f^{(n)} \xrightarrow{\mathbb{G}} f \Leftrightarrow \forall m \in \mathbb{N},\ \forall x \in \mathbb{G}^{(m)} \lim_{n\to\infty, n\geq m} f^{(n)}(x) = f(x).$$

To establish this convergence it will be required that the quantities $C(x)$ and $\sup_n D(n,x)$ are finite, where

$$C(x) := \mathbf{E}_{0,x}\left[\sup_{t\in[0,T+1]}\left(S^2_{t-}\sigma^2(S_{t-}) + \int_{\mathbb{R}} S^2_{t-} z^2 \nu(S_{t-}, \mathrm{d}z)\right)\right], \tag{4.36}$$

$$D(n,x) := \mathbf{E}_{0,x}\left[\sup_{t\in[0,T+1]}\left(\sum_{z\in\mathbb{J}^{(n)}(M^{(n)}_{t-})} z^2\Lambda^{(n)}\left(M^{(n)}_{t-}, M^{(n)}_{t-}+z\right)\right)\right]. \tag{4.37}$$

Here $\Lambda^{(n)}$ denotes the generator matrix of $M^{(n)}$ and $\mathbb{J}^{(n)}(x) = \mathbb{G}^{(n)} - x$ denotes the set of possible jump-sizes of $M^{(n)}$ starting from site x.

Theorem 4.3.1. *Let $M^{(n)}$ be a sequence of Markov chains that weakly converges to S and assume that ϕ is Lipschitz continuous. Assume that $C(x)$ and $\sup_n D(n,x)$ are finite for any x. The following hold true:*

(i) $V^{(n),M}(0,\cdot) \xrightarrow{\mathbb{G}} v^M(0,\cdot)$, *as $n \to \infty$ for any $M \in \mathbb{N}$.*

(ii) $V^{(n),M}(0,\cdot) \xrightarrow{\mathbb{G}} v(0,\cdot)$ *as* $\min\{n, M\} \to \infty$.

(iii) $V^{(n)}(0,\cdot) \xrightarrow{\mathbb{G}} v(0,\cdot)$ *if $n \to \infty$.*

Proof. First, the following claims are proven: For any x there exist a constant $\widetilde{C}(x)$ such that for all $M \in \mathbb{N}$

$$|v^M(0,x) - v(0,x)| \leq \frac{\widetilde{C}(x)}{\sqrt{M}}. \tag{4.38}$$

For any $n \in \mathbb{N}$ and $x \in \mathbb{G}^{(n)}$ there exist a constant $\widetilde{D}(n,x)$ such that for all $M \in \mathbb{N}$

$$|V^{(n),M}(0,x) - V^{(n)}(0,x)| \leq \frac{\widetilde{D}(n,x)}{\sqrt{M}}. \tag{4.39}$$

This proof is given only in the case when the underlying is given by S as the proof of the case that the underlying is a Markov chain is analogous.

Observe that the collection of stopping times of the form $\tau_M = \inf\{s \geq \tau : s \in \mathbb{T}\}$ for $\tau \in \mathcal{T}_{0,T}$ is equal to the set $\mathcal{T}^M_{0,T}$. By an application of the triangle inequality

$$\begin{aligned}
|v(0,x) - v^M(0,x)| &\leq \sup_\tau \mathbf{E}_{0,x}\left[\left|e^{-r\tau}\phi(S_\tau) - e^{-r\tau_M}\phi(S_{\tau_M})\right|\right] \\
&\leq \sup_\tau \mathbf{E}_{0,x}\left[\left|(e^{-r\tau} - e^{-r\tau_M})\phi(S_\tau)\right| + \left|e^{-r\tau_M}(\phi(S_\tau) - \phi(S_{\tau_M}))\right|\right] \\
&\leq \frac{1}{M}\cdot c(x) + K \cdot \sup_\tau \mathbf{E}_{0,x}\left[|S_\tau - S_{\tau_M}|\right]
\end{aligned}$$

holds, where the suprema are taken over the set $\mathcal{T}_{t,T}$ of stopping times taking values in the interval $[t,T]$, and where the last inequality follows by the triangle inequality and Lipschitz continuity of ϕ, as follows:

$$\sup_{\tau} \mathbf{E}_{0,x}[rTe^{-r\tau} |\phi(S_\tau)|] \leq rT(\phi(x) + 2Kx) := c(x),$$

where K is the Lipschitz constant. By the strong Markov property of S the supremum on the right-hand side can be estimated by

$$\sup_{\tau} \mathbf{E}_{0,x} \left[|S_\tau - S_{\tau_M}| \right] \leq \sup_{\tau} \mathbf{E}_{0,x} \left[\mathbf{E}_{0,S_\tau} \left[|S_0 - S_{\tau_M - \tau}| \right] \right]. \tag{4.40}$$

An application of the triangle-inequality yields the estimate

$$\begin{aligned} &\mathbf{E}_{0,s} \left[|S_0 - S_{\tau_M-\tau}| \right] \qquad (4.41) \\ &\quad\leq \mathbf{E}_{0,s} \left[\left| S_0 - e^{-\gamma(\tau_M-\tau)} S_{\tau_M-\tau} \right| \right] + \mathbf{E}_{0,s} \left[|e^{-\gamma(\tau_M-\tau)} - 1| S_{\tau_M-\tau} \right] := e_1(s) + e_2(s), \end{aligned}$$

for any non-negative s. An application of Doob's Optional Stopping theorem to the càdlàg martingale $e^{-\gamma t} S_t$ implies that $e_2(s)$ can be bounded by

$$\begin{aligned} e_2(s) &= \mathbf{E}_{0,s} \left[\left| 1 - e^{\gamma(\tau_M-\tau)} \right| e^{-\gamma(\tau_M-\tau)} S_{\tau_M-\tau} \right] \\ &\leq \frac{|\gamma T|}{M} \cdot e^{\gamma^+ T/M} \cdot \mathbf{E}_{0,s} \left[e^{-\gamma(\tau_M-\tau)} S_{\tau_M-\tau} \right] \\ &= \frac{|\gamma T|}{M} \cdot e^{\gamma^+ T/M} \cdot s, \end{aligned}$$

where $\gamma^+ = \max(0, \gamma)$. Another application of Doob's Optional Stopping Theorem implies that the following bound holds:

$$\begin{aligned} \mathbf{E}_{0,x} \left[e_2 \left(S_\tau \right) \right] &\leq \frac{|\gamma T|}{M} \cdot e^{\gamma^+ T/M} \cdot e^{\gamma^+ T} \cdot \mathbf{E}_{0,x} \left[e^{-\gamma\tau} S_\tau \right] \\ &= \frac{|\gamma T|}{M} \cdot e^{\gamma^+ T/M} \cdot e^{\gamma^+ T} \cdot x, \; x \in \mathbb{R}_+. \end{aligned} \tag{4.42}$$

By an application of Doob's L^2-inequality to the martingale $M' = \{M'_t = e^{-\gamma t} S_t\}_{t \in [0,T]}$

and an application of Dynkin's lemma

$$
\begin{aligned}
e_1(s) &\leq \mathbf{E}_{0,s}\left[\sup_{t:t<\frac{T}{M}}\left|e^{-\gamma t}S_t - S_0\right|\right] \\
&\leq 2\mathbf{E}_{0,s}\left[\left|e^{-\gamma T/M}S_{T/M} - S_0\right|^2\right]^{1/2} \\
&= 2\left(\mathbf{E}_{0,s}\left[\int_0^{T/M} e^{-2\gamma t}S_{t-}^2\left\{\sigma^2(S_{t-}) + \int_{\mathbb{R}} z^2\nu(S_{t-},\mathrm{d}z)\right\}\mathrm{d}t\right]\right)^{1/2},
\end{aligned}
$$

holds, for $s \in \mathbb{R}_+$. The Cauchy-Schwarz inequality and the strong Markov property yield

$$
\mathbf{E}_{0,x}[e_1(S_\tau)] \leq \mathbf{E}_{0,x}\left[e_1(S_\tau)^2\right]^{1/2} \leq \frac{2T^{1/2}}{M^{1/2}}e^{(-\gamma)^+T/M}\cdot C(x)^{1/2}, \qquad x \in \mathbb{R}_+, \tag{4.43}
$$

where $C(x)$ is defined in Equation (4.36). By combining Equations (4.40), (4.41), (4.42) and (4.43), it follows that Equation (4.38) holds with

$$
\widetilde{C}(x) = 2KT^{1/2}e^{(-\gamma)^+T/M}C(x)^{1/2} + K|\gamma T|e^{\gamma^+T/M}e^{\gamma^+T}\cdot x + c(x).
$$

Next turn to the proof of the three assertions. (i) By extending the probability space if necessary, assume that the processes S and $(M^{(n)})_n$ are all defined on a single probability space.

Denote by $\mathbf{H}$ the filtration generated by the process $\{S, M^{(n)}, n \in \mathbb{N}\}$ and by $\widetilde{\mathcal{T}}_{t,T}^M$ the collection of $\mathbf{H}$-stopping times taking values in the set $[t, T]$ intersected with the grid $\mathbb{T}$. Thus, write

$$
v^M(0,x) = \sup_{\tau\in\widetilde{\mathcal{T}}_{0,T}^M}\mathbf{E}_{0,x}\left[e^{-r\tau}\phi\left(S_\tau\right)\right], \qquad V^{(n),M}(0,x) = \sup_{\tau\in\widetilde{\mathcal{T}}_{0,T}^M}\mathbf{E}_{0,x}\left[e^{-r\tau}\phi\left(M_\tau^{(n)}\right)\right].
$$

for any $x \in \mathbb{G}^{(n)}$.

By the triangle inequality and the Lipschitz continuity of ϕ (with Lipschitz constant K)

$$
\begin{aligned}
&\left|v^M(0,x) - V^{(n),M}(0,x)\right| \\
&\quad \leq \sup_{\tau \in \widetilde{\mathcal{T}}_{0,T}^M} \mathbf{E}_{0,x}\left[e^{-r\tau}\left|\phi\left(M_\tau^{(n)}\right) - \phi\left(S_\tau\right)\right|\right] \leq K \sup_{\tau \in \widetilde{\mathcal{T}}_{0,T}^M} \mathbf{E}_{0,x}\left[\left|S_\tau - M_\tau^{(n)}\right|\right] \\
&\quad \leq K\mathbf{E}_{0,x}\left[\sup_{t \in \mathbb{T}}\left|M_t^{(n)} - S_t\right|\right] \qquad (4.44)
\end{aligned}
$$

holds, where in the last line is is used that any stopping time τ in the set $\widetilde{\mathcal{T}}_{0,T}^M$ takes values in the grid $\mathbb{T}$. As, by assumption, $M^{(n)}$ converges weakly to S in the Skorokhod topology as $n \to \infty$, it follows that $M_t^{(n)}$ converges to S_t in distribution as n tends to infinity, for any fixed $t \in \mathbb{T}$. The Skorokhod representation theorem implies that, for any given $t \in \mathbb{T}$, there exists a probability space carrying random variables $\widetilde{M}_t^{(n)}$, $n \in \mathbb{N}$, and $\widetilde{S}_t$ that have the same distribution as $M_t^{(n)}$ and S_t, respectively, such that $\widetilde{M}_t^{(n)}$ converges a.s. to $\widetilde{S}_t$ as $n \to \infty$. The uniform integrability of the collection $(M_t^{(n)}, S_t, t \in \mathbb{T}, n \in \mathbb{N})$ (which is in turn a consequence of the integrability condition $C(x) + \sup_n D(n,x) < \infty$) thus implies

$$
\forall m \in \mathbb{N}\ \forall x \in \mathbb{G}^{(m)} \lim_{n\to\infty, n \geq m} \mathbf{E}_x\left[\left|S_t - M_t^{(n)}\right|\right] = 0 \qquad \text{for any } t \in \mathbb{T}, \qquad (4.45)
$$

which implies that also the supremum in (4.44) converges to zero as $\mathbb{T}$ contains M elements. The proof of part (i) is completed by combining Equations (4.44) and (4.45).

(ii), (iii) The triangle inequality implies that the differences between $V^{(n)}(0,x)$ and $v(0,x)$ and $V^{(n),M}(0,x)$ and $v(0,x)$ can be estimated as

$$
\begin{aligned}
\left|V^{(n)}(0,x) - v(0,x)\right| &\leq \left|V^{(n)}(0,x) - V^{(n),M}(0,x)\right| + \left|V^{(n),M}(0,x) - v(0,x)\right| \\
\left|V^{(n),M}(0,x) - v(0,x)\right| &\leq \left|V^{(n),M}(0,x) - v^M(0,x)\right| + \left|v^M(0,x) - v(0,x)\right|. \qquad (4.46)
\end{aligned}
$$

Let $\epsilon > 0$ be arbitrary. By virtue of Equations (4.38) and (4.39) and the fact that $\sup_n D(n,x)$ is finite it follows that there exists an M_ϵ such that, for all $M \geq M_\epsilon$ and for all $n \in \mathbb{N}$,

$$\max\left\{\left|v^M(0,x) - v(0,x)\right|, \sup_{n\in\mathbb{N}}\left|V^{(n)}(0,x) - V^{(n),M}(0,x)\right|\right\} \leq \epsilon. \tag{4.47}$$

Fixing an M larger than M_ϵ, part (i) implies that there exists an N_ϵ such that

$$\left|V^{(n),M}(0,x) - v^M(0,x)\right| \leq \epsilon \text{ for all } n \geq N_\epsilon$$

holds. Combining this estimate with Equations (4.46) and (4.47) yields the estimates $|V^{(n)}(0,x) - v(0,x)| \leq 3\epsilon$ and $|V^{(n)}(0,x) - V^{(n),M}(0,x)| \leq 2\epsilon$. Since ϵ was arbitrary the statements in (ii) and (iii) follow. $\square$

4.4 Numerical examples

To provide an illustration of the effectiveness of the methods, in this section the results of the approximation of the value of the American put option by the free boundary approach (Algorithm 4.1, which shall be referred to as 'FB') and dynamic programming approach (Algorithm 4.2, which shall be referred to as 'DP') are reported and compared. The algorithm for the pricing of American options takes as input a Markov chain M that closely approximates the Feller process S which is constructed by suitably specifying its state space and generator matrix: the state space will be taken non-uniform with higher density in relevant areas (for example around the spot value S_0 and the strike K, in the case of a put option) and the generator matrix is chosen so as to match the first two instantaneous moments of S. The smallest and largest points of the state space are taken sufficiently small and large respectively to guarantee that the truncation error is negligible at the level of accuracy that is considered in the examples below (these levels were determined after some numerical experimentation). Along these lines an algorithm for the construction of a Markov chain was developed in Mijatović and Pistorius [62] (see Appendix 4.A.1) which will be deployed in the numerical illustrations below.

4.4.1 The American put option under GBM model

In a first benchmark example the classical Geometric Brownian Motion (GBM) will be studied.The GBM is a (time-homogeneous) diffusion process that evolves according to the SDE

$$\mathrm{d}S_t = S_t((r-d)\mathrm{d}t + \sigma \mathrm{d}W_t) \quad t \in (0,T], \quad S_0 = s > 0,$$

where W is a Brownian motion, σ is the volatility and r, d represent as before the interest rate and dividend yield. This model has an infinitesimal generator that acts on $f \in \mathbb{C}_c^2$ as

$$\mathcal{L}f = (r-d)xf'(x) + \frac{\sigma^2}{2}x^2 f''(x).$$

An approximating Markov chain can be generated by deploying the diffusion version of the general method described in Appendix 4.A The approximating Markov chain is then used to value the American put option

$$v = \sup_{\tau \in \mathcal{T}_{0,T}} \mathbf{E}_{0,x}\left[e^{-r\tau}(K - M_\tau)^+\right].$$

In Table 4.1 the values calculated by the free boundary method (FB) and Dynamic programming method (DP) are reported together with the results given in Carr [19], he states results with both the randomisation method and for a binomial tree method. In Figure 4.3 the absolute error of the DP method, as the number of exercise times (M) is increased, is given for a fixed size Markov chain. The slope of the line in Figure 4.3 is approximately -1, which corresponds to a linear decay of the error of the dynamic programming method (DP) in $1/M$ where M is the number of time-steps.

The change in execution times when varying the number of exercise times is very small. One explanation for this small change is that the bulk of the computation effort is in calculating the matrix exponential $\exp(\Delta\Lambda)$, and it appears that the time to calculate $\exp(\Delta\Lambda)$ is only marginally affected by the size of Δ, and decreasing Δ

	Size	$N = 200$	$N = 400$	$N = 800$
DP	$M = 3200$	8.3316	8.3359	8.3370
DP	$M = 6400$	8.3318	8.3361	8.3371
FB		8.3320	8.3363	8.3373
CR				8.3371
Binomial	2000	Time	Steps	8.3378

Table 4.1: Value of the at the money American put option with strike $S_0 = K = 100$ and maturity $T = 1$. The underlying is a GBM with parameter values taken from Carr [19] ($r = 0.1$, $\delta = 0$, $\sigma = 0.3$). The row CR refers to the randomisation algorithm of Carr [19] with 15 randomisation steps (using Richardson's extrapolation). The row Binomial refers to the outcome of a binomial tree algorithm with 2000 times steps given in Carr [19]. For the DP and FB methods the size parameter N is the size of the Markov chain and M is the number of exercise points in the DP method.

often results in slightly faster calculations.

Figure 4.4 shows the absolute error, for varying sizes of the states space of the Markov chain, for the FB and DP methods with a fixed number of exercise times. The outcomes of the FB method appear to converge slightly faster than those of the DP method, but at the expense of longer execution times. Figure 4.5 shows the execution times, for varying sizes of the states space of the Markov chain, for the FB and DP methods with a fixed number of exercise times. Figure 4.4 appears to show a quadratic speed of convergence in $1/N$ with N the cardinality of the state space $\mathbb{G}$.

4.4.2 The American put option under the CEV model

The constant elasticity of variance (CEV) model, introduced in Cox [25], is a generalisation of GBM. In the CEV model the volatility is a taken to be a power function. The CEV model evolves according to the SDE

$$\mathrm{d}S_t = S_t \left((r - d)\mathrm{d}t + \sigma (S_t/S_0)^\beta \mathrm{d}W_t \right).$$

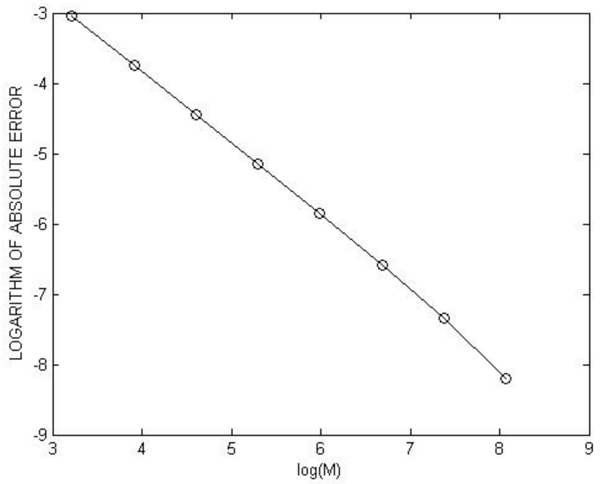

Figure 4.3: Displayed is the absolute error of the Dynamic programming method for a varying number of exercise times M for a fixed size ($N = 1600$) Markov chain. The underlying is a geometric Brownian motion with parameters $r = 0.1$, $\delta = 0$, $\sigma = 0.3$. The option parameters are fixed at $T = 1$, $K = 100$, $S0 = 100$. The reference value used to calculate the errors is the outcome of the method using $M = 12800$ exercise times.

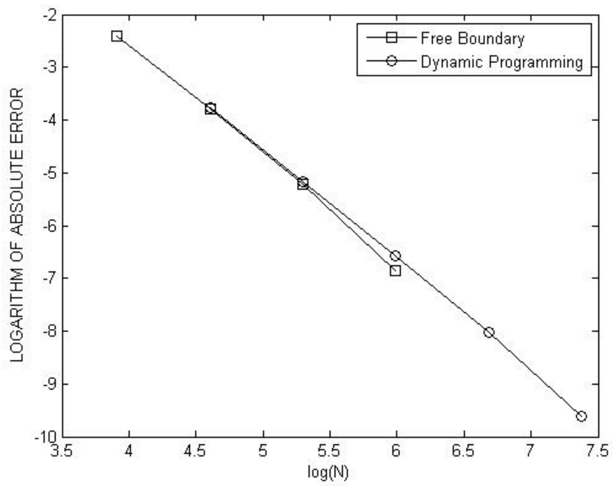

Figure 4.4: Displayed is the absolute error of the American put option for varying size Markov chain N for the free boundary method and the dynamical programming method with $M = 1600$ exercise times. The Geometric Brownian motion with parameters $r = 0.1$, $\delta = 0$, $\sigma = 0.3$ and option parameters $T = 1$,$K = 100$, $S0 = 100$ are used. The reference value is the outcome of the respective method using a larger Markov chain, for DP $N = 3200$ and for FB $N = 800$.

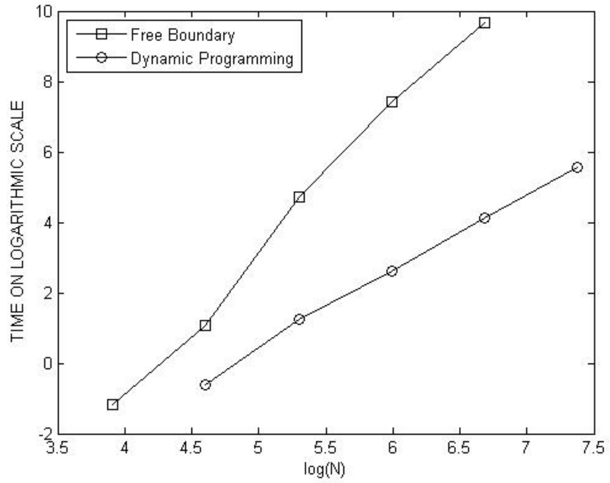

Figure 4.5: Displayed is the execution time, for varying size Markov chain N, of the free boundary method and the dynamical programming method with $M = 1600$ exercise times. The Markov chain approximates a Geometric Brownian motion with parameters $r = 0.1$, $\delta = 0$, $\sigma = 0.3$. The option parameters are $T = 1$, $K = 100$, $S0 = 100$.

where W_t is a Brownian motion. The resulting infinitesimal generator is

$$\mathcal{L}f = (r - d)xf'(x) + \frac{\sigma^2}{2}\left(\frac{x}{S_0}\right)^{2\beta} x^2 f''(x).$$

In Table 4.2 the results calculated using the DP and FB methods are reported together with those given in Wong and Zhao [78] using a finite difference method with an artificial boundary. Figure 4.6 shows the absolute error of the dynamic programming method for a varying number of exercise times, it appears to show a slightly better than linear convergence in $1/M$ where M is the number of exercise times. Just as for the GBM case, it is observed that the change in execution time when varying the number of exercise times is very small.

Figure 4.7 show the absolute error and Figure 4.8 the execution time for both methods for varying size Markov chains. The DP method is evaluated using 3200 exercise points. The slopes indicate a slightly better than quadratic convergence in $1/N$ where N is the cardinality of the state space.

	Size	$N = 400$	$N = 600$	$N = 800$
DP	$M = 1600$	4.6488	4.6490	4.6491
DP	$M = 3200$	4.6488	4.6491	4.6491
FB		4.6489	4.6491	4.6492
WZ				4.6489
Binomial				4.6491

Table 4.2: Numerical results for the American put option using the CEV model with model parameters $r = 0.05$, $q = 0$, $\sigma = 0.2$ and $\beta = -1/3$, and option parameters $S_0 = 100$, $K = 100$, $T = 0.5$. The parameters are taken from Wong and Zhao [78] and the WZ and Binomial values are retrieved from thier paper. The row WZ refers to results obtained using a finite difference scheem and the row Binomial refers to a binomial tree algorithm with 5000 time steps. For the Markov chain methods Size denots the size of the Markov chain.

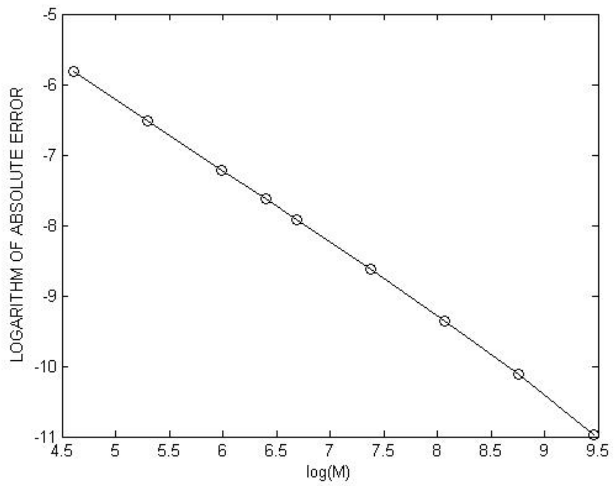

Figure 4.6: Using the CEV model the figure shows the absolute error with a varying number of exercise times M, for a problem with a fixed size Markov chain ($N = 1600$). The model parameters are $r = 0.05$, $q = 0$, $\sigma = 0.2$ and $\beta = -1/3$, and the option parameters $S_0 = 100$, $K = 100$, $T = 0.5$. The exactvalue is calculated using the same size Markov chain and $M = 51200$ exercise times.

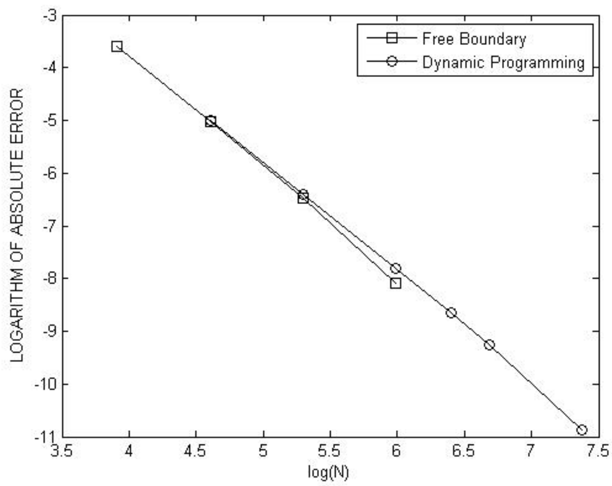

Figure 4.7: For the CEV model, the figure shows the absolute error for varying size Markov chains N for the free boundary method and the dynamic programming method with $M = 3200$ exercise times. The model parameters are $r = 0.05$, $q = 0$, $\sigma = 0.2$ and $\beta = -1/3$, and the option parameters $S_0 = 100$, $K = 100$, $T = 0.5$. The exactvalue is calculated using a Markov chain of size $N = 3200$ for DP and $N = 800$ for FB. The slope of the lines are slightly less than -2 showing an observed quadratic convergence.

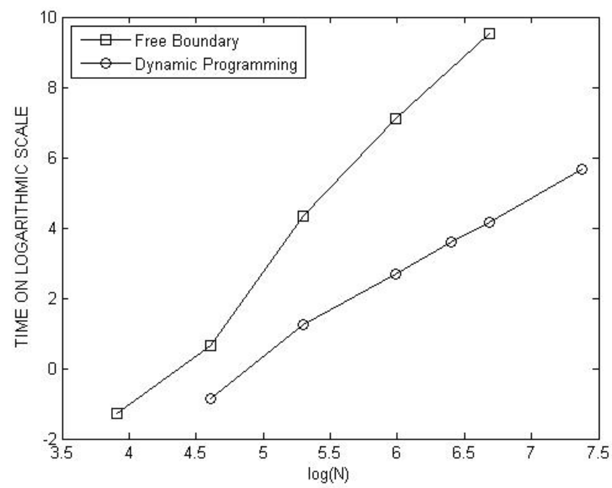

Figure 4.8: For the CEV model, the figure shows the execution time for varying size Markov chains N for the free boundary method and the dynamic programming method with $M = 3200$ exercise times. The model parameters are $r = 0.05$, $q = 0$, $\sigma = 0.2$ and $\beta = -1/3$, and the option parameters $S_0 = 100$, $K = 100$, $T = 0.5$.

4.4.3 The American put option under the Kou model

The Kou model, introduced in Kou [46], is a jump diffusion model. The process evolves according to the SDE

$$\frac{\mathrm{d}S_t}{S_t} = (r - d - \lambda\xi)\mathrm{d}t + \sigma \mathrm{d}W_t + \mathrm{d}L_t$$
$$L_t = \sum_{i=1}^{N_t} \left(e^{K_i} - 1\right)$$

where W_t is a Brownian motion and N_t is a Poison process with intensity $\lambda > 0$. The processes W and N and the collection of random variables $\{K_i\}_i$ are assumed to be mutually independent. The random variables K_i are distributed according to the double exponential density

$$f_K(k) = p\lambda_p e^{-\lambda_p k}\mathbf{I}(k \in (0,\infty)) + (1-p)\lambda_m e^{\lambda_m k}\mathbf{I}(k \in (-\infty, 0)),$$

with $\lambda_p > 0$, $\lambda_m > 0$ and $p \in [0,1]$. The parameter ξ is given by

$$\xi = \mathbf{E}\left[e^{K_1} - 1\right] = \frac{p\lambda_p}{\lambda_p - 1} + \frac{(1-p)\lambda_m}{\lambda_m + 1} - 1.$$

In Table 4.3 the outcomes of the two Markov chain methods are compared with those given in Kou and Wang [47, Table:3]. Note that, although the results are reported in Kou and Wang [47] for an interest rate equal to $r = 0.05$ the values calculated by the FB and DP methods match theirs by using $r = 0.06$. Perhaps this is due to a misprint in Kou and Wang [47].

Figures 4.9, 4.10 and 4.11 show the error plots and execution times for the Kou model. The exact value used to calculate the errors are taken from a large sized execution of the respective method. The slopes of the error plots indicate a linear convergence in the number of exercise times and quadratic convergence in the size of the Markov chain.

K	λ	λ_p	λ_m	FB	DP	Kou Bin.	Kou Approx
90	3	50	25	2.6709	2.6707	2.66	2.72
90	3	50	50	2.4568	2.4566	2.46	2.51
90	7	25	50	3.2282	3.2280	3.24	3.29
90	7	50	50	2.6662	2.6660	2.66	2.72
100	3	50	25	6.2700	6.2698	6.26	6.29
100	3	50	50	6.0120	6.0118	6.01	6.03
100	7	25	50	7.0524	7.0522	7.07	7.09
100	7	50	50	6.2891	6.2889	6.28	6.31
110	3	50	25	12.0559	12.0557	12.04	12.00
110	3	50	50	11.8442	11.8440	11.84	11.78
110	7	25	50	12.8296	12.8294	12.85	12.79
110	7	50	50	12.0928	12.0926	12.08	12.03

Table 4.3: Displayed are American put option prices under the Kou model. The final two columns are obtained from Kou and Wang [47, Table:3]. In all cases it is assumed that the spot is $S_0 = 100$, the maturity is $T = 1$, the interest rate is $r = 0.06$, the volatility is $\sigma = 0.2$ and the probability of an upward jump is $p = 0.6$, with the remaining parameters as given in the table. A Markov chain with state space of size $N = 400$ is employed, and for the dynamical programming algorithm $M = 3200$ exercise times is used.

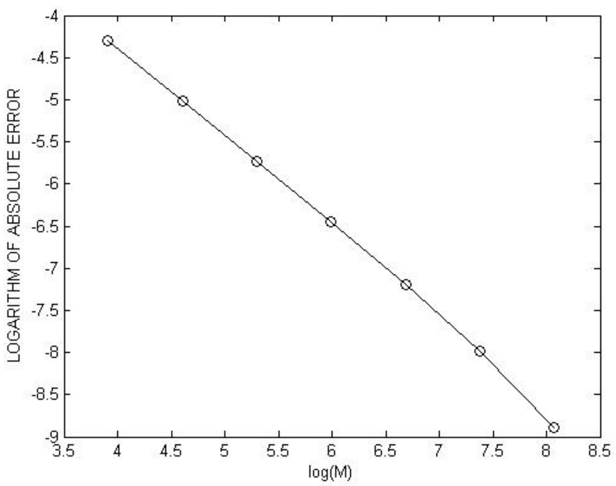

Figure 4.9: Displayed is the absolute error of the value of the American put option generated by the DP method for varying a number of exercise times M using a fixed size Markov chain. The underlying evolves according to the Kou model with the following parameters $r = 0.06$, $\sigma = 0.2$, $p = 0.6$, $\lambda = 7$, $\lambda_p = 50$ and $\lambda_m = 50$. The option parameters are $S_0 = 100$, $K = 100$ and $T = 1$. The referenve value is calculated using $M = 10000$ exercise times.

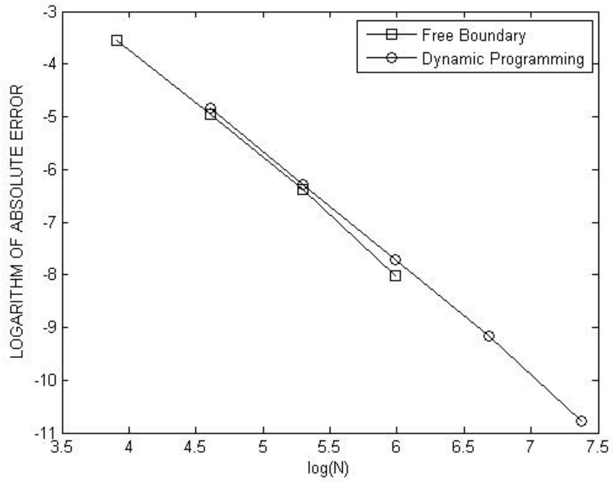

Figure 4.10: Displayed is the absolute error of the value of the American put option for varying size Markov chains N for the free boundary and the dynamic programming ($M = 3200$ exercise times) methods. The underlying evolves according to the Kou model with the following parameters $r = 0.06$, $\sigma = 0.2$, $p = 0.6$, $\lambda = 7$, $\lambda_p = 50$ and $\lambda_m = 50$. The option parameters are $S_0 = 100$, $K = 100$ and $T = 1$. The reference values are calculated using a Markov chain of size $N = 3200$ for DP and size $N = 800$ for FB

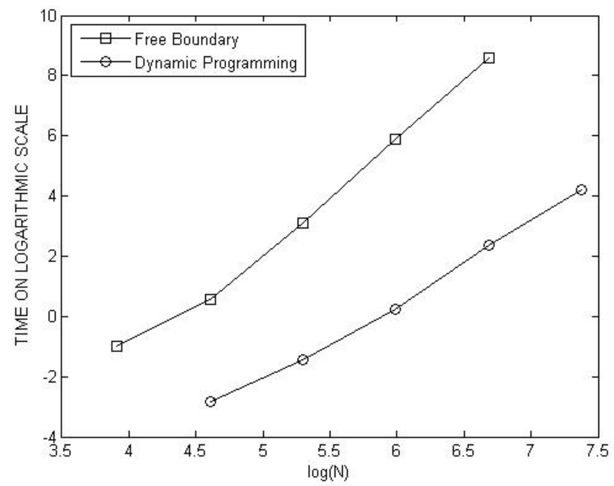

Figure 4.11: Displayed is the execution time of the calculation of the value of the American put option obtained by the free boundary and the dynamic programming ($M = 3200$ exercise times) method with varying size Markov chains N. The underlying evolves according to the Kou model with the following parameters $r = 0.06$, $\sigma = 0.2$, $p = 0.6$, $\lambda = 7$, $\lambda_p = 50$ and $\lambda_m = 50$. The option parameters are $S_0 = 100$, $K = 100$ and $T = 1$.

4.4.4 The American put option under the CEV-Kou model

In this section a local volatility jump diffusion is studied that will be referred to as the CEV-Kou model. Under this model the price process S follows the SDE

$$\begin{aligned}\frac{\mathrm{d}S_t}{S_t} &= \left(r - d - \lambda\xi(S_t/S_0)^\beta\right)\mathrm{d}t + (S_t/S_0)^\beta \mathrm{d}L_t\\ L_t &= \sigma W_t + \sum_{i=1}^{N_t}\left(e^{K_i} - 1\right)\end{aligned}$$

where W is a Brownian motion, N a Poisson process and the K_i are independent random variables following a double exponential distribution, given by

$$f_K(k) = p\lambda_p e^{-\lambda_p k}\mathbf{I}(k \in (0,\infty)) + (1-p)\lambda_m e^{\lambda_m k}\mathbf{I}(k \in (-\infty, 0))$$

with $\lambda_p > 0, \lambda_m > 0$ and $p \in [0,1]$. The parameter ξ is given by

$$\xi = \mathbf{E}\left[e^{K_1} - 1\right] = \frac{p\lambda_p}{\lambda_p - 1} + \frac{(1-p)\lambda_m}{\lambda_m + 1} - 1.$$

The processes W and N and the collection of random variables $\{K_i, i \in \mathbb{N}\}$ are assumed to be mutually independent.

The model under consideration is a combination of the Kou model proposed in Kou [46], which is obtained by setting $\beta = 0$ and is a geometric Lévy process with double exponential jumps, and the constant elasticity of variance model put forward in Cox [25], which is obtained by taking $\lambda = 0$ and is a diffusion with local volatility function given by a power.

The results of the Markov chain approximation algorithms using the FB and DP methods are reported in Table 4.4.

Figures 4.12, 4.13 and 4.14 show the convergence and execution plots. The slope in Figure 4.12 indicate linear convergence in the number of exercise times for DP. For the size of the Markov chain the slopes in Figure 4.13 indicate quadratic convergence

	Size	$N = 200$	$N = 400$
	$\beta = -1$		
DP	$M = 3200$	6.6926	6.6957
DP	$M = 6400$	6.6926	6.6958
FB		6.6927	6.6958
	$\beta = -3$		
DP	$M = 3200$	6.6576	6.6609
DP	$M = 6400$	6.6577	6.6610
FB		6.6578	6.6611

Table 4.4: The value of the American put option when the underlying is the CEV-Kou model using model parameters $r = 0.05$, $\sigma = 0.2$, $p = 0.3$, $\lambda_p = 50$, $\lambda_m = 25$ and $\lambda = 3$, β is given in the table. The option parameters are used $K = 100$, $S_0 = 100$ and $T = 1$.

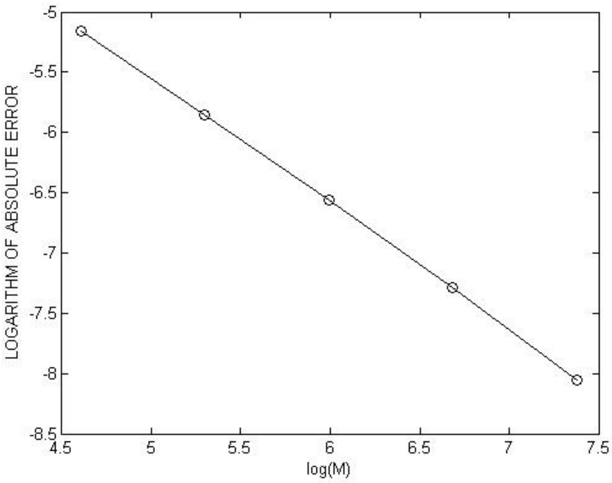

Figure 4.12: Displayed is the absolute error of the American put option values generated by the DP method for varying number of exercise times M using a Markov chain of fixed size ($N = 1600$). The Markov chain is an approximation of the CEV-Kou model with model parameters $r = 0.05$, $\sigma = 0.2$, $\beta = -1$, $p = 0.3$, $\lambda_p = 50$, $\lambda_m = 25$ and $\lambda = 3$. The option parameters are $K = 100$, $S_0 = 100$ and $T = 1$. As reference value is taken the outcome of the DP method with $M = 12800$ exercise times.

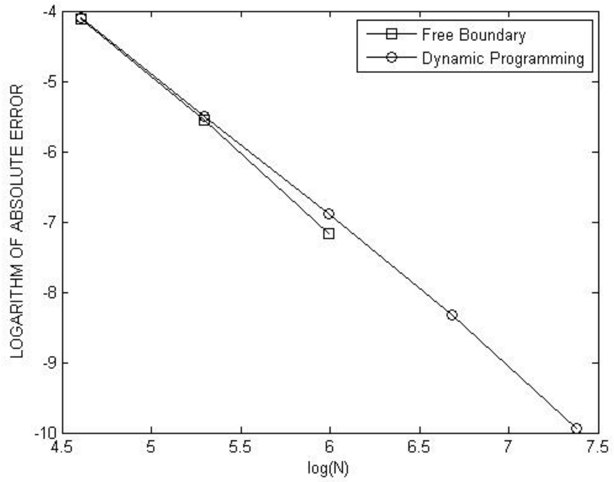

Figure 4.13: Displayed is the absolute error of the American put option values when varying the size of the Markov chain N for the FB and DP Markov chain methods. For the DP method $M = 6400$ exercise times was used. The Markov chain is an approximation of the CEV-Kou model with model parameters $r = 0.05$, $\sigma = 0.2$, $\beta = -1$, $p = 0.3$, $\lambda_p = 50$, $\lambda_m = 25$ and $\lambda = 3$. The option parameters are $K = 100$, $S_0 = 100$ and $T = 1$. The references values are calculated using a Markov chain of size $N = 3200$ for the DP method and $N = 800$ for the FB method.

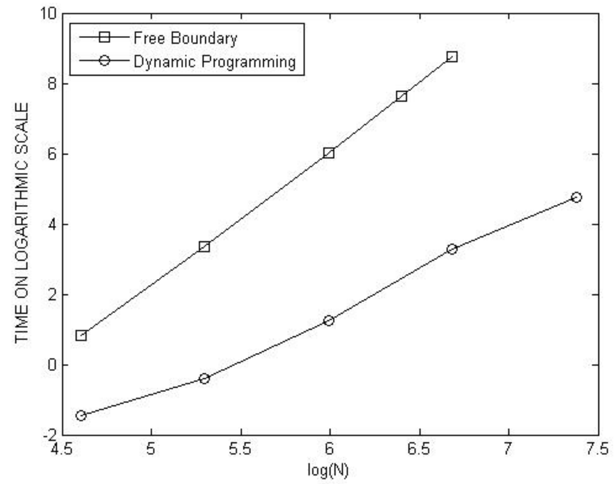

Figure 4.14: Displayed is the execution time for varying size Markov chains N for the free boundary and the dynamic programming ($M = 6400$ exercise times) methods. The underlying follows a CEV Kou model with model parameters $r = 0.05$, $\sigma = 0.2$, $\beta = -1$, $p = 0.3$, $\lambda_p = 50$, $\lambda_m = 25$ and $\lambda = 3$. The option parameters are $K = 100$, $S_0 = 100$ and $T = 1$.

for both methods.

4.4.5 The American put option under the CGMY model

The CGMY and closely related KoBoL models were developed independently in Carr et al. [22] and Boyarchenko and Levendorskii [16]. This model is an exponential Lévy model with infinite activity jumps of the form

$$S_t = se^{(r-d)t}\frac{e^{L_t}}{\mathbf{E}_0[e^{L_t}]}$$

where L has the Lévy density

$$k(y) = C\left(\mathbf{I}(y \in (-\infty, 0))\frac{e^{-G|y|}}{|y|^{Y+1}} + \mathbf{I}(y \in (0, \infty))\frac{e^{-My}}{|y|^{Y+1}}\right),$$

with C, G, $M > 0$ and $0 \leq Y < 2$. It is shown in Madan and Yor [57] that the process L is in law equal to the process $\{W_{Z_t} + \theta Z_t\}$, where $\theta = (G + M)/2$, W is a Brownian motion and Z is an independent Lévy subordinator Z with Laplace exponent ψ_Z given by

$$\mathbf{E}\left[e^{-uZ_t}\right] = e^{t\psi_Z(u)} = \exp\left(tC\Gamma(-Y)\left[2r(u)^Y\cos(\eta(u)Y) - M^Y - G^Y\right]\right)$$
$$r(u) = \sqrt{2u + GM}, \quad \eta(u) = \arctan\left(\frac{2\sqrt{2u - \theta^2}}{G + M}\right)$$

The subordinated formulation is used to create the approximating Markov chain following the algorithm in Section 4.A.1.

In Almendral and Oosterlee [4] the value of an American put option with maturity $T = 1$ under the CGMY model is computed using a finite difference scheme. In the case that the spot is $S_0 = 1$, the risk-free interest rate is $r = 0.1$ and the CGMY parameters are given by $C = 1$, $M = G = 5$ and $Y = 0.5$ a reference value of 0.112171 is reported in Almendral and Oosterlee [4, Table:3]. In Table 4.5 the values of this option obtained by using the FB and DP methods are given. Note that they agree

	Size	$N = 200$	$N = 400$
DP	$M = 1600$	0.11219	0.11217
DP	$M = 3200$	0.11220	0.11218
FB		0.11220	0.11218
AO			0.112171

Table 4.5: Numerical values of the American put option modelled by the CGMY process, modell parameters $r = 0.1$, $C = 1$, $M = G = 5$ and $Y = 0.5$ and option parameters $K = 1$, $S_0 = 1$ and $T = 1$. The row AO is retrived from Almendral and Oosterlee [4, Table:3]

with the reference value given above.

Figure 4.15 show the absolute error for the DP method as the number of exercise dates varies. The observed convergence is the usual linear. Figure 4.16 show the absolute error for the FB and DP methods as the size of the Markov chain is varied. Note that the CGMY approximation show lower observed convergence in the size of the Markov chain compared to previous examples. In this case the convergence is approximately linear, while the previous examples are approximately quadratic. It is interesting to note however that the error starts from a lower level compared to the previous examples. The relative error in the CGMY case is only 0.04% for a Markov chain of size 200.

4.4.6 The American put option under a local volatility Merton jump diffusion model

Let the jump process L be of compound Poison type. Let the intensity be λ and the jumps on the form $(e^K - 1)$, where K is a normal random variable with mean m and variance s^2. The process that evolves according to the SDE

$$\begin{aligned} \frac{\mathrm{d}S_t}{S_t} &= (r(t) - \lambda\xi)\,\mathrm{d}t + \Sigma(t, S_t)\mathrm{d}W_t + \mathrm{d}L_t \\ \xi &= \mathbf{E}\left[e^K - 1\right] = e^{m+s^2/2} - 1. \end{aligned}$$

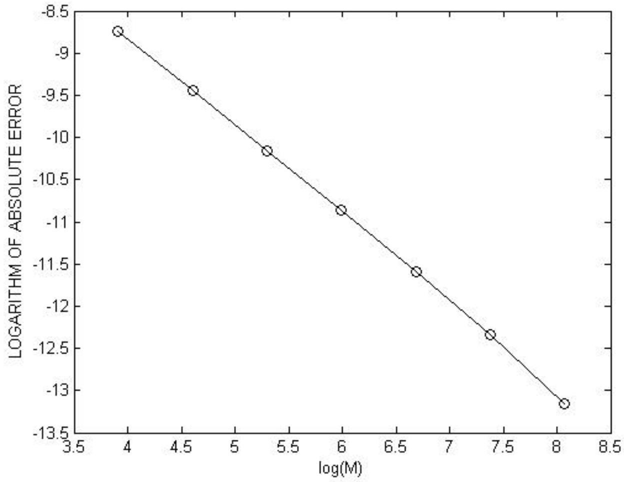

Figure 4.15: Display is the absolute error of the American put option values generated by the DP method for varying number of exercise times M using a Markov chain with state space of fixed size $N = 1600$. The Markov chain approximated the CGMY model with parameters $r = 0.1$, $C = 1$, $M = G = 5$ and $Y = 0.5$. The parameters of the option are $K = 1$, $S_0 = 1$ and $T = 1$. To calculate the reference value the DP method is used with $M = 15000$ exercise times.

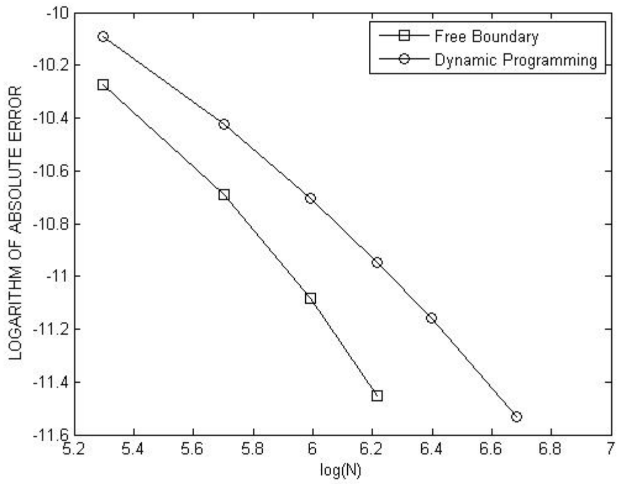

Figure 4.16: Display are the absolute errors of the American put option values generated by the free boundary and the dynamic programming ($M = 15000$ exercise times) methods, for varying sizes N of the state space of the Markov chain. The CGMY model was used with model parameters $r = 0.1$, $C = 1$, $M = G = 5$ and $Y = 0.5$ and option parameters $K = 1$, $S_0 = 1$ and $T = 1$. The reference value for the DP method is calculated using a Markov Chian of size $N = 2500$. The reference value for the FB method is calculated using a Markov chain of size $N = 1000$.

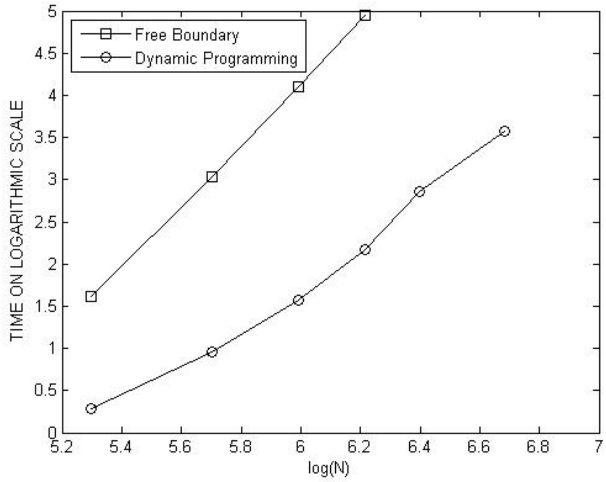

Figure 4.17: Display are the execution times of the American put option values generated by the free boundary and the dynamic programming ($M = 15000$ exercise times) methods, for varying sizes N of the state space of the Markov chain. The underlying was modelled by the CGMY process with parameters $r = 0.1$, $C = 1$, $M = G = 5$ and $Y = 0.5$ the option parameters were $K = 1$, $S_0 = 1$ and $T = 1$.

is then a local Lévy process. Let the non-constant instantaneous interest rate $r(t)$ and volatility $\Sigma(t, x)$ be given by

$$\begin{aligned}\Sigma(t,x) =& v(t)(x/S_0)^\beta \\ v(t) =& \theta + (\sigma_0 - \theta)e^{-kt} \\ r(t) =& r_0 + r_1 e^{a_0 t},\end{aligned}$$

then S can be viewed as a local volatility Merton jump diffusion.

This is a time-inhomogeneous Markov process with compensator of the associated jump-measure given by

$$\nu(\mathrm{d}y) = \frac{\lambda}{\sqrt{2\pi s^2}} \exp\left(-\frac{(\log(1+y) - m)^2}{2s^2}\right) \frac{\mathrm{d}y}{1+y}.$$

	Size	$N = 200$	$N = 400$	$N = 600$
	$n = 50$			
DP	$M = 8000$	5.8239	5.8249	5.8250
DP	$M = 12000$	5.8239	5.8249	5.8250
FB		5.8240	5.8249	5.8251
	$n = 100$			
DP	$M = 8000$	5.8290	5.8299	5.8301
DP	$M = 12000$	5.8290	5.8300	5.8301
FB		5.8291	5.8300	5.8302

Table 4.6: Numerical values for the American put option with the underlying modelled by a local volatility Merton jump diffusion. The model parameters used were $\sigma_0 = 0.25$, $\theta = 0.2$, $k = 10$, $\beta = -2$, $r_0 = 0.01$, $r_1 = 0.09$, $a_0 = -1$, $\lambda = 0.1008$, $m = -0.9144$ and $s = 0.4367$ the option parameters are $K = 100$, $S_0 = 100$ and $T = 0.5$. n is the size of the time discretisation and N is the size of the Markov chain and M is the number exercise times in for the DP method.

If Φ is the standard normal cumulative distribution function (cdf)

$$\begin{aligned}\int_a^b \nu(\mathrm{d}y) =& \lambda \left[\Phi\left((\log(1+b)-m)/s\right) - \Phi\left((\log(1+a)-m)/s\right)\right] \\ \int_{-1}^{\infty} y^2 \nu(\mathrm{d}y) =& \lambda \left[e^{2\left(m+s^2\right)} - 2e^{m+s^2/2} + 1\right]\end{aligned}$$

holds. These are used by the method in Section 4.A.1 to build the approximating Markov chains.

To handle the case of this time-inhomogeneous Markov process with the presented algorithms the time-dependent parameters are approximated by parameters that are piecewise constant in time. The DP and FB algorithms are then applied to each of the n time-sections that the parameters are constant, and the final results are obtained by multiplication using the Markov property, as in Mijatović and Pistorius [62].

In Table 4.6 the numerical outcomes are presented that were obtaind by applying the FB and DP methods to the approximating Markov chain.

Convergence results are shown in Figures 4.18 and 4.19 and execution time in Figure 4.20. The slope in Figure 4.18 is close to -1, and the slopes in Figure 4.19 is close

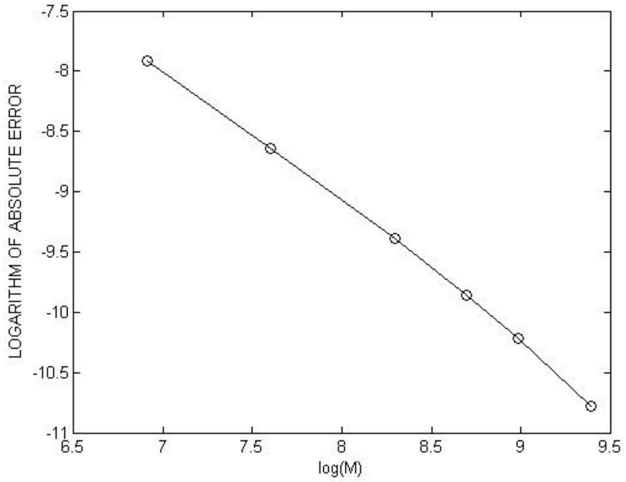

Figure 4.18: For a local volatility Merton jump diffusion the absolute error is given for a varying number of exercise times M with a fixed Markov chain size $N = 1600$ in the DP method. The model parameters are $\sigma_0 = 0.25$, $\theta = 0.2$, $k = 10$, $\beta = -2$, $r_0 = 0.01$, $r_1 = 0.09$, $a_0 = -1$, $\lambda = 0.1008$, $m = -0.9144$ and $s = 0.4367$ and the option parameters are $K = 100$, $S_0 = 100$ and $T = 0.5$. The exact value is calculated using $M = 36000$ exercise times.

to -2, showing the familiar linear convergence in exercise times and quadratic in size of Markov chain.

4.5 Conclusion

Two methods for pricing American options driven by a continuous-time Markov chain have been presented. The first method is an exact algorithm for pricing the American option by determining the optimal barrier. Numerical implementation of this method requires root-finding and matrix-exponentiation, for which efficient numerical algorithms are available.

The second method concerns valuation of Bermudan options. It is shown that the computational effort only marginally increases as function of the number of exercise times. As a consequence, this method can be used to accurately approximate the value of an American option by evaluating the corresponding Bermudan option with

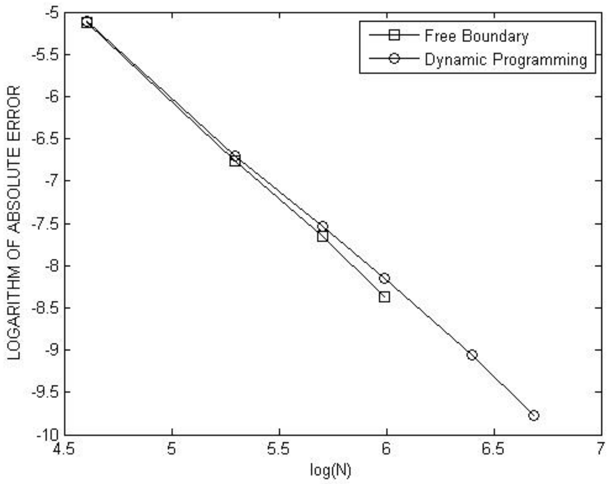

Figure 4.19: For a local volatility Merton jump diffusion the absolute error is given for varying sizes of the Markov chain (N) for the free boundary and the dynamic programming method, the number of exercise times are $M = 36000$ in the dynamic programming method. The model parameters are $\sigma_0 = 0.25$, $\theta = 0.2$, $k = 10$, $\beta = -2$, $r_0 = 0.01$, $r_1 = 0.09$, $a_0 = -1$, $\lambda = 0.1008$, $m = -0.9144$ and $s = 0.4367$ and the option parameters are $K = 100$, $S_0 = 100$ and $T = 0.5$. The exactvalues used are calulated using a Markov chain of size $N = 800$ for FB, and $N = 1600$ for DP

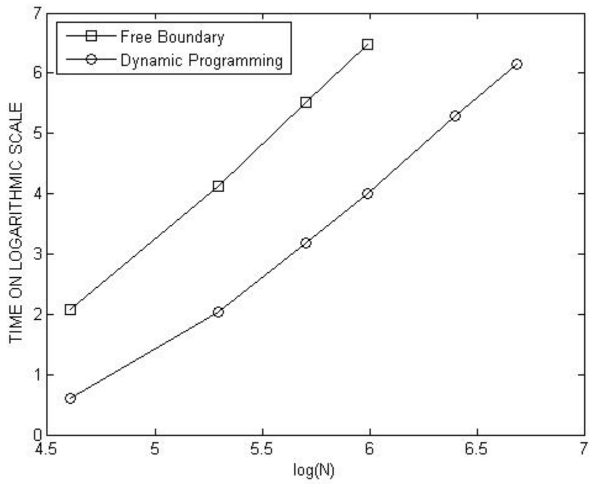

Figure 4.20: For the local volatility Merton jump diffusion the execution time is given for varying sizes of the Markov chain (N) for the free boundary and the dynamic programming methods, the number of exercise times are $M = 36000$ in the DP method. The model parameters are $\sigma_0 = 0.25$, $\theta = 0.2$, $k = 10$, $\beta = -2$, $r_0 = 0.01$, $r_1 = 0.09$, $a_0 = -1$, $\lambda = 0.1008$, $m = -0.9144$ and $s = 0.4367$ and the option parameters are $K = 100$, $S_0 = 100$ and $T = 0.5$.

sufficiently many exercise times. Convergence of a Bermudan option to an American option as the number of exercise times increase is shown.

A method for approximating local volatility with jumps method by a Markov chain is provided. For this approximation it is shown that the value of the American and Bermudan options converge as the size of the Markov chain increase. The method is illustrated by numerical examples.

Appendix

4.A A Markov chains method to value barrier options

This appendix is a summary of the paper Mijatović and Pistorius [62] on valuing barrier options by Markov chain approximation. The appendix is structured as follows. Appendix 4.A.1 contains a description of how to construct a Markov chain approximating a Markov process. Appendix 4.A.2 provides the method to value barrier options when the underlying is a Markov chain.

4.A.1 Construction of Markov chains approximating Markov processes

In this section an outline is given of the method, presented in Mijatović and Pistorius [62, Sec:5], for the construction of a Markov chain that approximates a given Markov process.

The Markov chain is determined by its generator matrix Λ. The idea behind the presented approximation is to build a matrix generator Λ that approximate the infinitesimal generator $\mathcal{L}$ of the given Markov process.

The first step in defining the approximating Markov chain (M_t) is to specify the state space $(\mathbb{G})$ of the Markov chain. The effectiveness of the pricing algorithm is highly dependent on the appropriate choice of the grid. Important points to consider are: resolution around important points, such as spot price and barrier levels, and controlling the truncation of the infinite state space of the process. To achieve this without letting the number of points in $\mathbb{G}$ get to large (would make the pricing method computationally expensive) a non uniform grid is suggested.

To be able to utilise specific characteristics of the process being approximated three different methods to build the approximating chain will be introduced in the following sections, one method for diffusion processes and two for Markov processes with discontinuous sample paths.

Diffusion process

In this section a method to approximate a diffusion process is described.

Consider a diffusion process S following the SDE

$$\mathrm{d}S_t = \gamma S_t \mathrm{d}t + \sigma(S_t) S_t \mathrm{d}W_t.$$

where W_t is a Brownian motion, $\gamma = r - d$ and the volatility function $\sigma(x)$ is such that S_t does not explode.

Let N be the size of the state space of the Markov chain $\mathbb{G}$. Define the boundary and interior of the grid by

$$\partial\mathbb{G} = \{x_1, x_N\} \text{ and } \mathbb{G}_o = \mathbb{G} \backslash \partial\mathbb{G},$$

where $\mathbb{G}$ is expected to be ordered so that x_1/x_N is the smallest/largest elements of $\mathbb{G}$ respectively.

Let Λ be the generator of the approximating chain M_t. For Λ to be a generator matrix

$$\sum_{y \in \mathbb{G}} \Lambda(x,y) = 0$$
$$\Lambda(x,y) \geq 0 \quad \forall y \in \mathbb{G} \backslash \{x\}$$

must hold for all $x \in \mathbb{G}^o$. The restrictions imposed on Λ to ensure M_t is an approximation of S_t is that the first and second instant moments of the chain match the moments of the process. That is for each $x \in \mathbb{G}^o$

$$\sum_{y \in \mathbb{G}} \Lambda(x,y)(y-x) = \gamma x,$$
$$\sum_{y \in \mathbb{G}} \Lambda(x,y)(y-x)^2 = (\sigma(x)x)^2.$$

For $x \in \partial\mathbb{G}$ absorbing boundary conditions, $\Lambda(x,y) = 0$ for all $y \in \mathbb{G}$, are imposed. This system of equations can typically be achieved by a tri-diagonal generator matrix. Tri-diagonality implies that the chain can only move to a neighbouring state, making a connection to the continuity of the sample paths of S_t. Note that the absorbing boundary is not natural to the process, therefore it is desirable to generate a grid such that the probability of reaching the boundaries in the timespan of the option is small.

Lévy subordinated diffusions

One way to build a Lévy process is to time change a Brownian motion by an independent Lévy subordinator. More generally if S_t is a Lévy subordinated diffusion, there exists a diffusion process S' with dynamic

$$\mathrm{d}S'_t = \gamma S'_t \mathrm{d}t + \sigma(S'_t)S'_t \mathrm{d}W_t$$

and a Lévy subordinator Z_t such that

$$S_t = e^{\mu t} S'_{Z_t}.$$

For the discounted asset price process $e^{-(r-d)t} S_t$ to be a martingale it is required that

$$\mu + \phi_Z(-\gamma) = r - d$$

where ϕ_Z is the Laplace exponent of Z, that is $\mathbf{E}\left[e^{-uZ_t}\right] = e^{t\phi_Z(u)}$, r is the rate of discounting and d is the dividend yield.

If the generator of S' is $\mathcal{L}'$ by Phillips theorem (see Sato [68, Thm:32.1]) the generator of S'_{Z_t} is

$$\mathcal{L}'_Z = \phi_Z(-\mathcal{L}')$$

and therefore the infinitesimal generator of S is

$$(\mathcal{L}f)(x) = \mu x f'(x) + (\mathcal{L}'_Z f)(x). \tag{4.48}$$

To construct a Markov chain approximating S_t Philips theorem will be applied to a Markov chain approximating the diffusion S'_t. First use the method in section 4.A.1 to build a generator matrix Λ' approximating S'. Then apply Philips theorem to this generator $\Lambda'_Z = \phi_Z(-\Lambda')$. To compute Λ'_Z decompose $\Lambda' = UDU^{-1}$, where D is a diagonal matrix and U is invertible, and define $\Lambda'_Z = U\phi(-D)U^{-1}$.

Finally to build the chain that approximate (4.48) drift is added. A tri-diagonal generator Λ_μ that satisfies

$$\sum_{y\in\mathbb{G}} \Lambda_\mu(x,y)(y-x) = (r-d)x - \sum_{z\in\mathbb{G}} \Lambda'_Z(x,z)(z-x) \quad \forall x \in \mathbb{G}^o$$

has the desired drift. The generator matrix Λ that approximate (4.48) is

$$\Lambda = \Lambda_Z' + \Lambda_\mu.$$

As in the diffusion case absorbing boundary conditions are imposed.

Jump processes with state-dependent characteristics

Let the process S have the infinitesimal generator

$$\mathcal{L}f(x) = \mathcal{L}_D f(x) + \mathcal{L}_J f(x)$$

where $\mathcal{L}_D f(x)$ is the generator of diffusion given by

$$\mathcal{L}_D f(x) = (r-d)xf'(x) + \frac{\sigma(x)^2 x^2}{2} f''(x)$$

and the generator of the jump part is given by

$$\mathcal{L}_J f(x) = \int_{(-1,\infty)} [f(x(1+y)) - f(x) - f'(x)xy]\nu(x, \mathrm{d}y),$$

where for every $x \in \mathbb{E}$, $\nu(x, \mathrm{d}y)$ is a (Lévy) measure with support in $(-1, \infty)$ such that

$$\int_{(-1,\infty)} \min\{y^2, |y|\}\nu(x, \mathrm{d}y) < \infty.$$

Sufficient conditions on σ and μ to guarantee the existence of a Feller process S corresponding to this generator were established in Kolokoltsov [45]. This process is approximated by first building a Markov chain that approximates the jump measure and then adding a Markov chain to ensure that the first two instantaneous moments are matched.

Let the state space of the Markov chain be $\mathbb{G}$. To approximate the jumps of the process a discretisation of the jump measure is employed. For any point in the interior $x \in \mathbb{G}^o$ the relative jump sizes of the chain is,

$$\mathbb{G}_x = \left\{ \frac{z}{x} - 1 : z \in \mathbb{G} \right\}.$$

Let $y_0 < \min\{y \in \mathbb{G}_x\}$ and define a function $\alpha_x : \mathbb{G}_x \bigcup \{y_0\} \to [-1, \infty]$ such that

$$\alpha_x(y_i) \in (y_i, y_{i+1}) \text{ and } i \in \{1, \ldots, N-1\}$$

and $\alpha_x(y_0) = -1, \quad \alpha_x(y_N) = \infty$. A natural choice for the interior α_x is the mid point between the y_is. Note that $\alpha_x(y_i)$ for $i \in \{0, 1, \ldots, N\}$ construct a partition of $(-1, \infty)$ of size N. Each partition contain one of the jump sizes of the chain. Using α_x define the generator of the jump part of the chain by

$$\begin{aligned} \Lambda_J(x, x(1+y_i)) &= \int_{\alpha_x(y_{i-1})}^{\alpha_x(y_i)} \nu(x, dy) \quad \text{where } i \in \{1, \ldots, N\} \text{ and } y_i \neq 0 \\ \Lambda_J(x, x) &= - \sum_{z \in \mathbb{G} \setminus \{x\}} \Lambda_J(x, z). \end{aligned}$$

That is the jump intensity of the chain is obtained by integrating the Lévy measure over the corresponding part of the partition. As usual set $\Lambda_J(x, y) = 0$ for $x \in \partial \mathbb{G}$ to impose absorbing boundary conditions.

Next find a tri-diagonal generator matrix Λ_c such that $\Lambda_c + \Lambda_J$ match the first two moments of S_t, this requires

$$\begin{aligned} &\sum_{z \in \mathbb{G}} \Lambda_c(x, z) = 0 \quad \text{ and } \Lambda_c(x, z) + \Lambda_J(x, z) \geq 0 \, \forall z \in \mathbb{G} \setminus \{x\}, \\ &\sum_{z \in \mathbb{G}} \Lambda_c(x, z)(z - x) = (r - d)x - \sum_{z' \in G} \Lambda_J(x, z')(z' - x), \\ &\sum_{z \in \mathbb{G}} \Lambda_c(x, z)(z - x)^2 = x^2 \left(\sigma(x)^2 + \int_{-1}^{\infty} y^2 \nu(x, dy) \right) - \sum_{z' \in \mathbb{G}} \Lambda_J(x, z')(z' - x)^2 \end{aligned}$$

to be fulfilled. As usual impose absorbing boundary conditions. Finally the total generator matrix is

$$\Lambda = \Lambda_J + \Lambda_c.$$

4.A.2 A method to value barrier contracts

The aim of this section is to provide a brief description of the method introduced in Mijatović and Pistorius [62] to price barrier type contracts when the underlying is a Markov chain.

Consider a contract with value function

$$V(t,x) = \mathbf{E}_{t,x}\left[e^{-r(T-t)}g(S_T)\mathbf{I}(\tau_A(t) > T) + \epsilon^{-r(\tau_A(t)-t)}h(S_{\tau_A(t)})\mathbf{I}(\tau_A(t) \leq T)\right],$$

where $\tau_A(t) = \inf\{s \geq t : S_s \in A\}$. That is an option that is knocked out if the process hits A.

Assume M_t is a Markov chain approximating S_t (see Appendix 4.A.1 on how to construct M_t) and let Λ be the generator matrix of the chain M_t with domain $\mathbb{G}$. The set $\mathbb{G}$ can be split into a continuation set

$$\widehat{\mathbb{G}} = \{x \in \mathbb{G} : x \in A^c\}$$

and a knock-out set $\widehat{\mathbb{G}}^c = \mathbb{G} \backslash \widehat{\mathbb{G}}$. Let

$$\tau(t) = \inf\left\{s \geq t : M_s \notin \widehat{\mathbb{G}}\right\}$$

be the first exit time from $\widehat{\mathbb{G}}$. Of importance to the method is the ease with which the generators of the chain killed $\left(\widehat{M_t}\right)$ and absorbed $\left(\widetilde{M_t}\right)$ upon hitting $\widehat{\mathbb{G}}$ can be calculated. Their generators are given by

$$\widetilde{\Lambda}_r(x,y) = \begin{cases} \Lambda(x,y) - r & \text{if } x \in \widehat{\mathbb{G}}, x = y, \\ \Lambda(x,y) & \text{if } x \in \widehat{\mathbb{G}}, x \neq y, \\ 0 & \text{if } x \in \widehat{\mathbb{G}}^c, y \in \mathbb{G}, \end{cases} \tag{4.49}$$

$$\widehat{\Lambda}(x,y) = \Lambda(x,y) \quad \text{if } x \in \widehat{\mathbb{G}}, y \in \widehat{\mathbb{G}}.$$

The main theorem in Mijatović and Pistorius [62, Thm:1] states that.

Theorem 4.A.1. *For any $T > t$, $x \in \mathbb{G}$ and $r \geq 0$ and any function $\phi : \mathbb{G} \to \mathbb{R}$ it holds that*

$$\mathbf{E}_{t,x}\left[e^{-r((T-\wedge\tau(t)-t)}\phi(M_{T\wedge\tau(t)})\right] = \left(\exp\left((T-t)\widetilde{\Lambda}_r\right)\phi\right)(x).$$

In particular, for $\psi : \widehat{\mathbb{G}} \to \mathbb{R}$ and $\xi : \mathbb{G} \to \mathbb{R}$ with $\xi(x) = 0$ for $x \in \widehat{\mathbb{G}}$

$$\mathbf{E}_{t,x}\left[\psi(M_T)\mathbf{I}(\tau(t) > T)\right] = \left(\exp\left((T-t)\widehat{\Lambda}\right)\psi\right)(x) \text{ for any } x \in \widehat{\mathbb{G}}$$
$$\mathbf{E}_{t,x}\left[e^{-r(\tau(t)-t)}\xi(M_\tau)\mathbf{I}(\tau(t) \leq T)\right] = \left(\exp\left((T-t)\widetilde{\Lambda}_r\right)\xi\right)(x) \text{ for any } x \in \mathbb{G}$$

holds.

Remark 4.A.1. As matrix exponentiation is a smooth operation the value function $V(t,x)$ is smooth in t.

References

[1] F. Abergel and A. Jedidi. A mathematical approach to order book modeling. In F. Abergel, B. K. Chakrabarti, A. Chakraborti, and M. Mitra, editors, *Econophysics of Order-driven Markets, Proceedings of Econophys-Kolkata V*, pages 93–107. Springer, 2011.

[2] J. Ahn and M. Song. Convergence of the trinomial tree method for pricing European/American options. *Applied Mathematics and Computation*, 189:575–582, 2007.

[3] W. Allegretto, G. Barone-Adesi, and R. J. Elliott. Numerical evaluation of the critical price and American options. *The European Journal of Finance*, 1(1): 69–78, 1995.

[4] A. Almendral and C. W. Oosterlee. Accurate evaluation of European and American options under the CGMY process. *SIAM Journal on Scientific Computing*, 29(1):93–117, 2007.

[5] A. Almendral and C. W. Oosterlee. On American options under the variance gamma process. *Applied Mathematical Finance*, 14(2):131–152, 2007.

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

[7] V. Bally, G. Pagés, and J. Printems. A quantization tree method for pricing and hedging multi-dimensional American options. *Mathematical Finance*, 15(1): 119–168, 2005.

[8] G. Barone-Adesi and R. J. Elliott. Approximations for the values of American options. *Stochastic Analysis and Applications*, 9(2):115–131, 1991.

[9] G. Barone-Adesi and R. E. Whaley. Efficient analytic approximation of American option values. *The Journal of Finance*, 42(2):301–320, 1987.

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

[11] A. B. Bhatt and V. S. Borkar. Occupation measures for controlled Markov processes: Characterization and optimality. *The Annals of Probability*, 24(3): 1531–1562, 1996.

[12] T. Björk. *Arbitrage Theory in Continuous Time*. Oxford University Press, 2nd edition, 2004.

[13] F. Black and M. Scholes. The pricing of options and corporate liabilities. *Journal of Politcal Economy*, 81(3):637–654, 1973.

[14] S. Boyarchenko and S. Levendorskii. *Irreversible Decisions under Uncertainty: Optimal Stopping Made Easy* . Springer, 2007.

[15] S. Boyarchenko and S. Levendorskii. American options in regime-switching Lévy models with non-semibounded stochastic interest rates. In *American Control Conference, 2008*, pages 1023–1028, 2008.

[16] S. I. Boyarchenko and S. Levendorskii. *Non-Gaussian Merton-Black-Scholes Theory*. World Scientific Publishing Co Pte Ltd, 2002.

[17] M. J. Brennan and E. S. Schwartz. The valuation of American put options. *The Journal of Finance*, 32(2):449–462, 1977.

[18] M. Broadie and P. Glasserman. Pricing American-style securities using simulation. *Journal of Economic Dynamics and Control*, 21:1323–1352, 1997.

[19] P. Carr. Randomization and the American put. *The Review of Financial Studies*, 11(3):597–626, 1998.

[20] P. Carr and J Crosby. A class of Lévy process models with almost exact calibration to both barrier and vanilla fx options. *Quantative Finance*, 10(10): 1115–1136, 2010.

[21] P. Carr, R. Jarrow, and R. Myneni. Alternative characterizations of American put options. *Mathematical Finance*, 2(2):87–106, 1992.

[22] P. Carr, H. Geman, D. B. Madan, and M. Yor. The fine structure of asset returns: An empirical investigation. *Journal of Business*, 75(2):305–332, 2002.

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

[24] R. Cont and P. Tankov. *Financial Modelling With Jump Processes*. Chapman & Hall/CRC, 2004.

[25] J. Cox. Notes on option pricing I: Constant elasticity of variance diffusions. *Journal of Portfolio Management*, 22(15–17), 1996.

[26] J. C. Cox, S. A. Ross, and M. Rubinstein. Options pricing: A simplified approach. *Journal of Financial Economics*, 3:125–144, 1979.

[27] J. C. Cox, J. E. Ingersoll, and S. A. Ross. A theory of the term structure of interest rates. *Econometrica*, 53(2):385–407, 1985.

[28] C. Cuchiero, D. Filipović, and J. Teichmann. Affine models. In *Encyclopedia of Quantitative Finance*. John Wiley & Sons, Ltd, 2010.

[29] C. Cuchiero, M. Keller-Ressel, and J. Teichmann. Polynomial processes and their applications to mathematical finance. *Finance and Stochastics*, 16(4):711–740, 2012.

[30] D. Davydov and V. Linetsky. Pricing options on scalar diffusions: An eigenfunction expansion approach. *Operations Research*, 51(2), 2003.

[31] Y. d'Halluin, P. A. Forsyth, and G. Labahn. A penalty method for American options with jump diffusion processes. *Numerische Mathematik*, 97:321–352, 2004.

[32] N. El Karoui. Les aspects probabilistes du controle stochastique. In P. Bickel, N. El Karoui, and M. Yor, editors, *Ecole d'Eté de Probabilités de Saint-Flour IX-1979*, volume 876 of *Lecture Notes in Mathematics*, pages 73–238. Springer, 1981.

[33] R. J. Elliot and P. E. Kopp. *Mathematics of Financial Markets*. Springer, 1999.

[34] B. Eriksson and M. Pistorius. A method of moments approach to pricing double barrier contracts driven by a general class of jump diffusions. *International Journal of Theoretical and Applied Finance*, 14(7):1139–1158, 2011.

[35] S. N. Ethier and T. G. Kurtz. *Markov Processes: Characterization and Convergence*. Wiley, 1986.

[36] H. Geman and M. Yor. Pricing and hedging double-barrier options: A probabilistic approach. *Mathematical Finance*, 6(4), 1996.

[37] R. Geske and H. E. Johnson. The American put option valued analytically. *The Journal of Finance*, 39(5):1511–1524, 1984.

[38] K. Helmes, S. Röhl, and R. H. Stockbridge. Computing moments of the exit time distribution for Markov processes by linear programming. *Operations Research*, 49(4), 2001.

[39] A. Hirsa and D. B. Madan. Pricing American options under variance gamma. *Jounal of Computational Finance*, 7(2):63–80, 2003.

[40] S. D. Jacka. Optimal stopping and the American put. *Mathematical Finance*, 1 (2):1–14, 1991.

[41] J. Jacod and A. N. Shiryaev. *Limit theorems for stochastic processes*, volume 288 of *A Series of Comprehensive Studies in Mathematics*. Springer-Verlag, 2nd edition, 2003.

[42] H. E. Johnson. An analytical approximation for the American put price. *Journal of Financial and Quantative Analysis*, 1983.

[43] N. Jua and R. Zhong. An approximate formula for pricing American options. *The Journal of Derivatives*, 7(2):31–40, 1999.

[44] I. J. Kim. The analytic valuation of American options. *Review of Financial Studies*, 3(4):547–72, 1990.

[45] V. N. Kolokoltsov. The Lévy-Khintchine type operators with variable Lipschitz continuous coefficients generate linear or nonlinear Markov processes and semigroups. *Probability Theory and Related Fields*, 151:95–123, 2011.

[46] S. G. Kou. A jump diffusion for option prices. *Management Science*, 48(8): 1086–1101, August 2002.

[47] S. G. Kou and H. Wang. Option pricing under a double exponential jump diffusion model. *Management Science*, 50(9):1178–1192, September 2004.

[48] T. G. Kurtz and R. H. Stockbridge. Existence of Markov controls and characterization of optimal Markov controls. *Siam Journal on Control and Optimization*, 36(2), 1998.

[49] H. J. Kushner. Numerical methods for stochastic control problems in finance. In *Mathematics of Derivative Securities*, number 15 in Publications of the Newton Institute, pages 504–527. Cambridge University Press, 1997.

[50] H. J. Kushner and P. G. Dupuis. *Numerical Methods for Stochastic Control Problems in Continuous Time*. Springer, 2nd edition edition, 2000.

[51] D. Lamberton. Error estimates for the binomial approximation of American put options. *The Annals of Applied Probability*, 8(1):206–233, 1998.

[52] D. Lamberton and M. Mikou. The critical price for the American put in an exponential Lévy model. *Finance and Stochastics*, 12(4):561–581, 2008.

[53] D. Lamberton and M. Mikou. The smooth-fit property in an exponential Lévy model. *Journal of Applied Probability*, 49(1):137–149, 2012.

[54] J. B. Lasserre, T. Prieto-Rumeau, and M. Zervos. Pricing a class of exotic options via moments and SDP relaxations. *Mathematical Finance*, 16(3), 2006.

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

[56] R. Lord, F. Fang, F. Bervoets, and C. W. Oosterlee. A fast and accurate FFT-based method for pricing early-exercise options under Lévy processes. *SIAM Journal on Scientific Computing*, 29:93–117, 2007.

[57] D. Madan and M. Yor. Representing the CGMY and Meixner Lévy processes as time changed Brownian motions. *The Journal of Computational Finance*, 12(1): 27–47, 2008.

[58] D. B. Madan, P. Carr, and E. Chang. The variance gamma process and option pricing model. *European Finance Review*, 2:79–105, 1998.

[59] R. A. Maller, D. S. Solomon, and A. Szimayer. A multinomial approximation for American option prices in Lévy process models. *Mathematical Finance*, 16 (4):613–633, 2006.

[60] H. P. McKean. Appendix: A free boundary problem for the heat equation arising from a problem in mathematical economics. *Industrial Management Review*, 6 (2):32–39, 1965.

[61] R. C. Merton. Theory of rational option pricing. *The Bell Journal of Economics and Management Science*, 4(1):141–183, 1973.

[62] A. Mijatović and M. Pistorius. Continuously monitored barrier options under Markov processes. *Mathematical Finance*, 23:1–38, 2013.

[63] G. N. Milstein, O. Reiß, and J. Schoenmakers. A new Monte Carlo method for American options. *International Journal of Theoretical and Applied Finance*, 7 (5):591–614, 2004.

[64] A. Pelsser. Pricing double barrier options using Laplace transforms. *Finance and Stochastics*, 4(1), 2000.

[65] G. Peskir and A. N. Shiryaev. *Optimal stopping and free-boundary problems.* Birkhäuser, 2006.

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

[67] L. C. G. Rogers. Dual valuation and hedging of Bermudan options. *SIAM Journal on Financial Mathematics*, 1:604–608, 2010.

[68] K. Sato. *Lévy processes and infinitely divisible distributions.* Cambridge University Press, 1999.

[69] J. Schoenmakers, J. Huang, and J. Zhang. Optimal dual martingales, their analysis and application to new algorithms for Bermudan products. *SIAM Journal on Financial Mathematics*, 4(1):86–116, 2013.

[70] W. Schoutens. *Lévy Processes in Finance: Pricing Financial Derivatives.* Wiley, 2003.

[71] A. Sepp. Analytical pricing of double-barrier options under a double-exponential jump diffusion process: Applications of Laplace transform. *International Journal of Theoretical and Applied Finance*, 7(2), 2004.

[72] A. N. Shiryaev. *Optimal Stopping Rules*. Springer, 2007.

[73] J. A. Shohat and J. D. Tamarkin. *The Problem of Moments*. American Mathematical Society, 1943.

[74] A. Szimayer and R. A. Maller. Finite approximation schemes for Lévy processes and their application to optimal stopping problems. *Stochastic Processes and their Applications*, 117:1422–1447, 2007.

[75] J. Toivanen. Numerican valuation of European and American options under Kou's jump-diffusion Model. *SIAM Journal on Scientific Computing*, 30(4): 1949–1970, 2008.

[76] I. R. Wang, J. W. L. Wan, and P. A. Forsyth. Robust numerical valuation of European and American Options under the CGMY process. *Journal of Computational Finance*, 10(4):31–69, 2007.

[77] T. Weithers. *Foreign Exchange: A Practical Guide to the FX Markets*. Wiley, 2006.

[78] H. Y. Wong and J. Zhao. An artificial boundary method for American option pricing under the CEV model. *SIAM Journal on Numerical Analysis*, 46(4): 2183–2209, 2008.

[79] U. Wystup. *FX Options and Structured Products*. Wiley, 2007.