



TECHNISCHE
UNIVERSITÄT
WIEN
Vienna University of Technology

Diplomarbeit

Pricing American options via the Cross-Entropy method

ausgeführt am Institut für
Wirtschaftsmathematik

an der

Technischen Universität Wien

unter Anleitung von

Ao. Univ. Prof. Dipl.-Ing. Dr. techn. Friedrich Hubalek

in Zusammenarbeit mit

Professor Dr.scient Søren Asmussen

durch

Christoph Hoeggerl

Rudolf von Alt-Platz 1/4, 1030 Wien

10.Mai 2010



TECHNISCHE
UNIVERSITÄT
WIEN
Vienna University of Technology

Diploma thesis

Pricing American options via the Cross-Entropy method

handed in at the

Dept. of Mathematical Methods in Economics

at the

Vienna University of Technology

under the supervision of

Supervisor at the University of Technology, Vienna:

Ao. Univ. Prof. Dipl.-Ing. Dr. techn. Friedrich Hubalek

Supervisor at Aarhus University:

Professor dr.scient Søren Asmussen

Author:

Christoph Hoeggerl

Rudolf von Alt-Platz 1/4, 1030 Wien

May 10, 2010

Acknowledgements

At the beginning of my master thesis on "Pricing American options via the Cross - Entropy Method" I would like to thank all those who have made this work possible.

My profound thanks to Professor Søren Asmussen, who more than a year ago kindly agreed to supervise my master thesis. Your books were a wonderful challenge, the more so as I was more into mathematics for computer science at first and started relatively late to study financial matters in earnest. Thank you very much for your advice, patience, encouragement and wisdom.

It was, in fact, thanks to Professor Friedrich Hubalek, whose lectures on stochastic processes as well as on risk theory had roused my interest in the subject in the first place, that I had the invaluable opportunity of going to Denmark and meeting this top expert in the field for which I had gradually developed a deep interest during the course of my studies. A heartfelt thank you for all you have done for me, especially for your paving my way to Aarhus University, Denmark.

My stay in Denmark, however, was even more than that as it also gave me an insight into a different culture and allowed me to experience a truly international community. Having been able to get to know people from all over the world and to establish close contact with them, which has marked me deeply, is an experience which I will never forget in my life.

Last but not least I would like to thank my parents as well as my girlfriend Fani Lamprianidou for their encouragement and support, especially in times when it was difficult to carry on.

Thank you to all those people whose guidance, friendship and moral support encouraged me to believe in what I was doing.

Christoph Hoeggerl
Aarhus // Vienna, May 2010

Contents

1. Introduction	6
2. Financial Markets	8
2.1. Bonds	8
2.2. Stocks	8
2.3. Derivatives	8
2.3.1. Futures and Forwards	8
2.3.2. Options	9
2.3.3. Swaps	11
2.4. Market Models	11
2.4.1. Black - Scholes Model	11
2.4.2. Exponential Lévy Model	13
3. Survey on Pricing American Options	15
3.1. Parametric Approximations	16
3.2. Random Tree Methods	17
3.3. State Space Partitioning	19
3.4. Stochastic Mesh Methods	21
3.5. Regression-Based Methods	22
4. Methods	24
4.1. Risk-Neutral Pricing in the Black - Scholes Model	24
4.1.1. Stock and Portfolio under the Risk-Neutral Measure	25
4.1.2. Arbitrage and Market Completeness	27
4.2. Pricing and Hedging in Incomplete Markets	28
4.3. The American Option as a Stopping-Time Problem	30
4.3.1. About the Existence of an Optimal Stopping Time	32
4.4. Free-Boundary Problems and their Connection to Stopping-Time Problems . .	34
4.4.1. A Dirichlet Problem and the Solution	34
4.5. Simulation of Asset Developments	36
4.5.1. Simulation of a Brownian Motion	36
4.5.2. Simulation of a Normal Inverse Gaussian Levy Processes	37
4.6. The Cross - Entropy Method	37
4.6.1. Rare Event Simulation	38
4.6.2. Combinatorial Optimization Problems	40
4.6.3. Smoothed Updating	41

Contents

4.7. Adjustment of the CE Method for Pricing American Options	41
4.7.1. Properties of the Optimal Exercising Boundary	42
4.7.2. The Choice of Basis Functions	43
4.7.3. Pricing American Put Options	43
4.7.4. Initial Parameters	44
4.7.5. Pricing Basket Options	46
5. Results	47
5.1. Standard Black - Scholes Model	47
5.1.1. American-Style Vanilla Put Option	47
5.1.2. Basket Option	56
5.2. Exponential Lévy Model	60
6. Conclusion & Further Work	68
A. Algorithms	69
A.1. Cross - Entropy Algorithm, Approximation of the Optimal Exercising Boundary with Step Functions	69
A.2. Cross - Entropy Algorithm, Approximation of the Optimal Exercising Boundary with a Linear Combination of Basis Functions	75
A.3. Generation of a Geometric Brownian Motion	80
A.4. Generation of an Exponential Lévy Process with NIG Increments	80

1. Introduction

Within the last decades the growth of the financial markets was accompanied by an increasing number of complex options giving market participants the opportunity to hedge their portfolio risk as well as to speculate more freely. Since American-style options can be exercised at any point prior and up to maturity, they are a lot more flexible than European-style options and therefore it is crucial to be able to determine the value of an American option as accurately as possible.

The objective of this dissertation is to investigate the efficiency of applying the Cross - Entropy method to the combinatorial optimization problem of determining the price of an American put option with finite horizon in the standard Black - Scholes as well as in the exponential Lévy model with normal inverse Gaussian increments.

The Cross - Entropy method, deriving its name from the cross entropy (or Kullback - Leibler) distance, was introduced in 1997 by Rubinstein as an iterative method involving the following two steps:

1. Generation of a sample of random data according to a specified random mechanism.
2. Updating the parameters of the random mechanism, on the basis of the data, in order to produce a "better" sample in the next iteration.

The significance of the Cross - Entropy concept is that it defines a precise mathematical framework for deriving fast, and in some sense "optimal" updating rules.

Since the price of an American put option can be written as the maximisation problem over all admissible stopping times within the interval $[t, T]$, where t is the current time and T the maturity date of the option, it is possible to apply the Cross Entropy method.

It is also known that the optimal stopping time is the first time when the asset price process drops beneath the boundary separating the continuation set C from the stopping set D . The continuation set C consists of the points (t, x) for which continuing is more profitable than exercising, whereas the stopping set contains the points (t, x) for which immediate exercise is as profitable as not exercising the option at that time.

So far no closed form solution for the optimal exercising boundary has been found, which makes numerical methods for pricing American put options so important.

In the algorithm that was used to obtain the option price, the optimal exercising boundary was approximated by either step functions or a linear combination of basis functions. The step heights or the coefficients of the basis functions in the linear combination respectively will be described by a multivariate normal distribution.

In every simulation step of the Cross - Entropy method the mean vector and the covariance matrix will be updated, which is supposed to lead to better approximations of the optimal

1. Introduction

exercising boundary.

Since the Cross - Entropy method has so far been successfully applied to various optimization problems within the past years, this method may be expected to produce fairly accurate results here, too.

2. Financial Markets

Nowadays, a variety of instruments for trade in financial markets is available, yet this dissertation will focus on securities. Contrary to other financial contracts, a security confers a financial claim to its owner. Dividing the group of securities into the broadest possible sub-groups, we obtain the economic classification of bonds, stocks and derivatives, as discussed in Cvitanic and Zapatero [12, p.3], for example.

2.1. Bonds

When purchasing a bond the owner acquires the right to a predetermined and fixed payment at a future predetermined date, called maturity. Since a bond has a guaranteed payoff at maturity, which is known in advance, it is referred to as a risk-free security. The reason for obtaining a bond is that the person has some purchasing power that he would prefer to delay. Saving money for retirement would be a good example why someone would buy a bond [12, p.3-5].

2.2. Stocks

A stock, however, entitles its owner to a proportion of any distributed profit by the firm issuing the stock. In the case that the firm liquidates, an adequate part of the company belongs to the stockholder [12, p.7]. Due to the randomness of the dividend payments and the absence of a guaranteed nominal value, buying a stock and selling it at a later time might result in a profit or a loss respectively.

2.3. Derivatives

The payoff of a derivative is dependent on another financial variable called the underlying. Two parties agree upon a rule, depending on the value of the underlying, as to when one party receives a payment from the other party [12, p.9].

Several types of financial instruments satisfy the aforementioned characteristic.

2.3.1. Futures and Forwards

Futures and forwards are contracts where one party agrees to purchase the underlying at a future predetermined date at a predetermined price from the other party. The other party in return commits itself to deliver the underlying at that date at the price that has been agreed

2. Financial Markets

upon. The only difference between futures and forwards is the way the payments are made. Forward contracts permit payment only at the expiration date in exchange for the underlying asset, while futures allow more complex exchanges occurring in stages up to maturity [12, p.10].

There are various reasons for trading in futures. Speculation on the direction of the price of the underlying is one possibility and hedging one's risk is another. We talk about hedging the position in a derivative if we determine the initial capital and the portfolio process $\Delta(t)$ so that the portfolio value at the exercising time T is exactly the payout of the option [35, p.218].

A farmer who fears bankruptcy due to an unexpected drop in the price of wheat, may, for example, enter a futures contract permitting him to sell his wheat at a predetermined future date at a predetermined price. The disadvantage, though, is that the price of wheat might go up and he will still only get the price that has been agreed upon.

2.3.2. Options

A call option is a contract between two parties, giving the buyer the right but not the obligation to obtain an agreed upon quantity of the underlying asset at a fixed predetermined value K , called the strike price. This leads to a payoff of $(S_t - K)^+$ when we stop at the time t , where the $+$ means that we only exercise when the payoff is non negative. A put option, however, gives the buyer the right but not the obligation to sell an agreed-upon quantity of the underlying asset according to the amount of the strike price K for each asset. Hence, the payoff at t is $(K - S_t)^+$. Cvitanic and Zapatero. point out [12, p.16] that hedging one's risk by obtaining an option offers an interesting investment possibility. Possessing a stock and a put option on the stock is equivalent to a portfolio where the loss in the stock price is limited.

Vanilla Options

The two standard styles of options, often referred to as plain Vanilla options, are European and American options in both the put and the call case. The style of the option defines when the option can be exercised. European-style options can only be exercised at maturity, while in the case of the American-style option the owner has the right to exercise at any time prior and up to the expiring date [12, p.13].

Shiryaev points out in [34, p.20] that with the growth of the financial market also the number of types of options has highly increased over the past decades. They are used to hedge the risk of portfolios and to speculate. Since American-style options can be exercised more freely, they are also a lot more flexible than European-style options and therefore it is crucial to be able to determine the value of an American option as accurately as possible.

In order to allow further conclusions concerning American options, the following definitions, taken from Shreve [35, p.51, 53, 74], are needed:

Definition 2.1 Assume Ω is a non empty set and T a fixed positive number. Further suppose that for every $t \in [0, T]$ exists a σ -algebra $\mathcal{F}(t)$ and that for each $s \leq t$ every set in

2. Financial Markets

$\mathcal{F}(s)$ is also in $\mathcal{F}(t)$. We will call the collection of σ -algebras $(\mathcal{F}(t))_{0 \leq t \leq T}$ a filtration.

The σ -algebra $\mathcal{F}(t)$ represents the information available at time t .

Definition 2.2 Consider a probability space (Ω, \mathcal{F}, P) and a filtration $(\mathcal{F}(t))_{0 \leq t \leq T}$. A stochastic process $X(t)$ is called adapted if $X(t)$ is $\mathcal{F}(t)$ -measurable for all $0 \leq t \leq T$.

This means that the information contained in the σ -algebra $\mathcal{F}(t)$ is sufficient to calculate the value of X at t .

Definition 2.3 Consider the probability space (Ω, \mathcal{F}, P) and the filtration $(\mathcal{F}(t))_{0 \leq t \leq T}$, where T is a fixed positive number. An adapted and integrable stochastic process $(M(t))_{0 \leq t \leq T}$ is called a martingale if

$$E[M(t) \mid \mathcal{F}(s)] = M(s) \quad \text{for all } 0 \leq s \leq t \leq T \quad (2.1)$$

A martingale has neither a tendency to rise nor to fall in the mean.

An adapted stochastic process $(M(t))_{0 \leq t \leq T}$ is called a submartingale if

$$E[M(t) \mid \mathcal{F}(s)] \geq M(s) \quad \text{for all } 0 \leq s \leq t \leq T \quad (2.2)$$

A submartingale has no tendency to fall and might even rise in the mean.

Analogously, an adapted stochastic process $(M(t))_{0 \leq t \leq T}$ is called a supermartingale if

$$E[M(t) \mid \mathcal{F}(s)] \leq M(s) \quad \text{for all } 0 \leq s \leq t \leq T \quad (2.3)$$

A supermartingale has no tendency to rise and might even fall in the mean.

In the case of a non-dividend paying American call option it is optimal to wait until maturity to decide whether to exercise or not. The reason for this is that the payoff function $(e^{-rt}(S(t) - K)^+)_{t \geq 0}$ is a submartingale under the risk neutral measure and thus rises in expectation. The submartingale property of the payoff function derives from $(e^{-rt}S(t))_{t \geq 0}$ being a martingale and $e^{-rt}K$ non-increasing as t increases, since $r > 0, t \geq 0$. Jensen's inequality used on the convex function $(e^{-rt}(S(t) - K)^+)$ just increases the upward trend. Hence, the expected payoff will be higher the later the option is exercised and therefore the price coincides with the price of a European call option on the same underlying and the same expiration date.

This argument does not work for the American put option, though, since the payoff function $(e^{-rt}(K - S(t))^+)_{t \geq 0}$ is a supermartingale. Applying Jensen's inequality in this case creates again an upward trend interfering with the supermartingale property [35, p. 363].

Basket Options

The difference between a Basket option and a Vanilla option is that, in the case of the Basket option, the underlying is driven by the (weighted) sum of multiple assets instead of a single option, as described by Asmussen and Glynn in [2, p.9].

2.3.3. Swaps

A contract, in which two parties agree on an exchange of cash flow with different features is called a swap. A swap can be imagined as exchanging interest rates on two different types of bonds. Generally only the interest rate payments are exchanged and not the principal. The most common swap is, where one party pays a fixed interest rate in exchange for a floating interest rate from the other party. We talk about fixed interest rates when the payments are predetermined and constant, whereas floating interest rates involve that after each payment the following payment is reset according to a rule that has previously been agreed upon [12, p.18]. Frequently the exchanged interest rates correspond to bonds denominated in different currencies. To be able to distinguish them, swaps that exchange only the interest rate are called interest-rate swaps and the ones also exchanging the currency are called currency swaps. The party that receives the floating interest rate is the buyer of the swap. Once again major interest when purchasing a swap is in speculation or hedging one's risk [12, p.18].

2.4. Market Models

Whenever working with a mathematical model, researchers have to balance computational performance and the adjustment of the model to reality. It can be helpful to work with extremely simple models at the beginning to get a better understanding of the qualitative behaviour of the modelled objects. Many basic conclusions obtained by studying simple models can then be translated into more sophisticated models [12, p.53].

The reason for using continuous-time models to approximate discrete trading is mainly mathematical convenience. With a small number of parameters it is already possible to model complex price dynamics. Differential calculus, only available in continuous-time, permits to find explicit solutions for many standard pricing and investment problems [12, p.62].

2.4.1. Black - Scholes Model

In order to describe the Black - Scholes model the definition of a Brownian motion, as stated in [35, p.94], is necessary:

Definition 2.4 *Consider the filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}(t))_{t \geq 0}, P)$. Further suppose that for each $\omega \in \Omega$ there exists a continuous function $W(t)$ for $t \geq 0$ with $W(0) = 0$ depending on ω . Then $W(t)$ is called a Brownian motion if for all $0 = t_0 < t_1 < \dots < t_m$ the increments*

$$W(t_1) - W(t_0), W(t_2) - W(t_1), \dots, W(t_m) - W(t_{m-1}) \quad (2.4)$$

are independent and each of these increments is normally distributed with

$$E[W(t_{i+1}) - W(t_i)] = 0 \quad (2.5)$$

$$Var[W(t_{i+1}) - W(t_i)] = t_{i+1} - t_i. \quad (2.6)$$

2. Financial Markets

The first to use Brownian motion to model stock markets was Louis Bachelier in 1900, but his idea was largely ignored by the other mathematicians and economists of his time. In the 1950s Brownian motion was finally introduced to the economic literature by Paul Samuelson. Robert C. Merton, Samuelson's student, developed the model further and in 1973 Fischer Black and Myron Scholes published a paper responsible for the great success of the model in the following years.

Since the model uses only a small number of parameters the price dynamics are greatly simplified. A bigger number of parameters would allow us to match the model and the available data better, but at the same time the parameter estimates are less precise. Another problem is the so-called "overfitting", where the model fits the existing data very well, but does not work with future outcomes of the underlying process [12, p.74].

In the Black - Scholes model we are given a risk-free asset B referred to as a bond or a bank account, possessing the following dynamics

$$dB(t) = rB(t)dt, \quad B(0) = 1 \quad (2.7)$$

where $r \geq 0$ is the constant interest rate. It is called risk-free because the payoff is pre-determined, since it does not have a stochastic component.

Additionally, we have a risky security, called a stock, driven by

$$dS(t) = \mu S(t)dt + \sigma S(t)dW(t), \quad S(0) = x \quad (2.8)$$

whereas μ is the long-term development of the asset price, referred to as drift, and σ determines how volatile the asset price is. Both variables are kept constant in the standard Black - Scholes model and there are also no dividends being paid. Solving the equation (2.8) the stock price at t can be written as

$$S(t) = x \exp \left\{ \sigma W(t) + \left(\mu - \frac{1}{2} \sigma^2 \right) t \right\}, \quad (2.9)$$

where $S(0)$ is the initial stock value. $S(t)$ changes according to a lognormal distribution, since the logarithm of $S(t)$ is normally distributed. The other assumptions imposed by Black and Scholes in [5, p.640] on the model are

1. No transaction costs or taxes
2. No restrictions on short selling

Short selling a stock consists in selling a borrowed stock, owned by someone else, and returning it at a later time [12, p.9].

The limitations to the model, coming from the small number of parameters, are due to the fact that the user is exposed to unexpected risks of all kinds. Firstly, extreme movements of the asset price are underestimated by the underlying normal distribution leading to so-called tail risks, discussed by Cont and Tankov in [11, p.5]. The assumption of instant and costless trading yielding liquidity risk, which is the risk that the asset cannot be traded quickly enough

2. Financial Markets

to prevent loss. Other problems are the assumptions of continuous trading, leading to gap risk, as well as the assumption of a stationary process yielding volatility risk.

In addition to the aforementioned problems the Black - Scholes model has the problem that the trajectories of the Brownian motion $t \mapsto W(t)$, contrary to the actual evolution of log prices, are continuous in t and therefore neglect abrupt movements in the price.

Relevant for the investor are also the returns (i.e. the increments of the log-prices), where the increments of the Brownian motion $W(t)$ have, in fact, the same variance as the empirical returns. The difference, though, is that the amplitude for the Brownian motion increments is roughly invariant over time, whereas the returns, observed in the market, exhibit frequent large peaks in the amplitude, corresponding to the jumps in the price and leading to the tail risks mentioned above [11, p.2-7].

2.4.2. Exponential Lévy Model

The exponential Lévy model is an approach to take the abrupt changes in price into account. It is the simplest Markov model with jumps leading to highly variable returns with realistic tail behaviour without non-stationarity, choosing extreme parameters for the volatility coefficient or adding unobservable random factors [11, p.6]. Mandelbrot was the first to introduce (non-Gaussian) Lévy processes to financial econometrics in [25], when he proposed stable Lévy processes as models for cotton prices.

The following definition is taken from Cont and Tankov [11, p.68]:

Definition 2.5 *A cadlag (i.e. right continuous with left limits) stochastic process $(X(t))_{t \geq 0}$ on (Ω, \mathcal{F}, P) with values in \mathbb{R}^d such that $X(0) = 0$ is called a Lévy process if it possesses the following properties:*

1. *Independent increments: For every increasing sequence of times t_0, t_1, \dots, t_n the random variables $X(t_0), X(t_1) - X(t_0), \dots, X(t_n) - X(t_{n-1})$ are independent.*
2. *Stationary increments: The law of $X(t+h) - X(t)$ depends only on h and not on t .*
3. *Stochastic continuity: $\forall \epsilon > 0 : \lim_{h \rightarrow 0} P(|X(t+h) - X(t)| \geq \epsilon) = 0$.*

Consider $(\Omega, \mathcal{F}, (\mathcal{F}(t))_{0 \leq t \leq T}, P)$ with finite time horizon T , $L(t)$ a Lévy process and the exponential model for the asset price dynamics

$$S(t) = S(0) \exp(L(t)) \quad (2.10)$$

Along with Benth et al. [3] and Rydberg [32] we will confine ourselves to the so-called NIG-Lévy processes, where the increments of the Lévy process follow a normal inverse Gaussian distribution.

The random variable $X \sim NIG(\alpha, \beta, \mu, \delta)$ is said to be normal inverse Gaussian distributed with the parameters α, β, μ and δ if the density can be written as

$$p(x; \alpha, \beta, \mu, \delta) = \frac{\delta \alpha}{\pi} \exp \left\{ \delta \sqrt{(\alpha^2 - \beta^2)} + \beta(x - \mu) \right\} \frac{K_1(\alpha \sqrt{\delta^2 + (x - \mu)^2})}{\sqrt{(\delta^2 + (x - \mu)^2)}}, \quad (2.11)$$

2. Financial Markets

where $x \in \mathbb{R}$, $\mu \in \mathbb{R}$, $\delta > 0$, $0 \leq |\beta| \leq \alpha$ and K_1 is the modified Bessel function of the third kind with index 1

$$K_1(y) = \frac{y}{4} \int_0^\infty \frac{1}{t^2} \exp \left[- \left(t + \frac{y^2}{4t} \right) \right] dt, \quad y \in \mathbb{R}. \quad (2.12)$$

To point out the advantages and shortcomings of representing the log prices as Lévy processes, we oppose, in the same manner as in Cont and Tankov [11, p.227], the empirical properties of asset returns with statistical properties of NIG-Lévy processes.

Therefore we need to explain the following properties satisfied by the log prices, observed in the market.

Absence of autocorrelation: The (linear) autocorrelation of asset returns are often insignificant except on an intraday time scale.

Semi - heavy tails: the distribution of returns displays a semi - heavy tail.

Volatility clustering: If a large change in the log price occurs there is a tendency that it is followed by a large change of either sign. Likewise small changes are often followed by small changes [25]. This leads to autocorrelation functions for the absolute returns, or their square, that are positive, significant and slowly decaying, although the returns themselves are uncorrelated.

Log - prices	NIG-Lévy processes
Absence of autocorrelation in increments	Satisfied by NIG-Lévy processes
Semi - heavy tails	Satisfied by NIG-Lévy processes
Jumps in price trajectories	Satisfied by NIG-Lévy process
Volatility clustering	Not satisfied since the increments are independent
Positive autocorrelation in absolute returns	Not satisfied since the increments are independent

Table 2.1.: Advantages and disadvantages of NIG-Lévy processes

3. Survey on Pricing American Options

In this chapter we will discuss various approaches to pricing American put options. The reason for omitting call options is that the price of an American call option is equivalent to its European pendant as already mentioned in 2.3.2.

To price an American put option it is necessary to find the optimal exercising time, as pointed out in chapter 4. Hence, by solving the optimal stopping problem, an optimal stopping rule can be obtained. This chapter discusses several approaches on pricing American options and examines their strengths and weaknesses.

With the emergence of computers in the last decades also sampling-based computational methods have become increasingly popular as a tool for researchers as well as practitioners in various areas ranging from statistics, finance, probability over economics and operations research to biology, chemistry and physics. Therefore the number of books and articles dealing with simulation and Monte Carlo methods is vast. Let us just mention some of the classical references as Hammersley & Handscomb [20], Rubinstein [29], Ripley [28], Fishman [15] or Glasserman [19].

Recently financial mathematics has become one of the main fields of application for stochastics, due to the fact that they are based on a well-established theory, but with upcoming challenges from a theoretical as well as from a computational point of view [2, p.6]. The pricing of options, where the different developments of the underlying are simulated, is a typical example. Practically it is not possible to take every path of the underlying into account and thus a predetermined number of paths are generated using stochastic simulation.

Unless stated differently, the base reference for this survey is the chapter about pricing American options in Glasserman [19], where a very detailed summary can be found. The main concept is that all simulation methods for pricing American options are affected by sources of bias. The estimation of the price is biased high if information about the future is included in the process of deciding when to exercise. Low bias, however, is a result of using a suboptimal exercising rule.

The following definition, taken from [26, p.27], is important for further conclusion:

Definition 3.1 *A stopping time is a measurable function $\tau : \Omega \rightarrow [0, \infty]$, that satisfies $\tau < \infty$ P – a.s. and for all sets of the form $\{\tau \leq t\}$ we have $\{\tau \leq t\} \in \mathcal{F}(t)$ for all $t \geq 0$.*

As discussed more extensively in chapter 4, the value of the American put option can be described by

$$V(x) = \sup_{\tau} E_x [e^{-r\tau} (K - X(\tau))^+], \quad (3.1)$$

3. Survey on Pricing American Options

where the expectation is taken under the risk neutral measure. The supremum considered in (3.1) is taken over all admissible stopping times in $[0, T]$, stopping times that take values between 0 and the maturity T and are adapted to the natural filtration of X given by $\mathcal{F}_t^X = \sigma(X_s : 0 \leq s \leq t)$ for $t \geq 0$. We suppose that the stochastic process $(X(t))_{t \geq 0}$ is an \mathbb{R}^d -valued Markov process recording the information on the prices of the underlying asset and all the other financial variables. To obtain the Markov property it may be necessary to augment the state vector to include variables such as the stochastic volatility. The supremum in (3.1) is obtained by an optimal stopping time

$$\tau = \inf \{t \geq 0 : X_t \leq b(t)\} \quad (3.2)$$

for some optimal exercising boundary $b(t)$, as pointed out in Glasserman [19, p.422]. Properties of the optimal exercising boundary will be discussed in more detail in section 4.7.1.

Working with a finite number of exercise times m , as necessary in computer simulations, a characterization of the American option value through dynamic programming comes by naturally. Denoting the payoff function at the exercise date t_i with h_i , we obtain the following recursion

$$V_m(x) = h_m(x) \quad (3.3)$$

$$V_{i-1}(x) = \max \{h_{i-1}(x), E[V_i(X_i) | X_{i-1} = x]\}, \quad i = 1, \dots, m, \quad (3.4)$$

where the payoff function is already augmented to contain the discount factor. The first equation points out that at maturity the option is exercised immediately. Should the payoff be negative, the option is not exercised and the payoff is 0. At all other time steps t_{i-1} the decision to exercise depends on the continuation value

$$C_{i-1}(x) = E_x[V_i(X_i) | X_{i-1} = x] \quad \text{for } i = 1, \dots, m-1. \quad (3.5)$$

If at t_{i-1} the continuation value is larger than the immediate payoff h_{i-1} the option is not exercised at t_{i-1} . If the immediate payoff h_{i-1} at t_{i-1} exceeds the continuation value C_{i-1} , however, the option is exercised.

3.1. Parametric Approximations

Rather than finding the best approximation for the value of the American put option, this method contents itself with finding the most accurate value within a parametric class of stopping rules τ_θ , $\theta \in \Theta$. Thereby the optimal stopping problem is reduced to an optimization problem, that can be solved more easily and written as

$$V_0^\theta = \sup_{\theta \in \Theta} E[h_{\tau(\theta)}(X_{\tau(\theta)})] \quad (3.6)$$

Since the stopping times used are a subclass of all the admissible stopping times in (3.1) it follows that the optimal value, within the parametric class, V_0^θ satisfies

$$V_0^\theta \leq V_0, \quad (3.7)$$

3. Survey on Pricing American Options

where V_0 denotes the true value of the American put and thereby that a consistent estimator of V_0^θ underestimates V_0 .

The subsequent algorithm produces an estimator with a low bias

1. Simulate n_1 samples $\mathbf{X}^{(j)} = (X_0^{(j)}, X_1^{(j)}, \dots, X_m^{(j)})$, $j = 1, \dots, n_1$ of the Markov chain.
2. Find the parameter $\hat{\theta}$ maximizing the following expression

$$\hat{V}_0^{\hat{\theta}} = \frac{1}{n_1} \sum_{j=1}^{n_1} h_{\tau^{(j)}(\hat{\theta})} \left(X_{\tau^{(j)}(\hat{\theta})}^{(j)} \right), \quad (3.8)$$

whereas $\tau^{(j)}(\theta)$ denotes the exercising time of the j th replication for the parameter θ .

3. Generate another n_2 independent replications of the Markov Chain (X_0, X_1, \dots, X_m) . Use the stopping rule $\tau(\hat{\theta})$ found in 2. to calculate

$$\hat{V}_0^{\hat{\theta}} = \frac{1}{n_2} \sum_{j=n_1+1}^{n_1+n_2} h_{\tau^{(j)}(\hat{\theta})} \left(X_{\tau^{(j)}(\hat{\theta})}^{(j)} \right) \quad (3.9)$$

To determine the in-sample maximum $\hat{\theta}$ in step 2, the future development of the underlying asset price is taken into account. Thus the estimator in step 2 is biased high. Together with the low bias obtained by using a suboptimal stopping time, the algorithm produces an estimator with unpredictable bias. To avoid that, an additional simulation step is added, where the stopping rule from step 2 is used on an independent sample offsetting the high bias in step 2 and resulting in a low biased estimator.

The most difficult part of this algorithm and therefore focus for optimization, is to find a parameter maximizing step 2. Within the last 15 years there have been various distinct approaches, ranging from the reduction of the optimization problem in step 2 to a sequence of one-dimensional searches as the ones by Andersen in [1], over the estimation of derivatives with respect to the parameters in order to find an optimal parameter, like Fu and Hu in [17], up to the description of the exercise region at each time step by two parameters like Garcia in [18].

3.2. Random Tree Methods

The strength of this algorithm is that it merely assumes more than the ability to simulate paths of the underlying Markov chain. As result we obtain a high as well as a low biased estimator, each converging to the true value of the American option. The major drawback of this method, though, is that the computational effort increases exponentially with the number of exercise dates making it unsuitable for simulations with more than 5 exercise dates.

Having two estimators, one converging to the true value from above and one from below, permits the creation of a confidence interval. Assume that $\hat{V}_n(c)$ and $\hat{v}_n(c)$, where c is the branching parameter, are sample means of n independent replications of the Markov

3. Survey on Pricing American Options

chain and as estimators of V_0 biased high and low respectively. By generating a confidence interval for $E(\hat{V}_n(c))$ with halfwidth $H_n(c)$,

$$\hat{V}_n(c) \pm H_n(c) \quad (3.10)$$

and one for $E(\hat{v}_n(c))$ with halfwidth $L_n(c)$ for $\hat{v}_n(c)$ respectively, a confidence interval for V_0 can be constructed as follows

$$(\hat{v}_n(c) - L_n(c), \hat{V}_n(c) + H_n(c)) \quad (3.11)$$

By increasing the number c of successor nodes as well as the number of independent replications n , the confidence interval can be tightened at the cost of computational effort.

By means of this method the high estimator is acquired through applying backward induction to the random tree. The tree structure is obtained by choosing a branching parameter $c \geq 2$ which determines the number of successor nodes for each node, as can be observed in Figure 3.1. All successor nodes of $X^i(t)$ are generated by the conditional law of X_{t+1} given $X(t) = X^i(t)$.

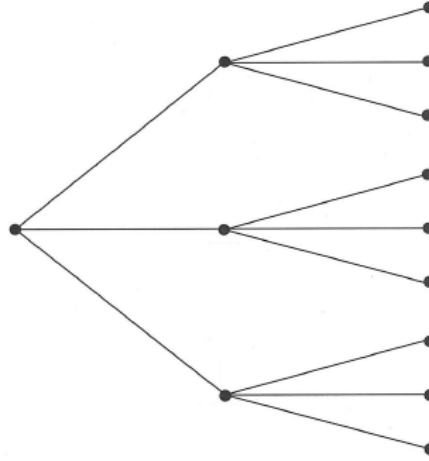


Figure 3.1.: Random Tree with branching parameter $c = 3$

Using the backward induction recursion, the discounted value of the American option is

$$V_m \equiv h_m \quad (3.12)$$

$$V_i(x) = \max \{ h_i(x), E[V_{i+1}(X_{i+1}) \mid X_i = x] \}, \quad i = 1, \dots, m-1 \quad (3.13)$$

Considering the structure of the tree, the high estimator at each node $X_i^{j_1 \dots j_i}$ ($j_1 \dots j_i$ denotes the path in the tree from the initial state on) is

$$\hat{V}_m^{j_1 \dots j_m} = h_m(X_m^{j_1 \dots j_m}) \quad (3.14)$$

3. Survey on Pricing American Options

for the terminal nodes and

$$\hat{V}_i^{j_1 \dots j_i} = \max \left\{ h_i(X_i^{j_1 \dots j_i}), \frac{1}{c} \sum_{j=1}^c \hat{V}_{i+1}^{j_1 \dots j_i j} \right\} \quad (3.15)$$

for the rest of the nodes, working backwards through the tree. To construct a confidence interval we fix the branching parameter c and produce n replicates of the random tree. The sample mean $\bar{V}_0(n, c)$ of \hat{V}_0 as well as the sample standard deviation $s_V(n, c)$ are obtained by using the n replications of the tree. This way we can write the $(1 - \delta)$ confidence interval for $E(\hat{V}_0)$ as

$$\bar{V}_0(n, c) \pm z_{\delta/2} \frac{s_V(n, c)}{\sqrt{n}}, \quad (3.16)$$

whereas $z_{\delta/2}$ is the $(1 - \frac{\delta}{2})$ quantile of the normal distribution.

As in all Monte Carlo methods the high bias is removed by separating the decision to exercise from the value of continuation.

One way to achieve this is to use the following estimators suggested in Broadie and Glasserman [7].

$$\hat{v}_m^{j_1 j_2 \dots j_m} = h_m(X_m^{j_1 j_2 \dots j_m}), \quad (3.17)$$

$$\hat{v}_i^{j_1 j_2 \dots j_i} = \frac{1}{c} \sum_{k=1}^c \hat{v}_{ik}^{j_1 j_2 \dots j_i}, \quad (3.18)$$

with

$$\hat{v}_{ik}^{j_1 j_2 \dots j_i} = \begin{cases} h_i(X_i^{j_1 j_2 \dots j_i}), & \text{if } \frac{1}{c-1} \sum_{j=1, j \neq k}^c \hat{v}_{i+1}^{j_1 j_2 \dots j_i j} \leq h_i(X_i^{j_1 j_2 \dots j_i}) \\ \hat{v}_{i+1}^{j_1 j_2 \dots j_i k}, & \text{else.} \end{cases} \quad (3.19)$$

Sample mean $\bar{v}_0(n, c)$ and sample standard deviation $s_v(n, c)$ can be acquired in the same manner as for the high estimator resulting in a confidence interval

$$\bar{v}_0(n, c) \pm z_{\delta/2} \frac{s_v(n, c)}{\sqrt{n}} \quad (3.20)$$

for $E(\hat{v}_0)$. Thereby a confidence interval for V_0 can be constructed like in (3.11).

Optimization techniques for the Random Tree method range from Depth First Processing implementation, over reduction of the needed storage space up to Pruning and Variance Reduction techniques used by Broadie et al. in [8].

3.3. State Space Partitioning

Contrary to the random tree method, in this method the states are not generated randomly, but the whole state space of the underlying Markov chain (X_0, X_1, \dots, X_m) is partitioned in advance. At each possible exercising time t_i the state space of X_i is partitioned into

3. Survey on Pricing American Options

b_i subsets A_{i1}, \dots, A_{ib_i} , starting with $b_0 = 1$ and $A_{01} = \{X_0\}$ at the initial time 0. Then transition probabilities are defined

$$p_{jk}^i = P(X_{i+1} \in A_{i+1,k} \mid X_i \in A_{ij}), \quad (3.21)$$

for all $j = 1, \dots, b_i$, $k = 1, \dots, b_{i+1}$ and $i = 0, \dots, m-1$, where the transition probability is 0 if $P(X_i \in A_{ij}) = 0$. Thereby the payoff at each state at time t_i can be calculated for each $i = 1, \dots, m$ and $j = 1, \dots, b_i$ by

$$h_{ij} = E[h_i(X_i) \mid X_i \in A_{ij}] \quad (3.22)$$

Hence, the option price can be estimated by V_{01} , following the backward recursion

$$V_{ij} = \max \left\{ h_{ij}, \sum_{k=1}^{b_{i+1}} p_{jk}^i V_{i+1,k} \right\}, \quad (3.23)$$

starting with $V_{mj} = h_{mj}$ for all $i = 0, 1, \dots, m-1$, $j = 1, \dots, b_i$.

By simulating a large number of replications of the underlying Markov chain (X_0, X_1, \dots, X_m) the transition probabilities p_{jk}^i and average payoffs h_{ij} can be estimated. Hence the estimate for V_{ij} can be defined as follows

$$\hat{V}_{ij} = \max \left\{ \hat{h}_{ij}, \sum_{k=1}^{b_{i+1}} \hat{p}_{jk}^i \hat{V}_{i+1,k} \right\}, \quad (3.24)$$

for $j = 1, \dots, b_i$, $i = 0, 1, \dots, m-1$. Accordingly, \hat{V}_{01} is the estimate for V_{01} . The strong law of large numbers ensures that, by increasing the number of replications of the Markov chain, the estimates \hat{p}_{jk}^i and \hat{h}_{ij} converge to p_{jk}^i and h_{ij} . Since the mapping from \hat{p}_{jk}^i and \hat{h}_{ij} to \hat{V}_{ij} is continuous, it follows that \hat{V}_{01} converges to V_{01} , resulting in a consistent estimator.

As in the case of the parametric approach before, an additional simulation step is needed to obtain an estimator that is biased low, otherwise no conclusions about the bias of the estimator could be obtained. Again the first simulation phase is used to receive a stopping rule

$$\tau = \min \{0 \leq i \leq m : h_i(X_i) \geq V_{iJ_i}\}, \quad (3.25)$$

where J_i is the subset at t_i containing X_i . In the case that the inequality is never satisfied τ is m . Using this stopping rule at the second simulation phase it is ensured that the estimator will be biased low, because no stopping rule exceeds the payoff of the optimal stopping rule. The considerable drawback of this approach is that the partitioning of the state space, main challenge of this method, has to be fitted to the form of the payoff and therefore cannot be applied to numerous different American options. Also the a priori finding of partitions is not convenient for high-dimensional state spaces as the computational requirement increases exponentially with the dimension of the state space.

3.4. Stochastic Mesh Methods

The stochastic mesh method is another Monte Carlo approach to approximating the price of an American put option. Main difference to the Random Tree method is that the estimator for the option value at each node at time step t_i uses all successive states from time step t_{i+1} instead of just a subset. The method's name derives directly from its major advantage over the Random Tree method. By using all states at the consecutive time step a mesh structure is obtained. This way the number of states at each time step is kept constant and the exponential growth making the random tree method inefficient is omitted.

A simple way to construct the mesh is to sample a number of independent paths of the underlying Markov chain (X_0, X_1, \dots, X_m) with the same starting point. Then each node at each time step and all nodes at the consecutive time step are connected, leading to the mesh structure. Thereafter weights W_{jk}^i (i.e. Likelihood ratio weights) are allocated to each connection, giving the probability that being at X_{ij} , which denotes the j th node at time t_i , the next state will be $X_{i+1,k}$. Figure 3.2, taken from Glasserman [19, p.444], demonstrates the construction of the stochastic mesh.

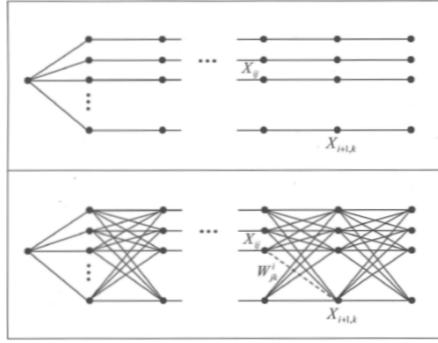


Figure 3.2.: Construction of the stochastic mesh

The actual pricing is similar to the Random Tree method. First the value at the terminal nodes at time step m is set to the associated payoff

$$\hat{V}_{mj} = h_m(X_{mj}) \quad \forall j = 1, \dots, b, \quad (3.26)$$

where b is the number of states at each time step. Then moving backwards through the mesh the estimator is obtained by

$$\hat{V}_{ij} = \max \left\{ h_i(X_{ij}), \frac{1}{b} \sum_{k=1}^b W_{jk}^i \hat{V}_{i+1,k} \right\} \quad (3.27)$$

The value at the root node is defined by the formula

$$\hat{V}_0 = \frac{1}{b} \sum_{k=1}^b \hat{V}_{1k}, \quad (3.28)$$

3. Survey on Pricing American Options

assuming that it is not possible to exercise at t_0 , otherwise the maximum over (3.28) and $h_0(X_0)$ has to be considered. The way the value at each node in (3.27) is determined looks superficially very much like the one in the state space partitioning method, but the major difference is that nodes used in the recursion are randomly sampled, whereas in the former method they are fixed subsets of the state space.

The following 3 conditions, imposed on the mesh, provide all necessities for an inductive argument showing that the estimator \hat{V}_0 from (3.28) is biased high.

1. $\{\mathbf{X}_0, \dots, \mathbf{X}_{i-1}\}$ and $\{\mathbf{X}_{i+1}, \dots, \mathbf{X}_m\}$ are independent, given \mathbf{X}_i for all $i = 1, \dots, m - 1$, where $\mathbf{X}_i = (X_{i1}, \dots, X_{ib})$ consists of all nodes at time step t_i , representing the state of the mesh at that time. This condition is naturally satisfied when the mesh is constructed by the independent path method mentioned above.
2. Each weight W_{jk}^i is a deterministic function of \mathbf{X}_i and \mathbf{X}_{i+1} .
3. $\frac{1}{b} \sum_{k=1}^b E[W_{jk}^i V_{i+1}(X_{i+1,k}) | \mathbf{X}_i] = C_i(X_{ij})$, implying that if the true option values are known at t_{i+1} , the true continuation value at time step t_i can be determined this way.

As in the previous methods the low estimator for the option price is obtained by using the strategy to exercise the option the first time the immediate payoff equals or exceeds the value of continuation. To be able to define the continuation value though, the weights W_{jk}^i at each step t_i have to be extended from X_{i1}, \dots, X_{ib} to every point in the state space, resulting in a weight function $W_k^i(x)$. Then the continuation value can be defined as

$$\hat{C}_i(x) = \frac{1}{b} \sum_{k=1}^b W_k^i(x) \hat{V}_{i+1,k} \quad (3.29)$$

Now the aforementioned strategy can be written as the stopping rule

$$\hat{\tau} = \min \left\{ i : h_i(X_i) \geq \hat{C}_i(X_i) \right\} \quad (3.30)$$

Simulating a new path, independent from the ones used to construct the mesh, this stopping rule yields the estimator

$$\hat{v} = h_{\hat{\tau}}(X_{\hat{\tau}}) \quad (3.31)$$

for the true option price. Since no stopping rule can be better than the optimal one, the estimator is indeed biased low.

Main drawback of this method is that the transition densities, necessary for the maximum likelihood weights, may not exist. To by-pass this problem the weights can be selected by solving a constrained optimization problem as was done by Broadie et al. in [9].

3.5. Regression-Based Methods

The aim of this method is to estimate the continuation value through regression. A big advantage of this method is that it can be applied to many different settings and is relatively

3. Survey on Pricing American Options

fast. It is also the method, where knowledge as well as intuition can be exploited to the full extent, when it comes to choosing the regression functions, which are responsible for the accuracy of the algorithm. Therefore the success of the algorithm is strongly dependent on the choice of these basis functions.

The regression method can be easily used in combination with a stochastic mesh framework, where it is used to determine the mesh weights. There the continuation value is written as a linear combination of basis functions $\psi_r : \mathbb{R}^d \mapsto \mathbb{R}$

$$\mathbb{E}[V_{i+1}(X_{i+1}) \mid X_i = x] = \sum_{r=1}^M \beta_{ir} \psi_r(x), \quad (3.32)$$

where β_{ir} are constants for all $r = 1, \dots, M$. Equivalently, the equation (3.32) can be written as

$$C_i(x) = \beta_i^\top \psi(x), \quad (3.33)$$

where $\beta_i^\top = (\beta_{i1}, \dots, \beta_{iM})$ and $\psi(x) = (\psi_1(x), \dots, \psi_M(x))^\top$. Under the assumption that the equation (3.32) holds, the vector β_i has the form

$$\beta_i = (\mathbb{E}[\psi(X_i)\psi(X_i)^\top])^{-1} \mathbb{E}[\psi(X_i)V_{i+1}(X_{i+1})] \equiv B_\psi^{-1} B_{\psi V} \quad (3.34)$$

Supposing that the values $V_{i+1}(X_{i+1,j})$ are known for $i = 1, \dots, b$, the least square estimator of β_i is given by

$$\hat{\beta}_i = \hat{B}_\psi^{-1} \hat{B}_{\psi V}, \quad (3.35)$$

whereas \hat{B}_ψ and $\hat{B}_{\psi V}$ represent the sample counterparts of the aforementioned (3.34). In practice, though, the value V_{i+1} also needs to be estimated by \hat{V}_{i+1} . The estimate of the continuation value then becomes

$$\hat{C}_i(x) = \hat{\beta}_i^\top \psi(x) \quad (3.36)$$

The whole algorithm to price American options can therefore be written in the following steps

1. Simulate b independent paths of the underlying Markov chain
2. Set value at terminal nodes to $\hat{V}_{mj} = h_m(X_{mj})$ for $j = 1, \dots, b$
3. Use backward induction for $i=m-1, \dots, 1$
 - a) Use regression to determine the estimators $\hat{\beta}_i$
 - b) Calculate $\hat{V}_{ij} = \max \{h_i(X_{ij}), \hat{C}_i(X_{ij})\}$, $j = 1, \dots, b$
4. Set $\hat{V}_0 = \frac{1}{b} \sum_{k=1}^b \hat{V}_{1k}$

The low estimator is once more obtained by following the stopping strategy to exercise when the immediate payoff is at least as high as the continuation value, obtained before from an independent simulation.

In Egorova and Ivanova [14] further numerical methods are described in detail.

4. Methods

4.1. Risk-Neutral Pricing in the Black - Scholes Model

In this section we want to discuss the basic conditions necessary to acquire a unique price for every derivative security in the standard Black - Scholes model. In order to price a derivative security we have to change the considered probability measure to the risk neutral measure.

Theorem 4.1 *The Itô integral*

$$\int_0^t \Delta(u) dW(u), \quad 0 \leq t \leq T \quad (4.1)$$

is a martingale [35, p.134].

Theorem 4.2 *(Girsanov)*

Suppose we are given a probability space (Ω, \mathcal{F}, P) , $(W(t))_{0 \leq t \leq T}$ is a Brownian motion and $(\mathcal{F}(t))_{0 \leq t \leq T}$ a filtration for this Brownian motion. Let $(\Theta(t))_{0 \leq t \leq T}$ be an adapted process and define further

$$Z(t) = \exp\left\{-\int_0^t \Theta(u) dW(u) - \frac{1}{2} \int_0^t \Theta^2(u) du\right\} \quad (4.2)$$

$$\tilde{W}(t) = W(t) + \int_0^t \Theta(u) du \quad (4.3)$$

Assuming that

$$E \int_0^T \Theta^2(u) Z^2(u) du < \infty \quad (4.4)$$

and setting $Z = Z(T)$ we obtain: $EZ = 1$ and under the probability measure

$$\tilde{P}(A) = \int_A Z(\omega) dP(\omega), \quad \text{for all } A \in \mathcal{F} \quad (4.5)$$

the process $(\tilde{W}(t))_{0 \leq t \leq T}$ is a Brownian motion.

We will refer to \tilde{P} as the risk neutral measure, since it turns the discounted expected payoff of any derivative into a martingale. Further, since Z satisfies per definition $P(Z > 0) = 1$ the probability measures P and \tilde{P} agree on which sets have probability 0 - i.e. which asset price developments are possible [35, p.210].

4. Methods

4.1.1. Stock and Portfolio under the Risk-Neutral Measure

Let us consider the usual price process dynamic

$$dS(t) = \mu(t)S(t)dt + \sigma(t)S(t)dW(t), \quad 0 \leq t \leq T \quad (4.6)$$

and the discount process

$$D(t) = \exp \left\{ - \int_0^t R(s)ds \right\} \quad (4.7)$$

then the discounted price process $D(t)S(t)$ is

$$D(t)S(t) = S(0) \exp \left\{ \int_0^t \sigma(s)dW(s) + \int_0^t (\mu(s) - R(s) - \frac{1}{2}\sigma(s)^2)ds \right\} \quad (4.8)$$

One way to determine the differential of $D(t)S(t)$ is Itô's product rule, proved in Shreve [35, p.168]. Since the theorem is stated for Itô processes let us first give the definition of an Itô process as in Shreve [35, p.143]:

Definition 4.1 *Let $W(t)$, $t \geq 0$, be a Brownian motion, and let $\mathcal{F}(t)$, $t \geq 0$, be an associated filtration. An Itô process is a stochastic process of the form*

$$X(t) = X(0) + \int_0^t \delta(u)dW(u) + \int_0^t \theta(u)du, \quad (4.9)$$

$$(4.10)$$

where $X(0)$ is nonrandom and $\delta(u)$ and $\theta(u)$ are adapted stochastic processes.

This allows us to state Itô's product rule as follows.

Theorem 4.3 *(Itô's product rule)*

Suppose we have two Itô processes $X(t)$ and $Y(t)$. Then

$$d(X(t)Y(t)) = X(t)dY(t) + Y(t)dX(t) + dX(t)dY(t) \quad (4.11)$$

Therefore we have

$$d(D(t)S(t)) = \sigma(t)D(t)S(t)[\Theta(t)dt + dW(t)], \quad (4.12)$$

where the market price of risk is defined by

$$\Theta(t) = \frac{\mu(t) - R(t)}{\sigma(t)}. \quad (4.13)$$

Since the differential of the discounted price process in (4.12) has a dt term, $D(t)S(t)$ is not a martingale under P .

4. Methods

Using Girsanov's theorem to change the probability measure from P to the risk neutral measure \tilde{P} , where we use the market price of risk from (4.13) to obtain the Brownian motion \tilde{W} , the discounted price process $D(t)S(t)$ can be written as

$$D(t)S(t) = S(0) + \int_0^t \sigma(u)D(u)S(u)d\tilde{W}(u) \quad (4.14)$$

and according to (4.1) is a martingale. The change of measure entails a change in the mean rate of return of the stock from $\mu(t)$ to $R(t)$. However, the volatility, telling us which price paths are possible, stays the same [35, p.217].

The portfolio consists of shares in the stock as well as in the bond and the changes in the portfolio value can be described by

$$dX(t) = \Delta(t)dS(t) + R(t)[X(t) - \Delta(t)S(t)]dt, \quad (4.15)$$

where $\Delta(t)$ represents the shares of stock at each time t and $X(0)$ is the initial capital. Investing and borrowing money is done at the interest rate $R(t)$. After plugging in (4.6) for $dS(t)$, using the notation (4.13) and regrouping the terms, (4.15) becomes

$$dX(t) = R(t)X(t)dt + \Delta(t)\sigma(t)S(t)[\Theta(t)dt + dW(t)] \quad (4.16)$$

Let us consider the discounted portfolio value now. It can be seen that

$$d(D(t)X(t)) = \Delta(t)\sigma(t)D(t)S(t)[\Theta(t)dt + dW(t)] \quad (4.17)$$

$$= \Delta(t)d(D(t)S(t)) \quad (4.18)$$

$$= \Delta(t)\sigma(t)D(t)S(t)d\tilde{W}(t) \quad (4.19)$$

Hence the discounted portfolio value is a martingale and changes in its value are solely due to changes in the discounted stock price [35, p.217].

In order to derive the value of a derivative with payoff $V(T)$, assumed to be $\mathcal{F}(T)$ measurable, at maturity T , we need to hedge the position in the derivative - determine the initial capital $X(0)$ and the portfolio process $\Delta(t)$ leading to $X(T) = V(T)$ almost surely [35, p.218]. When this is accomplished the fact that $D(t)X(t)$ is a martingale under the risk neutral measure \tilde{P} implies

$$D(t)X(t) = E_{\tilde{P}}[D(T)X(T) | \mathcal{F}(t)] = E_{\tilde{P}}[D(T)V(T) | \mathcal{F}(t)] \quad (4.20)$$

with $V(t) \stackrel{\text{def}}{=} X(t)$ the price of the derivative - the capital necessary to hedge the short position in the derivative with payoff $V(T)$.

Theorem 4.4 (Martingale Representation Theorem)

Consider the probability space (Ω, \mathcal{F}, P) , the Brownian motion $(W(t))_{0 \leq t \leq T}$ and the filtration $(\mathcal{F}(t))_{0 \leq t \leq T}$ generated by the Brownian motion. Any stochastic process $(M(t))_{0 \leq t \leq T}$ that is a martingale with respect to $(\mathcal{F}(t))_{0 \leq t \leq T}$ has the representation

$$M(t) = M(0) + \int_0^t \Gamma(u)dW(u), \quad 0 \leq t \leq T \quad (4.21)$$

where $(\Gamma(u))_{0 \leq u \leq T}$ is an adapted process [35, p.221].

4. Methods

This theorem is crucial to proving the existence of a hedging portfolio and uncovers that the only source of uncertainty in this context is the Brownian motion.

Corollary 4.1 *If we restrict the filtration $(\mathcal{F}(t))_{0 \leq t \leq T}$ in the Girsanov theorem to the one generated by the Brownian motion $(W(t))_{0 \leq t \leq T}$ the same results still apply. Additionally, we acquire that if $(\tilde{M}(t))_{0 \leq t \leq T}$ is a martingale under the risk neutral measure \tilde{P} and $(\tilde{\Gamma}(u))_{0 \leq u \leq T}$ an adapted process, $\tilde{M}(t)$ can be written, as done in [35, p.222], as*

$$\tilde{M}(t) = \tilde{M}(0) + \int_0^t \tilde{\Gamma}(u) d\tilde{W}(u) \quad 0 \leq t \leq T \quad (4.22)$$

Taking into account that $D(0)V(0) = V(0)$, the discounted price process $D(t)V(t)$, which is a martingale under \tilde{P} , can be written in the form

$$D(t)V(t) = V(0) + \int_0^t \tilde{\Gamma}(u) d\tilde{W}(u) \quad 0 \leq t \leq T \quad (4.23)$$

Integrating (4.17) we obtain the second representation of the discounted portfolio value

$$D(t)X(t) = X(0) + \int_0^t \Delta(u)\sigma(u)D(u)S(u)d\tilde{W}(u), \quad 0 \leq t \leq T \quad (4.24)$$

Since we want $X(t) = V(t)$ for all t , it is necessary that $X(0) = V(0)$ and to choose the portfolio process $(\Delta(t))_{0 \leq t \leq T}$ according to

$$\Delta(t) = \frac{\tilde{\Gamma}(t)}{\sigma(t)D(t)S(t)}, \quad 0 \leq t \leq T \quad (4.25)$$

This way a hedge in the short position in the derivative with payoff $V(T)$ at T is possible under the following assumptions:

1. The volatility $\sigma(t)$ is not zero.
2. The filtration $(\mathcal{F}(t))_{0 \leq t \leq T}$ is generated by the Brownian motion $(W(t))_{0 \leq t \leq T}$.

4.1.2. Arbitrage and Market Completeness

Definition 4.2 *An arbitrage possibility is a portfolio value process $X(t)$ satisfying $X(0) = 0$ and*

$$P(X(T) \geq 0) = 1, \quad P(X(T) > 0) > 0, \quad (4.26)$$

for some $T > 0$.

The definition above is taken from Shreve [35, p.230] and implies that it is possible that we start with 0 capital and without any risk of losing money end up with a profit with positive probability, which is a feature in market models economists ideally like to have absent with the argument that any arbitrage possibility will vanish due to an increase in interest from market participants, leading to an adjustment of the prices [35, p.230].

4. Methods

Theorem 4.5 *(First fundamental theorem of asset pricing)*

There is no arbitrage possibility in the market model if it has a risk-neutral probability measure.

Definition 4.3 *A market model is said to be complete if every derivative security can be hedged.*

Theorem 4.6 *(Second fundamental theorem of asset pricing)*

Consider a market model that has a risk-neutral probability measure. The model is complete if and only if the risk-neutral probability measure is unique.

Proofs of the fundamental theorems of asset pricing as well as the definition of a complete market model are offered in [35, p.230ff].

The fundamental theorems state that if there is a unique process $\Theta(t)$ solving the market price of risk equation (4.13), then there exists also a unique risk neutral measure and hence the market model does not permit arbitrage and is complete.

4.2. Pricing and Hedging in Incomplete Markets

As pointed out in section 4.1 the standard Black - Scholes model is an example of a complete market. Hence, any option can be perfectly replicated by a self-financing strategy involving the underlying and cash. In markets like that, options are redundant since they are replaceable by trading in the underlying.

In real markets, though, perfect hedging is impossible and options enable market participants to hedge risks that cannot be hedged by trading in the underlying only. This way, an option allows the market participant to transfer his risk, which was the purpose for the creation of derivative markets in the first place [31].

In discontinuous price models riskless replication is an exception rather than the rule, but this even makes them fit reality better.

In incomplete markets, as well as in the real markets, a perfect hedge does not exist, because of the so-called residual risk. Thus, when pricing options in incomplete markets, the goal is to approximate a target payoff with a trading strategy and not to find the cost of replicating it [11, p.320].

In order to approximate an option price we have to take the residual risk into account by specifying a way of measuring and subsequently minimizing it. Hence, the value of an option in an incomplete market is the cost of the hedging strategy with an additional risk premium. Obviously, we obtain different hedging approaches when choosing different ways of measuring the risk, like super-hedging, utility maximization and mean-variance hedging [11, p.320].

If only interested in arbitrage-free pricing of options, without taking the residual risk into account, one can choose any equivalent martingale measure as self-consistent pricing rule.

4. Methods

However, the obtained option price does not correspond to a specific hedging strategy anymore, contrary to the Black - Scholes case.

In a market model without Gaussian component it is not possible to obtain an equivalent martingale measure by changing the drift. Hence, we have to alter the distribution of jumps, leading to a greater variety of equivalent measures [11, p.310].

It is also important to realize that a P-Lévy process $(X(t))_{t \geq 0}$ does not have to be a Q-Lévy process anymore after changing the measure from P to Q, since the increments might neither be independent nor stationary [11, p.292].

In the case of an exponential Lévy model, however, it is possible to remain in the class of exponential Lévy processes using the Esscher transform to change to a risk neutral measure.

Cont and Tankov point out in [11, p.80] that, due to the Lévy-Itô decomposition, the distribution of every Lévy process is uniquely defined by a vector γ , a positive definite matrix A and a positive measure ν . The Lévy-Itô decomposition states that every Lévy process $(X(t))_{t \geq 0}$ can be represented by the sum of a Gaussian component and a discontinuous component, where the first is accountable for the continuous movement of the process $(X(t))_{t \geq 0}$ and the second incorporates the jumps of the process $(X(t))_{t \geq 0}$. The compendium of the three quantities to (A, ν, γ) is called characteristic triplet, whereas γ represents the drift vector and A the covariance matrix of the Gaussian component of $(X(t))_{t \geq 0}$ (but note that there is also a contribution to the overall drift coming from a possible compensation). The measure ν is called Lévy measure and is defined as in Cont and Tankov [11, p.76] as

Definition 4.4 Consider the Lévy process $(X(t))_{t \geq 0}$ on \mathbb{R}^d . The measure ν on \mathbb{R}^d defined by

$$\nu(A) = E[\#\{t \in [0, 1] : \Delta X(t) \neq 0, \Delta X(t) \in A\}], \quad A \in \mathcal{B}(\mathbb{R}^d), \quad (4.27)$$

is called the Lévy measure of X , where $\nu(A)$ is the expected number of jumps per unit time, where the jump size $\Delta X(t)$ belongs to A .

Theorem 4.7 Consider the Lévy process $(X(t))_{t \geq 0}$. If for $\theta \in \mathbb{R}^d$ we have

$$\int_{|x| \geq 1} e^{\theta x} \nu(dx) < \infty, \quad (4.28)$$

it follows that the measure

$$\tilde{\nu}(dx) := e^{\theta x} \nu(dx) \quad (4.29)$$

is a Lévy measure and for any Lévy process $(X(t))_{t \geq 0}$ on \mathbb{R}^d with characteristic triplet (A, ν, γ) the process with characteristic triplet $(A, \tilde{\nu}, \gamma)$ is also a Lévy process called Esscher transform of X [11, p.110].

Using the Theorems 33.1 and 33.2 from Sato [33, p.218] we obtain an equivalent probability, where the Lévy process $(X(t))_{t \geq 0}$ with characteristic triplet (A, ν, γ) is changed to a Lévy process with the new triplet $(A, \tilde{\nu}, \tilde{\gamma})$, where

$$\tilde{\gamma} = \gamma + \int_1^1 x(e^{\theta x} - 1) \nu(dx). \quad (4.30)$$

4. Methods

Theorem 4.8 (Absence of arbitrage in exp-Lévy models)

Considering the Lévy process (X, P) , the exponential Lévy model $S(t) = e^{(rt+X(t))}$ is arbitrage-free, if the trajectories of X are neither almost surely increasing nor almost surely decreasing. Therefore an equivalent martingale measure Q exists, such that $(e^{-rt}S(t))_{0 \leq t \leq T}$ is a Q -martingale, where r is the interest rate.

The corresponding Radon-Nikodym derivative to the measure change is

$$\frac{dQ|_{\mathcal{F}_t}}{dP|_{\mathcal{F}_t}} = \frac{e^{\theta X(t)}}{E[e^{\theta X(t)}]}, \quad (4.31)$$

as pointed out in [11, p.310].

This shows that exponential Lévy models are arbitrage-free, but since exponential Lévy models are also incomplete market models there exists more than one equivalent martingale measure.

Although having many possible ways of pricing an option might seem to be undesirable for a market model, it is a realistic property shared with real markets [11, p.316].

In the Black - Scholes model option prices are uniquely determined once a model is estimated from historical events and often does not coincide with traded option prices, whereas the flexibility of an incomplete market allows the exponential Lévy model to reproduce the market prices correctly [11, p.316].

For further reading we refer to Delbaen and Schachermayer [13], Kabanov in [22] and Cerny and Shiryaev [10], where no-arbitrage theorems and relation with equivalent martingale measures are discussed at length.

4.3. The American Option as a Stopping-Time Problem

To determine the unique price of an American put option we need the following definitions, stated in Peskir and Shiryaev [26, p.91ff], allowing us to rewrite the problem.

Definition 4.5 Suppose we are given a phase space (E, \mathcal{E}) , a family of probability spaces $(\Omega, \mathcal{F}, (\mathcal{F}(t))_{t \geq 0}, P_x, x \in E)$, where each P_x is a probability measure on (Ω, \mathcal{F}) and a stochastic process $X = (X(t))_{t \geq 0}$ where each $X(t)$ is $\mathcal{F}(t)/\mathcal{E}$ - measurable. If the following properties are fulfilled

1. the function $P(t, x; B) = P_x(X(t) \in B)$ is \mathcal{E} -measurable in x
2. $P(0, x; E \setminus \{x\}) = 0$, for $x \in E$
3. for all $s, t \geq 0$ and $B \in \mathcal{E}$, the Markov property holds:

$$P_x(X(t+s) \in B | \mathcal{F}(t)) = P(s, X(t); B) \quad P - a.s. \quad (4.32)$$

4. for any $\omega \in \Omega$ and $h > 0$ there exists $\omega' \in \Omega$ such that $X(t+h, \omega) = X(t, \omega')$ for all $t \geq 0$.

4. Methods

the process X is said to be a (time-homogeneous) Markov process on $(\Omega, \mathcal{F}, (\mathcal{F}(t))_{t \geq 0}, P_x, x \in E)$ with the time independent transition function $P(t, x; B)$.

In order to define a strong Markov process we need to state the definition of a stopping time σ -algebra, as in [27, p.5], first

Definition 4.6 Let τ be a stopping time. The stopping time σ -algebra is defined as

$$\mathcal{F}(\tau) = \{A \in \mathcal{F} : A \cap \{\tau \leq t\} \in \mathcal{F}(t) : t \geq 0\}. \quad (4.33)$$

Definition 4.7 If a stochastic process satisfies in addition to being a Markov process also the property

$$P_x(X(\tau + s) \in B \mid \mathcal{F}(\tau)) = P_{X_\tau}(X(s) \in B) \quad P_x - a.s. \text{ on } \{\tau < \infty\} \quad (4.34)$$

for all stopping times τ , it is called a strong Markov process.

It is very convenient to assume that the process $(X(t))_{t \geq 0}$ is a strong Markov process defined on $(\Omega, \mathcal{F}, (\mathcal{F}(t))_{t \geq 0}, P_x)$ and taking values in a measurable space (E, \mathfrak{B}) with $E = \mathbb{R}^d$ and \mathfrak{B} the Borel σ -algebra on \mathbb{R}^d . Should the Markov property be violated somehow, an augmentation of the state vector can often establish this property [19, p. 422]. We presume that the process $(X(t))_{t \geq 0}$ starts at x under P_x , is right-continuous and left-continuous over stopping times (i.e. if $\tau_n \uparrow \tau$ are stopping times, then $X_{\tau_n} \rightarrow X_\tau$ P_x a.s. as $n \rightarrow \infty$). In addition we assume that $(\mathcal{F}(t))_{t \geq 0}$ is right-continuous, ensuring that the first entry time to open and closed sets are stopping times. It is also indispensable to presume that $x \mapsto P_x(F)$ is measurable $\forall F \in \mathcal{F}$, the mapping $x \mapsto E_x(Z)$ is measurable for each bounded or non-negative random variable Z .

The major difference between a European and an American option is that in the case of an American option the option can be exercised at any given time prior and up to maturity. Since the exercising time is unknown in advance, the price of an American option differs from (4.20), where the exercise time is predetermined. In order to hedge the short position in an American option it is necessary to find the optimal stopping time that would maximize the payoff for the option holder. Until the optimal exercising time the American put option is a martingale, afterwards the option has a downward drift, making it a supermartingale. Consequently, the short position can be hedged in the usual way as long as the option price is a martingale and after the optimal exercising time the owner of the short position can take money off the table [35, p.354]. It is shown in [35, p.359] that the price the writer of the American put option has to demand in order to fulfil the contract is at least

$$V(x) = \sup_{\tau} E_x[e^{-r\tau}(K - X(\tau))^+], \quad (4.35)$$

where the expectation is taken under the risk-neutral measure. The supremum considered in (4.35) is taken over all admissible stopping times τ of X within $[0, T]$. The admissible set of stopping times we are considering in (4.35) consists of all possible exercising strategies using only the available information up to that point of the Markov process X , i.e. τ is

4. Methods

adapted to the natural filtration of X , \mathcal{F}_t^X [35, p.340]. $V(x)$ is the smallest initial capital allowing us to hedge a short position in the American put option [35, p.368].

The maturity date T can either be ∞ or take values in $(0, \infty)$. If we assume that $T < \infty$ we have the slightly different optimal stopping problem:

$$V(t, x) = \sup_{0 \leq \tau \leq T-t} \mathbb{E}_{t,x} [e^{-r\tau} (K - X(t + \tau))^+] \quad (4.36)$$

Other than that the rest of time $T - t$ changes when the initial state changes in its first argument, no argument is seriously affected by this change so that the following results are applicable in both cases. Another side effect of the finite horizon is that we get the terminal condition $V(T, x) = e^{-rT} (K - X(T))^+ \forall x \in E$, making sure that the first-entry time to the stopping set D will be finite.

To guarantee the existence of the expected value in (4.35), we have to ensure that

$$\mathbb{E}_x \left[\sup_{0 \leq t \leq T} |(e^{-r\tau} (K - X(t))^+)| \right] < \infty, \quad (4.37)$$

for all $x \in E$ and $T \in [0, \infty]$ [26, p.35].

Taking into account that the stochastic process we are considering is a strong Markov process and that stopping times are only using the information up to the present point, it is plausible that we are capable to decide optimally at any time to stop and exercise or to continue. Hence, it is possible to split the state space E into the two complementary sets C and D denoting the continuation set and the stopping set respectively. These sets are determined by the deterministic functions $V(x)$, representing the value of the option at x , and the gain function at x

$$G(x) = (K - x)^+. \quad (4.38)$$

The idea is to stop at the first time $t \geq 0$ when exercising the option is as profitable as continuing [26, p.36]. This is reflected by the form of the continuation set C and the stopping set D

$$C = \{x \in E : V(x) > G(x)\} \quad (4.39)$$

$$D = \{x \in E : V(x) = G(x)\} \quad (4.40)$$

Hence, as stopping time we consider the first entry-time

$$\tau_D = \inf \{t \geq 0 : X(t) \in D\} \quad (4.41)$$

4.3.1. About the Existence of an Optimal Stopping Time

The following definitions, taken from Bourbaki [6, p.166] as well as from Peskir and Shiryaev [26, p.37], are necessary to draw conclusions as to the existence of an optimal stopping time.

Definition 4.8 A real valued function f is said to be lower semi-continuous (lsc) in x_0 if $\forall \epsilon > 0 : \exists \delta > 0 : \forall x \in |x - x_0| < \delta : f(x) > f(x_0) - \epsilon$.

4. Methods

Definition 4.9 A real valued function f is said to be upper semi-continuous (usc) in x_0 if $\forall \epsilon > 0 : \exists \delta > 0 : \forall x \in |x - x_0| < \delta : f(x) < f(x_0) + \epsilon$.

Definition 4.10 A measurable function $F : E \rightarrow \mathbb{R}$ is called superharmonic if

$$E_x F(X_\sigma) \leq F(x) \quad (4.42)$$

for all stopping times σ and $\forall x \in E$, where it is assumed that the left hand-side is well defined and finite, i.e. $F(X_\sigma) \in L^1(P_x)$.

On the one hand, the necessary conditions for the existence of an optimal stopping time are

Theorem 4.9 Under the assumption that there exists an optimal stopping time τ_* in (4.35) for all $x \in E$, it is possible to show the following:

- The value function V is the smallest superharmonic function which dominates the gain function G on E .

Assuming additionally that V is lsc and G is usc we obtain:

- The stopping time τ_D satisfies $\tau_D \leq \tau_*$ P_x - a.s. $\forall x \in E$ and is optimal in (4.35).
- The stopped process $(V(X(t \wedge \tau_D)))_{t \geq 0}$ is a right-continuous martingale under P_x for every $x \in E$.

On the other hand, we have the following sufficient conditions:

Theorem 4.10 Assuring the existence of the expected value in (4.35) by (4.37), supposing that the smallest superharmonic function \hat{V} dominating G on E exists and that \hat{V} is lsc and G is usc we obtain by setting $D = \{x \in E : \hat{V}(x) = G(x)\}$ that:

- If $P_x(\tau_D < \infty) = 1 \forall x \in E$, then $\hat{V} = V$ and τ_D is optimal in (4.35)
- If $P_x(\tau_D < \infty) < 1$ for some $x \in E$, then there is no optimal stopping time with probability 1 in (4.35).

This gives us the following Corollary, which is especially useful in the finite horizon case, where it guarantees the existence of an optimal stopping time.

Corollary 4.2 (Existence of an optimal stopping time)

In the infinite horizon case we observe the stopping time problem (4.35) with the general assumption (4.37). The function V is lsc and the function G is usc. If $P_x(\tau_D < \infty) = 1 \forall x \in E$, then τ_D is optimal in (4.35). However, if $P_x(\tau_D < \infty) < 1$ for some $x \in E$, then there is no optimal stopping time with probability 1 in (4.35).

In the finite horizon case we consider (4.36) with a similar condition to (4.37) providing the existence of the expected value. Simply assuming that V is lsc and G is usc is enough to ensure that τ_D is optimal.

Proofs of these theorems can be found in Peskir and Shiryaev [26, p.37ff].

Note that if $x \mapsto E_x(e^{-r\tau}(K - X(\tau))^+)$ is continuous (or lsc) for every stopping time τ then $x \mapsto V(x)$ is lsc, giving us the means to apply the corollary from above.

4.4. Free-Boundary Problems and their Connection to Stopping-Time Problems

As we saw in the last section the stopping problem (4.35) is directly connected with the problem of finding the smallest superharmonic function \hat{V} dominating the gain function G from (4.38) on E . We have also seen that in this case τ_D , representing the first-entry time into the stopping set D , is optimal. Hence we can write for $x \in E$

$$V(x) = E_x G(X(\tau_D)) \quad (4.43)$$

Due to the Markovian structure of X any function of the form (4.43) is related to a deterministic equation.

In the case of the standard Black - Scholes model, where the process X is a geometric Brownian motion and therefore continuous, (4.43) is connected to a partial differential equation.

However, when we consider the exponential Lévy model, the process X has jumps and (4.43) is therefore related to a more general partial integro-differential equation.

To be able to point out the just aforementioned connection some additional definitions, from Peskir and Shiryaev [26, p.101, 129] are needed first.

Definition 4.11 *If $X = (X(t))_{t \geq 0}$ is a time-homogeneous Markov process defined on $(\Omega, \mathcal{F}, (\mathcal{F}(t))_{t \geq 0}, P_x)$ taking values in the state space (E, \mathcal{B}) we can define the characteristic operator as follows*

$$\mathbb{L}_X F(x) = \lim_{U \downarrow x} \frac{E_x f(X(\tau(U))) - f(x)}{E_x \tau(U)} \quad (4.44)$$

where U is a set containing x , $\tau(U)$ is the exiting time of U and $U \downarrow x$ stands for making U smaller and smaller until it shrinks to the point x itself in the limit.

Definition 4.12 *The boundary ∂C is said to be regular if each of its points satisfies $P_x(\sigma_D = 0) = 1$, with $\sigma_D = \inf \{t > 0 : X(t) \in D\}$. In words, the boundary is said to be regular if starting at a boundary point we immediately enter the interior of D .*

4.4.1. A Dirichlet Problem and the Solution

The definition of a Dirichlet problem can be described as in Folland [16, p.106]:

Definition 4.13 *A Dirichlet problem is to find a function solving a certain (partial) differential equation on the inside of a given region while having predetermined values on the boundary of that region.*

Consider

$$F(x) = E_x M(X(\tau_D)) \quad (4.45)$$

4. Methods

where the function $M : \partial C \rightarrow \mathbb{R}$, is assumed to be continuous. Using the strong Markov property of X it is shown in [26, p. 130] that F solves the Dirichlet problem

$$\mathbb{L}_X F = 0, \text{ in } C \quad (4.46)$$

$$F|_{\partial C} = M \quad (4.47)$$

and has the following properties, stated in [26, p.147]:

1. If ∂C is a regular boundary for D , then F is continuous on $\overline{C} = \partial C \cup C$:

$$F|_D = M|_D \quad (4.48)$$

2. If X is a diffusion, like in the standard Black - Scholes model, the smooth fit condition

$$\frac{\partial F}{\partial x}|_{\partial C} = \frac{\partial M}{\partial x}|_{\partial C} \quad (4.49)$$

has to be satisfied.

In the case of the exponential Lévy model, where X has jumps and no diffusion component, the smooth fit condition has to be replaced by the continuous fit condition

$$F|_{\partial C} = M|_{\partial C} \quad (4.50)$$

In the case of the American put option we have to apply the preceding results on the discounted version of the gain function G . Replacing the process X by its killed version \tilde{X} - representing the process, where the sample paths of X are killed at the constant discounting rate $r \geq 0$ - we can reduce

$$V(x) = \mathbb{E}_x[e^{-r\tau_D} G(X(\tau_D))] \quad (4.51)$$

to the Dirichlet problem

$$V(x) = \mathbb{E}_x[G(\tilde{X}(\tau_D))] \quad (4.52)$$

where $V(x)$ solves

$$\mathbb{L}_X V = rV \text{ in } C, \quad (4.53)$$

As already mentioned above, the continuation set C and the stopping set D are complementary sets, i.e. $x \in C$ or $x \in D$. Additionally, we have that V is lsc and G is continuous (usc) implying that C is an open and D is a closed set [26, p. 36]. Therefore the closure of C , ∂C , is a natural boundary between the two sets for any $t \in [0, T]$.

Since the continuation value $V(t, x)$ is non-increasing in t , we have

$$\forall t_1 \leq t_2 : t_1, t_2 \in [0, T], x \geq 0 \text{ fixed} : (t_2, x) \in C \Rightarrow (t_1, x) \in C \quad (4.54)$$

as well as

$$\forall t_1 \leq t_2 : t_1, t_2 \in [0, T], x \geq 0 \text{ fixed} : (t_1, x) \in D \Rightarrow (t_2, x) \in D. \quad (4.55)$$

4. Methods

Together with the fact that the function $t \mapsto b(t)$ is non-decreasing, as shown in [26, p.384] for the standard Black - Scholes model and in [23, p.574] for the exponential Lévy model, we can rewrite the optimal stopping time from (4.41) as

$$\tau = \inf\{t \geq 0 : x \leq b(t)\} \quad (4.56)$$

In the case where $T = \infty$, we are talking about an option with infinite horizon. There it is possible to derive the following closed form solution for the asset price in the standard Black - Scholes model,

$$V(x) = \begin{cases} \frac{\sigma^2}{2r} \left(\frac{K}{1+\sigma^2/2r} \right)^{1+2r/\sigma^2} x^{-2r/\sigma^2}, & \text{if } x \in [b, \infty), \\ K - x, & \text{if } x \in (0, b]. \end{cases} \quad (4.57)$$

as done in [26, p. 377], where $b = \frac{K}{1+\sigma^2/2r}$ is the constant optimal exercising boundary.

Is the horizon of the problem finite, though, we obtain an additional $\frac{\partial}{\partial t}$ term in (4.53),

$$V_t + \mathbb{L}_X V = rV \text{ in } C, \quad (4.58)$$

which gets in the way of solving the problem explicitly, even in the Black - Scholes model. In this case, expressing V in terms of ∂C , the closure of the continuation set, leads to a nonlinear Volterra integral equation for ∂C . So far there is no explicit solution to this kind of integrals which makes numerical methods so important, as mentioned in Peskir and Shiryaev [26, p.146].

In (4.52) we see that in order to determine the price of the American put option, it is necessary to estimate the optimal exercising boundary as accurately as possible. In the next sections we point out possibilities as to how to simulate the price of the underlying for the standard Black - Scholes model as well as the exponential Lévy model and then discuss the algorithm used for the estimation of the optimal exercising boundary.

4.5. Simulation of Asset Developments

There are many different approaches for the simulation of the random movement of an asset price. For the two most popular approaches, already mentioned in 2.4, we will discuss possibilities to generate random asset paths, necessary for the pricing of derivatives using Monte Carlo methods.

4.5.1. Simulation of a Brownian Motion

Let $(W(t))_{t \geq 0}$ denote a standard Brownian Motion. First the time line has to be partitioned into equidistant intervals $t_n^h = nh$ for $n = 0, 1, 2, \dots$ of length h . Using $\Delta_n^h W = W(t_n^h) - W(t_{n-1}^h)$ and taking into account that the time grid is equidistant, the generation of the increments $\Delta_n^h W$ as i.i.d. $\mathcal{N}(0, h)$ leads to the straightforward simulation of W_n^h by

$$W_n^h = W(t_n^h) = \Delta_1^h W + \dots + \Delta_n^h W. \quad (4.59)$$

4. Methods

as explained in Asmussen and Glynn [2, p.276].

4.5.2. Simulation of a Normal Inverse Gaussian Levy Processes

In order to simulate paths for the Lévy process $L(t)$ it is necessary to generate random numbers of the increment process, which we assume to be normal inverse Gaussian. The fact that the normal inverse Gaussian family of distributions is infinitely divisible, which is pointed out in Benth et al. [3, p.846], implies

$$L(t + \Delta t) - L(t) \sim L(\Delta t) \sim NIG(\alpha, \beta, \mu, \delta) \quad (4.60)$$

Therefore we can proceed as in the case of the Brownian motion and simulate the normal inverse Gaussian random numbers along an equidistant time grid $t_n^h = nh$ for $n = 0, 1, 2, \dots$ of length h . To obtain $L_n^h = L(t_n^h)$ use the following property

$$L_n^h = L(t_n^h) = \Delta_1^h L + \dots + \Delta_n^h L, \quad (4.61)$$

where $\Delta_n^h L = L(t_n^h) - L(t_{n-1}^h)$

This way it all boils down to generating a normal inverse Gaussian random variable X with the parameters $\alpha, \beta, \mu, \delta$, where $0 \leq |\beta| \leq \alpha$, $\mu \in \mathbb{R}$ and $\delta > 0$. We are following the algorithm used by Benth et al. [3, p.847], which was first introduced by Rydberg [32, p.897].

Since the normal inverse Gaussian distribution is defined as a normal variance-mean mixture, with the inverse Gaussian distribution as the mixing density, we can write

$$X = \mu + \beta Z + \sqrt{Z} Y, \quad (4.62)$$

where Z is sampled from the mixing density with parameters δ^2 and $\alpha^2 - \beta^2$ and $Y \sim \mathcal{N}(0, 1)$. This leaves us with the task of generating inverse Gaussian random numbers, which can be accomplished by sampling first from a $\chi^2(1)$ random variable V , setting

$$W = \xi + \frac{\xi^2 V}{2\delta^2} - \frac{\xi}{2\delta^2} \sqrt{4\xi\delta^2 V + \xi^2 V^2} \quad (4.63)$$

and then letting

$$Z = W 1_{\{U_1 \leq \frac{\xi}{\xi+W}\}} + \frac{\xi^2}{W} 1_{\{U_1 \geq \frac{\xi}{\xi+W}\}}, \quad (4.64)$$

where $U_1 \sim \text{Uniform}(0, 1)$ and $\xi = \frac{\delta}{\sqrt{\alpha^2 - \beta^2}}$.

4.6. The Cross - Entropy Method

In order to give an introduction to the Cross - Entropy Method the structure set up by Rubinstein and Kroese in [30] seems to be best suited.

Descendent from a variance minimization problem for rare-event simulation, it was transformed to become a randomized optimization method, where the primary variance minimization approach was exchanged by the problem of minimizing the Kullback-Leibler distance,

4. Methods

also called Cross - Entropy(CE). The CE between the two probability density functions (pdf) g and h is defined by

$$\mathcal{D}(g, h) = \mathbb{E}_g \ln \frac{g(\mathbf{X})}{h(\mathbf{X})} = \int g(\mathbf{x}) \ln g(\mathbf{x}) \mu(d\mathbf{x}) - \int g(\mathbf{x}) \ln h(\mathbf{x}) \mu(d\mathbf{x}) \quad (4.65)$$

and satisfies $\mathcal{D}(g, h) \geq 0$ and $\mathcal{D}(g, h) = 0$ if and only if $h(x) = g(x)$. Hence, thinking of the Cross - Entropy as a measurement for the "distance" between two densities proves to be useful, even though it violates the symmetry property, $\mathcal{D}(g, h) \neq \mathcal{D}(h, g)$, a distance usually has to possess.

In rare event simulations the Cross - Entropy method is combined with importance sampling, a variance reduction method, simulating the problem with a different density, making these rare events more likely. Often it is convenient to choose a density belonging to the same distribution family, where only the reference parameter differs.

Although, in general, it is difficult to obtain these reference parameters via time-consuming variance minimization methods, in the case of the Cross - Entropy method these parameters can be obtained by a simple and fast adaptive procedure.

It is important to point out that the Cross - Entropy method terminates with probability 1 in a finite number of steps and produces a consistent (i.e. $\lim_{n \rightarrow \infty} \Pr(|t_n - \theta| \geq \epsilon) = 0$, where t_n is a sequence of estimators and θ the original parameter) and asymptotically normal (i.e. $\sqrt{n}(t_n - \theta) \rightarrow \mathcal{N}(0, \sigma)$) estimator for the optimal reference parameter.

4.6.1. Rare Event Simulation

The event $l := \mathbb{P}_{\mathbf{u}}(S(\mathbf{X}) \geq \gamma) = \mathbb{E}_{\mathbf{u}} I_{\{S(\mathbf{X}) \geq \gamma\}}$, where S is a real valued function, is considered to be rare, when the probability that the event occurs is less than 10^{-5} .

In this case, simply sampling from the given density like in the crude Monte Carlo approach

$$\frac{1}{N} \sum_{i=1}^N I_{\{S(\mathbf{X}_i) \geq \gamma\}} \quad (4.66)$$

is very ineffective on account of the sample-size needed to determine an accurate estimation of the probability of the rare event. Hence, using importance sampling to simulate under a different density, increasing the probability of the former rare event, enables us to reduce the sample size significantly.

If $f(x; u)$ is the original density and $g(x)$ the density of the importance sampling, then we obtain the following estimator for l

$$\hat{l} = \frac{1}{N} \sum_{i=1}^N I_{\{S(\mathbf{X}_i) \geq \gamma\}} \frac{f(\mathbf{X}_i; \mathbf{u})}{g(\mathbf{X}_i)} \quad (4.67)$$

where $\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_N$ are sampled under the density g . It would be optimal if g had the form

$$g^*(\mathbf{x}) = \frac{I_{\{S(\mathbf{x}) \geq \gamma\}} f(\mathbf{x}; \mathbf{u})}{l} \quad (4.68)$$

4. Methods

giving us the wanted probability already after just one simulation step. Unfortunately this density uses the unknown parameter l .

The idea now is to chose a density belonging to the same density family as $f(\mathbf{x}; \mathbf{u})$ with a minimal Kullback-Leibler distance

$$\mathcal{D}(g^*, f) = \mathbb{E}_{g^*} \ln \frac{g^*(\mathbf{X})}{f(\mathbf{X}; \mathbf{v})} = \int g^*(\mathbf{X}) \ln g^*(\mathbf{X}) d\mathbf{x} - \int g^*(\mathbf{X}) \ln f(\mathbf{X}; \mathbf{v}) d\mathbf{x} \quad (4.69)$$

between g^* and $f(\mathbf{X}; \mathbf{v})$. Minimizing the Cross - Entropy equals maximizing the second integral in (4.69). When we also insert g^* from (4.68) we get

$$\max_{\mathbf{v}} \int \frac{I_{\{S(\mathbf{x}) \geq \gamma\}} f(\mathbf{x}; \mathbf{u})}{l} \ln f(\mathbf{x}; \mathbf{v}) d\mathbf{x} \quad (4.70)$$

equalling the program

$$\max_{\mathbf{v}} D(\mathbf{v}) = \max_{\mathbf{v}} \mathbb{E}_{\mathbf{u}} I_{\{S(\mathbf{X}) \geq \gamma\}} \ln f(\mathbf{X}; \mathbf{v}) \quad (4.71)$$

leading to the optimal reference parameter \mathbf{v} .

The problem with this approach is that the indicator function $I_{\{S(\mathbf{X}) \geq \gamma\}}$ will, just like in the Crude Monte Carlo case, still be 0 for most of the samples. If, however, we calculated in a first step $P_{\mathbf{u}}(S(\mathbf{X}) \geq \gamma(t))$ for $\gamma(t) < \gamma$ the indicator function $I_{\{S(\mathbf{X}_i) \geq \gamma(t)\}}$ would be $\neq 0$ more often. Thus, a multilevel algorithm adapting the parameter $\gamma(t)$ and hence $v(t)$ at every step, achieves what we are looking for.

By reapplying importance sampling on (4.71) we obtain an additional likelihood ratio term in the program

$$\max_{\mathbf{v}} D(\mathbf{v}) = \max_{\mathbf{v}} \mathbb{E}_{\mathbf{w}} I_{\{S(\mathbf{X}) \geq \gamma\}} W(\mathbf{X}; \mathbf{u}, \mathbf{w}) \ln f(\mathbf{X}; \mathbf{v}) \quad (4.72)$$

where

$$W(\mathbf{X}; \mathbf{u}, \mathbf{w}) = \frac{f(\mathbf{x}; \mathbf{u})}{f(\mathbf{x}; \mathbf{w})} \quad (4.73)$$

is the likelihood ratio between $f(\cdot; \mathbf{u})$ and $f(\cdot; \mathbf{w})$.

Since the expected value in (4.72) is unknown a so-called stochastic program or stochastic counterpart to (4.72) is necessary for the actual calculation

$$\max_{\mathbf{v}} \hat{D}(\mathbf{v}) = \max_{\mathbf{v}} \frac{1}{N} \sum_{i=1}^N I_{\{S(\mathbf{X}_i) \geq \gamma\}} W(\mathbf{X}_i; \mathbf{u}, \mathbf{w}) \ln f(\mathbf{X}_i; \mathbf{v}) \quad (4.74)$$

As \hat{D} is usually convex and differentiable with respect to \mathbf{v} in applications, the solution \mathbf{v} to (4.74) can be obtained by solving

$$\frac{1}{N} \sum_{i=1}^N I_{\{S(\mathbf{X}_i) \geq \gamma\}} W(\mathbf{X}_i; \mathbf{u}, \mathbf{w}) \nabla \ln f(\mathbf{X}_i; \mathbf{v}) = \mathbf{0}, \quad (4.75)$$

with ∇g denoting the gradient of g . For the multilevel algorithm the equation (4.74) has to be adjusted to

$$\max_{\mathbf{v}} \hat{D}(\mathbf{v}) = \max_{\mathbf{v}} \frac{1}{N} \sum_{i=1}^N I_{\{S(\mathbf{X}_i) \geq \hat{\gamma}(t)\}} W(\mathbf{X}_i; \mathbf{u}, \hat{\mathbf{v}}_{t-1}) \ln f(\mathbf{X}_i; \mathbf{v}) \quad (4.76)$$

4. Methods

Summing up the previous results, the following algorithm can be used to determine the rare event probability estimate.

Algorithm for Rare-Event Simulation with the CE Method

1. Define $\hat{\mathbf{v}}_0 = \mathbf{u}$. Set $t = 1$.
2. Generate a sample $\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_N$ from the density $f(\cdot; \mathbf{v}_{t-1})$ and compute the sample $(1 - \rho)$ -quantile $\hat{\gamma}(t) = S_{(\lceil (1-\rho)N \rceil)}$ provided $\hat{\gamma}(t)$ is less than γ . Otherwise set $\hat{\gamma}(t) = \gamma$.
3. Use the same sample $\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_N$ to solve the stochastic program (4.76). Denote the solution by $\hat{\mathbf{v}}(t)$.
4. If $\hat{\gamma}(t) < \gamma$, set $t = t + 1$ and reiterate from Step 2. Else proceed with Step 5.
5. Estimate the rare-event probability l with

$$\hat{l} = \frac{1}{N} \sum_{i=1}^N I_{\{S(\mathbf{x}_i) \geq \gamma\}} W(\mathbf{X}_i; \mathbf{u}, \hat{\mathbf{v}}_T) \quad (4.77)$$

where T denotes the final number of iterations.

4.6.2. Combinatorial Optimization Problems

Assume that we have a maximization problem over the finite state space \mathcal{X} , where S is a real valued performance function

$$S(\mathbf{x}^*) = \gamma^* = \max_{\mathbf{x} \in \mathcal{X}} S(\mathbf{x}) \quad (4.78)$$

To be able to use the previous section, we need to rewrite the problem in form of an estimation problem using the indicator functions $I_{\{S(\mathbf{x}) \geq \gamma\}}$ for various level $\gamma \in \mathbb{R}$. Additionally, we are going to consider a family of discrete densities $\{f(\cdot; \mathbf{v}, \mathbf{v} \in \mathcal{V})\}$, where \mathcal{V} is the parameter set. Then the associated stochastic problem has the following form

$$l(\gamma) = P_{\mathbf{u}}(S(\mathbf{X}) \geq \gamma) = \sum_{\mathbf{x}} I_{\{S(\mathbf{x}) \geq \gamma\}} f(\mathbf{x}; \mathbf{u}) = E_{\mathbf{u}} I_{\{S(\mathbf{x}) \geq \gamma\}} \quad (4.79)$$

with a certain $\mathbf{u} \in \mathcal{V}$.

The transformation into the associated stochastic problem is based on the fact that for γ approaching γ^* from below, the probability of the sample $S(\mathbf{X}_i)$ being bigger than γ decreases up to the point of making it a rare-event. Hence it is possible to estimate $l(\gamma)$ by (4.77) with the reference parameter \mathbf{v}^* determined via (4.74), where the likelihood ratio term $W(\mathbf{X}; \mathbf{u}, \mathbf{w})$ is set to 1. The likelihood ratio term is not important, because the original density parameter was arbitrary anyway. Thus the algorithm adapted to optimization problems is the following

4. Methods

CE Algorithm for Optimization Problems

1. Define $\hat{\mathbf{v}}_0 = \mathbf{u}$. Set the level counter $t = 1$.
2. Generate a sample $\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_N$ from the density $f(\cdot; \mathbf{v}_{t-1})$ and compute the sample $(1 - \rho)$ -quantile $\hat{\gamma}(t)$ of the performances.
3. Use the same sample $\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_N$ and solve the stochastic program (4.76) with $W = 1$. Denote the solution by $\hat{\mathbf{v}}(t)$
4. If for some $t \geq d$, say $d = 5$,

$$\hat{\gamma}(t) = \hat{\gamma}_{t-1} = \dots = \hat{\gamma}_{t-d}$$

then stop; otherwise set $t = t + 1$ and reiterate from Step 2.

4.6.3. Smoothed Updating

To prevent 0s and 1s in the parameter vectors, which are unwanted due to the fact that once the entry is 0 or 1, it most often remains like this forever, we can adapt the updating rule which simply replaces $\mathbf{v}(t - 1)$ by $\mathbf{v}(t)$ to

$$\mathbf{v}(t) = \alpha \hat{\mathbf{w}}(t) + (1 - \alpha) \mathbf{v}(t - 1)$$

where $\hat{\mathbf{w}}(t)$ is the vector derived in (4.76), ensuring a smoothed updating procedure without 0s and 1s.

4.7. Adjustment of the CE Method for Pricing American Options

The Cross - Entropy Method for Combinatorial Optimization seeks to solve problems of the form

$$S(\mathbf{x}^*) = \gamma^* = \max_{\mathbf{x} \in \mathcal{X}} S(\mathbf{x}) \quad (4.80)$$

by following the algorithm already mentioned in section 4.6

1. Define $\hat{\mathbf{v}}_0 = \mathbf{u}$. Set the level counter $t = 1$.
2. Generate a sample $\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_N$ from the density $f(\cdot; \mathbf{v}_{t-1})$ and compute the sample $(1 - \rho)$ -quantile $\hat{\gamma}(t)$ of the performances.
3. Use the same sample $\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_N$ and solve the stochastic program (4.76) with $W = 1$. Denote the solution by $\hat{\mathbf{v}}(t)$
4. If for some $t \geq d$, say $d = 5$,

$$\hat{\gamma}(t) = \hat{\gamma}_{t-1} = \dots = \hat{\gamma}_{t-d}$$

then stop; otherwise set $t = t + 1$ and reiterate from Step 2.

4. Methods

As pointed out before in section 4.4, the problem of determining the fair price of an American put option with finite horizon can be rewritten as

$$V(t, x) = \sup_{0 \leq \tau \leq T-t} \mathbb{E}_{t, x} [e^{-r\tau} (K - X(t + \tau))^+], \quad (4.81)$$

where the supremum is taken over all admissible stopping times τ of X in $[0, T - t]$. It is shown in Peskir and Shiryaev [26, p.123] and also brought up in section 4.4 that (4.81) can be rewritten as

$$V(t, x) = \mathbb{E}_{t, x} [e^{-r\tau_D} (K - X(t + \tau_D))^+] \quad (4.82)$$

implying that the highest expected payoff will be obtained by exercising according to the strategy

$$\tau_D = \inf \{0 \leq s \leq T - t : ((t + s), X(t + s)) \in D\} \quad (4.83)$$

$$= \inf \{0 \leq s \leq T - t : X(t + s) \leq b(t + s)\}, \quad (4.84)$$

where $D = \{(s, x) \in [t, T] \times (0, \infty) : V(t + s, x) = G(t + s, x)\}$ and $b(t)$ is the optimal exercising boundary on $[0, T]$.

4.7.1. Properties of the Optimal Exercising Boundary

Since we aim to simulate the optimal exercising boundary with the Cross - Entropy algorithm, it is necessary to learn more about its properties.

In Peskir and Shiryaev [26, p.384], the case where the underlying is driven by a geometric Brownian motion is studied. They give proof of the following important properties:

1. $0 < b^* \leq b(0+) < K$,
2. $b(T-) = K$,
3. $t \mapsto b(t)$ non decreasing and convex on $[0, T]$,

where b^* is the constant optimal exercising boundary for the perpetual American option on the same underlying with the same parameters and K is the strike price.

In the case of an exponential Lévy model similar conclusions as those above apply, as demonstrated by Lambertson and Mikou in [23, p.574], where, under the assumption that

$$\nu((-\infty, 0)) > 0, \quad (4.85)$$

we know that the optimal exercising boundary $b(t)$ has to satisfy

1. $0 \leq b(t) < K$ for all $t \in [0, T]$,
2. $t \mapsto b(t)$ non decreasing and continuous on $[0, T]$.

4. Methods

The major differences between the two cases are that in the exponential Lévy model we only know that the boundary will always be positive and continuous, whereas in the standard Black - Scholes model we know that the boundary will always be positive and at least as large as the constant value of the perpetual American put option b^* as well as convex. Contrary to the Black - Scholes model, the optimal exercising boundary in the exponential Lévy model is not continuous on the whole interval $[0, T]$, since Levendorskii showed in [24, p.533] that at the maturity date T , the boundary has a discontinuity.

4.7.2. The Choice of Basis Functions

Up to now there has existed no closed form solution for the optimal exercising boundary. Therefore, in this algorithm, the optimal boundary function is approximated by basis functions. At the very beginning we used a 2-step function, satisfying the necessary conditions mentioned above, to estimate the optimal exercising boundary $b(t)$ on $[0, 1]$. The random steps were generated by a bivariate normal distribution with $\mu = \mathbf{0}$ and $\Sigma = I_2$. Gradually the program was refined to any maturity T and a 64-step boundary function, accompanied by a generalization from a bivariate to a multivariate normal distribution with $\mu = \mathbf{0}$ and $\Sigma = I_n$.

After that the approximating step functions were replaced by linear combinations of the basis functions

$$t^n, t^{n-1}, \dots, t, 1 \quad \text{for } n \geq 0 \quad (4.86)$$

There the multivariate normal distributed parameters, optimized by the Cross Entropy algorithm, were the coefficients α_k , $k = 0, \dots, n + 2$ of the linear combination

$$\tilde{b}(t) = \sum_{k=0}^n \alpha_{n-k} t^k \quad (4.87)$$

The basis functions were altered once more to

$$t^{2n}, t^{2n-2}, \dots, t^2, 1 \quad (4.88)$$

in the hope of achieving an approximation of the optimal exercising boundary that would fit the properties from 4.7.1 better.

When sticking to the algorithm above, the program carries out the subsequent steps

4.7.3. Pricing American Put Options

1. Initialize μ_0, Σ_0 and set the level counter $t = 1$.
2. Generate a large number of paths according to the price dynamics of the Black - Scholes or exponential Lévy model, respectively. These paths represent the various

4. Methods

developments of the underlying asset.

This way we do not actually compute the price of the option in the model, but within the empirical distribution of the simulated paths.

3. Generate a sample $\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_N$ of optimal exercising boundary parameters, which are $\mathcal{N}(\cdot; \mu_{t-1}, \Sigma_{t-1})$ distributed. Each parameter set $\mathbf{X}_i = (\mathbf{X}_{i1}, \mathbf{X}_{i2}, \dots, \mathbf{X}_{ik})$ represents a boundary.
 - 3.1. In the case where the optimal exercising boundary is approximated by step functions with k steps, the distribution of the steps is $\mathcal{N}_k(\cdot; \mu_{t-1}, \Sigma_{t-1})$.
 - 3.2. When the optimal exercising boundary is approximated by a linear combination of k basis functions, their coefficients in the linear combination are $\mathcal{N}_k(\cdot; \mu_{t-1}, \Sigma_{t-1})$ distributed.
4. Compute the sample $(1 - \rho)$ -quantile $\hat{\gamma}(t)$ of the payoffs of each \mathbf{X}_i . The boundary parameters for which the payoff is $\geq \hat{\gamma}(t)$ are the elite sample \mathcal{E} .
5. Use the same sample $\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_N$ to update the parameters according to (4.76) and denote the solution by $\mu_t = (\mu_{t1}, \mu_{t2}, \dots, \mu_{tk})$, $\Sigma_t = (\sigma_t^{ij})_{1 \leq i, j \leq k}$:

$$\mu_{ti} = \frac{1}{|\mathcal{E}|} \sum_{\mathbf{X}_l \in \mathcal{E}} \mathbf{X}_{li} \quad \text{for all } i = 1, \dots, k \quad (4.89)$$

$$(4.90)$$

$$\sigma_t^{ij} = \frac{1}{|\mathcal{E}| - 1} \sum_{\mathbf{X}_r \in \mathcal{E}} (\mathbf{X}_{ri} - \mu_{ti})(\mathbf{X}_{rj} - \mu_{tj}) \quad \text{for all } i, j = 1, \dots, k \quad (4.91)$$

$$(4.92)$$

6. If for some $t \geq d$, say $d = 5$,

$$\hat{\gamma}(t) = \hat{\gamma}_{t-1} = \dots = \hat{\gamma}_{t-d}$$

then stop; otherwise set $t = t + 1$ and reiterate from Step 2.

4.7.4. Initial Parameters

Since the initial parameters μ_0, Σ_0 influence the time needed for the algorithm to converge, it was necessary to test various possibilities until satisfying initial parameters were found. In the case where we approximated the optimal exercising boundary by means of a step function, we ended up with the following iterative method for their determination:

- Start with a 2 step boundary function where the initial parameters for the bivariate normal distribution are

$$\mu_0 = \left(b^*, \frac{1}{2}b^* + \frac{1}{2}K \right), \quad (4.93)$$

4. Methods

$$\Sigma_0 = \mathbf{I}_2 K^2. \quad (4.94)$$

with b^* the constant optimal exercising boundary for the perpetual American option with the same parameters, \mathbf{I}_2 the 2 dimensional identity matrix and K the strike price. The vector of expectations μ_0 , mentioned in (4.93), was chosen, because the optimal exercising boundary is increasing in t and satisfies $0 < b^* \leq b(0-) as well as $b(T-) = K$. Moreover, the covariance matrix Σ of a multivariate normal distribution has to be positive definite, which a diagonal matrix with positive entries obviously satisfies. The fact that the initial covariance matrix Σ_0 is a diagonal matrix also implies that the initial steps are uncorrelated in the first round of simulations in the algorithm 4.7.3 until the parameters get updated. The reason for taking Σ_0 instead of the identity matrix \mathbf{I}_2 is that this way the standard deviation is equal to the strike price K , implying that about 68% of the random numbers simulated will be in the interval$

$$[b^* - K, b^* + K] \quad (4.95)$$

for the first boundary step and in the interval

$$\left[\frac{1}{2}b^* - \frac{1}{2}K, \frac{1}{2}b^* + \frac{3}{2}K \right] \quad (4.96)$$

for the second.

- To obtain the initial parameters μ_0 and Σ_0 for a 4-step boundary function, we can thus use the final values of the 2-step boundary parameters, $\tilde{\mu}_1$ and $\tilde{\mu}_2$, in the following way

$$\mu_0 = \left(\tilde{\mu}_1, \frac{1}{2}\tilde{\mu}_1 + \frac{1}{2}\tilde{\mu}_2, \tilde{\mu}_2, \frac{1}{2}\tilde{\mu}_2 + \frac{1}{2}K \right). \quad (4.97)$$

The initial covariance matrix, a diagonal matrix with positive entries, is determined as follows

$$\text{diag}(\Sigma_0) = \left[\left(\frac{\mu_0^{(1)} - \mu_0^{(2)}}{2} \right)^2, \dots, \left(\frac{\mu_0^{(3)} - \mu_0^{(4)}}{2} \right)^2, \left(\frac{\mu_0^{(4)} - K}{2} \right)^2 \right]. \quad (4.98)$$

The reason for taking Σ_0 instead of the identity matrix \mathbf{I}_2 is that this way the standard deviation is half the distance between two consecutive boundary steps.

- This procedure can be iterated arbitrarily, where in each iteration the number of steps for the boundary function is doubled.

In the case where we approximated the optimal exercising boundary by a polynomial function an iterative method is not possible, since changing the degree of the polynomial renders the parameters of the former step useless. Hence, we simply used

$$\mu_0 = \left[\frac{1}{2}b^* + \frac{1}{2}K, \dots, \frac{1}{2}b^* + \frac{1}{2}K \right], \quad (4.99)$$

4. Methods

where b^* is the value of the perpetual American put option in the standard Black - Scholes model and 0 in the exponential Lévy model, and

$$\Sigma_0 = \mathbf{I}_n \left(\frac{K}{T^n} \right)^2, \quad (4.100)$$

where $n + 1$ is the number of basis functions used. The square root of the additional factor in (4.100) is a rough estimation of the necessary size of the coefficient accompanying the basis function with the highest degree to keep its value at maturity T below the strike price K .

For the parameter update the smoothed updating procedure from 4.6.3 with $\alpha = 0.6$ was used.

In the course of the research we discovered an article by Zhang and Fu [36] in which the authors used the Cross - Entropy method to obtain comparable results for their own algorithm, called MRAS. Unfortunately, a comparison between the Cross - Entropy algorithm developed there and in this dissertation is not reasonable, since we used an iterative method to obtain results for a 64 step boundary, whereas in [36] the authors used only 6 exercising times making an application of our 64-step boundary algorithm impossible.

4.7.5. Pricing Basket Options

The pricing of basket options of American style with arbitrary weights demands a simulation of the paths for each of the assets in the basket. As found by data analysis, the correlation between the Brownian motions driving the asset prices is about 0.4. The geometric Brownian motions, embodying the asset price movements, have to have the same drift, but may vary in the volatility. As proposed by Asmussen and Glynn in [2, p.49] correlated normal random variables can be simulated by

$$X_1 = \sigma_1(\sqrt{1 - |\rho|}Y_1 + \sqrt{|\rho|}Y_3) \quad (4.101)$$

and

$$X_2 = \sigma_2(\sqrt{1 - |\rho|}Y_2 + \sqrt{|\rho|}Y_3), \quad (4.102)$$

where Y_1, Y_2 and $Y_3 \sim \mathcal{N}(0, 1)$ and thus give the means to simulate correlated Brownian motions. To price the basket option the same algorithm as for single underlying options can be used with the slight difference that the price dynamics at each time step t_i are described by the (weighted) sum of the underlying assets.

5. Results

In this chapter we want to demonstrate the results gathered with the help of the adapted Cross - Entropy algorithm described in section 4.7.3, where we used the standard Black - Scholes model from 2.4.1 as well as the exponential Lévy model with normal inverse Gaussian increments from 2.4.2 as market models.

5.1. Standard Black - Scholes Model

In the Black-Scholes model, the algorithm using step functions as well as the one using a polynomial to approximate the optimal exercising boundary were applied to American-style vanilla put options as well as Basket options of that type.

5.1.1. American-Style Vanilla Put Option

In the following graphs we want to present the change in the 4-step boundary function, when the boundary parameters are updated according to the Cross - Entropy algorithm. As mentioned before, the algorithm calculates the discounted payoff for each path with respect to the optimal stopping time, depending on the present step function. The parameters of the step function, approximating the optimal exercising boundary $b(t)$, are updated to maximize the average discounted payoff. With the change of the step function, the optimal stopping time may change as well, which can be observed in Figures 5.1 to 5.4, where the following parameters were used: $r = 5\%$, $\sigma = 25\%$, $T = 5$, $K = 120\%$ and an elite sample size of $\rho = 5\%$ of the total $N = 1.000$ simulations for each step of the Cross - Entropy method.

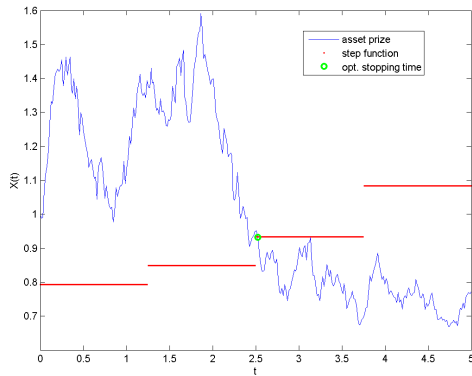


Figure 5.1.: Simulation-step 1

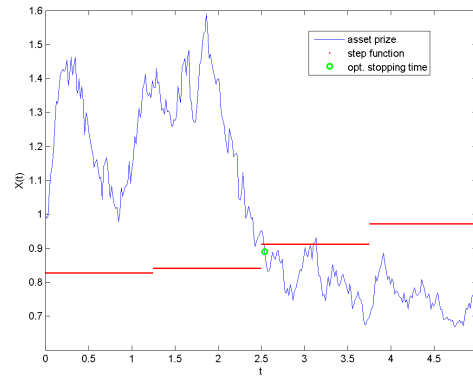


Figure 5.2.: Simulation-step 2

5. Results

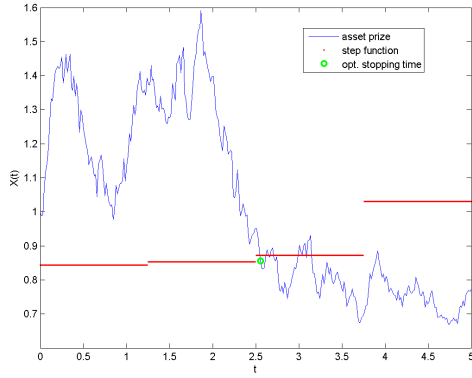


Figure 5.3.: Simulation-step 3

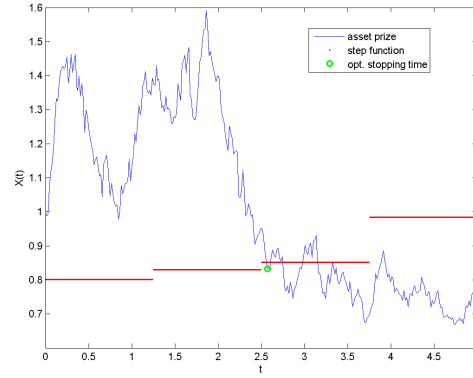


Figure 5.4.: Simulation-step 4

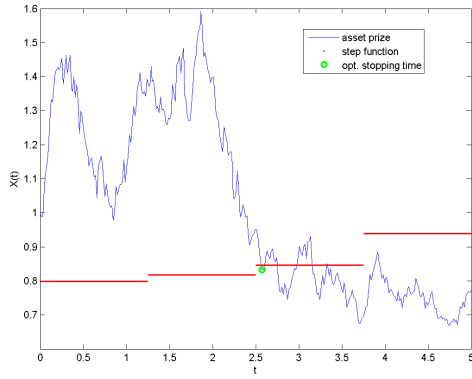


Figure 5.5.: Simulation-step 5

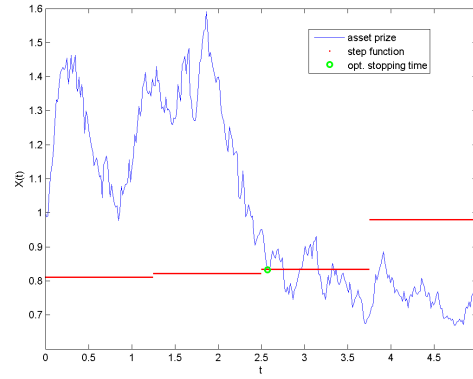


Figure 5.6.: Simulation-step 6

Shape of the 64-Step Boundary Function

Obviously, the number of steps for the approximating boundary function influences the accuracy of the algorithm directly. The more steps we allow, the more accurate the estimation of the price for the American option will be. However, increasing the number of steps will also increase the computational effort needed to come by a result. Therefore it is necessary to find an appropriate number of steps, which we set to 64. We also observed between one and two distinct bumps in disagreement with the known convex shape of the boundary function, which are most probably caused by the fact that we are working with a numerical method, where rounding errors as well as the limited number of simulation paths take their toll.

5. Results

A typical approximation of the optimal exercising boundary by a 64-step function takes a shape like in Figure 5.7.

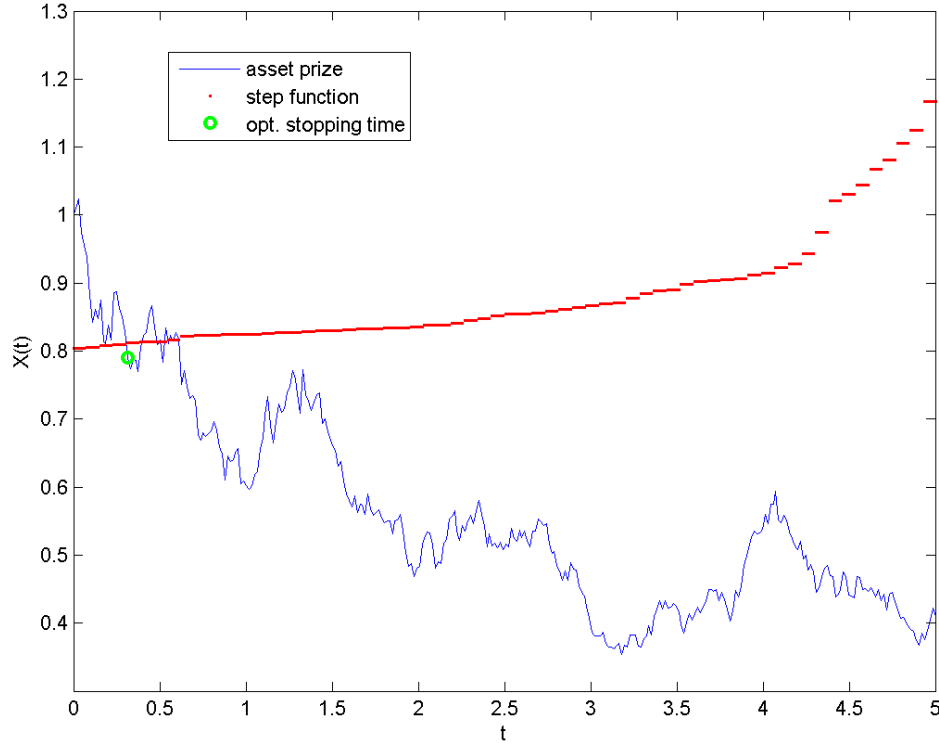


Figure 5.7.: 64-step boundary function

Overview

Considering the Black - Scholes model we used the constant interest rate $r = 4\%$ or $r = 5\%$ respectively, the volatility $\sigma = 25\%$ and an initial asset price $X(0) = 1$. We varied the strike price from $K = 80\%$ to $K = 120\%$ of the initial value in 10% steps and the maturity date, which we measured in years, from $T = 1$ to $T = 5$, where we assumed a year to consist of 64 weeks for mathematical convenience.

As representatives of all possible asset price developments, a fixed number of 3.000 paths was used. This number proved meaningful, since the computational effort increases linearly in the number of paths, whereas the accuracy of the option price could not be increased noticeably by using up to 5 times the number of paths.

Additionally, we set the sample size for updating the Cross - Entropy parameters to 1.000, whereas the actual sample used for the update of the boundary parameters, called elite sample, were the 5% of the parameters that obtained the highest payoffs.

5. Results

With these parameters we obtained the following results for the estimate of the American put option price:

	T=1	T=2	T=3	T=4	T=5
K=0.8	0.017	0.036	0.048	0.058	0.071
K=0.9	0.042	0.064	0.082	0.092	0.103
K=1.0	0.083	0.107	0.124	0.138	0.153
K=1.1	0.140	0.162	0.176	0.194	0.195
K=1.2	0.215	0.233	0.241	0.255	0.265

Table 5.1.: price estimate, when $r = 4\%$ and 64-step boundary

	T=1	T=2	T=3	T=4	T=5
K=0.8	0.015	0.032	0.041	0.050	0.055
K=0.9	0.040	0.059	0.073	0.085	0.092
K=1.0	0.080	0.100	0.118	0.125	0.134
K=1.1	0.139	0.152	0.172	0.179	0.186
K=1.2	0.212	0.226	0.234	0.246	0.244

Table 5.2.: price estimate, when $r = 5\%$ and 64-step boundary

Since our step boundary function $\tilde{b}(t)$ is only an approximation of the optimal exercising boundary $b(t)$, the stopping time

$$\tilde{\tau} = \inf \left\{ 0 \leq s \leq T - t : X_{t+s} \leq \tilde{b}(t + s) \right\} \quad (5.1)$$

is suboptimal, leading to an estimate for the American option price that is biased low. Also the results in the tables 5.1 and 5.2 are consistent in themselves, since increasing either the maturity T or the strike price K leads to a higher option price. The price for an option with interest rate $r = 4\%$ exceeds the one for an option with interest rate $r = 5\%$, since a lower interest rate r demands more initial money to allow to hedge the same risk. This can be seen when looking at the payoff of the American option for a fixed τ

$$e^{(-r\tau)}(K - X_\tau)^+, \quad (5.2)$$

Since the expression in (5.2) is decreasing as a function of r we have

$$e^{(-r_2\tau)}(K - X_\tau)^+ < e^{(-r_1\tau)}(K - X_\tau)^+, \quad \text{for } r_2 > r_1 \geq 0. \quad (5.3)$$

Applying the supremum over all admissible stopping times on both sides of this inequality does not change the inequality and hence our results coincide with the theory.

5. Results

Approximation of the optimal exercising boundary with basis functions

Alternatively we can simulate $b(t)$ using a linear combination of the basis functions

$$t^n, t^{n-1}, \dots, t, 1, \quad \text{for } n \geq 0, \quad (5.4)$$

where the coefficients in the linear combination are updated with the Cross - Entropy method. In order to be able to compare the results obtained with the different methods, we used the same parameters as above (interest rate $r = 4\%$ or 5% , volatility $\sigma = 25\%$, initial value 1, maturity $T = 1, 2, 3, 4, 5$, strike price $K = 80\%, 90\%, 100\%, 110\%, 120\%$ of the initial value, elite sample $\rho = 5\%$ and simulation sample size $N = 1.000$). As mentioned above in 4.7.2 the basis functions used for the approximation were $t^4, t^3, t^2, t, 1$.

As an example of the development of the polynomial boundary function we present the following graphs, where $r = 4\%$, $\sigma = 25\%$, $T = 4$, $K = 80\%$, $\rho = 5\%$ and $N = 1.000$.

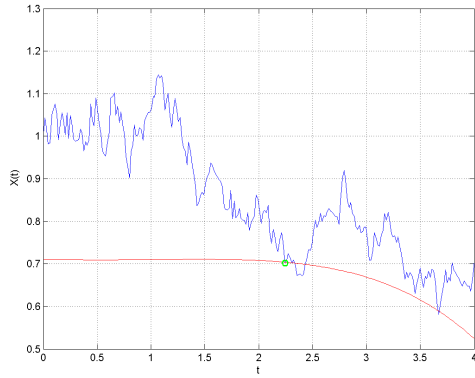


Figure 5.8.: Simulation-step 1



Figure 5.9.: Simulation-step 2



Figure 5.10.: Simulation-step 3



Figure 5.11.: Simulation-step 4

5. Results



Figure 5.12.: Simulation-step 5



Figure 5.13.: Simulation-step 6

Shape of the Approximating Boundary Function

As can be seen in Figure 5.8, the initial parameters lead to a boundary function that obviously contradicts the properties in 4.7.1. Within the next 5 steps of the Cross - Entropy method the polynomial boundary function changes so much that it resembles a function with the properties mentioned in 4.7.1.

The final approximation of the optimal exercising boundary in most cases has a similar shape as the one pictured in Figure 5.14 below.



Figure 5.14.: Polynomial boundary function

5. Results

Overview

To increase the accuracy of the algorithm a couple of options are available. Obviously, we could increase the number of paths used to capture all asset developments or the number of simulations in each step of the Cross - Entropy method, i.e. the elite sample ρN .

Alternatively, we could increase the number of basis functions used to approximate the optimal exercising boundary $b(t)$ or simply use different basis functions that fit the demanded properties for the optimal exercising boundary $b(t)$ better, as suggested in 4.7.2.

In Tables, 5.3 and 5.4, we present the results obtained by using the basis functions $t^4, t^3, t^2, t, 1$.

	T=1	T=2	T=3	T=4	T=5
K=0.8	0.016	0.035	0.046	0.059	0.068
K=0.9	0.042	0.065	0.083	0.093	0.098
K=1.0	0.084	0.109	0.126	0.138	0.147
K=1.1	0.143	0.164	0.182	0.191	0.205
K=1.2	0.219	0.231	0.249	0.257	0.265

Table 5.3.: price, when $r = 4\%$ paths 3.000

	T=1	T=2	T=3	T=4	T=5
K=0.8	0.014	0.032	0.041	0.052	0.058
K=0.9	0.040	0.062	0.073	0.089	0.092
K=1.0	0.080	0.104	0.118	0.131	0.133
K=1.1	0.137	0.162	0.169	0.182	0.192
K=1.2	0.216	0.223	0.237	0.246	0.253

Table 5.4.: price, when $r = 5\%$ paths 3.000

An augmentation of the number of asset price paths from 3.000 to 5.000 yielded the following results presented in Tables 5.5 and 5.6, where the minimal changes in the price estimates can be explained by the fact that the payoffs are path-dependent and since the number of paths is limited minor fluctuations in the price estimates may occur.

	T=1	T=2	T=3	T=4	T=5
K=0.8	0.016	0.035	0.047	0.060	0.065
K=0.9	0.042	0.064	0.083	0.090	0.101
K=1.0	0.082	0.111	0.129	0.132	0.149
K=1.1	0.142	0.166	0.179	0.194	0.200
K=1.2	0.215	0.233	0.246	0.261	0.265

Table 5.5.: price, when $r = 4\%$ paths 5.000

5. Results

	T=1	T=2	T=3	T=4	T=5
K=0.8	0.015	0.031	0.043	0.052	0.060
K=0.9	0.038	0.060	0.074	0.084	0.092
K=1.0	0.081	0.105	0.116	0.128	0.134
K=1.1	0.136	0.159	0.171	0.181	0.190
K=1.2	0.212	0.228	0.235	0.247	0.252

Table 5.6.: price, when $r = 5\%$ paths 5.000

In the case where we increased the number of basis functions from

$$t^4, t^3, t^2, t, 1 \quad (5.5)$$

to

$$t^8, t^7, t^6, t^5, t^4, t^3, t^2, t, 1 \quad (5.6)$$

the comparison of the results in Tables 5.3 and 5.4 to 5.5 and 5.6 lead to the same conclusion as above.

	T=1	T=2	T=3	T=4	T=5
K=0.8	0.015	0.036	0.049	0.060	0.063
K=0.9	0.043	0.065	0.080	0.091	0.099
K=1.0	0.084	0.111	0.127	0.133	0.145
K=1.1	0.141	0.169	0.180	0.193	0.201
K=1.2	0.216	0.233	0.245	0.257	0.267

Table 5.7.: price, when $r = 4\%$ and degree is increased

	T=1	T=2	T=3	T=4	T=5
K=0.8	0.016	0.032	0.043	0.054	0.058
K=0.9	0.039	0.061	0.074	0.087	0.092
K=1.0	0.077	0.103	0.120	0.123	0.135
K=1.1	0.136	0.155	0.172	0.183	0.183
K=1.2	0.211	0.230	0.233	0.240	0.252

Table 5.8.: price, when $r = 5\%$ and degree is increased

The changes in the option price estimates accompanying a change in the basis functions from

$$t^4, t^3, t^2, t, 1 \quad (5.7)$$

to

$$t^8, t^6, t^4, t^2, 1, \quad (5.8)$$

5. Results

where the actual number of basis functions is unaltered, are presented in Tables 5.9 and 5.10 below.

	T=1	T=2	T=3	T=4	T=5
K=0.8	0.015	0.036	0.049	0.060	0.063
K=0.9	0.043	0.065	0.080	0.091	0.099
K=1.0	0.084	0.111	0.127	0.133	0.145
K=1.1	0.141	0.169	0.180	0.193	0.201
K=1.2	0.216	0.233	0.245	0.257	0.267

Table 5.9.: price, when $r = 4\%$ and degrees are even

	T=1	T=2	T=3	T=4	T=5
K=0.8	0.015	0.032	0.043	0.054	0.058
K=0.9	0.039	0.061	0.074	0.087	0.092
K=1.0	0.077	0.103	0.120	0.123	0.135
K=1.1	0.136	0.155	0.172	0.183	0.183
K=1.2	0.211	0.230	0.233	0.240	0.252

Table 5.10.: price, when $r = 5\%$ and degrees are even

Once more no detectable improvement of the option price estimates is noticeable, which leads us to the conclusion that a deflection from the original parameters causing an increase in the computational effort is not worthwhile.

Further we know that the stopping time, using the approximation $\tilde{b}(t)$ of the optimal exercising boundary $b(t)$

$$\tilde{\tau} = \inf \left\{ 0 \leq s \leq T - t : X_{t+s} \leq \tilde{b}(t+s) \right\} \quad (5.9)$$

is suboptimal, and therefore yields an estimate for the American option price that is biased low.

As for step functions the results in Tables 5.3 to 5.10 are consistent in themselves, since increasing either the maturity T or the strike price K leads to a higher option price. Again we can observe that the price for an option with interest rate $r = 4\%$ is always higher than the one with an interest rate $r = 5\%$.

Comparison between the two Algorithms

The results obtained by means of the algorithm and presented in Tables 5.1 and 5.2 as well as in 5.3 and 5.4 show that the approach using a basis function is significantly faster than the one using a step function, since obtaining the initial parameters through an iterative method is not necessary. However, the step function approximating the optimal exercising boundary fits the known properties of the optimal exercising boundary mentioned in section 4.7.1,

5. Results

better. Note also that the initial parameters in the polynomial approach have to be chosen in a way that the approximating boundary function does not exceed the strike price at any time prior to maturity T . Consequently, the coefficients of the basis functions grow smaller and smaller as the maturity T increases, which could lead to problems when the precision of the program is insufficient. Another disadvantage of the approach using a polynomial function is that in rare occasions we obtain a final approximation of the optimal exercising boundary that clearly violates the properties mentioned in 4.7.1. The following example using the parameters $r = 5\%$, $\sigma = 25\%$, initial asset price $X(0) = 1$, maturity $T = 3$ and strike price $K = 1$ demonstrates such a boundary.

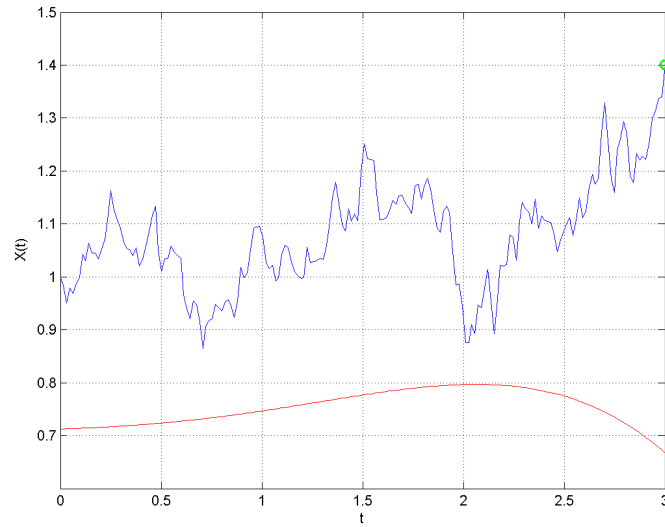


Figure 5.15.: Polynomial boundary function

The comparison of the results showed that the different estimates were within a 5% range of each other. Since each model is a simplification of the real markets, the pricing of an option is always linked to model risk, making it impossible to determine the exact value of an option. With the Cross - Entropy approach we are, however, able to give a good idea of an estimate for the American put option.

5.1.2. Basket Option

In order to be able to use the same algorithm as for the simple Vanilla American put option we assumed that the same boundary conditions are valid for Basket options. Unfortunately, there is very little known about Basket options, since the theory behind them is more complicated than in the case of an option with one underlying asset. The stopping criterion was also expected to depend on both underlying asset prices and not just on the sum of them. For this reason the postulation of a boundary depending only on the sum is clearly very crooked and therefore the corresponding lower bound obtained by a numerical procedure

5. Results

must be expected to be much cruder than in the standard Black - Scholes model with one underlying asset.

With that in mind, the results obtained can only be seen as rough estimates.

Contrary to the Vanilla American put option, the graphs in Figures 5.16 to 5.19 not only show the development of one asset price, but the paths of both asset prices and their arithmetic mean representing the evolution of the driving component of the Basket option.

In the example we used the following parameters: interest rate $r = 4\%$, which obviously has to be the same for both assets, whereas the volatilities can differ and were chosen to be $\sigma_1 = 25\%$ and $\sigma_2 = 30\%$. The initial values were $X_1(0) = 0.8$ and $X_2(0) = 1.1$, the maturity set to $T = 5$ and the strike price $K = 1.1$.

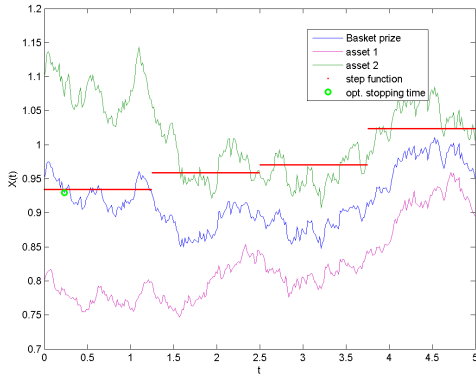


Figure 5.16.: Simulation-step 1

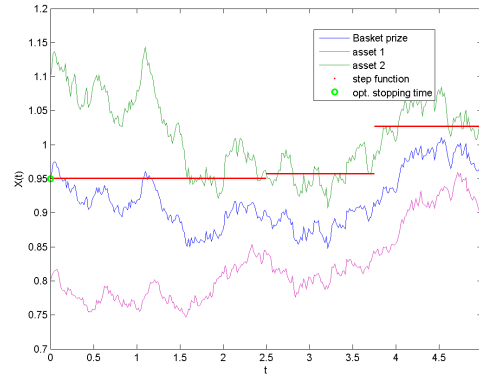


Figure 5.17.: Simulation-step 2

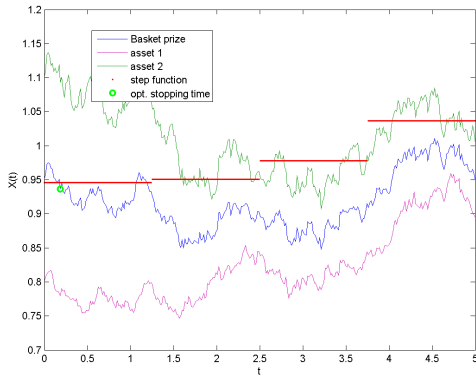


Figure 5.18.: Simulation-step 3

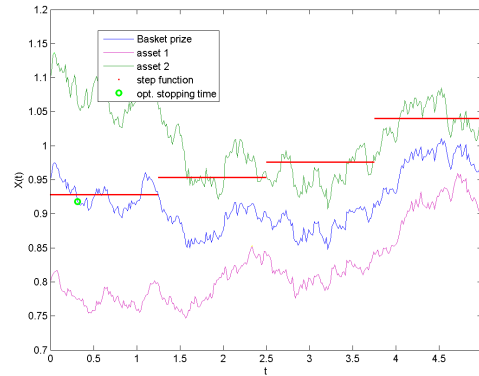


Figure 5.19.: Simulation-step 4

In Figures 5.16 to 5.19 above, it can be clearly seen how the asset prices influence the dynamics of the underlying. We can also observe that the changes in the asset price with the higher volatility fluctuate notably more.

5. Results

Shape of the 64-Step Boundary Function for Basket Options

The final approximation of the optimal exercising boundary, in the case of a Basket of two underlyings, has a shape varying between the one in Figure 5.20 and the one in Figure 5.21 below.

The first graph depicts the approximative boundary for a strike price of $K = 0.8$, an interest rate of $r = 4\%$, volatilities of $\sigma_1 = 25\%$ and $\sigma_2 = 30\%$ for the two underlying assets with a correlation of 0.4 for their Brownian motions, initial values of 0.8 and 1.1 respectively and a maturity of $T = 5$ years. For the same parameters the graph in Figure 5.21 displays an approximation of the optimal exercising boundary with a strike price of $K = 1.1$.

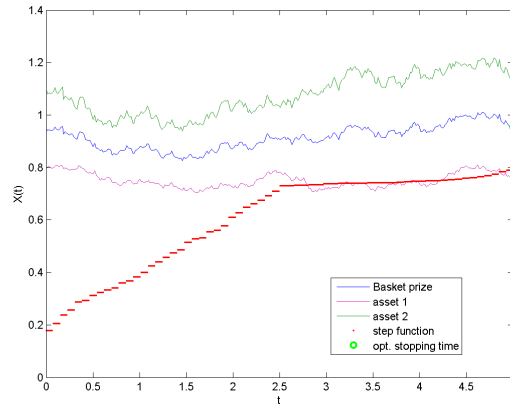


Figure 5.20.: Polynomial boundary function, $K = 0.8$

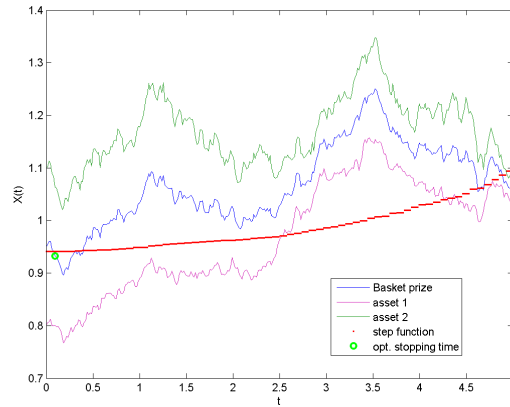


Figure 5.21.: Polynomial boundary function, $K = 1.1$

Contrary to Figure 5.21, where the boundary approximation looks similar to the one for American options on a single underlying, the boundary in Figure 5.20 differs from our ex-

5. Results

pectations, since it is obviously not convex.

Overview

In Tables 5.11 and 5.12 the price estimates for an interest rate of $r = 4\%$ or 5% respectively, maturities ranging from $T = 1$ to $T = 5$ years and strikes of $K = 0.8, 0.9, 1.0, 1.1, 1.2$ are displayed. The geometric Brownian motions driving the underlying assets used the same drift r and different volatilities $\sigma_1 = 25\%$ and $\sigma_2 = 30\%$. The Brownian motions were assumed to have a correlation of 0.4 and 3.000 paths per underlying represented all possible asset developments.

	T=1	T=2	T=3	T=4	T=5
K=0.8	0.000	0.000	0.001	0.003	0.005
K=0.9	0.006	0.013	0.018	0.021	0.024
K=1.0	0.055	0.059	0.065	0.067	0.070
K=1.1	0.149	0.149	0.150	0.150	0.150
K=1.2	0.249	0.250	0.250	0.250	0.250

Table 5.11.: price, when $r = 4\%$

	T=1	T=2	T=3	T=4	T=5
K=0.8	0.000	0.000	0.001	0.001	0.002
K=0.9	0.005	0.009	0.012	0.014	0.016
K=1.0	0.051	0.054	0.056	0.058	0.059
K=1.1	0.149	0.149	0.150	0.150	0.147
K=1.2	0.249	0.250	0.250	0.250	0.250

Table 5.12.: price, when $r = 5\%$

Obviously, the calculations performed are not precise enough for the estimation of the price of a Basket with two underlyings when the duration of the contract is only $T = 1$ or $T = 2$ years and has a strike price of $K = 0.8$ for the used parameters.

We can also observe that there is no difference in the option price between interest rate $r = 4\%$ and $r = 5\%$ for a strike of $K = 1.1$ or $K = 1.2$. There the algorithm even obtains the same price for all five maturities $T = 1, 2, 3, 4, 5$.

The results obtained suggest that, without additional information about the optimal exercising boundary, the application of the Cross - Entropy algorithm in the present form is not reasonable.

5. Results

5.2. Exponential Lévy Model

In this section we want to present the collected results obtained by the application of the Cross - Entropy algorithm on the problem of pricing American put options in the exponential Lévy model with normal inverse Gaussian increments.

First of all it must be said that the exponential Lévy process with normal inverse Gaussian increments is a pure jump process, implying that there are no continuous parts in the evolution of the asset price. For better legibility, however, the obtained function values were connected by straight lines similar to the Black - Scholes model.

An obvious disadvantage is that we only have function values on a discrete grid, hence the actual jump size is impossible to determine. The example below used the parameters suggested by Rydberg in [32, p.906] for the Deutsche Bank (i.e. $r = 4\%$, $\alpha = 75.49$, $\beta = -4.089$, $\mu = 0$, $\delta = 0.012$, initial value 1, $N = 1.000$, $\rho = 5\%$, with 3.000 different asset paths) to demonstrate the change in the asset price and in the boundary approximation.

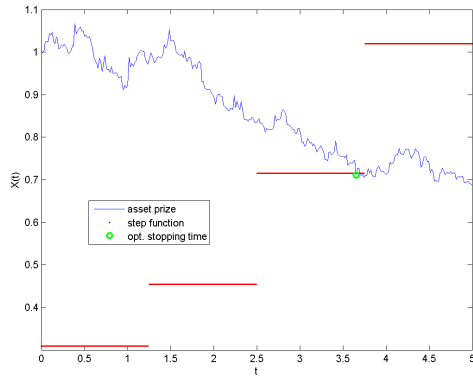


Figure 5.22.: Simulation-step 1

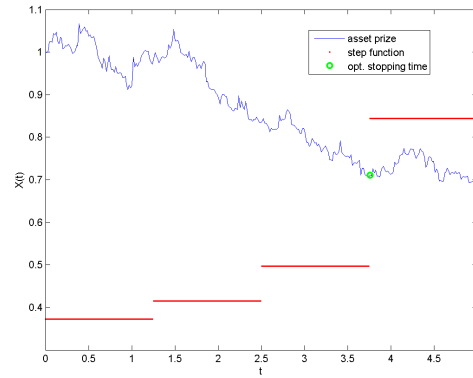


Figure 5.23.: Simulation-step 2

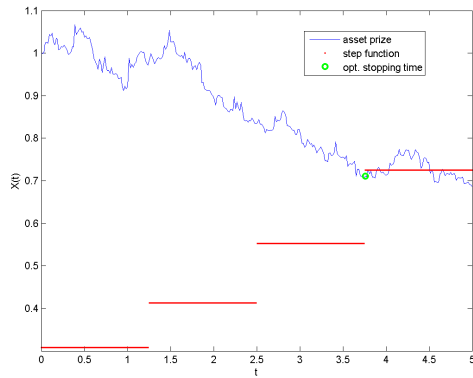


Figure 5.24.: Simulation-step 3

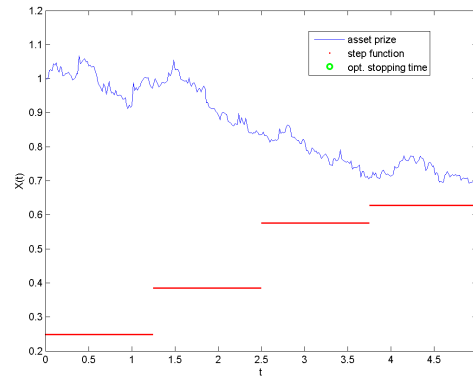


Figure 5.25.: Simulation-step 4

5. Results

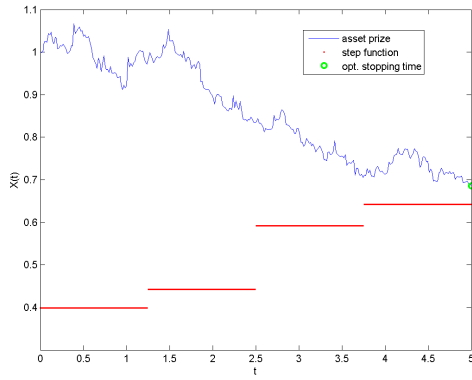


Figure 5.26.: Simulation-step 5

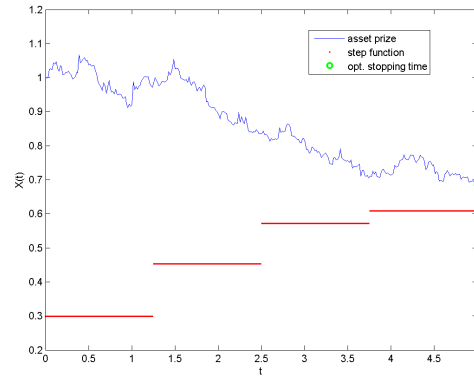


Figure 5.27.: Simulation-step 6

Shape of the 64-step boundary function

A 64-step boundary function was used to obtain good estimates for the prices of American put options, where the underlying is driven by an exponential Lévy model with normal inverse Gaussian increments. Figure 5.28 depicts the general shape of the 64-step boundary approximation acquired with the Cross - Entropy algorithm and is consistent with the properties in 4.7.1. Note also that a big change in the price evolution between two consecutive time steps corresponds to a 'real' jump in the asset price.

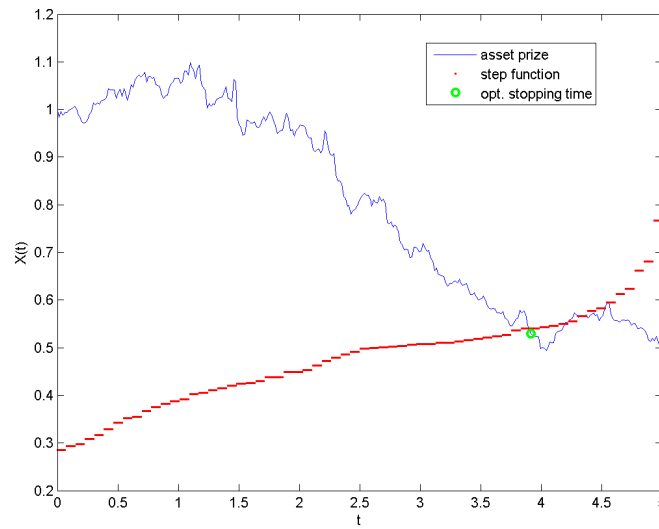


Figure 5.28.: 64-step boundary function

5. Results

Overview

Considering the exponential Lévy model with normal inverse Gaussian increments we used the constant interest rate $r = 4\%$ or $r = 5\%$ respectively, an initial asset value of 1, maturities $T = 1, 2, 3, 4, 5$ years and strikes of $K = 80\%, 90\%, 100\%, 110\%, 120\%$. The parameters $\alpha = 75.49$, $\beta = -4.089$, $\mu = 0$ and $\delta = 0.012$ describe the increment process, where α is the tail heaviness, β the skewness, μ the location and δ the scale. Once more 3.000 paths were used to simulate all possible asset price developments. Additionally, we set the sample size for updating the Cross - Entropy parameters to 1.000, whereas the actual sample used to update the boundary parameters, called elite sample, were the 5% of the parameters that obtained the highest payoffs.

With these parameters we obtained the following results for the estimates of the American put option prices in the exponential Lévy model with normal inverse Gaussian increments:

	T=1	T=2	T=3	T=4	T=5
K=0.8	0.001	0.008	0.020	0.034	0.047
K=0.9	0.012	0.036	0.056	0.077	0.095
K=1.0	0.058	0.084	0.112	0.138	0.155
K=1.1	0.132	0.167	0.187	0.213	0.230
K=1.2	0.226	0.249	0.267	0.290	0.303

Table 5.13.: price estimate, when $r = 4\%$ and 64-step boundary

	T=1	T=2	T=3	T=4	T=5
K=0.8	0.001	0.008	0.020	0.031	0.047
K=0.9	0.012	0.033	0.055	0.074	0.088
K=1.0	0.057	0.088	0.112	0.134	0.146
K=1.1	0.135	0.162	0.182	0.201	0.215
K=1.2	0.223	0.248	0.265	0.273	0.290

Table 5.14.: price estimate, when $r = 5\%$ and 64-step boundary

As for the Black - Scholes model the step boundary function $\tilde{b}(t)$ is only an approximation of the optimal exercising boundary $b(t)$, the stopping time

$$\tilde{\tau} = \inf \left\{ 0 \leq s \leq T - t : X_{t+s} \leq \tilde{b}(t + s) \right\} \quad (5.10)$$

is therefore suboptimal, leading to an estimate for the American option price that is biased low. Also the results in Tables 5.13 and 5.14 are consistent in themselves, since increasing either the maturity T or the strike price K leads once more to a higher option price. The price for an option with interest rate $r = 4\%$ exceeds the one for an option with interest rate $r = 5\%$, since a lower interest rate r demands more initial money to allow to hedge the same risk. Tables 5.13 and 5.14 satisfy this property overall, except for $T = 1$, $K = 1.1$

5. Results

and $T = 2$, $K = 1.0$, where the price estimates are higher for an interest rate of $r = 5\%$. This difference in the price estimates can be explained by the limited amount of asset paths used.

Approximation of the optimal exercising boundary with basis functions

As in the standard Black - Scholes model the optimal exercising boundary $b(t)$ was simulated using a linear combination of the following basis functions

$$t^n, t^{n-1}, \dots, t, 1, \quad \text{for } n \geq 0, \quad (5.11)$$

where the coefficients in the linear combination were updated with the Cross - Entropy method. To be able to compare the results between the two approaches we used the same parameters as before ($r = 4\%$ or $r = 5\%$ respectively, an initial asset value of 1, maturities $T = 1, 2, 3, 4, 5$ years, strikes of $K = 80\%, 90\%, 100\%, 110\%, 120\%$, 3.000 asset evolutions, 1.000 simulations of the Cross - Entropy with an elite sample of 5%, tail heaviness $\alpha = 75.49$, skewness $\beta = -4.089$, location $\mu = 0$ and scale parameter $\delta = 0.012$).

As mentioned above in 4.7.2, the basis functions used for the approximation of the optimal exercising boundary $b(t)$ were $t^4, t^3, t^2, t, 1$.

Additionally, we increased the number of asset evolutions from 3.000 to 10.000 in a second run to get an idea of the sensitivity of the price estimates towards the number of asset evolutions.

As an example of the development of the polynomial boundary function as well as the evolution of the asset in an exponential Lévy model we present the following graphs, where $r = 5\%$, $T = 5$, $K = 120\%$, $\rho = 5\%$ and $N = 1.000$.

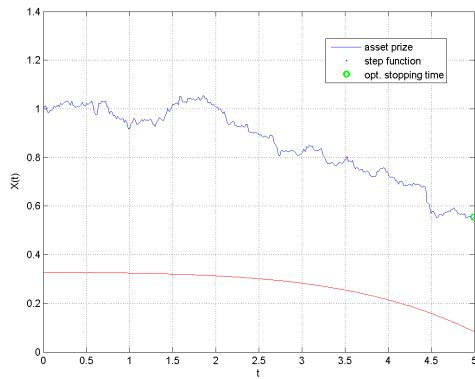


Figure 5.29.: Simulation-step 1

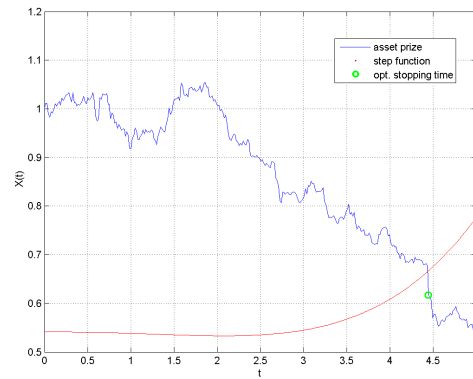


Figure 5.30.: Simulation-step 2

5. Results

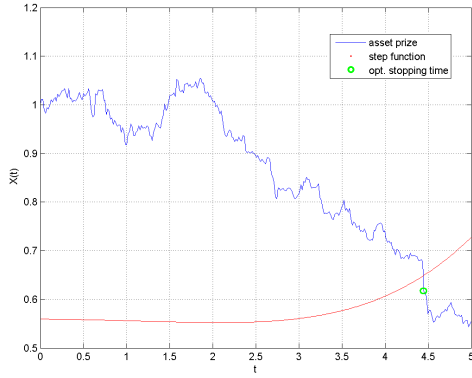


Figure 5.31.: Simulation-step 3

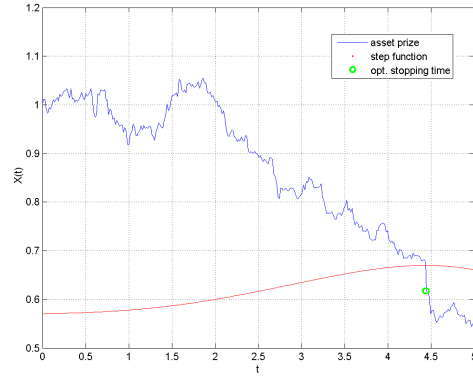


Figure 5.32.: Simulation-step 4

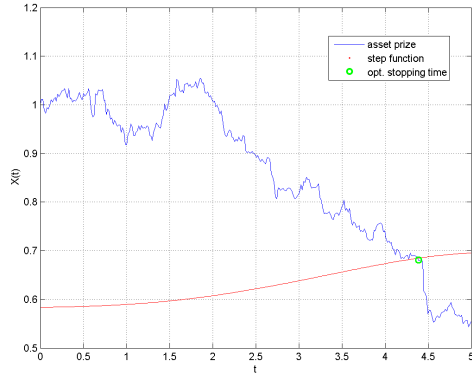


Figure 5.33.: Simulation-step 5

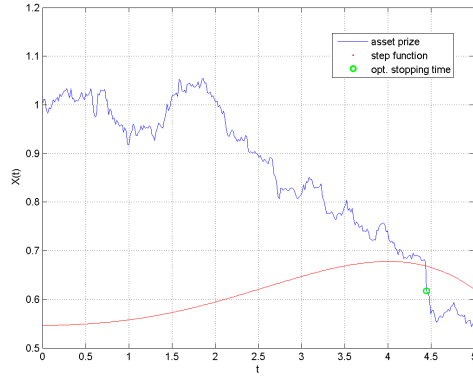


Figure 5.34.: Simulation-step 6

Shape of the approximating boundary function

Figures 5.29 to 5.34 demonstrate that despite the fact that the initial boundary clearly contradicts the properties in 4.7.1 the final approximation fits the necessary conditions, besides having a slight downward turn close to maturity. This downward turn originates in the boundary condition that the optimal exercising boundary has to stay beneath the strike price for the whole interval $[0, T]$.

The final approximation of the optimal exercising boundary, in most cases, has a similar shape as the one pictured in Figure 5.35 below, where the following parameters were used: interest rate $r = 5\%$, $T = 5$, strike price $K = 1.2$ and initial value 1.

5. Results

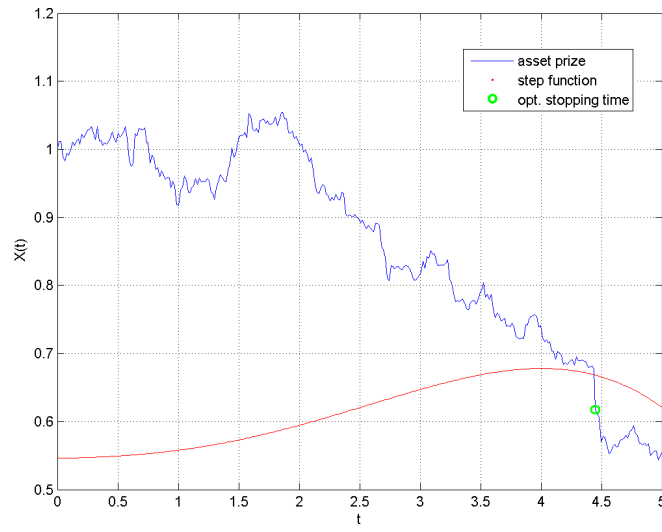


Figure 5.35.: Polynomial boundary function

Overview

In Tables 5.15 and 5.18 we present the results obtained by using the basis functions $t^4, t^3, t^2, t, 1$ for 3.000 as well as for 10.000 paths.

	T=1	T=2	T=3	T=4	T=5
K=0.8	0.001	0.008	0.018	0.032	0.047
K=0.9	0.013	0.034	0.056	0.076	0.096
K=1.0	0.056	0.089	0.116	0.137	0.156
K=1.1	0.133	0.166	0.188	0.210	0.222
K=1.2	0.228	0.253	0.272	0.293	0.302

Table 5.15.: price, when $r = 4\%$ paths 3.000

5. Results

	T=1	T=2	T=3	T=4	T=5
K=0.8	0.001	0.008	0.018	0.032	0.045
K=0.9	0.013	0.033	0.055	0.072	0.089
K=1.0	0.056	0.086	0.111	0.134	0.148
K=1.1	0.132	0.160	0.183	0.201	0.218
K=1.2	0.226	0.247	0.263	0.279	0.292

Table 5.16.: price, when $r = 5\%$ paths 3.000

	T=1	T=2	T=3	T=4	T=5
K=0.8	0.001	0.008	0.020	0.033	0.047
K=0.9	0.012	0.035	0.056	0.077	0.094
K=1.0	0.058	0.088	0.113	0.136	0.158
K=1.1	0.134	0.165	0.188	0.210	0.229
K=1.2	0.227	0.248	0.274	0.289	0.302

Table 5.17.: price, when $r = 4\%$ paths 10.000

	T=1	T=2	T=3	T=4	T=5
K=0.8	0.001	0.008	0.019	0.031	0.044
K=0.9	0.012	0.034	0.055	0.073	0.091
K=1.0	0.056	0.087	0.111	0.131	0.149
K=1.1	0.132	0.160	0.184	0.202	0.218
K=1.2	0.224	0.247	0.265	0.280	0.292

Table 5.18.: price, when $r = 5\%$ paths 10.000

5. Results

Once more no detectable improvement of the option price estimates is noticeable, which leads us to the conclusion that a deflection from the original parameters causing an increase in the computational effort is not worthwhile.

As pointed out in the sections above, it is not possible for the boundary approximation to achieve more accurate estimates of the option price than the optimal exercising boundary itself. Hence, the estimates obtained above are biased low. As for step functions the results in Tables 5.15 to 5.18 are consistent in themselves, since increasing either the maturity T or the strike price K leads to a higher option price. Again we can observe that the price for an option with interest rate $r = 4\%$ is higher than the one with an interest rate $r = 5\%$.

Comparison between the two algorithms

In this model the downward turn of the polynomial boundary approximation close to maturity occurs more often than in the Black - Scholes model. This is clearly an unwanted feature of this approach, since the optimal exercising boundary is known to be non-decreasing on $[0, T]$, as pointed out in 4.7.1 and mentioned in the sections above. Since the only change in the program between the two models is the generation of the asset evolutions, the program using a polynomial approach for the approximation of the optimal exercising boundary $b(t)$ terminates faster again, whereas the step function boundary satisfies the properties of the optimal exercising boundary, mentioned in 4.7.1, better. Also in this model the results obtained by the two methods are within a 5% range of each other, giving us reason to believe that the obtained estimates are sensible.

6. Conclusion & Further Work

The results obtained in section 5 show that both algorithms yield useful estimates of the American put option with a single underlying asset.

By comparing the price estimates for a fixed strike price in the two models it can be noticed that the estimates in the standard Black - Scholes model exceed the ones in the exponential Lévy model under the given set of parameters for a strike price below and up to the initial value of the asset. The price estimates in the exponential Lévy model, on the contrary, exceed the ones in the standard Black - Scholes model under the same set of parameters for a strike price above the initial asset price. Since a higher option price entails a higher average payoff, which in turn corresponds to a lower average asset evolution, the price estimates observed suggest that the asset evolution in the exponential Lévy model is less volatile than in the standard Black - Scholes model.

We could also observe that the final approximation in the step function approach fitted the known properties of the optimal exercising boundary in 4.7.1 better, since the polynomial approximation exhibited a downward turn close to the maturity date. This downward turn comes from the condition that the optimal exercising boundary has to stay beneath the strike price at all times $t \in [0, T]$. It was also pointed out in the results that this downward drift occurred more often in the exponential Lévy model than in the standard Black - Scholes model, which is caused by the more pessimistic parameter approach in the former model.

The program using the polynomial approach, however, terminated much faster than the one using the step functions, since in the latter an iterative method, described in 4.7.4, was used to find the initial parameters of the 64-step function.

In both models various changes in the parameters have been implemented, but neither a change in the number of asset paths nor in the sample size of the Cross - Entropy method yielded a notable improvement in the price estimates in either model. In the case where the optimal exercising boundary was approximated with a polynomial we also increased the number of the basis functions used, but once more no noticeable improvement was observable.

The results obtained for the Basket options in 5.1.2 enforced our guess that in this case further research on the optimal exercising boundary is necessary to allow a reasonable application of the Cross - Entropy algorithm. It would make sense that the optimal exercising strategy depends on both underlying assets and not only their sum.

Further work following up this dissertation would be to investigate more general stochastic optimal control problems in finance, as the ones in Björk [4], and their solution via the Cross - Entropy method.

A. Algorithms

For the sake of completeness we state the various algorithms developed for the pricing of the American put option. The code presented in the next sections is written in MATLAB.

A.1. Cross - Entropy Algorithm, Approximation of the Optimal Exercising Boundary with Step Functions

```
1 % INPUT:
2 % r          interest rate
3 % sigma_gbm  volatility gbm
4 % startT, endT  alg. determ. opt. prices for maturities fr. startT to endT
5 % x          initial asset value
6 % startK, endK  alg. determines opt. prices for strike prices from startK
7 %            to endK
8 % rho        rho*N is the number of elite samples
9
10 % functions needed: gbm.m
11
12
13 % CONNECTION between timestep and actual index of timestep
14 % (index-1)/stepsize=timestep
15
16 function [] = final_version_pao_steps(r,sigma_gbm,startT,endT,x,startK,...
17     endK,N,rho,nob_exp,number_of_gbm_paths)
18
19 % final_version_pao_steps(0.05,0.25,1,5,1,0.8,1.2,1000,0.05,6,3000)
20
21 for T=startT:endT,
22     for K=startK:0.1:endK,
23         tic;
24
25         % to save date in filename
26         cdate = date;
27         % display format
28         format short
29
30         % stepsize for Gbm
31         n= T*64-1;
32
33         % saves the parameters from the last round
34         mu2=[];
35
36         % tailor the folder and the file name
37         testfname= sprintf('...txt',...)
```

A. Algorithms

```

38 fid=fopen(testfname, 'a');
39 fprintf(fid, '... ', ...);
40
41 % calculate for 2 steps then for 4 .. until we reach 2^nob
42 for m=1:nob_exp,
43
44     % payoff vector
45     payoff=zeros(N,1);
46
47     % variable that checks if last 5 payoffs the same
48     checkpayoff = 0;
49
50     % counts in which round of the CE while loop we are
51     roundcount = 0;
52
53     % set gamma 0 at each boundary step
54     gamma = [];
55
56     % simulate Gbm paths
57     % timesteps of the paths of Gbm
58     t = (0:1:n)/n;
59     t=t*T;
60     % initialize matrix for Gmb paths
61     % gbm generates vec with n+1 entries, since for 0 also value needed
62     Gbm = zeros(number_of_gbm_paths,n+1);
63     % simulate the gbm path once
64     for k=1:number_of_gbm_paths,
65         Gbm(k,:) = gbm(n, r,sigma_gbm,T,x);
66     end
67
68
69     % initialize boundary
70
71     % const. optim. exercising boundary perpetual American put case
72     % optimal exercising boundary in the finite case is always above
73     perpetual_opt_boundary = K/(1+sigma_gbm^2/(2*r))
74
75     % for the plot
76     perpet_vec=[];
77     for k=1:n+1,
78         perpet_vec = [perpet_vec perpetual_opt_boundary];
79     end
80
81     % initialize boundary, value at every time of Gbm
82     b = zeros(N,n+1);
83
84     % linear function starting at the const perpetual optimal exercise bound
85     % ends at the strike price
86     mu=[];
87     if (m==1)
88         for k=0:1,
89             mu=[mu (1-k/2)*perpetual_opt_boundary+k*1/2*K];
90         end
91     % initial standard deviation is strike price

```

A. Algorithms

```

92     SIGMA=eye(2)*K^2;
93 else
94     % last is a arithmetic middle between last of former step and K
95     for k=1:2^(m-1)-1,
96         mu = [mu mu2(1,k) 1/2*mu2(1,k)+1/2*mu2(1,k+1)];
97     end
98     mu=[mu mu2(1,2^(m-1)) 1/2*mu2(1,2^(m-1))+1/2*K];
99     SIGMA=eye(2^m);
100    for k=1:2^m-2,
101        SIGMA(k,k)= (mu(1,k+1)-mu(1,k))^2/4;
102    end
103    SIGMA(2^m-1,2^m-1)=SIGMA(2^m-2,2^m-2);
104    SIGMA(2^m,2^m)=SIGMA(2^m-2,2^m-2);
105 end
106
107 mu
108 SIGMA
109 fprintf(fid,'m = % d\r\n\r\n',m);
110 fprintf(fid,'mu=\r\n');
111 dlmwrite(testfname,mu,'-append','delimiter','\t','newline','pc',...
112     'precision',3, 'roffset',1 );
113 fprintf(fid,'\r\n SIGMA=\r\n');
114 dlmwrite(testfname,SIGMA,'-append','delimiter','\t','newline','pc',...
115     'precision',3, 'roffset',1 );
116
117 mnrn=[];
118 % smoothed updating parameter between 0.3 and 0.9
119 smoothed_upd_param = 0.6;
120
121 % MAIN
122 % Cross entropy evaluates the optimal boundaries
123 while (checkpayoff ==0 ),
124
125     roundcount = roundcount+1
126     fprintf(fid,'\r\n\r\n roundcount = % d\r\n',roundcount);
127
128     % for each sample calculate the payoff
129     for k=1:N,
130
131         % begin calculation of one special payoff
132         mnrn(k,:)=mvnrnd(mu,SIGMA);
133         boundary_correct = 0;
134         % numbers of simulation for a correct boundary
135         simnumber_for_correct_boundary = 0;
136
137         % check if boundary has the right shape: increasing >0, <=K
138         while (boundary_correct ==0)
139
140             % assume boundary correct, check all properties
141             % if one is not satisfied new boundary
142             boundary_correct =1;
143
144             simnumber_for_correct_boundary = simnumber_for_correct_boundary +1;
145

```

A. Algorithms

```

146     % value of boundary for each step of Gbm
147
148     % starts above perpetual opt value and below K
149     if (mnrn(k,1)<perpetual_opt_boundary || mnrn(k,1)>K)
150         boundary_correct=0;
151     end
152
153     if (boundary_correct ==1)
154         for j=2:2^m,
155             if (mnrn(k,j)<mnrn(k,j-1) || (mnrn(k,j) > K))
156                 boundary_correct = 0;
157                 % if boundary incorrect look directly for new one
158                 break
159             end
160         end
161     end
162
163     % boundary not correct, determine new one
164     if (boundary_correct ==0)
165         mnrn(k,:)= mvnrnd(mu,SIGMA);
166     end
167
168 end
169
170 % simnumber_for_correct_boundary
171 % mnrn(j,:)
172
173 % make boundary function out of the steps
174 for j=1:(2^m-1),
175     for i=1:floor((n+1)/2^m),
176         b(k,(j-1)*floor((n+1)/2^m)+i)=mnrn(k,j);
177     end
178 end
179
180 for j=1:(n+1)-(2^m-1)*floor((n+1)/2^m),
181     b(k,(2^m-1)*floor((n+1)/2^m)+j)=mnrn(k,2^m);
182 end
183 % b(k,:)
184
185 % checkout stoppingtime for each path
186 % initialization
187 stopping_time_index = ones(number_of_gbm_paths,1)';
188 stopping_time = ones(number_of_gbm_paths,1);
189
190 for j=1:number_of_gbm_paths,
191     % find first hitting time
192     % take into account that boundary changes
193     % if no value goes below the current boundary until maturity
194     % exercise anyway
195
196     % checkvar =1 tells me that stopping time was already found
197     checkvar = 0;
198
199     % find stopping time

```


A. Algorithms

```

200 while (checkvar == 0 && (stopping_time_index(j) < n+1))
201     if (Gbm(j, stopping_time_index(j)) <= b(k, stopping_time_index(j)))
202         checkvar = 1;
203     else
204         stopping_time_index(j) = stopping_time_index(j)+1;
205     end
206 end
207
208 % convert from index of Gbm to actual time
209 stopping_time(j) = (stopping_time_index(j)-1)*T/n;
210
211 % plot Gbm paths with current boundary
212 if (j==1 && k==1) % plots same paths with updated boundary
213     figure('visible','off');
214     plot(t, Gbm(j,:), 'b');
215     xlabel('t'); ylabel('X(t)');
216     hold on
217     plot(t, b(k,:), 'rs', 'MarkerEdgeColor','r', ...
218          'MarkerFaceColor','r', 'MarkerSize', 1);
219     plot(stopping_time(j), Gbm(j, stopping_time_index(j)), ...
220          'go', 'LineWidth', 2);
221     legend('asset', 'step funct.', 'opt. stopp. time', 'Location','Best');
222     hold off
223     % auto save figures
224     fname = sprintf('... ', ...);
225     saveas(gcf, fname);
226 end
227 end
228
229 for j=1:number_of_gbm_paths,
230     payoff(k) = payoff(k) + exp(-r*stopping_time(j))*...
231         max(0, (K-Gbm(j, stopping_time_index(j))));
232 end
233 payoff(k) = 1/number_of_gbm_paths*payoff(k);
234 % payoff(k) is performance of one sample
235 end
236
237
238 % begin UPDATE parameters
239 % gives the order statistic for payoff vector
240 [payoff_sorted, IX] = sort(payoff, 1);
241
242 % boundary coefficients belonging to those payoffs
243 % sort the multivariate normal parameters and update them
244 sorted_sample = mnrn(IX, :);
245
246 % get new level for CE
247 gamma(roundcount) = payoff_sorted(ceil((1-rho)*N));
248
249 SIGMA2 = zeros(2^m, 2^m);
250
251 % for the elite sample calculate the
252 % average expected value of first step of boundary
253

```

A. Algorithms

```

254 % for the coefficient of each power calculate average of the elite s.
255 for k=1:2^m,
256     mu2(1,k)=0;
257     for j=ceil((1-rho)*N):N,
258         mu2(1,k) = mu2(1,k)+ sorted_sample(j,k);
259     end
260     mu2(1,k) = 1/(floor(rho*N)+1)*mu2(1,k);
261     mu(1,k) = smoothed_upd_param*mu2(1,k) + ...
262         (1-smoothed_upd_param)*mu(1,k);
263 end
264
265 % same for their covariance matrix
266 for k=1:2^m,
267     for j=1:2^m,
268         for i=ceil((1-rho)*N):N,
269             SIGMA2(k,j) = SIGMA2(k,j)+ (sorted_sample(i,k)-mu(1,k))*...
270                 (sorted_sample(i,j)-mu(1,j));
271         end
272     end
273     % no -1 since we have floor()+1
274     SIGMA2(k,j) = 1/(floor(rho*N))*SIGMA2(k,j);
275     SIGMA(k,j) = smoothed_upd_param*SIGMA2(k,j) + ...
276         (1-smoothed_upd_param)*SIGMA(k,j);
277 end
278 end
279
280 % if last 5 payoffs the same 3 digits behind the comma, stop algorithm
281 if ( roundcount > 4 && ...
282     (floor(1000*gamma(roundcount))==floor(1000*gamma(roundcount-1)))&&...
283     (floor(1000*gamma(roundcount))==floor(1000*gamma(roundcount-2)))&&...
284     (floor(1000*gamma(roundcount))==floor(1000*gamma(roundcount-3)))&&...
285     (floor(1000*gamma(roundcount)) == floor(1000*gamma(roundcount-4))))
286     checkpoint=1;
287 end
288 end
289 % end main
290
291 % save parameterss to use it in the next simulation step where we have
292 % double as many boundary steps
293 mu2 = mu
294 SIGMA2 = SIGMA
295 levels_for_payoff = gamma'
296 roundcount
297 fprintf(fid, '\r\n mu2=\r\n');
298 dlmwrite(testfname, mu2, '-append', 'delimiter', '\t', 'newline', 'pc', ...
299     'precision', 3, 'roffset', 1 );
300 fprintf(fid, '\r\n SIGMA2=\r\n');
301 dlmwrite(testfname, SIGMA2, '-append', 'delimiter', '\t', 'newline', 'pc', ...
302     'precision', 3, 'roffset', 1 );
303 fprintf(fid, '\r\n levels_for_payoff=\r\n\r\n');
304 dlmwrite(testfname, levels_for_payoff, '-append', 'delimiter', '\t', ...
305     'newline', 'pc', 'precision', 3, 'roffset', 1 );
306 fprintf(fid, '\r\n roundcount = % d\r\n\r\n\r\n\r\n', roundcount);
307 end

```

A. Algorithms

```
308
309 fclose(fid);
310 % write the final price in extra file
311 pricelist=sprintf('...\prices.txt')
312 priceid=fopen(pricelist,'a');
313 fprintf(priceid,'... ',...);
314 fclose(priceid);
315 toc;
316 end
317 end
```

A.2. Cross - Entropy Algorithm, Approximation of the Optimal Exercising Boundary with a Linear Combination of Basis Functions

```
1 % INPUT:
2 % n stepsize for Gbm
3 % r drift parameter
4 % sigma_gbm volatility gbm
5 % T maturity date
6 % x starting point for Gbm
7 % K strike price
8 % M number of sample = number of average payoffs
9 % rho: rho*N number of elite samples
10 % the function gbm is needed
11
12 % CONNECTION between timestep and actual index of timestep
13 % (index-1)/stepsize=timestep
14
15 function [] = final_version_pao_poly_all(r,sigma_gbm,startT,endT,x,...
16     startK,endK,N,rho,polygrad,number_of_gbm_paths)
17 % final_version_pao_poly_all(0.04,0.25,1,5,1,0.8,1.2,1000,0.05,4,3000)
18
19 for T=startT:endT,
20     for K=startK:0.1:endK,
21         tic;
22         % degree has to be an even number
23         if (mod(polygrad,2) ~= 0)
24             polygrad=polygrad+1;
25             fprintf('polynomialdegree was not even, so added 1 degree')
26         end
27
28         % 64 weeks per year to compare results
29         n=T*64-1;
30         % in polynomial case set nob=#coefficients
31         nob=polygrad+1;
32
33         % tailor the folder and the file name
34         testfname= sprintf('...txt'...)
35         fid=fopen(testfname,'a');
36         fprintf(fid,'... ',...);
37
```

A. Algorithms

```

38 % payoff vector
39 payoff=zeros(N,1);
40
41 % variable that checks if last 5 payoffs the same
42 checkpayoff = 0;
43
44 % counts in which round of the CE while loop we are
45 roundcount = 0;
46
47 % simulate Gbm paths
48 % timesteps of the paths of Gbm
49 t = (0:1:n)'/n;
50 t=t*T;
51 % initialize matrix for Gbm paths
52 % gbm generates vec with n+1 entries, since for 0 also value needed
53 Gbm = zeros(number_of_gbm_paths,n+1);
54 %simulate the gbm path once
55 for i=1:number_of_gbm_paths,
56     Gbm(i,:) = gbm(n, r,sigma_gbm,T,x);
57 end
58
59 % initialize boundary
60 % boundary function is: sum of basis functions
61
62 % initialize boundary, value at every time of Gbm
63 b = zeros(N,n+1);
64
65 % const. optim. exercising boundary perpetual American put case
66 % optimal exercising boundary in the finite case is always above
67 perpetual_opt_boundary = K/(1+sigma_gbm^2/(2*r))
68
69 % coefficients
70 alpha=zeros(N,polygrad+1);
71
72 mu=alpha(1,:);
73
74 factor=K/T^polygrad;
75 SIGMA=eye(polygrad+1)*factor^2;
76
77 % constant value has to be at least as large as perpetual option value
78 SIGMA(polygrad+1,polygrad+1)= (1/2*perpetual_opt_boundary+1/2*K)^2/4;
79
80 fprintf(fid, 'mu=\r\n');
81 dlmwrite(testfname,mu,'-append','delimiter','\t','newline','pc',...
82     'precision',3, 'roffset',1 );
83 fprintf(fid, '\r\n SIGMA=\r\n');
84 dlmwrite(testfname,SIGMA,'-append','delimiter','\t','newline','pc',...
85     'precision',3, 'roffset',1 );
86
87
88 % MAIN
89 % Cross entropy evaluates the optimal boundaries
90 while (checkpayoff ==0 ),
91

```

A. Algorithms

```

92 roundcount = roundcount+1
93 % for each sample calculate the payoff
94 for j=1:N,
95
96     % begin calculation of one special payoff
97     alpha(j,:)=mvnrnd(mu,SIGMA);
98     boundary_correct = 0;
99     % numbers of simulation for a correct boundary
100     simnumber_for_correct_boundary = 0;
101
102     % check if boundary has the right shape: increasing >0, <=K
103     % value not less than last one
104     while (boundary_correct ==0)
105
106         % assume boundary correct, check all properties
107         % if one is not satisfied new boundary
108         boundary_correct =1;
109
110         simnumber_for_correct_boundary = simnumber_for_correct_boundary +1;
111
112         % value of boundary for each step of Gbm
113         b(j,:)=polyval(alpha(j,1:polygrad+1), t);
114
115         % check if boundary above perpet_val and below K at each step
116         if (boundary_correct ==1)
117             for k=1:n+1,
118                 if (b(j,k) < perpetual_opt_boundary || b(j,k) > K)
119                     boundary_correct = 0;
120                     % fprintf('the problem is index %d with value %d\r\n',k,b(j,k));
121                     break;
122                 end
123             end
124         end
125
126         % boundary not correct, determine new one
127         if (boundary_correct ==0)
128             alpha(j,:) = mvnrnd(mu,SIGMA);
129         end
130     end
131
132     % simnumber_for_correct_boundary
133
134     % checkout stoppingtime for each path
135     % initialization
136     stopping_time_index = ones(number_of_gbm_paths,1)';
137     stopping_time = ones(number_of_gbm_paths,1);
138
139     for i=1:number_of_gbm_paths,
140         % find first hitting time
141         % take into account that boundary changes
142         % if no value goes below the current boundary until maturity
143         % exercise anyway
144
145         % checkvar =1 tells me that stopping time was already found

```

A. Algorithms

```

146     checkvar = 0;
147
148     % find stopping time
149     while (checkvar == 0 && (stopping_time_index(i) < n+1))
150         if (Gbm(i, stopping_time_index(i)) <= b(j, stopping_time_index(i)))
151             checkvar = 1;
152         else
153             stopping_time_index(i) = stopping_time_index(i)+1;
154         end
155     end
156
157     % convert from index of Gbm to actual time
158     stopping_time(i) = (stopping_time_index(i)-1)*T/n;
159
160     % plot Gbm paths with current boundary
161     if ( i==1 && j==1)
162         figure('visible','off');
163         plot(t,Gbm(i,:), 'b'), title('Gbm mit boundaries'); xlabel('t'); grid
164         hold on
165         plot(t,b(j,:), 'r', stopping_time(i), ...
166             Gbm(i, stopping_time_index(i)), 'go')
167         hold off
168         fname = sprintf('...png', ...);
169         saveas(gcf, fname);
170     end
171 end
172
173 for p=1:number_of_gbm_paths,
174     payoff(j) = payoff(j) + ...
175         exp(-r*stopping_time(p))*max(0, (K-Gbm(p, stopping_time_index(p))));
176 end
177 payoff(j) = 1/number_of_gbm_paths*payoff(j);
178 % end payoff(j)
179 end
180
181 % begin UPDATE parameters
182 % gives the order statistic for payoff vector
183 [payoff_sorted, IX] = sort(payoff, 1);
184 % boundary coefficients belonging to those payoffs
185 sorted_sample=alpha(IX,:);
186
187 % get new level, normally (1-rho)*#sims, first turn rho sample of fewer
188 % simulations
189 gamma(roundcount)=payoff_sorted(ceil((1-rho)*N));
190
191 SIGMA = zeros(polygrad+1, polygrad+1);
192
193 % for the rho*N boundaries that brought the highest payoff calculate the
194 % average expected value of first step of boundary
195
196 % for the coefficient of each power calculate average of the elite s.
197 for q=1:polygrad+1,
198     mu(1,q)=0;
199     for i=ceil((1-rho)*N):N,

```

A. Algorithms

```

200     mu(1,q) = mu(1,q)+ sorted_sample(i,q);
201     end
202     mu(1,q) = 1/(floor(rho*N)+1)*mu(1,q);
203     end
204
205     % same for their covariance matrix
206     % could be optimized by just calculating the upper triangle and then
207     % transposing, diag zeros and adding to original
208     for q=1:polygrad+1,
209         for l=1:polygrad+1,
210             for i=ceil((1-rho)*N):N,
211                 SIGMA(q,l) = SIGMA(q,l)+ (sorted_sample(i,q)-mu(1,q))...
212                     *(sorted_sample(i,l)-mu(1,l));
213             end
214             SIGMA(q,l) = 1/(floor(rho*N))*SIGMA(q,l);
215         end
216     end
217
218
219     % if last 5 payoffs the same – 3 digits behind the comma, stop algorithm
220     if ( roundcount > 4 && ...
221         (floor(1000*gamma(roundcount))==floor(1000*gamma(roundcount-1))) &&...
222         (floor(1000*gamma(roundcount))==floor(1000*gamma(roundcount-2))) &&...
223         (floor(1000*gamma(roundcount))==floor(1000*gamma(roundcount-3))) &&...
224         (floor(1000*gamma(roundcount)) == floor(1000*gamma(roundcount-4))))
225         checkpayoff=1;
226     end
227
228     mu
229     SIGMA
230     end
231     % end main
232
233     levels_for_payoff = [];
234     levels_for_payoff = gamma'
235     roundcount
236     fprintf(fid, '\r\n mu=\r\n');
237     dlmwrite(testfname,mu,'-append','delimiter','\t','newline','pc',...
238         'precision',3, 'roffset',1 );
239     fprintf(fid, '\r\n SIGMA=\r\n');
240     dlmwrite(testfname,SIGMA,'-append','delimiter','\t','newline','pc',...
241         'precision',3, 'roffset',1 );
242     fprintf(fid, '\r\n levels_for_payoff=\r\n\r\n');
243     dlmwrite(testfname,levels_for_payoff,'-append','delimiter','\t',...
244         'newline','pc','precision',3, 'roffset',1 );
245     fprintf(fid, '\r\n roundcount = %d\r\n\r\n\r\n\r\n\r\n',roundcount);
246     fclose(fid);
247     % write the final price in extra file
248     pricelist=sprintf('...\prices.txt')
249     priceid=fopen(pricelist,'a');
250     fprintf(priceid, '...',...);
251     fclose(priceid);
252     toc;
253 end

```

254 **end**

A.3. Generation of a Geometric Brownian Motion

```

1 % geometric_brownian(N,r,alpha,T) simulates a geometric Brownian motion
2 % on [0,T] using N normally distributed steps and parameters r and alpha
3
4 function [X] = gbm(N,r,alpha,T,x)
5 % gbm(100,0.04,0.25,5,1)
6
7 % t is the column vector [0 1/N 2/N ... 1]
8 t = (0:1:N)'/N;
9 t = t*T;
10
11 % explanation:  $Z \sim N(0,1) \rightarrow X = \sqrt{h} * Z \sim N(0,h)$ , here  $h=T/n$ 
12 % S is running sum of  $N(0,1/N)$  variables
13 W = [0; cumsum(randn(N,1))]/sqrt(N);
14 W = W*sqrt(T);
15
16 Y = (r-(alpha^2)/2)*t + alpha * W;
17
18 X = x*exp(Y);

```

A.4. Generation of an Exponential Lévy Process with NIG Increments

```

1 % simulates normal inverse gaussian r.v., used as the distribution for the
2 % increments of the levy process
3 % S0 >0
4
5 function [S] = normalinversegauss(n,alpha,beta,mu,delta,T,start)
6 % normalinversegauss(100,75.49,-4.089,0,0.012,5,1)
7
8 % n=T*64-1;
9 % t is the column vector [0 1/N 2/N ... 1]
10 t = (0:1:n)'/n;
11 t = t*T;
12 xi=delta/sqrt(alpha^2 - beta^2);
13
14 L=[];
15 S=[];
16 U=[];
17 V=[];
18 W=[];
19 X=[];
20 Y=[];
21 Z=[];
22
23 for i=1:n,
24     V(i) = randn(1,1)^2;
25     W(i) = xi + xi^2*V(i) / (2*delta^2) - xi/(2*delta^2)*...
26         sqrt(4*xi*delta^2*V(i) + xi^2*V(i)^2);

```


A. Algorithms

```
27 U(i) = unifrnd(0,1);
28 if (U(i) < xi/(xi + W(i)))
29     Z(i) = W(i);
30 else
31     Z(i) = xi^2/W(i);
32 end
33 Y(i) = randn(1,1);
34 X(i) = mu + beta*Z(i) + sqrt(Z(i))*Y(i);
35 end
36
37 L = [0 cumsum(X)];
38 S = start*exp(L);
39 S = S';
```

Bibliography

- [1] Andersen, L. (2000) *A simple approach to the pricing of Bermudan swaptions in the multi-factor Libor Model*, Journal of Computational Finance 3 pp.5-32
- [2] Asmussen, S. & Glynn, P. W. (2007) *Stochastic Simulation. Algorithms and Analysis*, New York: Springer Science+Business Media
- [3] Benth, F. E., Groth, M. & Kettler, P.C. (2006) *A quasi-Monte Carlo algorithm for the normal inverse Gaussian distribution and valuation of financial derivatives*, International Journal of Theoretical and Applied Finance 9 (6), pp. 843-867
- [4] Björk, T. (2009) *Arbitrage theory in continuous time*, Oxford: Oxford University Press
- [5] Black, F. & Scholes, M. (1973) *The pricing of Options and Corporate Liabilities*, The Journal of Political Economy, Vol. 81, No. 3 pp.637-654
- [6] Bourbaki, N. (1961) *Éléments de mathématique 3 Topologie générale 1-4*, Paris VI: Hermann
- [7] Broadie, M. & Glasserman, P. (1997) *Pricing American-style securities by simulation*, Journal of Economic Dynamics and Control 21 pp.1323-1352
- [8] Broadie, M., Glasserman, P., & Jain, G. (1997) *Enhanced Monte Carlo estimates of American option prices*, Journal of Derivates 4:25-44
- [9] Broadie, M., Glasserman, P., & Ha, Z. (2000) *Pricing American options by simulation using a stochastic mesh with optimized weights*, Probabilistic Constrained Optimization: Methodology and Applications 32-50
- [10] Cerny, A. & Shiryaev, A. (2002) *Vector stochastic integrals and the fundamental theorems of asset pricing*, Proceedings of the Steklov Institute of mathematics, 237, pp.6-49
- [11] Cont, R. & Tankov, P. (2004) *Financial Modelling With Jump Processes*, Chapman & Hall/CRC
- [12] Cvitaní, J. & Zapatero, F. (2004) *Introduction to the Economics and Mathematics of Financial Markets*, The MIT Press
- [13] Delbaen, F. & Schachermayer, W. (1998) *The fundamental theorem of asset pricing for unbounded stochastic processes*, Mathematische Annalen, 312, pp.215-250

Bibliography

- [14] Egorova, V. & Ivanova, A. (2009) *High-order compact methods for the American Option pricing problem*, Halmstad University
- [15] Fishman, G. (1996) *Monte Carlo. Concepts, Algorithms and Applications*, Springer Verlag
- [16] Folland, G. (1976) *Introduction to partial differential equations*, Princeton, New Jersey: Princeton University Press
- [17] Fu, M.C. & Hu, J.-Q. (1995) *Sensitivity Analysis for Monte Carlo simulation of option pricing*, Probability in the Engineering and Information Sciences pp.417-446
- [18] Garcia, D. (2003) *Convergence and biases of Monte Carlo estimates of American option prices using a parametric exercise rule*, Journal of Economic Dynamics and Control 27 pp.1855-1879.
- [19] Glasserman, P. (2004) *Monte Carlo Methods in Financial Engineering*, New York: Springer Verlag
- [20] Hammersley, J.M. & Handscomb, D.C. (1964) *Monte Carlo Methods*, Methuen
- [21] Kabanov, Y. (2001) *Arbitrage theory*. In Jouini, E., Cvitanic, J. & Musiela, M. *Option pricing interest rates and risk management*, Handb. Math. Finance, Cambridge: Cambridge Univ. Press, pp.3-42
- [22] Jouini, E. & Pliska, S. (2001) *Option Pricing, Interest Rates and Risk Management*, Cambridge University Press: Cambridge
- [23] Lamberton, D. & Mikou, M. (2008) *The critical price for the American put in an exponential Lévy model*, Finance and Stochastics 12 (4), pp. 561-581
- [24] Levendorskii, S. (2004) *Early exercise boundary and option prices in Lévy driven models*, Quantitative Finance 4 (5), pp. 525-547
- [25] Mandelbrot, B. B. (1963) *The variation of certain speculative prices*, Journal of Business, XXXVI, pp. 392-417
- [26] Peskir, G. & Shiryaev, A. (2006) *Optimal Stopping and Free-Boundary Problems*, Basel: Birkhäuser Verlag
- [27] Protter, P. (2005) *Stochastic Integration and Differential Equations*, Berlin: Springer Verlag, pp.5
- [28] Ripley, B. (1987) *Stochastic Simulation*, New York: Wiley
- [29] Rubinstein, R.Y. (1981) *Simulation and the Monte Carlo Method*, Wiley
- [30] Rubinstein, R. Y. & Kroese, D. P. (2004) *The Cross - Entropy Method. A Unified Approach to Combinatorial Optimization, Monte-Carlo Simulation, and Machine Learning*, New York: Springer Science+Business Media

Bibliography

- [31] Ross, S. (1976) *Options and efficiency*, Quarterly Journal of Economics, 90, pp. 75-89
- [32] Rydberg, T. (1997) *The normal inverse Gaussian Lévy process: simulation and approximation*, Stochastic Models, 13:4, pp. 887-910
- [33] Sato, K. (1999) *Lévy processes and Infinitely Divisible Distributions*, Cambridge university Press: Cambridge, UK
- [34] Shiryaev, A. (1999) *Essentials of Stochastic Finance: Facts, Models, Theory*, Singapore: World Scientific Publishing Co. Ptl. Ltd.
- [35] Shreve, S. E. (2004) *Stochastic Calculus for Finance II Continuous-Time Models*, New York: Springer Verlag
- [36] Zhang, H. & Fu, M.C.(2006) *Applying model reference adaptive search to American-style option pricing*, Proceedings - Winter Simulation Conference , art. no. 4117674, pp. 711-718