

INAUGURAL - DISSERTATION

zur
Erlangung der Doktorwürde
der
Naturwissenschaftlichen-Mathematischen Gesamtfakultät
der
Ruprecht-Karls-Universität
Heidelberg

vorgelegt
von
Anamaria Bodea
aus Baia Mare, Rumänien

Tag der mündlichen Prüfung:

.....

Valuation of Swing Options in Electricity Commodity Markets

Gutachter:

Herr Prof. Dr. Dr. h.c. mult. Willi Jäger

.....

Abstract

Although electricity is considered to be a commodity, its price behavior is remarkably different from most other commodities or assets on the market. Since power can hardly be stored physically, the storage-based methodology, which is widely used for valuing commodity derivatives, is unsuitable for electricity. Therefore, new approaches are required to understand and reproduce its price dynamics. Concurrently, the demand for derivative instruments has grown and new types of contracts for energy markets have been introduced. Swing options, in particular, have attracted an increasing interest, offering more flexibility and reducing exposure to strong price fluctuations.

In this thesis, we propose a mean-reverting model with seasonality and double exponential jumps. It is able to accurately reproduce the behavior and main peculiarities of electricity's spot prices. With this model, we can characterize the swing option value as a solution to a partial integro-differential complementarity problem, which we solve numerically.

In the last part of the thesis, we present a more complex type of swing options, in which we also include variable electricity volumes in the contract. This formulation leads to a two-dimensional Hamilton-Jacobi-Bellman (HJB) equation. By applying the method of characteristics, this problem is simplified to a sequence of one dimensional HJB equations, which are solved numerically by using a similar approach as before.

Zusammenfassung

Strom wird zwar als ein Rohstoff angesehen, sein Preis-Verhalten unterscheidet sich aber erheblich von den meisten anderen Waren oder Vermögenswerten auf dem Markt. Da Strom physikalisch kaum gespeichert werden kann, ist die speicherbasierte Methodik, die weithin für die Bewertung von Finanzderivaten verwendet wird, für Strom ungeeignet. Deshalb sind neue Ansätze erforderlich, um die Preisdynamiken adäquat abzubilden und zu verstehen. Gleichzeitig ist die Nachfrage nach derivativen Instrumenten gestiegen, und neue Arten von Verträgen sind für Energiemärkte eingeführt worden. Insbesondere geraten Swing-Optionen immer mehr in den Fokus, da sie eine größere Flexibilität und geringere Anfälligkeit für starke Preisschwankungen bieten.

In dieser Arbeit schlagen wir einen Mittelwert-Modell mit Saisonalität und doppelt exponentiellen Sprüngen vor. Es kann das Verhalten und die wichtigsten Eigenschaften des Strompreises genau reproduzieren. Mit diesem Modell können wir den Wert der Swing-Option als Lösung eines Komplementaritätsproblems mit partieller Integro-Differenzialgleichung beschreiben, und das wird dann numerisch gelöst.

Im letzten Teil der Arbeit präsentieren wir einen komplexeren Typ von Swing-Optionen, der auch variable Strommengen als Vertragsbestandteil enthält. Diese Formulierung führt zu einer zweidimensionalen Hamilton-Jacobi-Bellman-Gleichung (HJB). Durch die Methode der Charakteristiken wird dieses Problem zu einer Abfolge von eindimensionalen HJB-Gleichungen vereinfacht, die numerisch durch einen ähnlichen Ansatz wie zuvor gelöst werden.

To my parents

Acknowledgment

The accomplishment of this thesis would not have been possible without the support of many people.

I wish to express my profound gratitude to my primary supervisor, Prof. Willi Jäger, for his confidence in my abilities, for the financial and personal support during the course of this work. I am very grateful to him for guiding me to the challenging area of energy markets, for trusting me and giving me the freedom and encouragement to realize my ideas. I would also like to thank my secondary advisor, Prof. Hans Georg Bock, for the helpful suggestions and for his enduring support within the Graduate School during the last years.

Special thanks to my mentor, Prof. Sebastian Sager, for his useful and constructive reviews, but also for his guidance and the helpful comments during our meetings in the last years. I am also indebted to my friend and mentor, Cristian Croitoru, for the patience to answering my questions from financial mathematics area, and for teaching me so many new things about computers and system administration.

I would also like to thank Dr. Elfriede Friedmann and Dr. Maria Neuss-Radu for their critical and useful suggestions, which improved considerably this work.

Many thanks to Nadi Serhan Aydin, for his remarks on the price modeling part and interesting discussions during his visit in Heidelberg. I am also very grateful to Dr. Cristoph Riesinger and Dr. Jari Toivanen, for helping me with the implementation of the numerical methods and for answering my questions fast and precisely.

I would like to acknowledge the German Research Council (DFG) for the financial support within the Heidelberg Graduate School of Mathematical and Computational Methods for the Sciences (HGS MathComp), and to the Interdisciplinary Center for Scientific Computing (IWR) of the University of Heidelberg. Many thanks to the Graduate School, for providing a stimulating environment for my research, for giving me the opportunity to work with many young talented researchers and to attend several conferences and lectures on various disciplines.

I would like to thank all my colleagues and the secretaries from the Applied Analysis and Simulation and Optimization groups, for a friendly working atmosphere and for helping me to improve my German. I am also grateful to all my friends, and especially all my Romanian friends, for making the time I spent in Heidelberg so enjoyable.

Finally, I wish to express my deepest gratitude to my family, none of my accomplishments would have been done without their support. My greatest appreciation is given to my boyfriend, Philipp, for his unconditioned love and endless patience, for sharing all the joy and sorrows with me in the last years. I am profoundly indebted to my parents and my sister, for their never-ending support, for understanding and encouraging me during the most difficult times. This thesis is dedicated to them.

I truly believe that all people whom I have not personally mentioned here are aware of my appreciation.

Contents

| | |
|--|-----------|
| Abstract | i |
| Zusammenfassung | i |
| Acknowledgment | ii |
| 1 Introduction | 1 |
| 2 Financial and mathematical backgrounds | 7 |
| 2.1 Introduction to financial derivatives | 7 |
| 2.2 Mathematical backgrounds | 10 |
| 2.2.1 Stochastic processes | 11 |
| 2.2.2 Itô's Lemma | 13 |
| 2.2.3 The Black-Scholes equation | 14 |
| 2.3 Optimal stopping in continuous time | 16 |
| 2.4 Introduction to jump-diffusion | 23 |
| 2.4.1 Itô's lemma for jump-diffusion processes | 26 |
| 2.4.2 Derivative pricing with jumps | 29 |
| 2.5 Viscosity solutions | 31 |
| 3 A new model for pricing swing options in the electricity market | 35 |
| 3.1 Introduction to electricity markets | 36 |
| 3.1.1 Characteristics of the electricity prices | 38 |
| 3.1.2 Literature review and modeling approaches | 41 |
| 3.2 The model for pricing swing options | 44 |
| 3.2.1 Risk neutral formulation | 45 |
| 3.2.2 Model calibration | 46 |
| 3.2.3 Simulation results | 51 |
| 4 Pricing swing options under the double exponential jump-diffusion model | 55 |
| 4.1 Introduction to swing options | 56 |
| 4.2 Numerical methods for pricing swing options | 57 |
| 4.3 Optimal multiple stopping and swing problem formulation | 59 |
| 4.4 Existence and uniqueness of the solution | 62 |
| 4.5 Numerical calculations | 75 |

| | | |
|----------|---|------------|
| 4.5.1 | The penalty method | 76 |
| 4.5.2 | The recursion formula for the integral term | 79 |
| 4.5.3 | Convergence analysis | 83 |
| 4.5.4 | Matrix iteration | 86 |
| 4.6 | Numerical results | 90 |
| 5 | Pricing swing options with power volume constrains | 95 |
| 5.1 | Introduction and swing problem formulation | 96 |
| 5.1.1 | Partial integro-differential complementarity formulation | 98 |
| 5.2 | Numerical calculations | 99 |
| 5.2.1 | Existence of solution for the penalized equation | 102 |
| 5.2.2 | Discretization | 103 |
| 5.2.3 | Convergence analysis | 106 |
| 6 | Conclusion | 111 |
| A | Solution to the SDE (3.6) | 113 |
| A.1 | Positive coefficient condition algorithm | 116 |
| B | The determination of the PIDCP with volume constraints | 117 |
| C | Abbreviations and notations | 119 |
| | Bibliography | 120 |

Chapter 1

Introduction

While most commodities may be properly modeled as traded securities, electricity presents new challenges to the derivative pricing discipline due to its non-storability. Thus, the characteristics of electrical generation and consumption are directly reflected in power spot prices: they indicate strong daily, weekly and annual seasonal patterns, and short-lived price deviations (jumps) with strong mean-reversion. The special features of the electricity prices brought on the need for new models and methodologies for pricing options.

There have been various mathematical models developed in the literature to simulate the electricity price dynamics. On the one hand, some of these processes are too complicated for pricing derivatives. The multitude of terms and additional parameters increase the complexity of the problem and drastically lengthen the computation times. On the other hand, many approaches in the literature of swing pricing use rather simple stochastic processes, like the Black-Scholes model, proposed by Black and Scholes [18]. This model was designed for the stock prices and it is not able to reproduce the mean-reversion, seasonality, price jumps or spikes, which are fundamental characteristics of electricity's behavior.

In this thesis, we model electricity prices as a sum of a deterministic seasonality function and a mean-reverting process with double exponential jumps. These types of jumps were introduced by Kou and Wang [66] for the modeling of interest rates. The double exponential distribution is able to capture the positive and negative jumps, which are also observed in electricity prices. The model we propose in this thesis includes two additional and important features of the power prices: mean-reversion and seasonality. Our simulation results show that the model can reproduce accurately the observed prices and that it can be used in valuating swing options.

Swing options are very complex derivatives, that have lately attracted an increasing interest. Developed for the energy markets, to reduce exposure to strong price fluctuations, swing options give its holder the right to buy electricity at different times during the contract period, providing flexibility in the quantity of energy to be purchased. The main complexity in pricing these options derives from the fact that the optimal decision of the option holder depends not only on the price, but also on the quantity of energy previously bought and the number of remaining

swing rights. Another challenge in pricing swing derivatives is that they have no explicit solution, so numerical methods are needed to approximate the solution. In the first part of the thesis, we study swing options that do not include electricity's volume flexibility. This type of options are more often analyzed in the literature. One of the earliest approaches to pricing such swing options was introduced by Thompson [98]. He proposed a lattice based algorithm to price take-or-pay contracts in a very simple framework. Jaillet, Ronnand and Tompaidis [60] extended his work by investigating a multi-level lattice method for swing options, using a mean-reverting process for the gas price. Kluge [65] studied the binomial forest approach, by pricing on a finite grid using a jump-diffusion model with seasonality. Due to the simplicity of the implementation, Monte Carlo methods also gained a lot of interest in the analysis of swing options, see Ibanez [54] for example. Dörr [38] and Cartea [25] used a least-squares Monte-Carlo algorithm to price swing derivatives. In Chapter 4 we take a closer look at these methods in the context of swing option pricing.

In this thesis, we compute swing option values by using the finite difference method. To our knowledge this method was first applied by Wegner [104] for valuating swing options, using a stochastic mean-reversion model with seasonality. Dahlgren and Korn [35] proposed later a more complicated type of swing option, considering different recovery times between exercise rights. They showed that the pricing problem can be transformed into a set of variational inequalities which can be discretized by using finite differences and solved by iterative methods. Kjaer [64] extended Dahlgren's approach by modeling electricity prices as a jump-diffusion process with seasonality. This formulation led to a partial integro-differential complementarity problem (PIDCP) that was solved numerically by using finite differences. The challenging part arises here due to the jumps in the price process, which introduce an integral term into the swing problem formulation. The discretization of this term leads to a dense matrix, which is computationally very expensive. Kjaer [64] applied the implicit-explicit Euler time discretization in which the diffusion part was treated implicitly, while the integral part was treated explicitly.

We present a different approach for the integral term: using the double exponential jumps in the model for the electricity prices, we are able to derive a recursion formula for this term. Toivanen [99] developed a similar formula for the Kou model, to study American option pricing. He showed that using this formula the integral term can be found with optimal computational cost and it can be used without any restrictions on non-uniform grids.

As the solution of the PIDCP is not smooth, we discuss the swing option valuation within the framework of viscosity solutions. Using the penalty method, we replace the PIDCP by a nonlinear equation which we discretize and then solve iteratively. The penalty method is able to handle nonlinearities or the early exercise¹ feature, and it has been successfully used in the context of American option pricing problems, see among others d'Halluin et. al. [48], Forsyth and Vetzal [42] or Toivanen [99]. This technique has never been used before in the context of swing option pricing. We prove that the finite difference scheme and the penalty iteration both

¹Buying or selling the asset prior to the option's expiration date.

converge to the unique viscosity solution of the PIDCP. We present the numerical results and examine the behavior of the swing option value in relation to changes in the model's parameters.

In the last part of the thesis, we investigate the structure and numerical methods for general swing contracts, which include electricity's volume flexibility. These options allow their holder to vary the quantity of energy they purchase at each exercise time. In this way they have more flexibility during the contract period, which is very convenient in the spiky electricity markets.

There are relatively few papers in the literature approaching this subject. Most of them use dynamic programming and tree-based methods to find the swing values. Lari-Lavassani et. al. [69] and later Jaillet et. al. [60] computed swing options with global constraints by using the binomial tree method in a mean-reverting model. Recently, Wahab et. al. [101], [102] developed a pentanomial lattice method to evaluate swing options in gas and electricity markets under a regime switching model for the spot price. Bardou et. al. [4] investigated a numerical integration method, the so-called optimal quantization, for pricing general swing options. Barrera-Esteve et. al. [10] considered gas swing options with changeable amounts and fixed exercise times. They worked under a forward market and extended the least-squares Monte-Carlo approach to swing options with variable volumes. A different approach was proposed by Keppo [63], who used a linear programming method to price swing contracts with local and global volume constraints, in a forward market.

We propose a new approach for the swing option problem with variable volume. We evaluate the price of the swing contract under the double exponential jump-diffusion model which we introduce in Chapter 3. Working on a spot, rather than a forward market, is a more realistic endeavor, since swing options are influenced by hourly price behavior.

We extend the numerical approach we have developed for swing options with re-fraction time, to the general swing contracts. Solving these options numerically is even more challenging due to the volume constraints in the problem's formulation. The introduction of these constraints adds a new state variable to the PIDCP. We show that by using the method of characteristics it is possible to approximate this PICDP by solving a series of one dimensional partial integro-differential equations. This idea comes from the optimal tree harvesting decision problem, and this approach has not been used previously in the pricing of swing options in electricity markets. Thus, we can apply the penalty method, discretize the resulting nonlinear problem by finite differences and solve it iteratively.

We also prove the existence of the viscosity solution to the penalized equation. Moreover, provided that a strong comparison result holds, we show that the finite difference scheme converges to the unique viscosity solution by verifying the stability, monotonicity and consistency of the scheme.

Contributions and results

The main objective of this thesis is to develop a robust numerical method for pricing swing options by using a model for electricity spot prices, that incorporates jumps, mean-reversion and seasonality.

The main results of this thesis can be summarized as follows:

- We propose a mean-reverting double exponential jump-diffusion model with seasonality for electricity spot prices, which is new in the context of swing option valuation. The seasonality function incorporates the trend observed in power prices, and the weekly, annual and semi-annual patterns. The calibration results show that the model is capable of fitting the market data accurately.
- The swing contracts studied in the first part of this thesis include a refraction period between two exercise rights. This means that the exercise dates are not fixed in the contract, which gives more flexibility to the option holder. This formulation is already a novelty in the literature of swing option pricing under a seasonal jump-diffusion model.
- By using the mean-reverting double exponential jump-diffusion model for the price, the swing pricing problem reduces to solving a partial integro-differential complementarity problem (PIDCP). As it is almost always impossible to find a classical solution to this problem, we prove all our results within the framework of viscosity solutions. Starting from the American options valuation theory, we are able to prove the existence and uniqueness of a viscosity solution for the PIDCP which arises in the valuation of swing options.
- We transform the PIDCP into a nonlinear integro-differential equation using the penalty method. The resulting penalized equation is discretized by finite differences and solved at each time step using the Newton iteration. This procedure is a new contribution to swing pricing literature. We determine the conditions under which the discretized problem is monotone, stable and consistent. According to Barles [5], these properties guarantee the convergence of the finite difference scheme to the unique viscosity solution of the pricing equation. We also show that the penalized iteration scheme is convergent. We implement this approach in MATLAB and compute the price of the swing option for different parameter values.
- Due to the jumps in the price model, we have to deal with a non-local integral term in the PIDCP formulation. We approximate this integral using a recursion formula, which can be efficiently used in the numerical approach described above.
- The second type of swing options we study in this thesis include energy volume constraints in their formulation. In this case, we have to deal with

a PIDCP with two state variables, the price and the quantity. Applying the method of characteristics, it is possible to approximate this PIDCP by solving a series of one dimensional partial integro-differential equations. In this way, we can extend the numerical approach we have developed for swing options with refraction times to these general swing contracts.

- We apply the penalty method and prove the existence of the viscosity solution in the penalized equation. Provided a strong comparison result holds, we show that the finite difference scheme converges to the unique viscosity solution by verifying the stability, monotonicity and consistency of the scheme.

Structure of the thesis

The rest of the thesis is organized as follows.

Chapter 2 gives an introduction to financial derivatives and option pricing. Then we present some mathematical notions needed during this thesis.

In Chapter 3 we introduce the mean-reverting double exponential jump diffusion model which we use for pricing swing options. We begin with a short introduction to electricity markets and present the characteristics of the prices. We show that the stochastic model we propose, is able to capture the main features observed in the electricity prices. We calibrate the model parameters using data from two European energy markets. The simulation results are presented in the end of the chapter and they show that the model can reproduce the observed prices accurately.

The valuation of swing contracts with a refraction period between two exercise rights is discussed in Chapter 4. These options give their holder more flexibility during the contract period, by not fixing the exercise rights beforehand. We show that the pricing problem under the double exponential jump diffusion model can be related to a partial integro-differential complementarity problem (PIDCP). We prove existence and uniqueness of the viscosity solution to this problem and then solve it numerically. We apply the penalty method, that transforms the PIDCP into a nonlinear integro-differential equation which is discretized by finite differences. The solution of this problem is then found iteratively at each timestep. We prove the convergence of the finite difference scheme and of the penalty iteration to the unique viscosity solution of the PIDCP. In the end we present the numerical results for this approach and test the behavior of the swing prices by varying the model parameters.

In Chapter 5 we focus on more complex swing options, which also include power volume constraints. Their formulation leads to a partial integro-differential complementarity problem with two state variables, the spot price and the quantity. We apply the method of characteristics and reduce this problem to a PIDCP with one state variable. In this way, we are able to apply the same numerical approach as in the previous chapter and solve the penalized problem iteratively.

In Chapter 6 the main results are recapitulated and conclusions are drawn.

Chapter 2

Financial and mathematical backgrounds

In this chapter we present the essential financial and mathematical concepts involved in the modeling and pricing of financial derivatives, in particular of swing options. The results are stated primarily without proofs, but the interested reader can find them in the references presented at the beginning of each section.

In this thesis we study the pricing of swing options in electricity markets. This topic leads to an optimal multiple stopping problem, that has a linear complementarity formulation. In this chapter we provide the main results from the optimal stopping theory and show its equivalence to an optimal stopping- linear complementarity problem. Our starting point are the American options, since they are a particular case of swing options.

In the second part of this chapter we present the mathematical formulation of the jump processes. The motivation for doing so comes from the model for electricity prices proposed in Chapter 3, which includes jumps in its formulation. We derive the main properties of the jump processes and present the methods for pricing and hedging in jump-diffusion markets. We then give the partial-integro complementarity formulation for American options. As the solution of this type of problems is generally known not to be smooth, the notion of viscosity solutions is required. Thus, we close this chapter by introducing some results from the viscosity solution theory.

2.1 Introduction to financial derivatives

In this section we will give a short introduction to financial markets and the most important types of derivative securities. For a more detailed description we refer to Hull [53], Neftci [81] and Seydel [94].

Financial transactions can occur in a huge variety, but they always need to take place in some financial markets. Financial markets allow people to trade securities (such as stocks and bonds) or commodities (such as precious metals or agricultural goods) at low transaction costs and at prices that reflect the efficient market hy-

pothesis.

The broadest way to characterize the financial markets is to distinguish between the exchanges (where the products are standardized) and over-the-counter (OTC) markets (where the investors can negotiate freely the terms of the contract).

The most important types of financial markets are categorized as follows.

- Capital markets, which consist of stock and bond markets
- Insurance markets, which facilitate the redistribution of various risks.
- Currency or foreign exchange markets, which facilitate the trading of foreign exchanges.
- Commodity markets, where physical assets as oil, gold or energy are traded.
- Derivatives markets, which provide instruments for the management of financial risk: options, forwards, futures and swaps.

We introduce next the basic concepts and definitions which we need in this thesis.

Definition 2.1.1. *A **derivative security**, or a **contingent claim**, is a financial instrument whose value is completely dependent on the price of the underlying asset in a fixed range of times within the interval $[0, T]$, where T is the expiry date.*

Remark 2.1.1. *The underlying asset refers to any market security such as stock, currency, bond or commodity.*

The major derivative securities in financial markets are options, forwards, futures contracts and swaps. We present here the main characteristics and definitions of these securities.

Definition 2.1.2. *An **option** is a financial tool that gives its holder the right, but not the obligation, to make a transaction at a given time for a given price. In particular, a **call option** allows its owner to buy and a **put option** to sell its underlying asset, at a certain time t , for a fixed strike price.*

There are two main groups of options on the market:

- Standard (vanilla) options, which are actively traded on an exchange, and whose value may be determined by looking up their price on the market (for example: European, American options);
- Exotic options, which are specially designed to fit the needs of the clients. There is no active market for them, so their value is computed by using a model to determine the premium (for example: lookback, basket, Asian or swing options). These types of derivatives are usually much more profitable than the vanilla products.

In this chapter we present properties and definitions of vanilla types of options. Exotic options, and in particular swing options, will be discussed in more detail in Chapter 4.

As stated in the definition, an option is an agreement between two parties: one party is the *writer*, often a bank, who fixes the terms of the option contract and sells the option; the other party is the *holder*, who purchases the option, paying the market price, which is the premium or option price (Seydel [94]).

We denote by V the price of the option, which depends on the underlying asset S and on the time t . T is known as the maturity or expiration of the option, and by K we denote the previously agreed price of the contract called strike or exercise price. It is important to note that the holder is not obligated to exercise (i.e. to buy or to sell the asset according to the terms of the contract).

In the case of a call option, at the maturity date, for spot prices below the strike price, the holder lets the option expire worthless, forfeits the premium, and buys the asset in the spot market. For asset prices greater than K , the holder exercises the option, buying the asset at K and has the ability to immediately make a profit equal to the difference between the two prices.

The payoff of a call option is plotted in Figure 2.1 and can be described mathematically as follows

$$V(S, t) = \max(S - K, 0) := (S - K)^+.$$



Figure 2.1: Payoff to Call (left) and Put Option (right).

In the case of a put option, the payoff function is plotted in Figure 2.1 and it can be written as

$$V(S, t) = \max(K - S, 0) := (K - S)^+.$$

It is not easy to compute the fair value of an option for $t < T$, as we will show later. But it is an easier task to determine the terminal value of V at expiration time T . This is the simplest option on the market, called European option.

Definition 2.1.3. *An option which can only be exercised at the maturity of the contract is called **European option**.*

A more complicated type of vanilla options is the American option.

Definition 2.1.4. *An **American option** can be exercised at any time prior to the maturity date $t \leq T$.*

The possibility of early exercise makes American options more valuable than similar European options. Later in this chapter we will give the mathematical formulation and main challenges in pricing American options.

The next types of derivative securities which we present in this section are the forwards and futures.

Definition 2.1.5. *A **forward contract**, or simply forward, is an agreement between two parties to buy or to sell an asset at a specified future point in time, for a certain price, which is agreed today.*

Definition 2.1.6. *A **futures contract**, or simply futures, is a standardized contract traded on exchange, to buy or sell asset at a certain date in the future, at a specified price.*

Although forwards and futures are similar, there are still some important differences between them: firstly futures are traded on exchange, whereas forwards contracts trade between individual institutions. Secondly, the cash flows of the two contracts occur at different times: forwards are settled once at maturity, while futures are daily marked with cash flows passing between the long and the short position to reflect the daily future price change (Clewlow et. al. [29]).

It should be emphasized that an option gives the holder the right, but not the obligation to take an action. This is what distinguishes options from forwards and futures, where the holder is obligated to buy or to sell the underlying asset.

We close this section with another important class of derivative securities, the swaps.

Definition 2.1.7. *A **swap** is the simultaneous selling and purchasing of cash flows involving various currencies, interest rates, and a number of other financial assets.*

The most usual swaps on the market are the currency swaps (exchange currencies) and interest rate swaps.

2.2 Mathematical backgrounds

In order to present the main results for stochastic processes, we recall at the beginning of this section some basic concepts and definitions, which can be found in several books, among others we mention Protter [89], Seydel [94] and Pham [85].

Definition 2.2.1. *A **filtration** on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ is an increasing family $\mathbb{F} = (\mathcal{F}_t)_{t \in \mathbb{T}}$ of σ -fields of $\mathcal{F} : \mathcal{F}_s \subset \mathcal{F}_t \subset \mathcal{F}$ for all $0 \leq s \leq t$.*

The time interval \mathbb{T} is either finite $\mathbb{T} = [0, T]$, $0 < T < \infty$, or infinite $\mathbb{T} = [0, \infty)$.

Remark 2.2.1. We say that a filtration $\mathbb{F} = (\mathcal{F}_t)_{t \in \mathbb{T}}$ satisfies the usual conditions if it is right continuous

$$\mathcal{F}_{t+} := \bigcap_{s \geq t} \mathcal{F}_s = \mathcal{F}_t, \quad \forall t \in \mathbb{T}$$

and if it is complete.

Definition 2.2.2. A process $(X_t)_{t \in \mathbb{T}}$ is **adapted** (with respect to \mathbb{F}) if for all $t \in \mathbb{T}$, X_t is \mathcal{F}_t -measurable.

\mathcal{F}_t can be interpreted as the available information up to time t , and if we want to characterize an event by its arrival time $\tau(\omega)$, occurred or not before time t , given the observation in \mathcal{F}_t , the notion of stopping time has to be introduced.

Definition 2.2.3. A random variable $\tau : \Omega \rightarrow (0, \infty]$ is a **stopping time** (with respect to the filtration \mathbb{F}) if for all $t \in \mathbb{T}$

$$\{\tau \leq t\} := \{\omega \in \Omega : \tau(\omega) \leq t\} \in \mathcal{F}_t.$$

We also note that if τ_1 and τ_2 are two stopping times, then $\tau_1 \wedge \tau_2$, $\tau_1 \vee \tau_2$ and $\tau_1 + \tau_2$ are stopping times.

Definition 2.2.4. An adapted process $(X_t)_{t \in \mathbb{T}}$ is a **supermartingale** if $\mathbb{E}[|X_t|] < \infty$ for all $t \in \mathbb{T}$, and

$$\mathbb{E}[X_t | \mathcal{F}_s] \leq X_s, \quad \text{a.s. for all } 0 \leq s \leq t, \quad s, t \in \mathbb{T}.$$

X is a **submartingale** if $-X$ is a supermartingale; X is a **martingale** if it is both a sub- and supermartingale, i.e. $\mathbb{E}[X_t | \mathcal{F}_s] = X_s$ a.s. .

2.2.1 Stochastic processes

Definition 2.2.5. A continuous **stochastic process** $S_t = S(\cdot, t)$, $t \geq 0$ is a family of random variables $S : \Omega \times [0, \infty) \rightarrow \mathbb{R}$ with $t \mapsto S(\omega, t)$ continuous for all $\omega \in \Omega$.

Definition 2.2.6. A **stochastic differential equation (SDE)** is an equation in which one or more of the terms is a stochastic process, thus resulting in a solution which is itself a stochastic process. A one factor SDE has the following form

$$dS_t = \mu(S_t, t)dt + \sigma(S_t, t)dW_t$$

or more precisely, written in integral form, we have

$$S_t = S_0 + \int_0^t \mu(S_u, u)du + \int_0^t \sigma(S_u, u)dW_u,$$

where μ is the drift of the stochastic process, σ is the volatility, and dW_t is a standard Wiener process.

One of the most important properties of stochastic differential equations is that their solutions exhibit the Markov property.

Definition 2.2.7. A continuous time stochastic process $S_t = S(t)$, ($t \geq 0$) is called a **Markov process** if it satisfies

$$\mathbb{P}(S(t_{n+1}) \in B | S(t_1) = s_1, \dots, S(t_n) = s_n) = \mathbb{P}(S(t_{n+1}) \in B | S(t_n) = s_n),$$

for all Borel subsets $B \in \mathbb{R}$, $0 < t_1 < \dots < t_n < t_{n+1}$, and all states $s_1, \dots, s_n \in \mathbb{R}$.

A stochastic process has the **Markov property** if at the present time, the future is independent of the past.

In other words, only the present value of a variable is relevant for predicting the future; the past history of the variable and the way that the present has developed from the past, are irrelevant.

These processes play an important role in the financial literature due to the fact that the stock prices are usually assumed to follow a Markov process.

We continue by introducing the Wiener process, which is a cornerstone of the financial modeling literature.

The Wiener process

Definition 2.2.8. A stochastic process $W = \{W_t, t \in \mathbb{R}_+\}$ is called a **Wiener process** or **Brownian motion** if the following conditions hold

- $W_0 = 0$.
- W_t has independent increments, i.e. if $r < s \leq t < u$ then $W_u - W_t$ and $W_s - W_r$ are independent stochastic variables.
- W_t is Normal distributed under \mathbb{P} with mean 0 and variance t .
- W_t is continuous in $t \geq 0$.

It follows that dW_t has the following properties, which are used extensively in the calculus of stochastic processes

- $E[dW_t] = 0$, $E[(dW_t)^2] = dt$.
- $Var[dW_t] = dt$.
- $(dW_t)^2 = dt$.
- $dW_t dt = 0$.
- $(dt)^2 = 0$.

The proof of these assumptions can be found for example in Levy [70].

2.2.2 Itô's Lemma

Consider a differentiable function of a stochastic variable $S = \{S_t, t \in \mathbb{R}_+\}$ that is driven by a Wiener process described by the equation

$$dS_t = \mu S_t dt + \sigma S_t dW_t, \quad (2.1)$$

where the drift μ and volatility σ are constant.

Let $f(S_t)$ be a smooth function of S_t . Thus if we vary S_t by an amount dS , then clearly f also varies by a small amount provided we are not close to singularities of f . From Taylor series expansion we have

$$df = f(S_t + dS_t) - f(S) = f'(S)dS_t + \frac{1}{2!}f''(S)(dS_t)^2 + \dots, \quad (2.2)$$

where the dots denote a reminder which is smaller than any of the terms we have retained. As the process S_t is described by the equation (2.1), it holds

$$(dS_t)^2 = \mu^2 S^2 dt^2 + \sigma^2 S^2 (dW_t)^2 + 2\mu\sigma S^2 dt dW_t = \sigma^2 S^2 (dW_t)^2 = \sigma^2 S^2 dt. \quad (2.3)$$

Next we substitute this result in (2.2) and retain only those terms which are at least as large as $\mathcal{O}(dt)$. Thus, using (2.1) we get

$$\begin{aligned} df &= \frac{dt}{dS_t}(\mu S_t dt + \sigma S_t dW_t) + \frac{1}{2}\sigma^2 S_t^2 \frac{d^2 f}{dS_t^2} dt = \\ &= \left(\mu S_t \frac{df}{dS_t} + \frac{1}{2}\sigma^2 S_t^2 \frac{d^2 f}{dS_t^2} \right) dt + \sigma S_t \frac{df}{dS_t} dW_t \end{aligned}$$

This result can be generalized by considering a function $f(S_t, t)$ which contains two variables S_t and t , of which S_t is a stochastic process

$$df(S_t, t) = \frac{\partial f}{\partial t} dt + \frac{\partial f}{\partial S_t} dS_t + \frac{1}{2} \left(\frac{\partial^2 f}{\partial S_t^2} dS_t^2 + \frac{2\partial^2 f}{\partial S_t \partial t} dS_t dt + \frac{\partial^2 f}{\partial t^2} \right).$$

We can set $(dt)^2 = 0$ and $dS_t dt = 0$. Therefore, the equation becomes

$$df(S_t, t) = \frac{\partial f}{\partial t} dt + \frac{\partial f}{\partial S_t} dS_t + \frac{1}{2} \frac{\partial^2 f}{\partial S_t^2} dS_t^2.$$

Now if we consider the process described by equation (2.1) and use (2.3) we arrive to *Itô's lemma*.

$$df(S_t, t) = \left(\frac{\partial f}{\partial t} + \mu S_t \frac{\partial f}{\partial S_t} + \frac{1}{2}\sigma^2 S_t^2 \frac{\partial^2 f}{\partial S_t^2} \right) dt + \sigma S_t \frac{\partial f}{\partial S_t} dW_t. \quad (2.4)$$

Two-dimensional Itô's formula

In the previous section we have established Itô's formula in the univariate case. In some cases the function $f(\cdot)$ may depend on more than one stochastic variable. Then a multivariate version of Itô's lemma has to be used.

As we only deal with bivariate examples in this thesis, we present the lemma for this framework, but the formula can be easily extended to higher-order systems. Suppose we have two stochastic processes S_t and X_t , and a continuous, twice-differentiable function, which we denote by $f(S_t, X_t, t)$. The the bivariate Itô's Lemma states

$$\begin{aligned} df(S_t, X_t, t) = & \frac{\partial f}{\partial t} dt + \frac{\partial f}{\partial S_t} dS_t + \frac{\partial f}{\partial X_t} dX_t + \\ & + \frac{1}{2} \frac{\partial^2 f}{\partial S_t^2} dS_t dS_t + \frac{1}{2} \frac{\partial^2 f}{\partial X_t^2} dX_t dX_t + \frac{\partial^2 f}{\partial S_t \partial X_t} dS_t dX_t. \end{aligned} \tag{2.5}$$

A proof for this formula can be found in Neftci [81].

Next we present one of the most important application of Itô's Lemma (2.4) in the financial literature, the Black-Scholes equation.

2.2.3 The Black-Scholes equation

The Black-Scholes formula transforms the pricing problem into a partial differential equation (PDE) with a final boundary condition.

Fisher Black, Myron Scholes and Robert Merton developed this revolutionary formula, and because of its theoretical and practical applications in the financial theory, they were rewarded with the Nobel Prize in Economics in 1997.

We now derive the fundamental partial differential equation (PDE) for pricing European options. Following the approach proposed by Black and Scholes, the following assumptions hold:

- The asset pays no dividends.
- There are no-arbitrage¹ possibilities (the market is arbitrage free).
- Options can be exercised only at maturity.
- Trading takes place continuously in time.
- The riskless interest rate and volatility are constant over time.
- There are no transactions costs and no taxes.

¹There are no possibilities that one can make a win without taking any risk.

We assume that the price process $S_t = S$ evolves according to a geometric Brownian motion

$$dS_t = \mu S_t dt + \sigma S_t dW_t, \quad (2.6)$$

where the drift μ and volatility $\sigma > 0$ are constant and W_t is a Wiener process. Now we consider a European call option, characterized by its payoff $(K - S_T)^+$ at maturity T and strike price K . Let $V = V(S, t)$ be the value of the call option, which is a function of the spot price at time t . By Itô's lemma (2.4) we get

$$\begin{aligned} dV &= \frac{\partial V}{\partial t} dt + \frac{\partial V}{\partial S} dS + \frac{1}{2} \sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} dt = \\ &= \left(\frac{\partial V}{\partial t} + \mu S \frac{\partial V}{\partial S} + \frac{1}{2} \sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} \right) dt + \sigma S \frac{\partial V}{\partial S} dW_t. \end{aligned}$$

At the expiration date of the option, we have $V(S_T, T) = (K - S_T)^+$, and for $t < T$, we consider a portfolio Π consisting of one unit of the derivative plus a number Δ_t of underlying stock sold. The portfolio value is then equal to

$$\Pi_t = V(S, t) - \Delta_t S \quad (2.7)$$

and its change in dynamics is given by

$$d\Pi_t = \left(\frac{\partial V}{\partial t} + \mu S \frac{\partial V}{\partial S} + \frac{1}{2} \sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} - \Delta_t \mu S \right) dt + \sigma S \left(\frac{\partial V}{\partial S} - \Delta_t \right) dW_t.$$

The random component in the evolution of the portfolio Π may be eliminated by choosing

$$\Delta_t = \frac{\partial V}{\partial S}.$$

This results in a portfolio with deterministic increment

$$d\Pi_t = \left(\frac{\partial V}{\partial t} + \frac{1}{2} \sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} \right) dt.$$

Now using the arbitrage free arguments, the rate of return of the riskless portfolio Π must be equal to the interest rate r

$$\frac{\partial V}{\partial t} + \frac{1}{2} \sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} = r\Pi.$$

Substituting, we have

$$\frac{\partial V}{\partial t} + rS \frac{\partial V}{\partial S} + \frac{1}{2} \sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} = rV. \quad (2.8)$$

This PDE is called the *Black-Scholes equation*, and together with the terminal condition $V(S, T) = (K - S_T)^+$ it is a linear parabolic Cauchy problem, whose solution is analytically known. Moreover, this formula can be computed as an expectation

$$V(S, t) = \mathbb{E}^{\mathbb{Q}} [e^{-r(T-t)} (K - S_T)^+ | S_t = S], \quad (2.9)$$

where $\mathbb{E}^{\mathbb{Q}}$ is the expectation under the risk neutral probability \mathbb{Q} .

Remark 2.2.2. In the mathematical literature, $rS\frac{\partial V}{\partial S}$ is the so-called convection term, $\frac{1}{2}\sigma^2S^2\frac{\partial^2 V}{\partial S^2}$ is the diffusion term and $-rV$ is the reaction term. The equation (2.8) is then a convection-diffusion PDE.

In the financial literature, $\frac{\partial V}{\partial S}$ denotes the option delta, $\frac{\partial^2 V}{\partial S^2}$ is the option gamma, and $\frac{\partial V}{\partial t}$ is the option theta. They are called in general the greeks and give information about different dimensions of risk in the option.

The Black-Scholes equation has an analytical solution only in a few cases, like for example a European call/put option. For most of the applications, this partial differential equation has no such solution. In these cases, the solution must be calculated numerically.

In the next section we take a closer look on the problem of American option valuation, for which the Black-Scholes equation does not have an analytical solution. We show that this problem can be related first to an optimal stopping- and then to a linear complementarity problem.

For the sake of completeness, we start with some basic properties and definitions from the optimal stopping theory.

2.3 Optimal stopping in continuous time

We start this section by summarizing the most important results and notions from the theory of optimal stopping in continuous time with a finite horizon.

The main references for this section are the work of Karatzas and Shreve [62], Musiela and Rutkowski [76], Lamberton [68] and Shiriyayev [96].

We consider as before a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ equipped with a filtration $\mathbb{F} = (\mathcal{F}_t)_{t \geq 0}$ which satisfies the usual conditions (see Remark 2.2.1).

We denote by \mathcal{T} the set of all stopping times with respect to the filtration \mathbb{F} and introduce the following subsets of \mathcal{T}

$$\mathcal{T}_{t,T} = \{\tau \in \mathcal{T} | \mathbb{P}(\tau \in [t, T]) = 1\}, \quad 0 \leq t \leq T < \infty \quad (2.10)$$

$$\mathcal{T}_{t,\infty} = \{\tau \in \mathcal{T} | \mathbb{P}(\tau \in [t, +\infty)) = 1\}, \quad t \geq 0. \quad (2.11)$$

We consider throughout this chapter that the process $Y = (Y_t)_{t \geq 0}$ is adapted, right-continuous and it satisfies

$$Y_t \geq 0 \quad \text{and} \quad \mathbb{E} \left(\sup_{t \geq 0} Y_t \right) < \infty, \quad \forall t \geq 0. \quad (2.12)$$

We introduce the following optimal stopping problem:

Problem (P) for $t \in [0, T]$ we want to find the value of

$$Z_0 = \sup_{\tau \in \mathcal{T}_{0,T}} \mathbb{E}[Y_\tau] \quad (2.13)$$

and an optimal stopping time $\tau_t^* \in \mathcal{T}_{0,T}$ for which the supremum in (2.13) is attained, if such an stopping time exists.

For each stopping time $\theta \in \mathcal{T}_{0,T}$ we introduce the random variable

$$Z_\theta = \operatorname{ess\,sup}_{\tau \in \mathcal{T}_\theta} \mathbb{E}[Y_\tau | \mathcal{F}_\theta]$$

where $\mathcal{T}_\theta = \{\tau \in \mathcal{T}_{0,T} : \tau \geq \theta\}$.

We introduce next two useful notions for this section:

Definition 2.3.1. *The process $(Y_t)_{t \geq 0}$ is called*

- **regular** if for every $\tau \in \mathcal{T}_{0,\infty}$, Y_τ is integrable and for every nondecreasing sequence $(\tau_n)_{n \in \mathbb{N}}$ of stopping times with $\tau = \lim_{n \rightarrow \infty} \tau_n$ we have

$$\lim_{n \rightarrow \infty} \mathbb{E}(Y_{\tau_n}) = \mathbb{E}(Y_\tau).$$

- **of class D** if the family $(Y_\tau)_{\tau \in \mathcal{T}_{0,\infty}}$ is uniformly integrable.

The main tool for solving optimal stopping Problem (P) is the concept of Snell envelope of the process Y .

Definition 2.3.2. *For $t \geq 0$ let $U_t = \operatorname{ess\,sup}_{\tau \in \mathcal{T}_{t,\infty}} \mathbb{E}[Y_\tau | \mathcal{F}_t]$ with the following properties:*

- The process $(U_t)_{t \geq 0}$ is a supermartingale.
- $\mathbb{E}[U_\tau] = Z_\tau$ a.s. for every $\tau \in \mathcal{T}_{0,T}$.
- U admits a right-continuous modification.

Then the right continuous modification of U is called **Snell envelope** of Y .

Remark 2.3.1. *If the essential supremum (ess sup) is taken over a countable number of random variables or over expectations, it coincides (a.s.) with the supremum. Therefore $U_0 = \sup_{\tau \in \mathcal{T}_{0,T}} \mathbb{E}[Y_\tau]$. Also by definition we have $U_T = Y_T$.*

Proposition 2.3.1. *The Snell envelope U is the smallest right-continuous supermartingale majorant of Y and for any $t \in [0, T]$, $\tau \in \mathcal{T}_{t,T}$ we have*

$$\mathbb{E}[U_\tau | \mathcal{F}_t] \leq U_t.$$

Proof: see Lamberton [68].

□

Next we present a characterization of optimal stopping times and an important result in the literature on stochastic process theory, the Doob-Meyer decomposition theorem.

Theorem 2.3.1. Characterization of optimal stopping times

A stopping time $\tau^ \in \mathcal{T}_{0,T}$ is optimal if and only if*

- $U_{\tau^*} = Y_{\tau^*}$ a.s. .
- The stopped supermartingale $\{U_{\tau^* \wedge t}, 0 \leq t \leq T\}$ is a martingale.

Proof: see Karatzas and Shreve [62].

□

Theorem 2.3.2. *If $Y = (Y_t)_{t \geq 0}$ satisfies (2.12) and Theorem 2.3.1, then*

$$\tau^* = \inf\{t \in [0, T] | Y_t = U_t\}.$$

Proof: see Lamberton [68].

□

Theorem 2.3.3. The Doob-Meyer decomposition

Let $U = (U_t)_{t \geq 0}$ be a right-continuous supermartingale of class D. There exists a martingale $(M_t)_{t \geq 0}$ and a nondecreasing right-continuous predictable process $A = (A_t)_{t \geq 0}$ with $A_0 = 0$, which are unique up to indistinguishability, and uniformly integrable, such that

$$U_t = M_t - A_t, \quad t \geq 0.$$

Moreover, if U is a regular process, the process A has continuous paths with probability one.

Proof: see Karatzas and Shreve [62].

□

The following theorem is a characterization of the Snell envelope in terms of its Doob-Meyer decomposition, and it is the result which establishes the relation between optimal stopping problems and variational inequalities, as we show later in this chapter.

Theorem 2.3.4. *Assume that the process $(Y_t)_{0 \leq t \leq T}$ is regular. Let $\hat{U} = (\hat{U}_t)_{0 \leq t \leq T}$ be a regular, right-continuous supermartingale of class D, with Doob-Meyer decomposition $\hat{U} = \hat{M} - \hat{A}$. \hat{U} is the Snell envelope of Y if and only if the following conditions are satisfied*

- $\hat{U} \geq Y$.
- $\hat{U}_T = Y_T$ a.s. .
- For every $t \in [0, T]$, $\hat{A}_t = \hat{A}_{\hat{\tau}_t}$ where $\hat{\tau}_t = \inf\{s \geq t | \hat{U}_s = Y_s\}$.

Proof: see Lamberton [68].

□

Next we turn our attention to American options, and show that they can be formulated as an optimal stopping and a free boundary value problem. One of the earliest works to examine the relationship between the early exercise feature and optimal stopping problems was the paper by McKean [77]. For a detailed presentation of this issue we also refer to Musiela and Rutkowski [76].

American options: from optimal stopping to free boundary value problems

We will only consider put options from now on during this thesis, but in the case of call options everything works in a similar way.

The price of an American put option at time $t \in [0, T)$ for a spot price $S_t \in \mathbb{R}_+$ is given by

$$V(S, t) = \sup_{\tau \in \mathcal{T}_{t,T}} \mathbb{E}^{\mathbb{Q}} [e^{-r(\tau-t)}(K - S_{\tau})^+ | S_t = S], \quad (2.14)$$

with $V(S, T) = (K - S_T)^+$.

As before, we denote by $\mathcal{T}_{t,T}$ the set of all stopping times in $[t, T]$ and we denote for the rest of the chapter by $\phi(S) = (K - S_t)^+$ the profit (or payoff) attached to exercising the option at time t .

During the lifetime of an American put option the investor chooses a time τ when to exercise. Since at time t he can only use the information \mathcal{F}_t , τ should be modeled as stopping time. Consequently, we have an optimal stopping problem and the buyer of the American option has to choose a stopping time in $\mathcal{T}_{t,T}$.

From the general results for optimal stopping we have $V(S, t) \geq \phi(S)$ and the optimal stopping time for problem (2.14) is

$$\tau_t^* = \inf\{l \in [t, T] \mid V(S_l^{S,t}, l) = \phi(S_l^{S,t})\}. \quad (2.15)$$

During this thesis we use the notation $S_l^{S,t}$ for S_l whenever we need to emphasize the dependence of the process S on its initial conditions.

Thus, the solution of the American option pricing is then implicitly determined by (2.14) and (2.15).

Remark that $V(0, t) = K$ and there exists a critical price $S_{fb}(t)$, below which the American put option should be exercised early

- If $S > S_{fb}$ then $V(S, t) > \phi(S)$.
- If $0 \leq S \leq S_{fb}$ then $V(S, t) = \phi(S)$.

Thus the domain $\mathbb{R}_+ \times [0, T)$ is divided by the optimal-stopping boundary $\{(S_{fb}(t), t), t \in (0, T]\}$ into the continuation region and the stopping region

$$\mathcal{C} = \{(x, t) \in \mathbb{R}_+ \times [0, T) \mid S > S_{fb}\}, \quad (2.16)$$

$$\mathcal{S} = \{(x, t) \in \mathbb{R}_+ \times [0, T) \mid S \leq S_{fb}\}. \quad (2.17)$$

In other words, S_{fb} marks the boundary between two regions: in one side one should hold the option and in the other side one should exercise it. S_{fb} is called *free boundary* and the problem of finding the value $V(S, t)$ for $S > S_{fb}$ is called **free boundary-value problem**.

The free boundary must be determined in addition to the option price, from the following conditions:

- The value function V is smooth in the continuation region \mathcal{C} and S_{fb} is a nondecreasing function on $[0, T)$ with

$$\lim_{t \rightarrow T} S_{fb}(t) = K,$$

- Furthermore, we have

$$\begin{aligned} \lim_{S \rightarrow S_{fb}} V(S, t) &= K - S_{fb}, \quad \forall t \in [0, T) & (2.18) \\ \lim_{S \rightarrow \infty} V(S, t) &= 0, \quad \forall t \in [0, T), \\ \lim_{S \rightarrow 0} V(S, t) &= \phi(S), \quad \forall t \in [0, T). \end{aligned}$$

- In order to ensure that the process $(e^{-r(T-t)}V(S_t, t))_{0 \leq t \leq T}$ is the Snell envelope of the discounted payoff process $(e^{-r(T-t)}\phi(S))_{0 \leq t \leq T}$, we impose

$$\frac{\partial V}{\partial t} + rS \frac{\partial V}{\partial S} + \frac{1}{2} \sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} - rV \leq 0,$$

so that we have a supermartingale. Taking in account the conditions of the Theorem 2.3.4 we have

$$\begin{aligned} V(S, t) &\geq (K - S)^+ \\ V(S_T, T) &= \phi(S_T) \end{aligned}$$

and in the continuation region \mathcal{C}

$$\frac{\partial V}{\partial t} + rS \frac{\partial V}{\partial S} + \frac{1}{2} \sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} - rV = 0.$$

- The next condition, known in the theory of optimal stopping as the *principle of smooth fit*, is referring to the differentiability of the option price across the boundary S_{fb}

$$\frac{\partial V}{\partial S}(S_{fb}, t) = -1.$$

Summarizing, we obtain the result characterizing the American put options as the solution of a free boundary value problem.

Theorem 2.3.5. *Let $V : \mathbb{R}_+ \times [0, T] \rightarrow \mathbb{R}$ be a nonincreasing and convex function and $S_{fb} : [0, T) \rightarrow \mathbb{R}$, $0 \leq S_{fb} \leq K$. Then the value function V of the American*

put option is the unique solution of the following free boundary value problem

$$\begin{aligned}
V(S, t) &= K - S_t, & \text{for } S < S_{fb} \\
\frac{\partial V}{\partial t} + rS \frac{\partial V}{\partial S} + \frac{1}{2} \sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} - rV &< 0, & \text{for } S < S_{fb} \\
V(S, t) &> (K - S_t)^+, & \text{for } S > S_{fb} \\
\frac{\partial V}{\partial t} + rS \frac{\partial V}{\partial S} + \frac{1}{2} \sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} - rV &= 0, & \text{for } S > S_{fb} \\
\lim_{S \rightarrow S_{fb}} V(S, t) &= K - S_{fb}, & \forall t \in [0, T) \\
\lim_{S \rightarrow S_{fb}} \frac{\partial V}{\partial S}(S_{fb}, t) &= -1, & \forall t \in [0, T) \\
\lim_{S \rightarrow \infty} V(S, t) &= 0, & \forall t \in [0, T) \\
\lim_{S \rightarrow 0} V(S, t) &= \phi(S), & \forall t \in [0, T) \\
V(S, T) &= (K - S_T)^+, & \forall S \in \mathbb{R}_+.
\end{aligned}$$

Proof: see McKean [77] and Jacka [58].

□

It is almost always impossible to find an explicit solution to a free boundary problem, so the way of solving is to construct robust numerical methods for their computation. The difficulty of dealing with free boundaries give rise to a reformulation of the problem in order to eliminate the explicit dependence on this boundary. A simple example of such a reformulation is the so-called **linear complementarity problem**, in the context of the obstacle problem. These problems have a linear complementation formulation which lead to efficient and accurate numerical solution schemes.

We introduce next the obstacle problem and its linear complementarity formulation. A detailed presentation of the problem can be found for in Wilmott [107] or in Seydel [94].

The obstacle problem and the linear complementarity formulation

An obstacle problem arises when an elastic string is held fixed at two ends, x_0 and x_1 , and passes over a smooth object, as showed in the next plot.

Let a function $g \in C^2(\mathbb{R})$ represent the obstacle, with $x \in \mathbb{R}$ and $g''(x) < 0$. Let $u \in C^1[x_0, x_1]$ be the stretched function and for simplicity let $u(x_0) = u(x_1) = 0$. Moreover, a and b are initially unknown and on the interval $[a, b]$, g and u coincide, and everywhere else $u > g$. Thus, we can write the obstacle formulation as a free boundary-value problem

$$\begin{aligned}
x_0 < x < a &\rightarrow u > g & \text{and} & u'' = 0 \\
a < x < b &\rightarrow u = g & \text{and} & u'' = g'' \\
b < x < x_1 &\rightarrow u > g & \text{and} & u'' = 0.
\end{aligned}$$

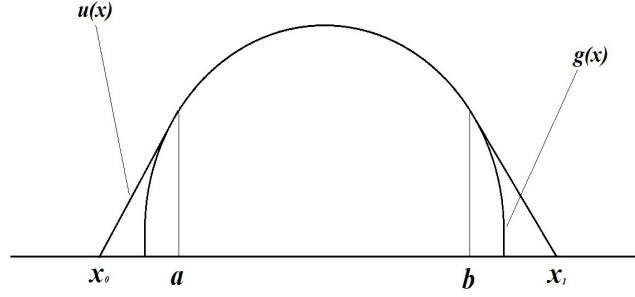


Figure 2.2: The obstacle problem.

This situation manifests complementarity in the following sense

$$\begin{aligned} \text{if } u > g & \quad \text{then} \quad u'' = 0 \\ \text{if } u = g & \quad \text{then} \quad u'' < 0. \end{aligned}$$

Hence the obstacle problem can be reformulated (equivalently) as a linear complementarity problem (LCP)

$$\begin{cases} u''(u - g) = 0 \\ u'' \geq 0 \\ u - g \geq 0 \\ u(x_0) = u(x_1) = 0. \end{cases} \quad (2.19)$$

Now we can state the linear complementarity formulation for an American put option, with $(S, t) \in \mathbb{R}_+ \times [0, T]$

$$\begin{cases} \frac{\partial V}{\partial t} + rS \frac{\partial V}{\partial S} + \frac{1}{2} \sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} - rV \leq 0 \\ V \geq \phi \\ (V - \phi) \left(\frac{\partial V}{\partial t} + rS \frac{\partial V}{\partial S} + \frac{1}{2} \sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} - rV \right) = 0 \\ V(T) = \phi(T) \end{cases}$$

where $\phi(S, t) = (K - S)^+$.

A proof of existence and uniqueness of the solution for this system is based on the work of Bensoussan and Lions [12], who proved regularity of the solutions to variational inequalities. Jaillet [59] showed later existence and uniqueness of the solution to the system of variational inequalities for American options. In Chapter 4 we give an overview of the numerical methods for linear complementarity problems.

As we have stated in the introduction, electricity prices present sometimes huge upward movements, followed by drops of almost the same amplitude. This behavior

is modeled in the financial literature as a jump process, and it will be introduced in the stochastic model for the electricity price.

In the next section we discuss general types of jump-diffusion models, derive the main properties and present the methods for pricing and hedging in these markets.

2.4 Introduction to jump-diffusion

Jump-diffusion models are a particular class of Lévy processes. We do not present here the general theory for Lévy processes, we rather concentrate on the specific aspects of jump-diffusion.

The main definitions and results presented in this section can be found in Runggaldier [93], Cont and Tankov [30] and Bermaud [15].

The following processes are defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with a filtration \mathcal{F}_t , to which the processes are adapted.

Definition 2.4.1. A *point process* describes events that occur randomly over time and it can be represented as a sequence of nonnegative random variables

$$0 = U_0 < U_1 < U_2 < \dots$$

where the generic U_n is the n -th instant of occurrence of an event.

This process may equivalently be represented via its associated counting process N_t , which counts the number of events up to and including time t , with $N_t < \infty$

$$N_t = n \quad \text{if } t \in [U_n, U_{n+1}), \quad n \geq 0 \quad \text{or} \quad N_t = \sum_{n \geq 1} 1_{\{U_n \leq t\}}.$$

Definition 2.4.2. A point process N_t is called **Poisson process** if

- $N_0 = 0$.
- N_t is a process with independent increments.
- $N_t - N_s$ is a Poisson random variable with a given parameter λ , called *intensity*.

A Poisson process has the following important properties:

- The probability of a jump in an interval Δ is $\lambda\Delta + \mathcal{O}(\Delta)$.
- The probability of two or more jumps in an interval of length Δ is $\mathcal{O}(\Delta)$.
- The Poisson distribution is given by $p(x, \lambda) = \frac{e^{-\lambda} \lambda^x}{x!}$, where $x = 0, 1, 2, \dots$ represents the number of occurrences of an event.

If we take a look at the above characterization of the Poisson process and at the definition of a Wiener process (Definition 2.2.8) we can state the following remarks.

- Both are processes with independent increments.
- The increments of a Wiener process are normally distributed, while those of a Poisson process are Poisson distributed.
- The Wiener process is the basic building block for processes with continuous trajectories, while the Poisson process is the basic block for processes with jumping trajectories.
- The Wiener process is itself a martingale, while a Poisson process is not.

Nevertheless, we show in the next theorem that the Poisson process can become a martingale, by subtracting a proper "mean".

Theorem 2.4.1. Martingale Property

Let N_t be a Poisson process with intensity λ . We define the **compensated Poisson process**

$$M_t := N_t - \lambda t. \quad (2.20)$$

Then M_t is a martingale.

Proof: let $0 \leq s < t$ be given. Because $N_t - N_s$ is independent of \mathcal{F}_s and has expected value $\lambda(t - s)$, we have

$$\begin{aligned} E[M_t | \mathcal{F}_s] &= E[M_t - M_s | \mathcal{F}_s] + E[M_s | \mathcal{F}_s] = \\ &= E[N_t - N_s - \lambda(t - s) | \mathcal{F}_s] + M_s = \\ &= E[N_t - N_s] - \lambda(t - s) + M_s = M_s. \end{aligned}$$

□

It is interesting to note also that

$$E[M_t] = 0 \quad \text{and} \quad E[M_t]^2 = \lambda t.$$

Next we define the notion of a marked point process. Let us first consider the representation (U_n, Y_n) , where we may interpret U_n as the n -th occurrence of some phenomenon and Y_n as an attribute or mark of this phenomenon.

Definition 2.4.3. A **marked point process** is a double sequence $(U_n, Y_n)_{n \geq 1}$ where

- U_n is a point process.
- Y_n is a sequence of \mathbb{R} -valued random variables.

We can now define an integer-valued random measure $m(dt, dy)$ associated to the marked point process $(U_n, Y_n)_{n \geq 1}$, which can be identified by the formula

$$m((0, t], A) = N_t(A) = \sum_{n \geq 1} 1_{\{U_n \leq t\}} 1_{\{Y_n \in A\}}, \quad (2.21)$$

where $A \in \mathcal{B}$, and \mathcal{B} is the Borel σ -field on U . We assume that there is a finite number of jumps during any finite time interval and that m is a compound Poisson process with intensity λ and $\nu(dy)$ is the probability measure under P of the i.i.d random variables U_1, U_2, \dots, U_n .

We define the compensated jump martingale of m by

$$\tilde{m}(dt, dy) = m(dt, dy) - \lambda\nu(dy) \quad (2.22)$$

and we have that $(\tilde{m}([0, t] \times A))_{0 \leq t \leq T}$ is a martingale.

We consider the following *general jump-diffusion process*

$$dS_t = \mu S_t dt + \sigma S_t dW_t + \int_{\mathbb{R}} \gamma(t, y) \tilde{m}(dt, dy), \quad (2.23)$$

where μ is the drift of the asset under \mathbb{P} and σ the constant volatility; (U_n) are the random time points of the jumps, and a jump at time U_n with amplitude Y_n corresponds to a proportional jump in the asset price of size $\gamma(Y_n)$, with Y_n distributed according to the probability law $\nu(dy)$. We assume that $\gamma \geq -1$ for the price to be real valued, and

$$\int_{\mathbb{R}} \gamma^2(y) \nu(dy) < \infty.$$

Remark that the last term in (2.23) can be also written as

$$\int_{\mathbb{R}} \gamma(t, y) m(dt, dy) = \gamma(t, Y_t) dN_t. \quad (2.24)$$

Thus, we obtain another way of representing a jump-diffusion process

$$dS_t = \mu S_t dt + \sigma S_t dW_t + \gamma(t, Y_t) dN_t \quad (2.25)$$

and for the particular case when $\gamma(t, Y_t) = Y_t$ we obtain

$$dS_t = \mu S_t dt + \sigma S_t dW_t + Y_t dN_t. \quad (2.26)$$

Remark 2.4.1. *In the financial literature one can find the models (2.23) or (2.25) written in the form*

$$dS_t = \mu S_t dt + \sigma S_t dW_t + dJ_t, \quad (2.27)$$

where $J_t = \sum_{n=1}^{N_t} Y_n$, or in the more general case $J_t = \sum_{n=1}^{N_t} \gamma(U_n, Y_n)$.

Furthermore, in models of the form (2.26) one may find the last term written as $(Y_t - 1)dN_t$, when at time $t = T_n$ the jump is given by $\Delta S_n = S_{T_n} - S_{T_n^-} = S_{T_n^-} Y_n = S_{T_n^-} Y_{T_n}$, so that $S_{T_n} = S_{T_n^-} (1 + Y_n)$.

We now introduce Itô's lemma for stochastic processes which include jumps.

2.4.1 Itô's lemma for jump-diffusion processes

Consider a general jump-diffusion process given by the following equation

$$dS_t = \mu dt + \sigma dW_t + J_t dN_t, \quad (2.28)$$

where $(W_t)_{t \geq 0}$ and $(N_t)_{t \geq 0}$ are Wiener and respectively Poisson process, mutually independent. Then the following lemma holds

Lemma 2.4.1. *Let $f(S_t, t)$ be a C^2 function of S_t and t . Then*

$$df(S_t, t) = \left(\frac{\partial f}{\partial t} + \mu \frac{\partial f}{\partial S_t} dt + \frac{1}{2} \sigma^2 \frac{\partial^2 f}{\partial S_t^2} \right) dt + \sigma \frac{\partial f}{\partial S_t} dW_t + [f(S_t + J_t) - f(S_t)] dN_t, \quad (2.29)$$

where all derivatives of f are evaluated at (S_t, t) , that is, just before any (potential) jump in $(t, t + dt]$.

Proof: the proof is a combination of Itô's lemma for Brownian motion (2.4) and the theory for a pure jump process. To simplify the computations, we set

$$dY_t = \mu dt + \sigma dW_t$$

and use the function $f(S_t)$ only on S_t , the general case being trivial to determine. Then we get

$$\begin{aligned} df &= f(S_{t+dt}) - f(S_t) = \\ &= \begin{cases} f(S_t + dY_t) - f(S_t) & \text{with probability } 1 - \lambda dt \\ f(S_t + J_t + dY_t) - f(S_t) & \text{with probability } \lambda dt. \end{cases} \end{aligned}$$

The Itô's Lemma 2.4 applied to the first line gives us

$$f'(S_t) dY_t + \frac{1}{2} f''(S_t) (dY_t)^2.$$

The second line, we can rewrite as

$$f(S_t + J_t + dY_t) - f(S_t + J_t) + f(S_t + J_t) - f(S_t).$$

We apply again Itô's lemma to the first term and we get

$$df = \begin{cases} f'(S_t) dY_t + \frac{1}{2} f''(S_t) (dY_t)^2 \\ f'(S_t + J_t) dY_t + \frac{1}{2} f''(S_t + J_t) (dY_t)^2 + f(S_t + J_t) - f(S_t) \end{cases}$$

with probabilities $1 - \lambda dt$ and λdt , respectively. In terms of dN_t we get

$$\begin{aligned} df &= f'(S_t) dY_t + \frac{1}{2} f''(S_t) (dY_t)^2 + [f(S_t + J_t) - f(S_t)] dN_t + \\ &\quad + \left[[f'(S_t + J_t) - f'(S_t)] dY_t + \frac{1}{2} [f''(S_t + J_t) - f''(S_t)] (dY_t)^2 \right] dN_t. \end{aligned}$$

The terms in the second line are all zero. Indeed, $dY_t dN_t$ and $(dY_t)^2 dN_t$ are all random variables with mean and variance 0.

$$\mathbb{E}(dW_t dN_t) = \mathbb{E}(dW_t) \mathbb{E}(dN_t) = 0$$

and

$$\text{Var}(dW_t dN_t) = \mathbb{E}(dW_t^2) \mathbb{E}(dN_t^2) = dt(\lambda dt + \lambda^2 (dt)^2) = 0.$$

Hence, $dW_t dN_t = 0$. Similarly we obtain $dt dN_t = 0$.

Using now the fact that μ and σ are adapted, the same conclusion follows for $dY_t dN_t$ and for $(dY_t)^2 dN_t$. Hence,

$$df(S_t) = f'(S_t) dY_t + \frac{1}{2} f''(S_t) (dY_t)^2 + [f(S_t + J_t) - f(S_t)] dN_t,$$

which is the result we have proven. □

Risk neutral and equivalent martingale measure

A key concept in the Black and Scholes formulation (Section 2.3) and in the financial mathematics is the absence of arbitrage, which states that there are no possibilities that one can make a win without taking any risk. This principle leads to a uniquely defined price, that is given in terms of an expectation value with respect to the equivalent martingale measure.

Definition 2.4.4. *A probability measure \mathbb{Q} is called an **equivalent martingale**, or a **risk neutral measure**, if \mathbb{Q} is equivalent to \mathbb{P} , and if under \mathbb{Q} the discounted asset prices are martingales.*

It is known that the existence of an equivalent martingale measure \mathbb{Q} guarantees that the market is arbitrage free. The inverse is not always true in the continuous time. A proof of this result can be found in Björk [17] and Duffie [39].

Another main assumption in the Black and Scholes model (Section 2.2.3) was the completeness of the market. We present next a result of a great importance in the mathematical finance theory, which gives the relation between market completeness and the equivalent martingale measure.

Theorem 2.4.2. *Suppose that a market is free of arbitrage. Then it is complete if and only if the equivalent martingale measure is unique.*

A proof of this result is given in Björk [17].

By introducing jumps in the model formulation, the market becomes incomplete, as there are two sources of risk and only one traded asset. Following Briani [20] we present next a result which ensures the existence of an equivalent martingale measure \mathbb{Q} in a jump-diffusion model (2.23).

Proposition 2.4.1. *The equivalent martingale measures are characterized by their Radon-Nikodym density with respect to \mathbb{P}*

$$\begin{aligned} \frac{d\mathbb{Q}}{d\mathbb{P}}|_{\mathcal{F}_t} = & \exp\left(-\int_0^t \theta_u dW_u - \frac{1}{2} \int_0^t \theta_u^2 du\right) \times \\ & \times \exp\left(\int_0^t \int_{\mathbb{R}} \ln(\gamma_u(y)) \nu(du, dy) - \int_0^t \int_{\mathbb{R}} (\gamma_u(y) - 1) \lambda \nu(dy) du\right), \end{aligned}$$

where θ_t and γ_t are two predictable processes such that

$$\mu - r = \theta_t \sigma + \lambda \int_{\mathbb{R}} \gamma(t, y) (1 - \gamma_t(y)) dy,$$

and

$$\gamma_t > 0 \quad \text{and} \quad \mathbb{E} \left[\frac{d\mathbb{Q}}{d\mathbb{P}} \right] = 1.$$

In Proposition 2.4.1, θ_t is interpreted as the *market price of diffusion risk* and γ_t as the *market price of jump risk*.

From Girsanov's theorem (see Øksendal and Sulem [82]), we know that there exists a probability measure \mathbb{Q} , equivalent to the original "risky" probability measure \mathbb{P} , such that the process

$$W_t^{\mathbb{Q}} = W_t + \int_0^t \theta_u du$$

is a \mathbb{Q} Brownian motion and ν is a $(\mathbb{Q}, \mathcal{F}_t)$ marked point process with predictable intensity $\lambda \gamma_t(y) \nu(y)$. Moreover, we have that the risky asset in (2.23) satisfies

$$dS_t = rS_t dt + \sigma S_t dW_t^{\mathbb{Q}} + \int_{\mathbb{R}} \gamma(t, y) [\nu(dt, dy) - \lambda \gamma_t(y) \nu(dy) dt]. \quad (2.30)$$

From this transformation we can remark that both the drift and intensity parameter will change, while the other parameters remain the same.

There are a few commonly used methods in the literature to approach the problem of pricing derivatives in incomplete markets:

- The idea of the first method is to obtain completeness of the market by adding as many new assets as the sources of uncertainty. Notice that in this case the dimension of the problem increases and may lead to complications;
- The second procedure relies on selecting an equivalent martingale measure corresponding to a "fair price". This can be done by minimizing the risk (calculate the so called minimal martingale measure), or by the super-hedging (based on the idea of "protecting from the worst" and find the maximal martingale measure);
- Looking at the economic point of view, the last method relies on selecting an unique price for derivatives by performing the maximization of utility at equilibrium.

We use the second approach and show that it is possible to characterize the fair price of derivatives as the solution to a partial integro-differential problem.

2.4.2 Derivative pricing with jumps

In Section 2.2.3 we have showed that the value of a European option can be computed by solving a partial differential equation (2.8) with boundary conditions. Following the approach presented by Cont and Tankov [30], we show that a similar result holds if the price is modeled by a jump-diffusion process and that in this case the value of a European option solves a second order *partial integro-differential equation (PIDE)*.

Consider a European option with maturity T and payoff $\phi(S_T)$ which satisfies the Lipschitz condition

$$|\phi(x) - \phi(y)| \leq c|x - y|$$

for $c > 0$ and $x, y \in \mathbb{R}_+$. This condition is verified by call and put options with $c = 1$. The value V of such an option is given by

$$V(S, t) = e^{-r(T-t)} \mathbb{E} [\phi(S_T) | S_t = S].$$

Theorem 2.4.3. PIDE for European options

Consider the dynamics of S_t given as a general jump-diffusion process under \mathbb{Q}

$$dS_t = \mu S_t dt + \sigma S_t dW_t + \int_{\mathbb{R}} \gamma(t, y) \tilde{m}(dt, dy).$$

If $\sigma > 0$, then the value of a European option with terminal payoff $\phi(S_T)$ is given by $V(S, t)$, where $V : \mathbb{R}_+ \times [0, T] \rightarrow \mathbb{R}$, $(S, t) \rightarrow V(S, t) = E[\phi(S_T) | S_t = S]$ is continuous on $\mathbb{R}_+ \times [0, T]$, smooth on $\mathbb{R}_+^* \times (0, T)$ and it verifies the partial integro-differential equation

$$\begin{aligned} \frac{\partial V}{\partial t} + \frac{1}{2} \sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + \mu S \frac{\partial V}{\partial S} - rV + \\ + \lambda \int_{\mathbb{R}} \left[V(S + \gamma(S, t, y), t) - V(S, t) - \gamma(S, t, y) \frac{\partial V}{\partial S} \right] \nu(dy) = 0, \end{aligned} \quad (2.31)$$

on $\mathbb{R}_+ \times [0, T)$ with the terminal condition $V(S, T) = \phi(S_T)$.

Proof: see Cont and Tankov [30].

□

In the next section we show that the American option valuation can be formulated as a free boundary value and a partial-integro complementarity problem in the case of a jump-diffusion model, using similar arguments as in a purely differential case (Section 2.3).

Partial-integro complementarity formulation for American options

As stated before, the value of an American put option is given by the supremum over all stopping times of the payoff at exercise

$$V(S, t) = \sup_{\tau \in \mathcal{T}_{t, T}} \mathbb{E}^{\mathbb{Q}} [e^{-r(\tau-t)} (K - S_{\tau})^+ | S_t = S].$$

Following Pham [84], Cont and Tankov [30] and the results from Section 2.3, we can give the following characterization of the American put option as a free boundary value problem, under a jump-diffusion model.

Proposition 2.4.2. Free boundary problem

Let $V : \mathbb{R}_+ \times [0, T) \rightarrow \mathbb{R}$ be a nonincreasing and convex function and $S_{fb} : [0, T) \rightarrow \mathbb{R}$, $0 \leq S_{fb} \leq K$. Moreover, assume that $\sigma > 0$, the intensity of jumps is finite and $r - \int (e^y - 1)\nu(y) \geq 0$. Then (V, S_{fb}) is the unique pair of continuous functions verifying the following conditions

$$\begin{aligned} V(S, t) &= K - S_t, & \text{for } S < S_{fb} \\ \frac{\partial V}{\partial t} + \mathcal{D}[V] + \mathcal{L}[V] - rV &< 0, & \text{for } S < S_{fb} \\ V(S, t) &> (K - S_t)^+, & \text{for } S > S_{fb} \\ \frac{\partial V}{\partial t} + \mathcal{D}[V] + \mathcal{L}[V] - rV &= 0, & \text{for } S > S_{fb} \\ \lim_{S \rightarrow S_{fb}} V(S, t) &= K - S_{fb}, & \forall t \in (0, T] \\ \lim_{S \rightarrow S_{fb}} \frac{\partial V}{\partial S}(S_{fb}, t) &= -1, & \forall t \in [0, T) \\ \lim_{S \rightarrow \infty} V(S, t) &= 0, & \forall t \in [0, T) \\ \lim_{S \rightarrow 0} V(S, t) &= \phi(S), & \forall t \in [0, T) \\ V(S, T) &= (K - S_T)^+ & \forall S \in \mathbb{R}_+, \end{aligned}$$

$$\begin{aligned} \text{where } \mathcal{D}[V] &= \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + \mu S \frac{\partial V}{\partial S} \\ \mathcal{L}[V] &= \lambda \int_{\mathbb{R}} \left[V(S + \gamma(S, t, y), t) - V(S, t) - \gamma(S, t, y) \frac{\partial V}{\partial S} \right] \nu(dy). \end{aligned} \tag{2.32}$$

Proof: see Pham [84].

□

Following similar arguments as in the Section 2.3, we can give the linear complementarity formulation for American put options, in the case when the underlying follows a jump-diffusion process.

Proposition 2.4.3. The Partial-Integro Complementarity Problem

The value of the American put option $V : \mathbb{R}_+ \times [0, T) \rightarrow \mathbb{R}$ satisfies the following partial-integro complementarity formulation

$$\begin{cases} \frac{\partial V}{\partial t} + \mathcal{D}[V] + \mathcal{L}[V] - rV = 0 \\ V(S, t) - \phi(S) \geq 0 \\ \left(\frac{\partial V}{\partial t} + \mathcal{D}[V] + \mathcal{L}[V] - rV \right) (V(S, t) - \phi(S)) = 0 \\ V(S, T) = \phi(S_T), \end{cases}$$

where $\mathcal{D}[V]$ and $\mathcal{L}[V]$ were defined in (2.32) and $\phi(S) = (K - S_t)^+$.

This partial-integro complementarity problem can be also written equivalently as a nonlinear Hamilton-Jacobi-Bellman equation on $\mathbb{R}_+ \times [0, T]$

$$\begin{cases} \min \left\{ -\frac{\partial V}{\partial t} - \mathcal{D}[V] - \mathcal{L}V + rV, V - \phi(S) \right\} = 0 \\ V(S, T) = \phi(S_T). \end{cases} \quad (2.33)$$

Following the usual mathematical theory, Proposition 2.4.3 could be proved by applying Itô's formula, if V were a smooth function. However, the solution $V(S, t)$ is known not to be $C^{2,1}(\mathbb{R}_+ \times [0, T])$, hence the notion of viscosity solution has to be introduced. Existence and uniqueness of the viscosity solution for this system of variational inequalities is given in Zhang [109] and involves the dynamic programming principle.

Using similar arguments, we prove in Chapter 4 existence and uniqueness theorems of the viscosity solution in the case of swing options.

In the next section we give a general introduction to viscosity solutions and a short summary of the main results needed to prove existence and uniqueness of the solution for the complementarity problems.

2.5 Viscosity solutions

Viscosity solutions gained an increasing interest in the last years in the financial and mathematical literature being an effective technique to obtain weak solutions for partial differential equations, when no classical solution can be shown to exist. Together with the results from Barles and Souganidis [6], the viscosity solution theory provides a framework for proving convergence for a wide range of numerical methods.

For a detailed overview of viscosity solutions see the classical article by Crandall, Ishii and Lions [33], Fleming and Soner [41] or Touzi [100].

In the jump-diffusion framework, viscosity solution theory has been studied in Alvarez and Tourin [1], Amadori [2], [3] and Cont and Tankov [30]. We follow these approaches in this section, and introduce the main definitions and results for the viscosity solutions in a jump-diffusion market.

We consider a general integro-differential operator of the form

$$F(x, t, v, Iv, \mathcal{D}v, \mathcal{D}^2v) = 0,$$

with $F : \mathbb{R}^n \times [0, T] \times \mathbb{R} \times \mathbb{R} \times \mathbb{R}^n \times \mathcal{M}(n) \rightarrow \mathbb{R}$ and $\mathcal{M}(n)$ is the set of symmetric $n \times n$ matrices, $\mathcal{D}v$, \mathcal{D}^2v are respectively the gradient and Hessian of v with respect to $x \in \mathbb{R}^n$, and Iv is an integral term given by

$$Iv(x, t) = \int_{\mathbb{R}^n} M[v(x + y, t, v(x, t))] \mu_{x,t}(dy),$$

where $\mu_{x,t}$ are positive measures and M is a Lipschitz continuous function, nondecreasing in the first argument.

The operator $F(x, t, v, Iv, \mathcal{D}v, \mathcal{D}^2v)$ satisfies the following properties

- $F(x, t, v, Iv, \mathcal{D}v, \mathcal{D}^2v)$ is *degenerate elliptic*, i.e.

$$F(x, t, v, I, p, X) \geq F(x, t, v, I, p, Y) \quad \text{whenever } X \leq Y \quad (2.34)$$

- F is *non-increasing* with respect to the non-local term Iv , i.e.

$$F(x, t, v, I, p, X) \geq F(x, t, v, J, p, X) \quad \text{whenever } I \leq J. \quad (2.35)$$

We introduce next a new class of admissible functions, which we will show to be very important for the stability of viscosity solutions.

Definition 2.5.1. *A function $f(z, l; y)$ has an **upper** (respectively **lower**) μ -bound at (x, t) if there exists a neighborhood $V_{x,t}$ of (x, t) and a function $\phi \in C(\mathbb{R}^n) \cap \mathcal{L}^1(\mathbb{R}^n; \mu_{x,t})$ such that*

$$\lim_{(z,l) \rightarrow (x,t)} \int_{\mathbb{R}^n} \phi(y) \mu_{z,l}(dy) = \int_{\mathbb{R}^n} \phi(y) \mu_{x,t}(dy)$$

$$f(z, l; y) \leq \phi(y) \text{ (respectively } \geq) \quad \mu_{z,l} - \text{a.e. } y, \text{ for all } (z, l) \in V_{x,t}.$$

In this thesis we only work with parabolic integro-differential equations, thus we define a general pricing equation in the following way

$$\frac{\partial v}{\partial t} + F(x, t, v, Iv, \mathcal{D}v, \mathcal{D}^2v) = 0 \quad (2.36)$$

$$v(x, 0) = g_0(x) \quad (2.37)$$

where g_0 is a given continuous function.

Before we give the definition of viscosity solution to (2.36)-(2.37) and show the importance of conditions (2.34)-(2.35), we present an equivalent definition of the notion of sub- and supersolutions in the classical sense

Definition 2.5.2. *Let $v \in C^2(\mathbb{R}^n \times [0, T])$. Then the following statements are equivalent*

- v is a **subsolution** (respectively **supersolution**) of (2.36)-(2.37),
- for all functions $\varphi \in C^2(\mathbb{R}^n \times [0, T])$ with $x_0 \in \mathbb{R}^n$ being a local maximum (respectively minimum) of $v - \varphi$, we have

$$\frac{\partial \varphi}{\partial t}(x_0, t) + F(x_0, t, v(x_0, t), Iv(x_0, t), \mathcal{D}\varphi(x_0, t), \mathcal{D}^2\varphi(x_0, t)) \leq 0$$

$$(\text{ respectively } \geq 0).$$

Next we introduce the following spaces of semicontinuous functions on $\mathbb{R}^n \times [0, T]$ (see Briani [20]).

USC_I – is the set of upper semicontinuous, locally bounded functions on $\mathbb{R}^n \times [0, T]$ such that $M[v(x+y, t), v(x, t)]$ has an upper μ – bound at any (x, t) . (2.38)

LSC_I – is the set of lower semicontinuous, locally bounded functions on $\mathbb{R}^n \times [0, T]$ such that $M[v(x+y, t), v(x, t)]$ has a lower μ – bound at any (x, t) . (2.39)

$\mathcal{C}_I = USC_I \cap LSC_I$. (2.40)

We can now give the main definitions of this section.

Definition 2.5.3. *A locally bounded function $v \in USC_I(\mathbb{R}^n \times [0, T])$ is a **viscosity subsolution** of (2.36)-(2.37) if for any test function $\varphi \in C^2(\mathbb{R}^n \times [0, T])$ and $x_0 \in \mathbb{R}^n$ being a local maximum of $v - \varphi$, we have*

$$\frac{\partial \varphi}{\partial t}(x_0, t) + F(x_0, t, v(x_0, t), Iv(x_0, t), \mathcal{D}\varphi(x_0, t), \mathcal{D}^2\varphi(x_0, t)) \leq 0.$$

*A locally bounded function $v \in LSC_I(\mathbb{R}^n \times [0, T])$ is a **viscosity supersolution** of (2.36)-(2.37) if for any test function $\varphi \in C^2(\mathbb{R}^n \times [0, T])$ and $x_0 \in \mathbb{R}^n$ being a local minimum of $v - \varphi$, we have*

$$\frac{\partial \varphi}{\partial t}(x_0, t) + F(x_0, t, v(x_0, t), Iv(x_0, t), \mathcal{D}\varphi(x_0, t), \mathcal{D}^2\varphi(x_0, t)) \geq 0.$$

*A function v is called **viscosity solution** of (2.36)-(2.37) if it is both a subsolution and a supersolution.*

An important consequence of the Definition 2.5.2 and Definition 2.5.3 is the following remark.

Remark 2.5.1. *Any classical solution is also a viscosity solution.*

By weakening the notion of solution in Definition 2.5.3, the set of solutions is enlarged, and one has to guarantee that uniqueness still holds. This can be done by establishing a comparison principle (or maximum principle), which is a key element in the viscosity theory. The comparison principle is mainly used to prove uniqueness and in some cases continuity of the viscosity solution.

Given the existence and uniqueness results, it is then possible to determine regularity for the solutions of the integro-differential and complementarity problems. We present these issues in more detail in the Chapter 4, in the context of swing options valuation.

Chapter 3

A new model for pricing swing options in the electricity market

In this chapter we present the mean-reverting double exponential jump-diffusion model which we use for the valuation of swing options. We start with an introduction to electricity markets and then explain its price formation and the factors which influence it. Two major European markets are described in detail, the Nord Pool and the European Energy Exchange (EEX), since we use spot data from these markets to calibrate our model.

Next, we describe the main characteristics of electricity prices and give a short overview of the existing models in the literature. A complete presentation of electricity markets and modeling approaches can be found for example in the works of Weron [105], Benth et. al. [14], Geman [44] and Burger et. al. [22].

The main part of this chapter is devoted to the mean-reverting double exponential jump-diffusion process which we propose for modeling electricity prices. This is a modification of the Kou model, introduced by Kou and Wang [66] for simulating the interest rates. Our model captures as well the mean-reversion and seasonal behavior observed in the electricity prices. The jump size follows an asymmetric double exponential distribution, which is able to reproduce the size and intensity of both negative and positive jumps. The seasonality function is periodic and deterministic, and incorporates weekly, annual and semi-annual patterns.

To calibrate the model, we have used one year historical daily prices from the Nord Pool and EEX markets. The simulation results show that the model is capable of fitting the market data accurately.

3.1 Introduction to electricity markets

In the past, the electricity market was strictly regulated and controlled by governments. Prices were set in advance by the authorities, reflecting simply the production costs of the energy delivered by state-owned companies. Therefore, electricity consumers were not exposed to any price risk. However, in the last two decades, governments have decided to liberalize the market in order to make the industry more competitive and efficient. Thus, electrical energy has become a good which can be traded in the form of delivery contracts on specialized exchanges, such as Nord Pool, European Energy Exchange (EEX), Amsterdam Power Exchange (APX), United Kingdom Power Exchange (UKPX), etc. .

Together with this deregulating process, electricity prices have become more volatile, exposing both producers and consumers to a higher price risk. In consequence, organized energy markets had to be created and new financial derivatives and methods to minimize the risk were developed.

So far, deregulated power markets have been mainly organized in two forms: power pools and power exchanges. We present their main characteristics as follows.

The power pool is an organized system where generators place their bids in terms of prices and quantities for each hour of the next day. The Transmission System Operator (TSO) acts as a single buyer and collects these bids, sorts them from the lowest to the highest price, in this way building the supply curve. There are two possible pool designs: either only the suppliers make bids on the pool and the system operator computes the expected demand; or, both buyers and sellers place bids on the pool and then the system operator builds a demand function, similarly to the supply curve. In both cases, intersecting the demand and the supply functions provides the *system marginal price* (SMP).

The power exchange (PX) is the major marketplace for electricity and it is usually owned by market participants like generators, distribution companies, traders and large consumers. Trading is realized through bilateral contracts, which have to be completed the day before delivery in order to give both market participants and the TSO enough time to arrange physical aspects of the delivery. Matching the supply and demand of electricity, renders the *market price* or *spot price*.

In this thesis we use daily average prices from the Nordic power market (Nord Pool) and the German-based European Power Exchange (EEX). Thus, we present these markets in more detail as follows.

The Nord Pool market

The Nordic commodity market for electricity, or Nord Pool, was founded in 1993 and is the first multinational power exchange in the world. Initially it was a Norwegian market for physical contracts, but in the years to follow it extended, and Sweden, Finland, Denmark and Estonia joined in.

The Nord Pool is located in Lysaker (Norway) and it is divided into one physical and one financial market:

- The physical market (called *Elspot*) handles power contracts and it is organized as a day-ahead market. This type of market is based on an auction trade system, where each morning players submit their bids for each of the 24 hours of the next day. At noon this market is closed for bids and the day-ahead price is derived for each hour of the next day. After the publication of the Elspot price, there is a physical intra-day market called *Elbas* for the areas Eastern Denmark, Finland and Sweden. Elbas provides continuous power trading 24 hours a day, and it allows market participants who had previously taken positions on Elspot to adjust these positions up to one hour before delivery.
- The financial market consists of different types of futures and forward contracts, but as well standardized options and arithmetic average options.

The European Energy Exchange - EEX

The European Energy Exchange was established in 2002 from the association of two German power exchanges, one based in Frankfurt (EEX) and the other in Leipzig (LEX). The resulting exchange (named EEX and located in Leipzig), became the biggest energy exchange in continental Europe, by number of players and generation capacity.

EEX includes the following markets: power spot market, power futures and options, coal futures and EU emission allowances.

- The power spot market includes a day-ahead and an intra-day market. The day-ahead market works similar as described above for the Nord Pool. Additionally, there is an auction for delivery in the Austrian Power Grid and in the Swissgrid. The intra-day market offers continuous trading 24 hours a day, seven days a week.
- The power futures market includes a variety of products like weekly, monthly, quarterly and yearly futures. They are based on the EEX spot market index Phelix (Physical Electricity Index) as the underlying asset. Phelix represents the average of all prices between 08:00 a.m. and 08:00 p.m. .
- EEX has started trading off emission allowances since 2005, on the basis of the EU Emission Trading Scheme (EU-ETS). The EU-ETS is the largest emissions trading system in the world and it is a market-based instrument of the EU climate policy. The emission allowances are traded on the EEX spot and derivatives market on a continuous basis.

On both exchanges, the average price over the entire day is called the *base load price*, the average over the most demanding hours is called *peak price* and the average over the remaining hours is called *off-peak price*.

Next we turn our attention to the electricity price and its characteristics. We can state in general terms that the price is given by the intersection of the aggregate demand and supply curve.

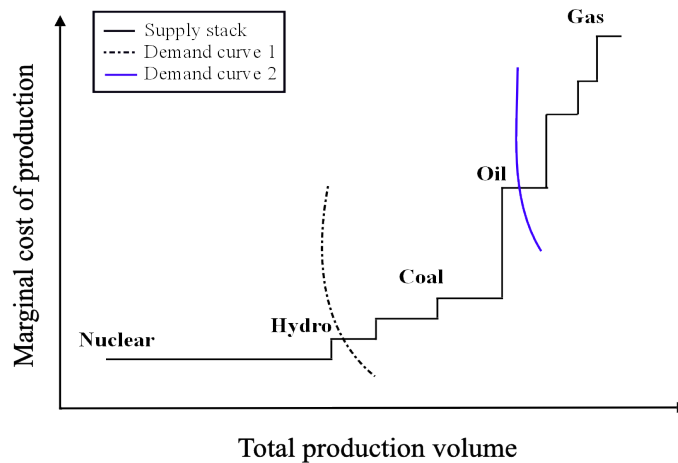


Figure 3.1: If the demand is low, power plants with lower production costs (nuclear, hydro), are used; if the demand is high, additional plants with higher production costs (oil, gas) are running, producing a huge effect on the price.

The graphic displayed in Figure 3.1 shows a schematic supply-demand curve. The supply and demand are affected by many factors, that will also act on the prices. For example, supply may be influenced by fluctuation of fuel prices (oil, gas) or the CO_2 prices (because of the emission allowances). Electricity demand on the other hand is very seasonal: it is not uniform during the week, it peaks during weekdays, working hours and is low during weekends and nights (due to the low industry activity). It has also yearly peaks in summer and winter due to the use of air conditioned, respectively heating systems.

In the next section we describe the main characteristics of the electricity spot prices, which should be taken in account for constructing a realistic price model.

3.1.1 Characteristics of the electricity prices

Electricity prices present new challenges to the commodity modeling discipline due to their non-storable¹ nature. This unique feature has strong implications in trading and in the behavior of prices. Thus, the dynamics of electricity spot is very complex, having daily, weekly and annual seasonal patterns, or short-lived price deviations (or jumps) with strong mean-reversion, a unique characteristic of the power market.

Another unusual feature of power prices is that the demand is inelastic to price deviations. This means that the consumers will buy electricity at any price in

¹Electricity can be partially stored by using hydro pumped storage power plants. However, considering that in most countries their capacity is quite small and that pumping generally leads to an energy loss of approximately 30%, it is reasonable to say that electricity is non-storable, at least not in an adequately efficient and conventional way and at sufficiently large volumes (Burger [22]).

order to fulfill their contracts, because it is an indispensable commodity. As a consequence, more expensive power plants are started in order to generate enough power, causing higher prices (see Figure 3.1). When demand decreases, so will do the prices. This permanent process leads mainly to the price dynamics observed below in Figure 3.2.

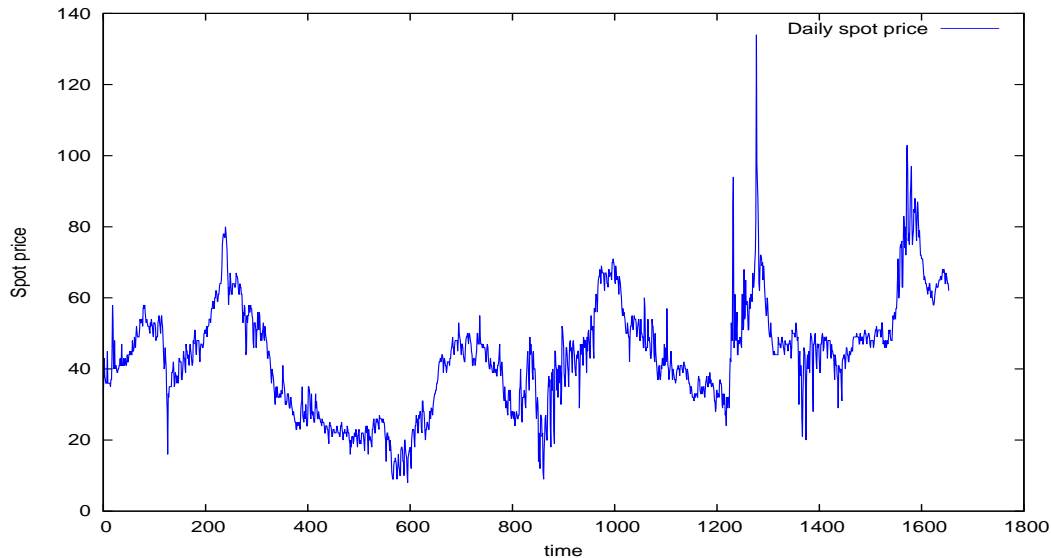


Figure 3.2: Nord Pool spot prices between 01/2006 and 03/2011.

Figure 3.2 shows five years Nord Pool electricity prices and their typical characteristics: mean-reversion, seasonality and huge jumps.

At the beginning of 2010, Nord Pool had one of the most difficult trading periods in its recent history. A number of factors combined caused these extreme prices. In the first months of the year, the Nordic region experienced a colder and dryer period than normal meteorological conditions. Thus, hydro reserves were 15% below their normal levels and the demand for heating increased significantly. Besides that, two of the four nuclear reactors in Sweden were off grid. All these factors determined the huge jumps and the price behavior observed in Figure 3.2.

The main characteristics of electricity prices are described in detail in the following part.

Mean-reversion is the typical property of commodity prices to fluctuate towards a price level (or a long-run mean), which may be viewed as the marginal cost of production. Even though prices can move far away from this mean, they will always be pulled back to it, by mean-reversion. This behavior is a direct consequence of the balance between supply and demand, and because weather is a dominant factor in the demand for electricity.

The **seasonality** of the electricity prices comes mainly from the seasonalities of demand and supply, which are influenced by economic activities and weather

conditions. The seasonal patterns of electricity's demand are among the most complicated, exhibiting three different types of seasonalities:

- Daily - due to the working and non-working hours.
- Weekly - where the price is lower on weekends and on national holidays due to lower industrial activity.
- Yearly - depending on the geographical region, the demand for electricity peaks in summer months due to humidity leading to extensive use of air conditioning, and in winter months due to the use of heating systems.

Even supply can have seasonal fluctuations, for example hydro units are heavily dependent on weather conditions, such as precipitation and melting snow, which varies from season to season.

Thus, part of the seasonality of electricity prices can be explained by the periodicity of the load, that is in some way predictable. In Section 3.2 we model seasonality as a deterministic function and present simulation results using real data.

The **jumps** are sudden large upward or downward price movements that are a unique characteristic of electricity price dynamics. The jumps are usually short-lived and soon as the phenomenon which caused them is over, the prices get back to a normal level, producing a **spike** in the commodity price process.

As we have shown before, in deregulated electricity market, prices are determined by the intersection of the demand and supply curves. Any unexpected event in either the supply or the demand side (like, for example, severe weather or political conditions, generation outages or transmission failures), will cause a spike in the price of electricity. The size of these jumps can be up to 30 times that of the normal price level. This extreme behavior is found only in the electricity prices.

Negative jumps and spikes can occur as well, though they are not so often and less obvious than the positive ones. They are usually the consequence of low demand hours, because the producers who generated too much electricity must find a way to compensate their long position. This would lead to prices below the production costs, and hence to negative jumps and spikes.

The spikes and extreme volatility observed in the electricity market also lead to non-Normal distribution of spot prices returns². Even though it has been generally assumed that the log-returns of commodity prices usually follow the Normal distribution, the plot in Figure 3.3 shows that for electricity, this assumption does not hold. If the log-returns were Normally distributed, the Quantile-Quantile³ plot of the returns would be close to a straight line, but fat-tails are clearly visible.

²The returns and log-returns are defined as $r_t = \ln\left(\frac{S_t}{S_{t-1}}\right)$, where S_t is the spot price.

³Quantiles are points taken at regular intervals from the cumulative distribution function of a random variable. A Quantile-Quantile plot (Q-Q plot) is a graphical method for comparing two probability distributions by plotting their quantiles against each other.

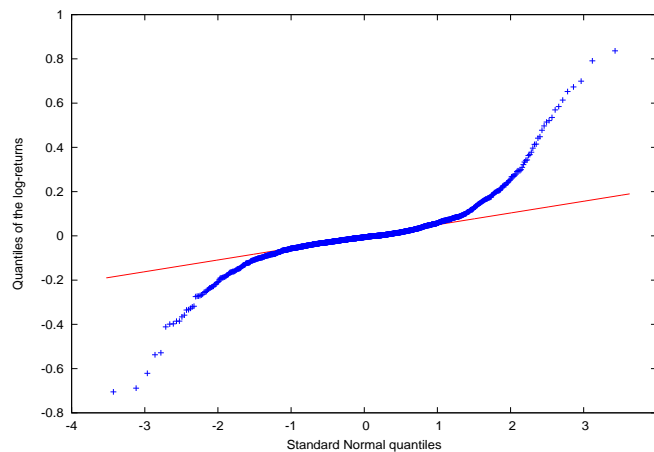


Figure 3.3: Quantiles of Nord Pool electricity log-returns versus standard Normal quantiles.

The plot in Figure 3.3 is generated by taking the cumulated distribution function (cdf) for the quantiles of the Nord Pool market data returns versus theoretical quantiles from a Normal distribution.

The motivation presented above shows that a simple Brownian Motion process is not suitable to model the electricity prices and emphasizes once again the importance of introducing jumps in the model.

In Section 3.2 we present a jump-diffusion model with seasonality, in which the jumps are described by a Poisson process (see Section 2.4) and the jumps sizes by an asymmetric double exponential distribution, which is able to reproduce accurately the size and intensity of both negative and positive jumps observed in the market data.

Taking in account the peculiarities of power prices outlined above, we show next that modeling the price behavior of electricity is a very challenging task.

3.1.2 Literature review and modeling approaches

Although there have been various approaches developed for the power price in the literature, there is so far no model which has been widely accepted as standard for electricity.

Generally, we can divide the electricity price models into two classes: spot-based and forward-based models. In this thesis we only present spot price models since they allow for a good mathematical description of the market. Moreover, the swing options are strongly influenced by the hourly-price behavior, which can be only captured by spot models.

One of the first attempts to model energy prices was a Brownian motion process, which is very popular in the financial modeling literature. However, given the motivation from the previous section it is clear that simple geometric Brownian motion is not well suited to model the electricity prices.

An extension of this model, which includes mean-reversion, is the Ornstein-Uhlenbeck (OU) process

$$dS_t = \alpha(\mu - S_t)dt + \sigma dW_t, \quad (3.1)$$

where α is the rate of mean-reversion and μ is the long-run mean, to which the process tends to revert.

A well known and commonly used process for electricity was proposed by Lucia and Schwartz [74], who modeled the price as an exponential OU process X_t plus a seasonal deterministic component f

$$\begin{aligned} S_t &= \exp(f(t) + X_t) \\ dX_t &= -\alpha X_t dt + \sigma dW_t. \end{aligned} \quad (3.2)$$

Spot price models can be divided into one- and multi-factor models. For single factor models the spot price is itself a Markov process, while in multi-factor models the price is a function of a multidimensional Markov process. Lucia and Schwartz [74] extended their model (3.2) to a two factor model, including another stochastic term Y_t . This process represents the long-term dynamics and is given by a Brownian motion with drift

$$\begin{aligned} S_t &= \exp(f(t) + X_t + Y_t) \\ dX_t &= -\alpha X_t dt + \sigma_X dW_X \\ dY_t &= \mu dt + \sigma_Y dW_Y \\ dW_X dW_Y &= \rho dt. \end{aligned}$$

Another two-factor diffusive mean-reverting process was introduced by Pilipovic [86], where the long-term mean changes over time

$$\begin{aligned} dS_t &= \alpha_1(L_t - S_t)dt + \sigma_1 S_t (dW_1)_t \\ dL_t &= \alpha_2 L_t dt + \sigma_2 L_t (dW_2)_t \end{aligned}$$

where S_t is the spot price, L_t is the equilibrium price, α_1 rate of mean-reversion and α_2, σ_2 drift, respectively volatility of the long-term equilibrium price.

A completely different approach was proposed by Barlow [9], where the electricity demand is assumed to be an OU process (X_t), and the supply is deterministic and nonlinear. The spot price process is then defined by the equilibrium between supply and demand

$$S_t = \begin{cases} f_\alpha(X_t), & 1 + \alpha X_t > \epsilon \\ \epsilon^{1/\alpha}, & 1 + \alpha X_t \leq \epsilon \end{cases} \quad (3.3)$$

$$dX_t = -\lambda(X_t - a)dt + \sigma dW_t \quad (3.4)$$

where $f_\alpha(x) = (1 + \alpha x)^{1/\alpha}$, $\alpha \neq 0$ and $f_0(x) = e^x$. When $\alpha = 0$ an exponential OU process is retrieved, and for $\alpha = 1$ one gets a normal OU process.

None of the processes presented until now included jumps in their formulation, so we introduce next the class of jump-diffusion models. The jumps are usually

modeled by a compound Poisson process (see Section 2.4) where the size and the time occurrence are independent.

The first type of jump-diffusion model was introduced by Merton [79]

$$dS_t = \mu S_t dt + \sigma S_t dW_t + dJ_t, \quad (3.5)$$

where $\{J_t, t \geq 0\}$ is a compound Poisson process with intensity λ and jump size distribution $J_t = \sum_{i=1}^{N_t} Z_i$. In this formula N_t is a Poisson process with intensity λ and $\{Z_i\}$ are i.i.d. jump magnitudes. The jump size is normal distributed, i.e. $J_t \sim \mathcal{N}(\mu_J, \sigma_J^2)$, with the density function

$$f_J(z) = \frac{1}{\sigma_J \sqrt{2\pi}} \exp \left\{ -\frac{(z - \mu_J)^2}{2\sigma_J^2} \right\}.$$

Kou and Wang [66] proposed the same type of jump-diffusion model, but the jump size is drawn from a double exponential distribution

$$f_J(z) = p\eta_1 e^{-\eta_1 z} \mathbb{1}_{\{z \geq 0\}} + q\eta_2 e^{\eta_2 z} \mathbb{1}_{\{z < 0\}},$$

where $\eta_1 > 1, \eta_2 > 0, p, q > 1, p + q = 1$ and p is the probability of the upward jumps and q is the probability of the downward jumps.

These two jump-diffusion models are very popular in financial modeling, since they offer solutions to European options in closed-form. However, they do not capture the mean-reversion and seasonal behavior observed in the electricity prices.

In the next section we present a similar model to the one introduced by Kou and Wang [66], which includes the specific electricity features.

Hambly, Howison and Kluge [52] proposed a model for the price consisting of three components: a deterministic periodic seasonality function, an OU process and a mean-reverting process with a jump component

$$\begin{aligned} S_t &= \exp(f(t) + X_t + Y_t) \\ dX_t &= -\alpha X_t dt + \sigma dW_t \\ dY_t &= -\beta Y_{t-} dt + J_t dN_t. \end{aligned}$$

The new parameter β represents the mean-reversion speed of the spike process. One important characteristic of this model is that it allows for separate mean-reversion speeds for both the diffusive and spike parts. Hambly, Howison and Kluge [52] also used this model for pricing swing options applying a method based on the tree approach.

Burger et al. [22] created the Spot Market Price Simulation (SMaPS) model, which uses ideas from fundamental market models and stochastic time-series theory

$$S_t = \exp(f(t, L_t/v_t) + X_t + Y_t).$$

This model can be considered as a three-factor model, where L_t represents the total system load (electricity demand), X_t the short-term price variations and Y_t

the long-term price variations. The other parameters in the model are the empirical merit order curve $f(t, L)$ and the average relative availability of power plants v_t .

We close this section with a different class of models for the electricity prices: the regime switching and Markov-switching models. According to these models, the price process can switch between two or multiple states or regimes. In a mean-reverting Markov-switching jump-diffusion model for example, some of the parameters are allowed to switch. This is done by introducing an extra state variable governed by a Markov process that allows changes in the parameters of the Poisson process. Thus, this non-linear mechanism is switching between normal and high price states. A transition matrix describes the probabilities of leaving and entering a new state. However, the estimation of these models becomes very complex with increasing number of states. For a detailed description of these models we refer to Deng [37], Hamilton [51] or Weron [105].

In the next section we introduce the mean-reverting double exponential jump-diffusion model with seasonality. This process is able to reproduce the main characteristics of the electricity price dynamics and has many advantages for pricing derivatives, which we discuss below.

3.2 The model for pricing swing options

In this section we present a stochastic model for electricity prices, which is new in the context of pricing swing options. This model is based on the Kou jump-diffusion model, developed by Kou and Wang [66]. Furthermore, our model includes mean-reversion and seasonality, two important features which can be observed in many electricity markets.

As discussed in the previous section, there is a multitude of price processes proposed in the literature. However, most of these models are quite complicated or contain many factors and they cannot be used for pricing complex derivatives, like swing options.

The mean-reverting double exponential jump-diffusion process which we propose in this thesis is simple to implement and it can reproduce the spikes observed in the power prices. It can also explain the leptokurtic feature⁴, the volatility smile⁵ and the asymmetric return distributions. These are characteristics which can also be observed in the electricity prices. Another advantage of using this jump distribution is that a recursion formula can be developed to approximate the integral term resulting from the jumps in the price model. This can be done with optimal computational complexity, and it has important properties for the numerical valuation of options. This issue is presented in more detail in the next chapter, in the

⁴The kurtosis is the degree of peakedness of a distribution, defined as a normalized form of the fourth central moment of a distribution. If the coefficient of kurtosis is greater than zero, it is said to be leptokurtic.

⁵The volatility surface often exhibits what are referred to as volatility smile and volatility skew. The distribution of price returns shows a higher kurtosis than the Normal distribution admits. This phenomenon is referred to as fat tails, and in turn generates the empirically observed smile.

context of swing contracts.

We assume that the spot price changes exponentially and it has two components: a deterministic seasonality function f and a mean-reverting process which incorporates jumps

$$\begin{aligned} S_t &= \exp(f(t) + X_t) \\ dX_t &= -\alpha X_t dt + \sigma dW_t + dJ_t. \end{aligned} \quad (3.6)$$

Recall that the jump size $J_t = \sum_{i=1}^{N_t} Z_i$ has an asymmetric double exponential distribution with the density

$$f_J(z) = p\eta_1 e^{-\eta_1 z} \mathbb{1}_{\{z \geq 0\}} + q\eta_2 e^{\eta_2 z} \mathbb{1}_{\{z < 0\}}, \quad (3.7)$$

where $p, q > 1$ are the probabilities of upward, respectively downward jumps. Moreover, $p + q = 1$, $\eta_1 > 1$ and $\eta_2 > 0$. Furthermore, we assume that all sources of randomness, W_t , J_t and Z_t are independent.

Applying Itô's formula with jumps (2.29) for $S_t = \exp(f(t) + X_t)$, we can rewrite the SDE for S_t as

$$dS_t = \alpha(\rho(t) - \ln S_t)S_t dt + \sigma S_t dW_t + S_t(e^{J_t} - 1)dN_t, \quad (3.8)$$

where the time-dependent mean-reverting level is given by

$$\rho(t) = \frac{1}{\alpha} \left(\frac{df(t)}{dt} + \frac{\sigma^2}{2} \right) + f(t). \quad (3.9)$$

The seasonality function $f(t)$ is a periodic and deterministic function which captures the trend, weekly, semi-annual and annual patterns observed in the electricity prices. Its exact formulation is presented later in this chapter.

The new density function of the jump process e^{J_t} is

$$f_{e^J}(z) = p\eta_1 z^{-\eta_1 - 1} \mathbb{1}_{\{z \geq 1\}} + q\eta_2 z^{\eta_2 - 1} \mathbb{1}_{\{0 < z < 1\}}. \quad (3.10)$$

3.2.1 Risk neutral formulation

The classical option pricing theory assumes in general that the market is complete. That means that there are enough tradable contracts available to hedge each factor separately. As we have showed in Section 2.4, in a complete market there exists a unique risk-neutral probability measure under which the price of any asset is equal to the expectation of its payout, discounted at the risk-free rate. However, as electricity is not storable, the no-arbitrage principle fails. Thus, the power market is incomplete and the equivalent martingale measure \mathbb{Q} is not unique.

It can be shown though, that there exists a set of equivalent risk neutral measures \mathbb{Q} , which may not be determined uniquely, but still keep the market arbitrage-free. It is also common in the literature to restrict this set of possible measures to a

subset of equivalent risk neutral measures which leave the structure of the jump process unchanged. In this thesis we assume that under the risk neutral measure, the jumps will still be generated by a Poisson process with a double exponential distribution.

Following the theory presented in the Chapter 2 and the fact that W_t , J_t and Z_t are independent, the dynamics of the spot price under the risk neutral measure \mathbb{Q} is given by

$$\begin{aligned} S_t &= \exp(f(t) + X_t) \\ dX_t &= (-\alpha X_t - \theta_t)dt + \sigma dW_t^{\mathbb{Q}} + dJ_t^{\mathbb{Q}}, \end{aligned} \quad (3.11)$$

where θ_t and γ_t represent the market price of diffusion respectively of jump risk. As before, $W_t^{\mathbb{Q}}$ is a \mathbb{Q} Brownian motion and $J_t^{\mathbb{Q}}$ is a Poisson process with intensity $\lambda(1 + \gamma_t)$. The parameters θ_t , γ_t and the jump size distribution J_t are determined by the particular choice of the measure \mathbb{Q} .

It can be proved that adding the market price of risk to the mean-reverting process results in another mean-reverting process (see Kluge [65]). This observation gives us the possibility to write the model under the risk neutral measure in the shape of the model (3.6) as follows

$$\begin{aligned} S_t &= \exp(f(t) + X_t) \\ dX_t &= -\hat{\alpha}X_t dt + \sigma dW_t^{\mathbb{Q}} + dJ_t^{\mathbb{Q}}. \end{aligned} \quad (3.12)$$

For simplicity, we use from now on in this thesis the original notations for the model parameters. Though, it is important to remark that they might have different values from the parameters under the real world measure \mathbb{P} .

Note that every asset that we consider along this thesis will be an expectation of its discounted payoffs under \mathbb{Q} and hence, the no-arbitrage principle will always hold.

Once we have specified the model under the risk neutral measure, we present in the next section the calibration results using historical market data.

3.2.2 Model calibration

In the next part of this chapter we focus on the calibration of the model parameters. To do this, we first extract the seasonality from the historical data. In the end we estimate the rest of the parameters from the deseasonalized prices.

Deseasonalization

The seasonality in the electricity prices arises mainly due to the seasonality in the demand, and it has daily, weekly and yearly patterns.

The seasonality function depends on the market we are working on: for example in the European markets electricity prices have a peak in the winter months, whereas in the US markets, prices are higher in summer than in winter. In any

case, the seasonality function has to be first calibrated from the historical data, in order to be able to correctly estimate the rest of the parameters.

In this thesis we propose a seasonality function composed of a weekly periodic part $f_{we}(t)$ and a combination of sine and cosine functions plus trend, which capture the annual pattern in the prices.

The deseasonalization is performed in three steps:

1. The annual sinusoidal function plus trend $f_{an}(t)$ is estimated from the historical spot prices.
2. The weekly periodicity $f_{we}(t)$ is modeled by the moving average technique and estimated from the spot data.
3. The deseasonalized spot prices are obtained by subtracting the seasonality function from the original spot, respectively log-spot prices

$$\begin{aligned} desS_t &= S_t - f(t) \quad \text{and} \\ des \ln S_t &= \ln S_t - f(t). \end{aligned}$$

where $f(t) = f_{we}(t) + f_{an}(t)$.

Next we describe these steps in more detail.

1. In order to determine the annual seasonality function, we fitted a combination of sine and cosine functions plus trend to the log-price series using the Gauss-Newton method for the nonlinear least-squares approach (see Bock [19]). Thus, the annual seasonality function is given by

$$f_{an}(t) = a + bt + \sum_{k=1}^6 [c(k) \cdot \sin(2kt\pi/365) + d(k) \cdot \cos(2kt\pi/365)] \quad (3.13)$$

After we have tested several scenarios and ran some tests, we found out that a combination of one sine and one cosine functions with a 4-, 6- and 12-months periodicity fits the best the data from Nord Pool, respectively EEX.

In equation (3.13), a might be viewed as the fixed costs of production, and the second term bt as the long-run linear trend in the total production costs. This trend can be also observed in Figure 3.2, where the spot prices show a clear upward tendency towards the end.

The next step of the deseasonalization process is the modeling and estimating of the weekly seasonality.

2. Electricity prices show a strong intra-week seasonality due to the different energy consumption between working days and the weekend. This behavior is represented in the Figure 3.4, where the hourly Nord Pool spot prices are plotted during a week (from Tuesday the 5th until Tuesday the 12nd of April).

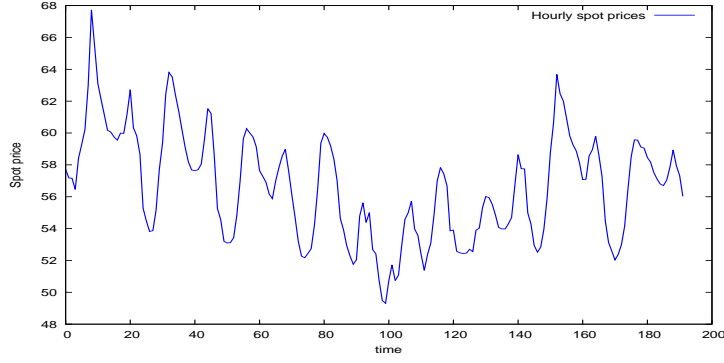


Figure 3.4: Hourly price behavior in the Nord Pool market, between 05/04/2011 and 12/04/2011.

In Figure 3.4 we can observe that the prices are lower on Saturday and Sunday and higher on the other week days, as a consequence of the weekly demand. This represents the intra-week seasonality.

In this thesis we model the weekly price behavior using the *moving average technique*. This method is generally used to analyze a set of data points by creating a series of averages of different subsets of the entire set. Given a series of numbers and a fixed subset size (7 in our case), the first element of the moving average is obtained by taking the average of the initial fixed subset of the number series. Then the subset is modified by "shifting forward" - that is excluding the first number of the series and including the next number following the original subset in the series. This creates a new subset of numbers, which is averaged. This process is repeated over the whole data series. The plot line connecting all the averages is the moving average.

The method is described in detail in Weron [105], so we only present the main steps here.

- For the vector of daily values (log-prices in our case) x_1, x_2, \dots, x_n we define the moving average filter

$$\hat{m}_t = \frac{1}{7}(x_{t-3}, \dots, x_{t-3}), \quad t = 4, \dots, n - 3.$$

- For each $k = 1, \dots, 7$ the average w_k of the deviations $\{(x_{k+7j} - \hat{m}_{k+7j}), 3 < k + 7j \leq n - 3\}$ is computed. Since these average deviations do not necessarily sum to zero, we estimate the seasonal component f_{we} as

$$f_{we}(k) = \hat{m}_k - \frac{1}{7} \sum_{i=1}^7 w_i, \quad k = 1, \dots, 7, \quad (3.14)$$

and $f_{we}(k) = f_{we}(k - 7)$ for $k > 7$.

3. In the end we compute the deseasonalized log-prices

$$f(t) = f_{an}(t) + f_{we}(t), \quad (3.15)$$

$$des \ln S_t = \ln S_t - f(t). \quad (3.16)$$

Table 3.1 presents the estimated parameters of the annual seasonality function (3.13), while Figure 3.5 shows the corresponding plots in the Nord Pool, respectively EEX markets. Figure 3.6 presents the same prices in the two markets, after the deseasonalization process.

| Parameter (NP) | Value |
|----------------|------------|
| a | 3.8218 |
| b | 0.2182e-03 |
| $c(1)$ | -0.1110 |
| $c(2)$ | -0.1713 |
| $c(3)$ | -0.0227 |
| $d(1)$ | -0.0907 |
| $d(2)$ | 0.1401 |
| $d(3)$ | -0.0246 |

| Parameter (EEX) | Value |
|-----------------|-------------|
| a | 4.0310 |
| b | -0.9834e-03 |
| $c(1)$ | 0.1133 |
| $c(2)$ | 0.0242 |
| $c(3)$ | -0.1306 |
| $d(1)$ | -0.0739 |
| $d(2)$ | 0.1742 |
| $d(3)$ | -0.0662 |

Table 3.1: The estimated parameters of the annual seasonality function using the Nord Pool and EEX data.

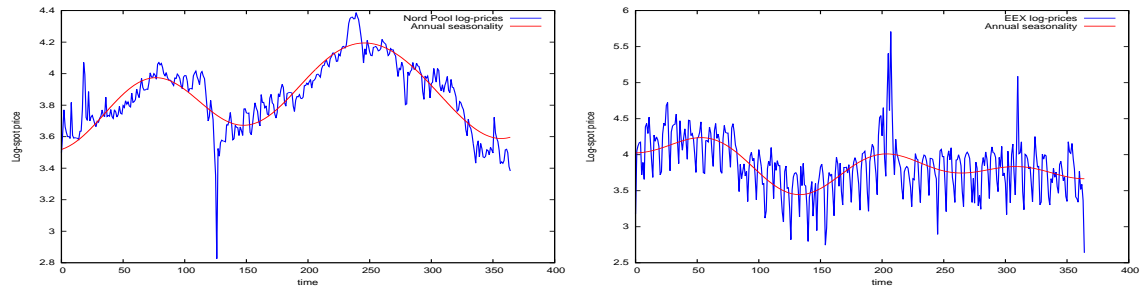


Figure 3.5: Nord Pool and EEX log-prices with the annual seasonality function.

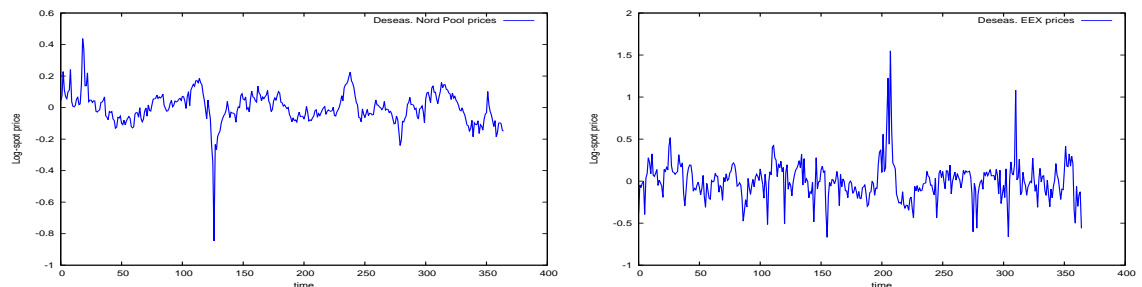


Figure 3.6: Nord Pool and EEX deseasonalized prices.

The model parameters are estimated from the deseasonalized data using Maximum Likelihood Estimation (MLE). We estimate the jump parameters separately,

using a filtering algorithm that we describe in the next section.

The calibration results are presented in Section 3.2.3, together with the model simulations.

The filtering algorithm

In order to estimate the jump parameters we apply a filtering algorithm proposed by Clewlow and Strickland [29]. The idea of this procedure is to first specify a threshold and use it to extract the jumps from the deseasonalized log-returns of the prices.

The log-returns of the one year Nord Pool data are plotted in Figure 3.7 and are computed using the following formula

$$r(t) = \ln \left(\frac{desS_t}{desS_{t-1}} \right). \quad (3.17)$$

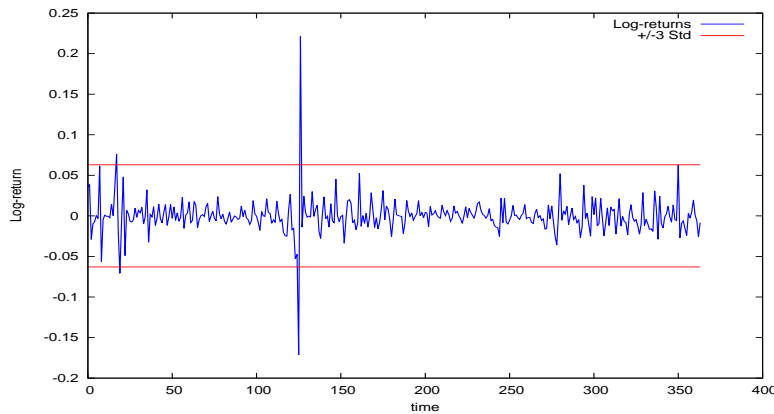


Figure 3.7: Log-returns for the daily Nord Pool data (01/01/2006 - 31/12/2006), with the highlighted level of ± 3 standard deviations.

The algorithm extracts the positive and negative jumps from the series of log-returns recursively. This procedure identifies as a jump any data which deviates in absolute value more than three standard deviations of the returns. On the second iteration, the standard deviation of the remaining series is computed again; those returns which are now greater (or smaller) than three times the standard deviation, are filtered again. This procedure is repeated until no new jumps are found.

Clewlow and Strickland [29] state that this recursive filtering ensures only the detection of high jumps and it prevents from considering usual price fluctuation as jumps.

The intensity of positive jumps is then determined by the number of positive jumps divided by the total number of observations. Similarly, the probability of positive jumps is computed as the number of positive jumps divided by the number of detected jumps. The probability of negative jumps and their intensity is found in the same way.

3.2.3 Simulation results

The last section of this chapter presents the simulation results based on historical daily average spot prices from Nord Pool and EEX, covering the period from January 2006 until December 2006.

The estimated parameters of the model introduced in (3.6) are given in Table 3.2.

| Parameter (NP) | Value | Parameter (EEX) | Value |
|----------------|--------|-----------------|--------|
| α | 80.21 | α | 238.10 |
| σ | 1.29 | σ | 4.927 |
| η_1 | 0.1114 | η_1 | 0.1082 |
| η_2 | 0.0982 | η_2 | 0.1344 |
| λ_1 | 0.0055 | λ_1 | 0.0136 |
| λ_2 | 0.0055 | λ_2 | 0.0082 |
| λ | 0.0110 | λ | 0.0218 |

Table 3.2: Estimation results using the Nord Pool and EEX data.

Figures 3.8 and 3.9 show the original prices during 2006, from Nord Pool and EEX, and the simulation results. Superimposed is the fitted seasonality function given in (3.15).

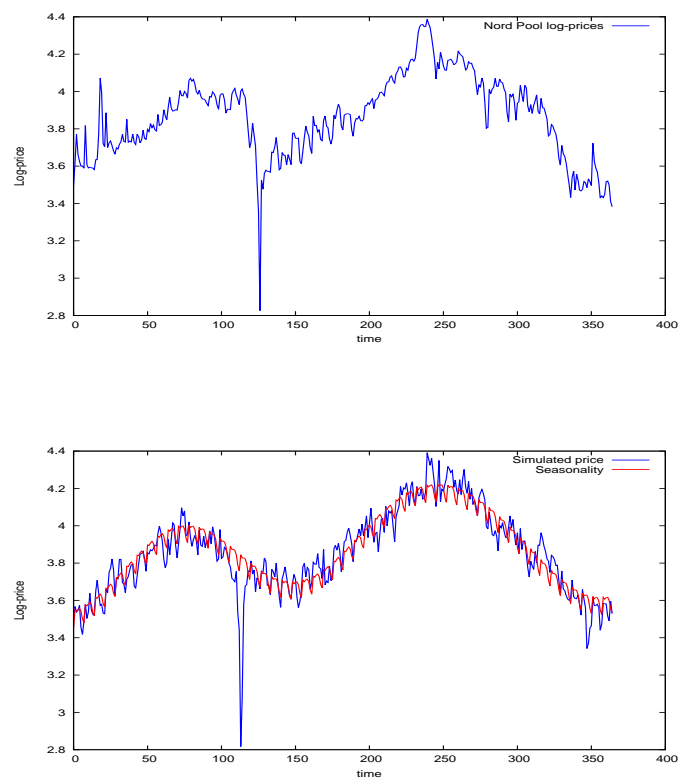


Figure 3.8: Nord Pool daily log-prices and the simulated prices with seasonality.

Another test to validate our model is to compare the first four empirical central moments of the real prices with those of the simulated data. The results are presented in Table 3.3 for the Nord Pool and in Table 3.4 for the EEX prices.

| | Nord Pool log-prices | Simulated log-prices |
|----------|----------------------|----------------------|
| Mean | 3.860 | 3.853 |
| Std. | 0.243 | 0.224 |
| Skewness | -0.277 | -0.148 |
| Kurtosis | 3.657 | 3.700 |
| Min | 2.825 | 2.816 |
| Max | 4.387 | 4.390 |

Table 3.3: Empirical moments for Nord Pool real data and from simulated data.

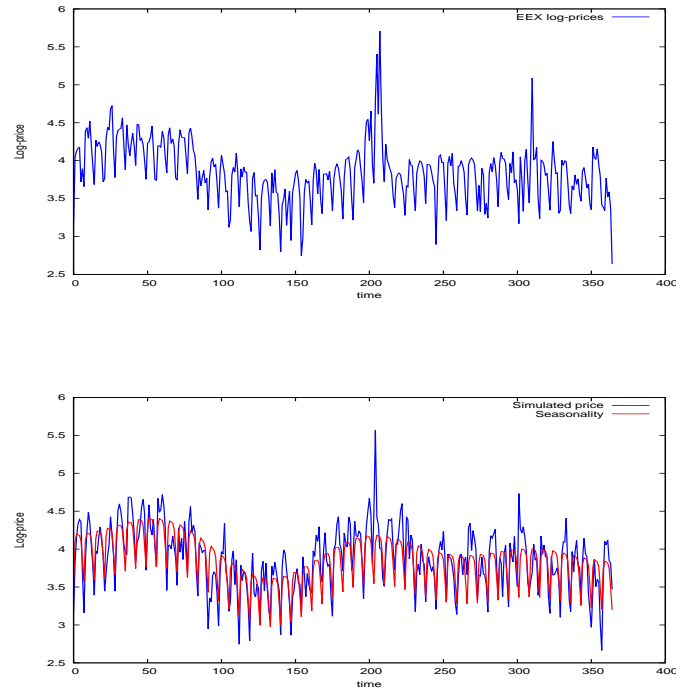


Figure 3.9: EEX daily log-prices and the simulated prices with seasonality.

| | EEX log-prices | Simulated log-prices |
|----------|----------------|----------------------|
| Mean | 3.848 | 3.866 |
| Std. | 0.386 | 0.408 |
| Skewness | 0.290 | -0.052 |
| Kurtosis | 5.006 | 3.397 |
| Min | 2.637 | 2.664 |
| Max | 5.708 | 5.563 |

Table 3.4: Empirical moments for EEX real data and from simulated data.

We can observe that the Nord Pool prices are lower than the prices from the EEX market. This can be explained by the hydro generation in the Nord Pool, which is cheaper than high efficient nuclear production used in the EEX market. The jumps are as well higher and more often in the EEX market, because in Germany, short-term regulating power is not widely available as in the hydro dominated systems. These jumps need a high mean-reversion rate in order to revert quickly to a normal level (see Table 3.2).

On the other hand, one can observe less strong seasonality in the EEX data compared to the Nord Pool prices. That is because the Nordic electricity market depends more on the hydro power plants, which are very seasonal.

Chapter 4

Pricing swing options under the double exponential jump-diffusion model

In this chapter we discuss the problem of pricing swing options in electricity markets, which are very complex derivatives that can be found in many different forms on the market.

We begin with a general characterization and then present the main challenges involved in pricing swing derivatives. In this chapter we study swing contracts which include a refraction period between two exercise times. This means that the exercise rights are not fixed beforehand. Consequently, the option holder has more flexibility during the contract period, which is very convenient in the spiky electricity markets. This formulation is to our knowledge a novelty in the literature of swing option pricing under a seasonal jump-diffusion process. A more complex type of swing options, which include also variable volume, is discussed in the next chapter.

Following the theory presented in Chapter 2 for American options, we show that the swing pricing problem can be formulated as an optimal multiple stopping problem and then as a partial integro-differential complementarity problem (PIDCP). We prove the existence and uniqueness of the viscosity solution for the PIDCP by using the mean-reverting double exponential jump-diffusion model.

We transform the partial integro-differential system into a nonlinear equation by applying the penalty method. This approach is new in the context of swing options valuation. The penalty method is able to handle nonlinearities and it has been shown to be very efficient for American pricing problems.

The resulting penalized partial integro-differential equation is discretized using the finite difference method, while the non-local integral term is approximated using a recursion formula.

We show that the monotonicity, stability and consistency conditions hold for the discrete problem. Thus, according to the results from Barles [5] and Briani et. al. [21], the finite difference scheme converges to the unique viscosity solution of the complementarity problem.

We solve the discrete penalized equation by using the generalized Newton itera-

tion, which can be proved to converge as well to the correct solution of the original PIDCP.

At the end of this chapter we present the numerical results for this approach and examine the behavior of the swing option value to changes in the model parameters.

4.1 Introduction to swing options

Due to the deregulation in the energy markets, the need for financial instruments which provide protection against extreme price fluctuations has excessively increased. As a consequence, energy exchanges have started to offer not only standard derivatives, but also so-called exotic contracts, which meet the needs for both, the producer and the consumer of the commodity. Swing options are one type of these contracts and they offer flexibility with respect to time and amount of energy delivered. As electricity is not storable, this timing flexibility is appreciated by the agents who are unable to control their electricity consumption. Consequently, these contracts represent an ideal hedging instrument.

A swing option is an agreement to buy or to sell electric energy, at a predetermined strike price (K), at different times (t_1, t_2, \dots, t_{N_s}) during the contract period ($[0, T]$). In some cases it also provides some flexibility in the quantity to be purchased or delivered: that means that a minimum (q_{min}) and a maximum (q_{max}) amount of power is specified for each swing action time, and as well for the whole contract period (Q_{min} and Q_{max}). The name swing comes from the volume flexibility, since the purchaser swings between the lower and the upper consumption boundaries. This flexibility is often further reduced by introducing penalty payments, therefore, swing options are also called take-or-pay options. These penalties are applied if the overall volume purchased during $[0, T]$ exceeds the predefined quantity in the contract.

Another characteristic of the swing options is the refraction (or recovery) time (δ_R), which separates two different exercise rights and avoids exercising all the rights at the same time.

In the case when there is no refraction time set in the contract, the problem of pricing swing options is simplified. We can see that if we consider two particular cases:

- One swing right - reduces the problem to an American option.
- Full-swing - the number of rights is equal to the number of exercise dates, then the value of the swing option is given by the value of a strip of European options expiring at the exercise dates t_1, \dots, t_{N_s} .

In this thesis we consider swing derivatives for which the exercise times are separated by a refraction period. In this way, the holder of the option has more flexibility to exercise his swing rights according to his own needs.

The pricing of such contracts rises several challenges. Firstly, one has to find an appropriate stochastic process for the electricity price. As we have showed in the

previous chapter, the double exponential jump-diffusion model with seasonality is able to reproduce the observed prices accurately. We show in this chapter, that this model can be used for pricing swing options, without introducing more complexity to the problem.

One of the main challenges in pricing swing contracts refers to the fact that the optimal decision of the option holder depends not only on the electricity price, but also on the quantity of energy already bought or sold, and on the number of remaining swing rights. Another problem comes from the fact that they have no explicit solution, so numerical methods are needed to approximate the solution.

4.2 Numerical methods for pricing swing options

Numerical methods constitute an important element in the pricing of swing derivatives, due to the fact that these options do not have an analytical solution.

In this section we give a short literature review on existing numerical methods for swing options.

Lattice method

This method was applied for the first time in the context of swing options, by Thompson [98]. Jaillet, Ronnand and Tompaidis [60] extended this approach and investigated a multi-level lattice method for swing contracts, using a mean-reverting process with seasonality for natural gas prices.

The lattice method is based on a multistage tree stochastic dynamic programming procedure. Firstly, the time domain is discretized and then the value of the option is computed backward in time, starting at maturity. At every time period, the maximum value is computed and added to the solution. When a swing right is used, the algorithm goes to a next tree, with one exercise right less.

The popularity of the lattice method in the literature comes from its simplicity and ease of implementation. However, in the case of more complicated models, the number of nodes increase exponentially, and thus, huge time and computation memory will be needed to build and solve the trees.

Monte Carlo simulation

Monte Carlo methods are widely used in the computational finance because they are conceptually very simple. Ibanez [54] proposed for the first time this technique for valuating options with multiple early exercise opportunities. He derived theoretical properties for this method and computed the optimal exercise frontier recursively to find the value of the option.

The Least Square Monte Carlo (LSMC) is a more interesting method to approximate the swing option value, and was introduced by Dörr [38]. This algorithm was developed by Longstaff and Schwartz [73] for American options, and gained an increasing interest in the swing option literature. (see among others Meinshausen

and Hambly [78], Figueroa [40], Løland and Lindqvist [72] or Bender [11]).

The algorithm uses dynamic programming, starting from the maturity date, as in the case of the lattice method. At each exercise opportunity, one has to compare the payoff from immediate exercise with the expected value from continuation (no exercise). The LSMC uses the least squares regression on a finite set of basis functions to approximate the continuation values. Discounting back and averaging these values for all paths, gives the price of the option.

The advantage of the method is that multifactor spot price models can be used, without increasing the complexity of the method. However, the drawback lies in the low accuracy and the low computation speed.

PDE-based methods

Finite difference and finite element methods are used in general to obtain numerical solutions to partial differential equations. These techniques are efficient, precisely, and can be applied to price various types of options. The difference between the two methods is only in the approximating procedure. Still, the finite difference methods have increased in popularity in the financial discipline, because they are easier to implement and give similar results as the finite element, but with less work and complexity.

One of the first approaches that introduced finite difference in the context of swing options was proposed by Wegner [104]. He approximated the option value as the solution of a partial differential equation (PDE) using a stochastic mean-reversion model with seasonality.

Dahlgren and Korn [35] proposed later a more complicated type of swing options, where different recovery times between exercise rights have been considered. Using the Black-Scholes model, he showed that the pricing problem can be transformed into a set of variational inequalities which can be then discretized using finite differences. Kjaer [64] extended the approach of Dahlgren [35] by introducing jumps and seasonality to the spot model. This formulation leads to a partial integro-differential complementarity problem (PIDCP) which is solved numerically by the finite difference method. Later in this thesis, we present more details about this method.

The finite element method was successfully applied for swing valuation, under the Black-Scholes model, in Wilhelm [106]. Her approach was based on the the Carmona and Touzi's work [23], which reduced the multiple stopping time problem to a sequence of single stopping time problems. In this way, Wilhelm [106] computed the swing option value by solving a sequence of European and American options using the finite element method. Recently, Kao and Wang [61] considered pricing of swing options under stochastic volatility, using the finite element method.

An advantage of the PDE-based methods is that the option price can be calculated for different initial spot prices, while Monte Carlo or tree methods are designed to compute the option value for only one initial spot price.

4.3 Optimal multiple stopping and swing problem formulation

In this section we give a short introduction to the optimal multiple stopping theory, following the approach proposed by Carmona and Touzi [23].

We show that the multiple stopping problem can be reduced to a sequence of optimal stopping problems which can be formulated (according to the results in Chapter 2) as a partial integro-differential complementarity problem.

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a complete probability space and $\{\mathcal{F}_t\}_{t \geq 0}$ a filtration satisfying the usual assumptions, as defined in Chapter 2. Also let $Y = (Y_t)_{t \geq 0}$ be a non-negative \mathbb{F} -adapted process satisfying (2.12) and the following conditions

- the filtration \mathbb{F} is left continuous and every \mathbb{F} -adapted martingale has continuous sample paths, (4.1)

- $\mathbb{E}[\sup_{0 \leq t \leq T} Y_t^p] < \infty$ for some $p > 1$. (4.2)

Let $N_s \geq 1$ be the total number of exercise rights given in the contract and $\delta_R > 0$ the refraction time which separates two successive exercise rights.

Denote by $\mathcal{T}_{t,T}$ the set of all \mathcal{F} -stopping times with values in $[t, T] \cup \{T_+\}$ and let $Y_{T_+} \equiv 0$. We also introduce the stopping times $\bar{\tau} = (\tau_1, \dots, \tau_{N_s})$ such that

$$\bar{\mathcal{T}}_{t,T} = \{\bar{\tau} = (\tau_1, \dots, \tau_{N_s}) \mid \tau_1 \leq T \text{ a.s.}, \tau_{k+1} - \tau_k \geq \delta_R, k = 1, \dots, N_s\}. \quad (4.3)$$

The optimal multiple stopping time problem is

$$Z = \sup_{\bar{\tau} \in \bar{\mathcal{T}}_{t,T}} \mathbb{E} \left[\sum_{i=1}^{N_s} Y_{\tau_i} \right]. \quad (4.4)$$

In other words, we need to compute the maximum expected reward Z and to find the optimal exercise strategy $\bar{\tau} = (\tau_1, \dots, \tau_{N_s})$ for which the supremum in (4.4) is attained.

We show that Z can be computed by solving N_s optimal stopping problems sequentially. We introduce the Snell envelopes

$$U^{(0)} \equiv 0 \quad \text{and} \quad U_t^{(i)} = \text{ess sup}_{\tau \in \mathcal{T}_{t,T}} \mathbb{E}[Y_\tau^{(i)} | \mathcal{F}_t], \quad (4.5)$$

where for each $i = 1, \dots, N_s$, the i -th exercise reward process $Y^{(i)}$ is given by

$$Y_t^{(i)} = \begin{cases} Y_t + \mathbb{E} \left[U_{t+\delta_R}^{(i-1)} | \mathcal{F}_t \right] & \text{for } 0 \leq t \leq T - \delta_R \\ Y_t & \text{for } t > T - \delta_R. \end{cases}$$

Following the theory from Section 2.3 we set

$$\tau_1^* = \left\{ t \geq 0 ; U_t^{(N_s)} = Y_t^{(N_s)} \right\}. \quad (4.6)$$

Observe that $\tau_1 \leq T$ and for $2 \leq i \leq N_s$, we define

$$\tau_i^* = \inf\{t \geq \delta_R + \tau_{i-1}^* ; U_t^{(N_s-i+1)} = Y_t^{(N_s-i+1)}\} \mathbb{1}_{\{\tau_{i-1}^* \leq T - \delta_R\}} + (T_+)^- \mathbb{1}_{\{\tau_{i-1}^* > T - \delta_R\}}, \quad (4.7)$$

with $\bar{\tau}^* = (\tau_1^*, \dots, \tau_{N_s}^*) \in \bar{\mathcal{T}}_{t,T}$.

Next we give the main result of this section, which shows that Z can be computed by solving inductively N_s single optimal stopping problems.

Theorem 4.3.1. *If we assume that the process Y satisfies the conditions (2.12), (4.1) and (4.2), then*

$$Z = U_0^{(N_s)} = \mathbb{E} \left[\sum_{i=1}^{N_s} Y_{\tau_i^*}^* \right], \quad (4.8)$$

where $(\tau_1^*, \dots, \tau_{N_s}^*)$ represents the optimal exercise strategy.

Proof: the proof of this theorem is based on mathematical induction and can be found in Carmona and Touzi [23] for the Black-Scholes framework and in Zeghal and Mnif [108] for a general jump-diffusion model. □

In the following we formulate the swing option problem using the optimal multiple stopping framework.

Let $v_k(s, t)$ be the value of the swing option with k exercise rights, at time t , with the starting asset value $S_t = s$ and maturity T . Then for $k = 1, \dots, N_s$, the value of the option is given by the supremum over the expected discounted payoff at each stopping time

$$v_k(s, t) = \sup_{\bar{\tau} \in \bar{\mathcal{T}}_{t,T}} \mathbb{E} \left[\sum_{k=1}^{N_s} e^{-r(\tau_k - t)} \Phi_k(S_{\tau_k}, \tau_k) \mid S_t = s \right], \quad (4.9)$$

where $\bar{\tau}$ and $\bar{\mathcal{T}}_{t,T}$ were defined in (4.3).

Using Theorem 4.3.1 the multiple-stopping time problem (4.9) can be reduced to a sequence of N_s optimal single stopping problems

$$v_k(s, t) = \sup_{\tau \in \mathcal{T}_{t,T}} \mathbb{E} \left[e^{-r(\tau - t)} \Phi_k(S_\tau, \tau) \mid S_t = s \right], \quad (4.10)$$

where the reward function is defined as¹

$$\Phi_k(s, t) = \begin{cases} \max \{ \Upsilon_k(s, t); v_k(S_{t+1}^{s,t}, t + 1) \} & \text{for } t + \delta_R \leq T \\ \phi_S(s) = \max(K - s, 0) & \text{for } t + \delta_R > T, \end{cases}$$

where $\Upsilon_k(s, t) = \phi_S(s, t) + \mathbb{E} \left[e^{-r\delta_R} v_{k-1}(S_{t+\delta_R}^{s,t}, t + \delta_R) \right]$.

¹We use the notation $S_l^{s,t}$ for S_l ($\forall l \in [0, T]$), whenever we need to emphasize the dependence of the process S on its initial condition.

That means that at time τ , the holder of the option has two choices: either a right is used and then he receives a payoff equal to $\phi_S(t)$. Then the holder owns a swing option with $k - 1$ remaining rights to exercise and the option cannot be exercised again before $\tau + \delta_R$. Alternatively, he decides not to exercise, and so he possesses an option with k opportunities left to exercise. In the view of profit maximizing agent, this choice depends on which action generates the largest value.

Assume from now on that the electricity price dynamics is given as in the previous chapter by

$$dS_t = \alpha(\rho(t) - \ln S_t)S_t dt + \sigma S_t dW_t + S_t(e^{J_t} - 1)dN_t,$$

starting at time $S_t = s$, with $\rho(t) = \frac{1}{\alpha} \left(\frac{df(t)}{dt} + \frac{\sigma^2}{2} \right) + f(t)$ and the seasonality function $f(t)$ defined in (3.15).

As in the case of the American option value (see Section 2.3), we can show that the swing option value can be characterized as a solution of the following partial integro-differential complementarity problem

$$\begin{cases} \left(\frac{\partial v_k}{\partial t} + \mathcal{D}[v_k] + \mathcal{L}[v_k] - rv_k \right) (v_k(s, t) - \Upsilon_k(s, t)) = 0 \\ \frac{\partial v_k}{\partial t} + \mathcal{D}[v_k] + \mathcal{L}[v_k] - rv_k \leq 0 \\ v_k(s, t) \geq \Upsilon_k(s, t) \\ v_k(s, T) = \phi_S(s) \end{cases} \quad (s, t) \in \mathbb{R}_+ \times [0, T] \quad (4.11)$$

where the differential and integral operators are defined as

$$\mathcal{D}[v_k] = \frac{1}{2} \sigma^2 s^2 \frac{\partial^2 v_k}{\partial s^2} + \alpha(\rho^*(t) - \ln s) s \frac{\partial v_k}{\partial s} \quad (4.12)$$

$$\mathcal{L}[v_k] = \lambda \int_{\mathbb{R}_+} [v_k(sz, t) - v_k(s, t)] f_{e^J}(z) dz. \quad (4.13)$$

Recall that the density function f_{e^J} has the following form

$$f_{e^J}(z) = p\eta_1 z^{-\eta_1-1} \mathbb{1}_{\{z \geq 1\}} + q\eta_2 z^{\eta_2-1} \mathbb{1}_{\{0 < z < 1\}} \quad (4.14)$$

and $\rho^*(t) = \rho(t) - \lambda\vartheta$ with $\vartheta \equiv \mathbb{E}[e^J - 1] = \frac{p\eta_1}{\eta_1-1} + \frac{q\eta_2}{\eta_2+1} - 1$. Also, for all $z > 0$ we have that

$$f_{e^J}(z) \geq 0 \quad \text{and} \quad \int_{\mathbb{R}_+} f_{e^J}(z) dz = 1. \quad (4.15)$$

We can write (4.11) in equivalent form

$$\begin{cases} \min \left\{ rv_k(s, t) - \frac{\partial v_k}{\partial t}(s, t) - \mathcal{D}[v_k] - \mathcal{L}[v_k], v_k(s, t) - \Upsilon_k(s, t) \right\} = 0 \\ v_k(s, T) = \phi_S(s). \end{cases} \quad (4.16)$$

Boundary conditions

Besides the terminal condition $v_k(s, T) = \phi_S(s)$ we need to provide additional boundary conditions in order to completely specify the problem (4.11).

Following Wilmott [107] and the results from Section 2.3, we impose the following initial and boundary conditions

$$v_k(s, 0) = 0 \quad (4.17)$$

$$\text{as } S \rightarrow 0 \text{ we set } \mathcal{D}[v_k] = \mathcal{L}[v_k] = 0 \quad (4.18)$$

$$\text{as } S \rightarrow +\infty \text{ we set } v_k(S, t) \rightarrow 0. \quad (4.19)$$

Moreover, for all $(s, t) \in \mathbb{R}^+ \times [0, T]$, the value of the swing option with no exercise rights is zero

$$N_s = 0 \quad \Rightarrow \quad v_k(s, t) = 0.$$

In the next section we prove that the swing value function is the unique continuous viscosity solution of the variational system of inequalities (4.16)-(4.19).

4.4 Existence and uniqueness of the solution

As the pricing inequation (4.16) is nonlinear, it is not possible to analyze the problem using the classical approach. Therefore, we present all the results in this chapter within the framework of viscosity (or weak) solutions.

In Section 2.5 we gave an introduction to viscosity solutions. This theory is very powerful because it does not require the solution to be smooth or continuous. In this context, it is also possible to formulate the PIDCP problem (4.11) for functions that are assumed to be only locally bounded. Combining these results with a strong comparison principle, one is able to prove as well the uniqueness of the viscosity solution.

To simplify the notations, let $mr(s, t) = \alpha(\rho^*(t) - \ln s)s$. We assume that there exists a constant $c > 0$ such that for all $s, y \in \mathbb{R}_+$ and $t \in [0, T]$ it holds

$$\int_{\mathbb{R}_+} |z|^2 f_{e^J}(z) dz < \infty \quad (4.20)$$

$$|mr(s, t) - mr(y, t)| \leq c|s - y| \quad (4.21)$$

$$|\Upsilon_k(s, t) - \Upsilon_k(y, t)| \leq c|s - y|. \quad (4.22)$$

Proposition 4.4.1. *For all $(s, t) \in \mathbb{R}^+ \times [0, T]$ and $k = 1, \dots, N_s$, there exists $c > 0$ such that*

$$|v_k(s, t) - v_k(y, t)| \leq c|s - y|. \quad (4.23)$$

The proof of the inequality (4.21) is trivial. We show inductively that (4.22) and (4.23) hold.

Proof: we prove these inequalities by mathematical induction on k . Using the Lipschitz property of the function $\phi_S(s, t)$ with respect to s , we get for $k = 1$

$$|\Upsilon_1(s, t) - \Upsilon_1(y, t)| = |\phi_S(s, t) - \phi_S(y, t)| \leq c|s - y|,$$

and $v_1(s, t)$ is, as stated before, the value of an American option, which is known to be Lipschitz

$$|v_1(s, t) - v_1(y, t)| \leq c|s - y|.$$

Second step: for $k = 2, \dots, N_s$ suppose that exists $c > 0$ such that Υ_{k-1} and v_{k-1} are Lipschitz in x and we prove that the properties are true for Υ_k and v_k

$$|\Upsilon_k(s, t) - \Upsilon_k(y, t)| \leq |\phi_S(s) - \phi_S(y)| + ce^{-r\delta_R} \mathbb{E} [|S_{\tau+\delta_R}^{s,\tau} - S_{\tau+\delta_R}^{y,\tau}|].$$

Since ϕ_S is Lipschitz in s , and taking in account that under the equivalent martingale measure \mathbb{Q} , the process $\{e^{-rt}S_t\}_{t \geq 0}$ is a martingale, we get that exists $c > 0$ such that

$$|\Upsilon_k(s, t) - \Upsilon_k(y, t)| \leq c|s - y|. \quad (4.24)$$

Now for v_k we have

$$\begin{aligned} |v_k(s, t) - v_k(y, t)| &\leq \left| \sup_{\tau \in \mathcal{T}_{t,T}} \mathbb{E} [\Upsilon_k(s, \tau)] - \sup_{\tau \in \mathcal{T}_{t,T}} \mathbb{E} [\Upsilon_k(y, \tau)] \right| \leq \\ &\leq \sup_{\tau \in \mathcal{T}_{t,T}} \mathbb{E} [|\Upsilon_k(s, \tau) - \Upsilon_k(y, \tau)|] \leq \\ &\leq c \sup_{\tau \in \mathcal{T}_{t,T}} \mathbb{E} [|S_{\tau}^{s,\tau} - S_{\tau}^{y,\tau}|] \leq \\ &\leq c|s - y|. \end{aligned}$$

□

Proposition 4.4.2. *For all $(s, t) \in \mathbb{R}_+ \times [0, T]$ and $k = 1, \dots, N_s$, there exists $c > 0$ such that*

$$|v_k(s, t)| \leq c(1 + |s|). \quad (4.25)$$

Proof: the linear growth condition (4.25) can be proved as above, by induction.

□

Next we focus on providing the continuity and viscosity properties of the value function using the dynamic programming principle.

Lemma 4.4.1. *Let (4.20)-(4.22) hold. For any $q \in [0, 2]$ there exists $C = C(q, c, T) > 0$ such that for all $s, y \in \mathbb{R}_+$, $t, l_1, l_2 \in [0, T]$ and $\tau \in \mathcal{T}_{l_1}$, the following inequations hold*

- $\mathbb{E}|S_{\tau}^{s,t}|^q \leq C(1 + |s|^q)$

- $\mathbb{E}|S_\tau^{s,t} - s|^q \leq C(1 + |s|^q)l_1^{\frac{q}{2}}$
- $\mathbb{E}[\sup_{0 \leq l_2 \leq l_1} |S_\tau^{s,t} - s|^q] \leq C(1 + |s|^q)l_1^{\frac{q}{2}}$
- $\mathbb{E}|S_\tau^{s,t} - S_\tau^{y,t}|^q \leq C|s - y|^2$
- $\lim_{l \rightarrow 0^+} \mathbb{E}[\sup_{h \in [t, t+l]} |S_h^{s,t} - s|^2] = 0$.

Proof: see Krylov [67] and Pham [85].

□

Theorem 4.4.1. *Dynamic Programming Principle*

For all $(s, t) \in \mathbb{R}_+ \times [0, T]$ and $\epsilon > 0$ define the stopping time

$$\tau^\epsilon = \inf \{l \in [t, T], v_k(S_l^{s,t}, l) \leq \Upsilon_k(S_l^{s,t}) + \epsilon\}. \quad (4.26)$$

Then for any stopping time $t \leq \theta \leq \tau^\epsilon$

$$v_k(s, t) = \mathbb{E}[e^{-r(\theta-t)} v_k(S_\theta^{s,t}, \theta)]. \quad (4.27)$$

Proof: see Pham [85].

□

In the proof of the existence theorem of the viscosity solution we use the following two results, which can be seen as equivalent formulations of the dynamic programming principle.

Proposition 4.4.3. For any $\theta \in \mathcal{T}_{t,T}$ and $\epsilon > 0$ we have

$$v(s, t) \geq \mathbb{E}[e^{-r(\theta-t)} v(S_\theta^{s,t}, \theta)] \quad (4.28)$$

$$v(s, t) - \epsilon \leq \mathbb{E}[e^{-r(\theta-t)} v(S_\theta^{s,t}, \theta)]. \quad (4.29)$$

Proof: see Pham [85].

□

Proposition 4.4.4. For any stopping time $\theta \in \mathcal{T}_{t,T}$ we have

$$v_k(s, t) = \sup_{\tau \in \mathcal{T}_{t,T}} \mathbb{E}[\mathbb{1}_{\{\tau < \theta\}} e^{-r(\tau-t)} \Upsilon_k(S_\tau^{s,t}) + \mathbb{1}_{\{\tau \geq \theta\}} e^{-r(\theta-t)} v_k(S_\theta^{s,t}, \theta)]. \quad (4.30)$$

Proof: see Karatzas and Shreve [62].

□

A decisive property in the proof of existence for viscosity solutions is the continuity of the value function, which is given in the next proposition.

Proposition 4.4.5. Continuity

Under the assumptions (4.20)-(4.22) the value function v_k is continuous on $\mathbb{R}_+ \times [0, T]$ and there exists $c > 0$ such that

$$|v_k(s, t) - v_k(s, l)| \leq c(1 + |s|)|t - l|^{\frac{1}{2}}, \quad (4.31)$$

for all $s, y \in \mathbb{R}_+$ and $t, l \in [0, T]$.

Proof: from (4.23) we have that $v_k(s, t)$ is Lipschitz in s , uniformly in t .

We proceed as before, by mathematical induction on k . To prove the continuity in t and the inequality (4.31) for $k = 1$, we follow the approach proposed by Pham in [83] in the case of a an optimal stopping time problem of a controlled jump-diffusion process.

Let $0 \leq t < l \leq T$ and for all $s \in \mathbb{R}_+$ we use Proposition 4.4.4. Thus, we have

$$\begin{aligned} v_1(s, t) - v_1(s, l) &= \sup_{\tau \in \mathcal{T}_{t, T}} \mathbb{E} [\mathbb{1}_{\{\tau < l\}} e^{-r(\tau-t)} \phi_S(S_\tau^{s, t}) + \mathbb{1}_{\{\tau \geq l\}} e^{-r(l-t)} v_1(S_l^{s, t}, l) - \\ &\quad - \mathbb{1}_{\{\tau \geq l\}} v_1(s, l) - \mathbb{1}_{\{\tau < l\}} v_1(s, l)] \leq \\ &\leq \sup_{\tau \in \mathcal{T}_{t, T}} \mathbb{E} [\mathbb{1}_{\{\tau < l\}} e^{-r(\tau-t)} (\phi_S(S_\tau^{s, t}) - \phi_S(s)) + \\ &\quad + \mathbb{1}_{\{\tau < l\}} \phi_S(s) - v_1(s, l) + \mathbb{1}_{\{\tau \geq l\}} e^{-r(l-t)} (v_1(S_l^{s, t}, l) - v_1(s, l)) \\ &\quad + \mathbb{1}_{\{\tau \geq l\}} (e^{-r(l-t)} - 1) v_1(s, l) + \mathbb{1}_{\{\tau < l\}} (e^{-r(\tau-t)} - 1) v_1(s, l)]. \end{aligned}$$

Using the estimates of Lemma 4.4.1 together with the following observations

- v_1 satisfies the Lipschitz continuity (4.23) and the growth condition (4.25),
- ϕ_S is Lipschitz in s ,
- $\phi_S(s) \leq v_1(s, t)$,
- $0 \leq 1 - e^{-rh} \leq r\sqrt{h}$,

we get

$$|v_1(s, t) - v_1(s, l)| \leq c(1 + |s|)|t - l|^{\frac{1}{2}}.$$

Second step: for $k = 2, \dots, N_s$, we assume that the inequality (4.31) holds for v_{k-1} . Then we apply the dynamic programming principle once again with $\theta = l$, and get

$$\begin{aligned} v_k(s, t) - v_k(s, l) &= \sup_{\tau \in \mathcal{T}_{t, T}} \mathbb{E} [\mathbb{1}_{\{\tau < l\}} e^{-r(\tau-t)} \Upsilon_k(S_\tau^{s, t}, \tau) + \mathbb{1}_{\{\tau \geq l\}} e^{-r(l-t)} v_k(S_l^{s, t}, l) - \\ &\quad - \mathbb{1}_{\{\tau \geq l\}} v_k(s, l) - \mathbb{1}_{\{\tau < l\}} v_k(s, l)] \leq \\ &\leq \sup_{\tau \in \mathcal{T}_{t, T}} \mathbb{E} [\mathbb{1}_{\{\tau < l-t\}} e^{-r(\tau-t)} (\Upsilon_k(S_\tau^{s, t}, \tau) - \Upsilon_k(s, l)) + \\ &\quad + \mathbb{1}_{\{\tau < l\}} e^{-r(\tau-t)} (\Upsilon_k(s, l) - v_k(s, l)) + \\ &\quad + \mathbb{1}_{\{\tau \geq l\}} e^{-r(l-t)} (v_k(S_l^{s, t}, l) - v_k(s, l)) + \\ &\quad + \mathbb{1}_{\{\tau \geq l\}} (e^{-r(l-t)} - 1) v_k(s, l) + \mathbb{1}_{\{\tau < l\}} (e^{-r(\tau-t)} - 1) v_k(s, l)]. \end{aligned}$$

Now using similar arguments as above for v_1 , we can show that (4.31) holds for v_k as well

$$|v_k(s, t) - v_k(s, l)| \leq c(1 + |s|)|t - l|^{\frac{1}{2}}.$$

□

For any $a \geq 0$ we define the following set of measurable functions with polynomial growth of degree a

$$C_a(\mathbb{R}_+ \times [0, T]) = \left\{ \varphi \in C^0(\mathbb{R}_+ \times [0, T]) \mid \sup_{\mathbb{R}_+ \times [0, T]} \frac{|\varphi(s, t)|}{1 + |s|^a} < +\infty \right\}. \quad (4.32)$$

Following the definition introduced in Section 2.5, we give next the definition of viscosity solution for the swing value problem.

Definition 4.4.1. *The function $v_k \in C^0(\mathbb{R}_+ \times [0, T])$ is a **viscosity subsolution (supersolution)** of (4.16) if*

$$\min\left\{r\varphi(s, t) - \frac{\partial\varphi}{\partial t}(s, t) - \mathcal{D}[\varphi] - \mathcal{L}[\varphi], \varphi(s, t) - \Upsilon_k(s, t)\right\} \leq 0 \quad (\geq 0), \quad (4.33)$$

whenever $\varphi \in C^2(\mathbb{R}_+ \times [0, T]) \cap C_2([\mathbb{R}_+ \times [0, T]])$ and $v_k - \varphi$ has a global maximum (minimum) at $(s, t) \in \mathbb{R}_+ \times [0, T]$.

The value function v_k is a **viscosity solution** of (4.16) if it is a sub- and super-solution.

Remark 4.4.1. *The condition $\varphi \in C_2(\mathbb{R}_+ \times [0, T])$ is sufficient to have a well-defined integral term in $\mathcal{L}[\varphi]$, (see Cont [31]).*

Before we present the existence theorem of the viscosity solution, we give a lemma which is useful for the proof.

Lemma 4.4.2. *Let $t, l \leq T$ and $\epsilon > 0$. Suppose that $v_k(s, t) - \Upsilon_k(s) > \epsilon$. Then*

$$\mathbb{Q}(\tau^\epsilon < l) \rightarrow 0 \quad \text{when } l \rightarrow 0,$$

where \mathbb{Q} is the risk neutral probability measure.

Proof: let $\eta > \epsilon$ such that $v_k(s, t) - \Upsilon_k(s) > \eta$ for all $k = 1, \dots, N_s$. First we show that

$$e^{-r\tau^\epsilon} v_k(S_{\tau^\epsilon}, \tau^\epsilon) - e^{-r\tau^\epsilon} \Upsilon_k(S_{\tau^\epsilon}) \leq \epsilon \quad \text{a.s. .}$$

For some sequence $t_n \downarrow \tau^\epsilon$, $e^{-rt_n} v_k(S_{t_n}, t_n) \leq e^{-rt_n} \Upsilon_k(S_{t_n}) + \epsilon$, for n large enough. From the results presented in Section 4.3 we know that S_{t_n} converges to S_{τ^ϵ} and we can find that

$$|v_k(S_{t_n}, t_n) - v_k(S_{\tau^\epsilon}, \tau^\epsilon)| \rightarrow 0 \quad \text{and} \quad \Upsilon_k(S_{t_n}) \rightarrow \Upsilon_k(S_{\tau^\epsilon}) \quad \text{a.s. when } n \rightarrow \infty.$$

Taking the limit we have

$$\begin{aligned} e^{-r\tau^\varepsilon} v_k(S_{\tau^\varepsilon}) &= \lim_{n \rightarrow \infty} e^{-rt_n} v_k(S_{t_n}, t_n) \leq \\ &\leq \lim_{n \rightarrow \infty} e^{-rt_n} \Upsilon_k(S_{t_n}) + \varepsilon = e^{-r\tau^\varepsilon} \Upsilon_k(S_{\tau^\varepsilon}) + \varepsilon \quad \text{a.s.} \end{aligned}$$

From the Proposition 4.4.5 we know that v_k is continuous and we get

$$e^{-rl} v_k(s, t+l) - e^{-rl} \Upsilon_k(s) > \eta, \quad \text{for } l \text{ small enough.}$$

Then following Roch [92] we can show that

$$\begin{aligned} \mathbb{Q}(\tau^\varepsilon < l) &\leq \mathbb{Q}(e^{-r\tau^\varepsilon} (v_k(s, \tau^\varepsilon) - \Upsilon_k(s)) + e^{-r\tau^\varepsilon} (\Upsilon_k(S_{\tau^\varepsilon}) - v_k(S_{\tau^\varepsilon}, \tau^\varepsilon)) > \eta - \varepsilon) \\ &\leq \mathbb{Q}(e^{-r\tau^\varepsilon} |v_k(s, \tau^\varepsilon) - v_k(S_{\tau^\varepsilon}, \tau^\varepsilon)| + e^{-r\tau^\varepsilon} |\Upsilon_k(S_{\tau^\varepsilon}) - \Upsilon_k(s)| > \eta - \varepsilon) \\ &\leq \mathbb{Q}(|S_{\tau^\varepsilon} - s| > c), \end{aligned}$$

for some constant $c > 0$. By the continuity in probability of the process S , we know that the last expression goes to zero when $s \rightarrow 0$.

□

Theorem 4.4.2. *Existence*

Under the assumptions (4.20)-(4.22) the value function v_k is a viscosity solution of (4.16)-(4.19).

Proof: we first prove that v_k is a supersolution of (4.16).

Let $(s, t) \in \mathbb{R}_+ \times [0, T]$ be a minimizer of $v_k - \varphi$, with $\varphi \in C^2(\mathbb{R}_+ \times [0, T]) \cap C_2(\mathbb{R}_+ \times [0, T])$ such that

$$0 = (v_k - \varphi)(s, t) = \min_{\mathbb{R}_+ \times [0, T]} (v_k - \varphi). \quad (4.34)$$

If $\theta \in \mathcal{T}_{t, T}$, then using (4.28) and (4.34) we get

$$v_k(s, t) \geq \mathbb{E}[e^{-r\theta} v_k(S_\theta^{s, t}, \theta + t)] \geq \mathbb{E}[e^{-r\theta} \varphi(S_\theta^{s, t}, \theta + t)].$$

Now applying Itô's formula (2.4) for $e^{-r\theta} \varphi(S_\theta^{s, t}, \theta + t)$ we have

$$v_k(s, t) \geq \varphi(s, t) + \mathbb{E} \left[\int_0^\theta e^{-rl} \left(-r\varphi(s, l) + \frac{\partial \varphi}{\partial l} + \mathcal{D}[\varphi(s, l)] + \mathcal{L}[\varphi(s, l)] \right) dl \right],$$

and by (4.34) we get

$$\mathbb{E} \left[\int_0^\theta e^{-rl} \left(-r\varphi(s, l) + \frac{\partial \varphi}{\partial l} + \mathcal{D}[\varphi(s, l)] + \mathcal{L}[\varphi(s, l)] \right) dl \right] \leq 0.$$

Dividing by θ , sending $\theta \rightarrow 0$ and by the mean-value theorem we obtain

$$r\varphi(s, t) - \frac{\partial \varphi}{\partial t} - \mathcal{D}[\varphi(s, t)] - \mathcal{L}[\varphi(s, t)] \geq 0.$$

Finally, taking in account that $v_k(s, t) \geq \Upsilon_k(s)$, we get the supersolution inequality.

In order to prove that v_k is a subsolution of (4.16) we use Theorem 4.4.1 and let $(s, t) \in \mathbb{R}_+ \times [0, T]$ and $\varphi \in C^2(\mathbb{R}_+ \times [0, T]) \cap C_2(\mathbb{R}_+ \times [0, T])$ such that

$$0 = (v_k - \varphi)(s, t) = \max_{\mathbb{R}_+ \times [0, T]} (v_k - \varphi). \quad (4.35)$$

As stated before, we have that $v_k(s, t) \geq \Upsilon_k(s, t)$. In the case $v_k(s, t) = \Upsilon_k(s, t)$, the inequality of subsolution is satisfied. We assume therefore that $v_k(s, t) > \Upsilon_k(s, t)$ and for $\varepsilon > 0$ let

$$v_k(s, t) - \Upsilon_k(s, t) > \varepsilon.$$

We define as previously in Theorem 4.4.1 the following stopping time

$$\tau^\varepsilon = \inf \{l \in [0, T - t], v_k(S_l^{s,t}, l) \leq \Upsilon_k(S_l^{s,t}) + \varepsilon\}.$$

For all $\theta > 0$ with $\theta_1 := \theta \wedge \tau^\varepsilon \leq \tau^\varepsilon$, we have

$$v_k(s, t) = \mathbb{E}[e^{-r\theta_1} v_k(S_{\theta_1}^{s,t}, t + \theta_1)] \quad (4.36)$$

and from (4.35) we get

$$0 \leq \mathbb{E}[e^{-r(\theta_1-t)} \varphi(S_{\theta_1}^{s,t}, t + \theta_1) - \varphi(s, t)].$$

Now applying Itô's formula (2.4) once again to $e^{-r\theta_1} \varphi(S_{\theta_1}, t + \theta_1)$ and dividing by θ we get

$$\begin{aligned} 0 &\leq \frac{1}{\theta} \mathbb{E} \left[\int_0^{\theta_1} e^{-r(l-t)} (-r\varphi(s, l) + \frac{\partial \varphi}{\partial l} + \mathcal{D}[\varphi(s, l)] + \mathcal{L}[\varphi(s, l)]) dl \right] \leq \\ &\leq \left\{ -r\varphi(s, t) + \frac{\partial \varphi}{\partial t}(s, t) + \mathcal{D}[\varphi(s, t)] + \mathcal{L}[\varphi(s, t)] \right\} \cdot \mathbb{E} \left[\frac{\theta_1}{\theta} \right]. \end{aligned}$$

Taking $\theta \rightarrow 0$ then from Lemma 4.4.2 it yields that $\mathbb{Q}[\tau^\varepsilon < \theta] \rightarrow 0$ and together with the Chebyshev's inequality² and the last inequality of Lemma 4.4.1 we get

$$rv_k(s, t) - \frac{\partial v}{\partial t}(s, t) - \mathcal{D}[v_k(s, t)] - \mathcal{L}[v_k(s, t)] \leq 0,$$

and thus, the subsolution inequality. □

Once we have proved the existence of the viscosity solution, the next step is to show that the problem (4.16)-(4.19) admits a unique viscosity solution. We do that by establishing a comparison principle.

²A general version of Chebyshev's inequality states that $\mathbb{P}(|x| \geq \varepsilon) \leq \frac{1}{\varepsilon^2} \mathbb{E}[x^2]$.

In proving the comparison result for viscosity solutions, we introduce the notion of *parabolic superjet* and *subjet* as defined in Lions [71]. Given $v_k \in C^0(\mathbb{R}_+ \times [0, T])$ and $(s, t) \in \mathbb{R}_+ \times [0, T)$, we define the parabolic superjet

$$\begin{aligned} \mathcal{P}^{2,+}v_k(s, t) = & \{(p_0, p_1, M) \in \mathbb{R} \times \mathbb{R} \times \mathbb{R} \mid v_k(y, l) - v_k(s, t) \leq p_0(y - s) + \\ & + p_1(l - t) + \frac{1}{2}M(y - s)^2 + o(|l - t| + |y - s|^2) \\ & \text{as } (y, l) \rightarrow (s, t)\}, \end{aligned} \quad (4.37)$$

its closure

$$\begin{aligned} \overline{\mathcal{P}}^{2,+}v_k(s, t) = & \left\{ (p_0, p_1, M) = \lim_{n \rightarrow \infty} (p_{0,n}, p_{1,n}, M_n) \text{ with } (p_{0,n}, p_{1,n}, M_n) \in \mathcal{P}^{2,+}v_k(s, t) \right. \\ & \left. \text{and } \lim_{n \rightarrow \infty} (s_n, t_n, v_k(s_n, t_n)) = (s, t, v_k(s, t)) \right\} \end{aligned} \quad (4.38)$$

and the parabolic subjet: $\mathcal{P}^{2,-}v_k(s, t) = -\mathcal{P}^{2,+}v_k(s, t)$,
with its closure $\overline{\mathcal{P}}^{2,-}v_k(s, t) = -\overline{\mathcal{P}}^{2,+}v_k(s, t)$.

Before introducing the comparison principle, we present two useful results for the proof. The first lemma is an equivalent formulation of the viscosity solution (see Definition 4.4.1) using the super- and subjet sets.

Lemma 4.4.3. *Let $v_k \in C_2(\mathbb{R}_+ \times [0, T])$ be a viscosity supersolution (resp. subsolution) of (4.16). Then for all $(s, t) \in \mathbb{R}_+ \times [0, T)$ and $(p_0, p_1, M) \in \overline{\mathcal{P}}^{2,-}v_k(s, t)$ (resp. $\overline{\mathcal{P}}^{2,+}v_k(s, t)$) there exists $\varphi \in C^2(\mathbb{R}_+ \times (0, T])$ such that*

$$\begin{aligned} \min \{ -p_1 + rv_k(s, t) - \mathcal{D}_{p_0}[s, t, p_0, M] - \mathcal{L}[v_k], v_k(s, t) - \Upsilon_k(s) \} & \geq 0 \\ & \text{(resp. } \leq 0) \end{aligned}$$

where

$$\mathcal{D}_{p_0}[s, t, p_0, M] = mr(s, t)p_0 + \frac{1}{2}\sigma^2s^2M \quad (4.39)$$

$$\mathcal{L}[v_k(s, t)] = \lambda \int_{\mathbb{R}_+} [v_k(sz, t) - v_k(s, t)] f_{e^J}(z) dz. \quad (4.40)$$

Proof: see Fleming and Soner [41].

□

The next theorem, known as "Theorem of sums", is a cornerstone of the theory of viscosity solutions and it is a key result in the proof of the comparison principle. We state the next theorem without proof and refer the reader to the original paper of Crandall et. al. [33].

Theorem 4.4.3. *Let \mathcal{O} be a locally compact subset of \mathbb{R} , $\underline{v}, -\bar{v} \in USC_I(\mathcal{O} \times (0, T))$ and $\varphi(s, y, t)$ a function that is twice continuous differentiable in (s, y) and once*

continuous differentiable in t . Moreover, let $(\bar{s}, \bar{y}) \in \mathcal{O} \times \mathcal{O}$ be a local maximum of the function

$$\underline{v}(s, t) - \bar{v}(y, t) - \varphi(s, y, t).$$

Then there exists $p_{1, \bar{s}}, p_{1, \bar{y}} \in \mathbb{R}$, $\kappa > 0$ and two symmetric matrices M_0 and M_1 such that

$$\begin{aligned} (D_{\bar{s}}\varphi(\bar{s}, \bar{y}, \bar{t}), p_{1, \bar{s}}, M_0) &\in \bar{\mathcal{P}}^{2,+} \underline{v}(\bar{s}, \bar{t}) \\ (-D_{\bar{y}}\varphi(\bar{s}, \bar{y}, \bar{t}), p_{1, \bar{y}}, M_1) &\in \bar{\mathcal{P}}^{2,-} \bar{v}(\bar{y}, \bar{t}), \end{aligned}$$

where $D\varphi(\bar{s}, \bar{y}, \bar{t})$ is the gradient vector, with $p_{1, \bar{s}} - p_{1, \bar{y}} = D_{\bar{t}}\varphi(\bar{s}, \bar{y}, \bar{t})$ and

$$-\left(\frac{1}{\kappa} + \|D_{\bar{s}, \bar{y}}^2\varphi(\bar{s}, \bar{y}, \bar{t})\|\right) \mathcal{I} \leq \begin{pmatrix} M_0 & 0 \\ 0 & -M_1 \end{pmatrix} \leq D_{\bar{s}, \bar{y}}^2\varphi(\bar{s}, \bar{y}, \bar{t}) + \kappa [D_{\bar{s}, \bar{y}}^2\varphi(\bar{s}, \bar{y}, \bar{t})]^2$$

where \mathcal{I} is the identity matrix and the norm of a symmetric 2×2 matrix A is defined as $\|A\| = \sup \{ | \langle A\xi, \xi \rangle | ; \xi \in \mathbb{R}^2, |\xi| \leq 1 \}$.

Theorem 4.4.4. Comparison Principle

Assume that (4.20)-(4.22) hold and let \underline{v} (resp. \bar{v}) be uniform continuous viscosity subsolution (respectively supersolution) of (4.16).

If $\underline{v}(s, T) \leq \bar{v}(s, T)$ for all $s \in \mathbb{R}_+$ then

$$\underline{v}(s, t) \leq \bar{v}(s, t), \quad \forall (s, t) \in \mathbb{R}_+ \times [0, T]. \quad (4.41)$$

Proof: firstly we can observe that it suffices to prove the inequality (4.41) for all $(s, t) \in \mathbb{R}_+ \times (0, T]$ due to the continuity of \underline{v} and \bar{v} in $t = 0$.

Following Pham [83] we introduce a function ϕ in $\mathbb{R}_+ \times (0, T]$ with $\epsilon, \alpha, \beta, \varrho > 0$

$$\phi(s, y, t) = \underline{v}(s, t) - \bar{v}(y, t) - \frac{\beta}{t} - \frac{1}{2\epsilon} |s - y|^2 - \varrho e^{\alpha(T-t)} (|s|^2 + |y|^2). \quad (4.42)$$

Since \underline{v} and $\bar{v} \in C_1(\mathbb{R}_+ \times [0, T])$, ϕ admits a maximum at $(\bar{s}, \bar{y}, \bar{t}) \in \mathbb{R}_+ \times \mathbb{R}_+ \times (0, T]$.

By classical arguments in the theory of viscosity solutions, from the continuity and Lipschitz continuity of \underline{v} and \bar{v} it can be shown that it $\exists c > 0$ such that

$$\frac{1}{\epsilon} |\bar{s} - \bar{y}|^2 \leq \omega(c\epsilon^{\frac{1}{2}}) \quad (4.43)$$

$$\varrho (|\bar{s}|^2 + |\bar{y}|^2) \leq c(1 + |\bar{s}| + |\bar{y}|) \quad (4.44)$$

$$|\bar{s}|, |\bar{y}| \leq c_\varrho, \quad (4.45)$$

where $\omega : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ is a modulus of continuity³ of \underline{v} and \bar{v} , and c_ϱ is a positive constant which depends on ϱ . Detailed proofs of these inequalities can be found in Pham [83].

³A function f admits ω as a *modulus of continuity* if and only if $|f(x) - f(y)| \leq \omega|x - y|$, i.e. ω measures the uniform continuity of functions. For example, the modulus $\omega(t) = kt$ describes the k -Lipschitz continuity, or the modulus $\omega(t) = kt^\alpha$ describes the Hölder continuity.

From (4.43)-(4.45) it follows that there exists a subsequence of $(\bar{s}, \bar{y}, \bar{t})$ which converges to $(s_0, s_0, t_0) \in \mathbb{R}_+ \times \mathbb{R}_+ \times [0, T]$ as $\epsilon \rightarrow 0^+$.

If $\bar{t} = T$ then from $\phi(s, s, t) \leq \phi(\bar{s}, \bar{y}, T)$ we get

$$\begin{aligned} \underline{v}(s, t) - \bar{v}(s, t) - \frac{\beta}{t} - 2\rho e^{\alpha(T-t)}|s|^2 &\leq \underline{v}(\bar{s}, T) - \bar{v}(\bar{s}, T) + \\ &+ \bar{v}(\bar{s}, T) - \bar{v}(\bar{y}, T) \leq \omega(|\bar{s} - \bar{y}|), \end{aligned}$$

where the second inequality follows from the uniform continuity of \bar{v} and from the assumption $\underline{v}(s, T) \leq \bar{v}(s, T)$.

By sending $\epsilon, \beta, \rho \rightarrow 0^+$ and using (4.43) we get that

$$\underline{v}(s, t) \leq \bar{v}(s, t).$$

Thus, for next part of the proof we assume that $\bar{t} < T$. Applying Theorem 4.4.3 to the function $\varphi(s, y, t)$ at the point $(\bar{s}, \bar{y}, \bar{t}) \in \mathbb{R}_+ \times \mathbb{R}_+ \times (0, T)$, there exist $p_{1, \bar{s}}, p_{1, \bar{y}}, M_0, M_1 \in \mathbb{R}$ such that

$$\begin{aligned} \left(\frac{1}{\epsilon}(\bar{s} - \bar{y}) + 2\rho e^{\alpha(T-\bar{t})}\bar{s}, p_{1, \bar{s}}, M_0 \right) &\in \bar{\mathcal{P}}^{2,+} \underline{v}(\bar{s}, \bar{t}) \\ \left(\frac{1}{\epsilon}(\bar{s} - \bar{y}) - 2\rho e^{\alpha(T-\bar{t})}\bar{y}, p_{1, \bar{y}}, M_1 \right) &\in \bar{\mathcal{P}}^{2,-} \bar{v}(\bar{y}, \bar{t}), \end{aligned}$$

where

$$p_{1, \bar{s}} - p_{1, \bar{y}} = -\frac{\beta}{\bar{t}^2} - \alpha \rho e^{\alpha(T-\bar{t})}(|\bar{s}|^2 + |\bar{y}|^2) 2\rho e^{\alpha(T-\bar{t})}, \quad (4.46)$$

and the symmetric 2×2 matrix $\begin{pmatrix} M_0 & 0 \\ 0 & -M_1 \end{pmatrix}$ satisfies

$$\begin{pmatrix} M_0 & 0 \\ 0 & -M_1 \end{pmatrix} \leq \frac{1}{\kappa} + \|D_{\bar{s}, \bar{y}}^2 \varphi(\bar{s}, \bar{y}, \bar{t})\| + D_{\bar{s}, \bar{y}}^2 \varphi(\bar{s}, \bar{y}, \bar{t}) + \kappa [D_{\bar{s}, \bar{y}}^2 \varphi(\bar{s}, \bar{y}, \bar{t})]^2, \quad (4.47)$$

where $\varphi(s, y, t) = -\frac{\beta}{t} - \frac{1}{2\epsilon}|s - y|^2 - \rho e^{\alpha(T-t)}(|s|^2 + |y|^2)$.

Thus, $D_{\bar{s}, \bar{y}}^2 \varphi(\bar{s}, \bar{y}, \bar{t}) = \frac{1}{\epsilon} \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} + 2\rho e^{\alpha(T-\bar{t})} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ and then we get that $\|D_{\bar{s}, \bar{y}}^2 \varphi(\bar{s}, \bar{y}, \bar{t})\| \leq \frac{2}{\epsilon} + 2\rho e^{\alpha(T-\bar{t})}$.

Thus, equation (4.47) becomes

$$\begin{aligned} \begin{pmatrix} M_0 & 0 \\ 0 & -M_1 \end{pmatrix} &\leq \left(\frac{1}{\epsilon} + \frac{\kappa}{\epsilon^2} + \frac{4\kappa\rho e^{\alpha(T-\bar{t})}}{\epsilon} \right) \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} + \\ &+ \left(\frac{1}{\kappa} + \frac{2}{\epsilon} + 4\rho e^{\alpha(T-\bar{t})} + 4\kappa\rho^2 e^{2\alpha(T-\bar{t})} \right) \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}. \end{aligned}$$

Without loss of generality consider $\kappa = \epsilon$ and we get

$$\begin{aligned} \begin{pmatrix} M_0 & 0 \\ 0 & -M_1 \end{pmatrix} &\leq \left(\frac{3}{\epsilon} + 4\rho e^{\alpha(T-\bar{t})} \right) \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} + \\ &+ \left(\frac{3}{\epsilon} + 4\rho e^{\alpha(T-\bar{t})} + 4\epsilon\rho^2 e^{2\alpha(T-\bar{t})} \right) \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}. \end{aligned} \quad (4.48)$$

Using Lemma 4.4.3 and the fact that \underline{v} and \bar{v} are viscosity solutions, it yields that

$$\min \left\{ -p_{1,\bar{s}} + r\underline{v}(\bar{s}, \bar{t}) - \mathcal{D}_{p_0} \left[\bar{s}, \bar{t}, \frac{1}{\epsilon}(\bar{s} - \bar{y}) + 2\rho e^{\alpha(T-\bar{t})}\bar{s}, M_0 \right] - \mathcal{L} [\underline{v}(\bar{s}, \bar{t})]; \right. \\ \left. \underline{v}(\bar{s}, \bar{t}) - \Upsilon_k(\bar{s}) \right\} \leq 0$$

and

$$\min \left\{ -p_{1,\bar{y}} + r\bar{v}(\bar{y}, \bar{t}) - \mathcal{D}_{p_0} \left[\bar{y}, \bar{t}, \frac{1}{\epsilon}(\bar{s} - \bar{y}) - 2\rho e^{\alpha(T-\bar{t})}\bar{y}, M_1 \right] - \mathcal{L} [\bar{v}(\bar{y}, \bar{t})]; \right. \\ \left. \bar{v}(\bar{y}, \bar{t}) - \Upsilon_k(\bar{y}) \right\} \geq 0$$

are satisfied, where \mathcal{D}_{p_0} and \mathcal{L} were defined in Lemma 4.4.3.

Subtracting these two inequalities and remarking that $\min(a, b) - \min(c, d) \leq 0$ implies either $a - c \leq 0$ or $b - d \leq 0$.

Thus, for $b - d$ we get

$$b - d = \underline{v}(\bar{s}, \bar{t}) - \Upsilon_k(\bar{s}) - \bar{v}(\bar{y}, \bar{t}) + \Upsilon_k(\bar{y})$$

and from the Lipschitz continuity of the function Υ_k (see (4.22)) together with (4.43), we have

$$\limsup_{\epsilon \rightarrow 0^+} (\underline{v}(\bar{s}, \bar{t}) - \bar{v}(\bar{y}, \bar{t})) \leq 0.$$

Consequently, we get that

$$\underline{v}(s, t) \leq \bar{v}(s, t).$$

Now using (4.46), for $a - c$ we have

$$r(\underline{v}(\bar{s}, \bar{t}) - \bar{v}(\bar{y}, \bar{t})) + \frac{\beta}{\bar{t}^2} + \alpha \rho e^{\alpha(T-\bar{t})} (|\bar{s}|^2 + |\bar{y}|^2) \leq \\ \leq \mathcal{D}_{p_0} \left[\bar{s}, \bar{t}, \frac{1}{\epsilon}(\bar{s} - \bar{y}) + 2\rho e^{\alpha(T-\bar{t})}\bar{s}, M_0 \right] - \mathcal{D}_{p_0} \left[\bar{y}, \bar{t}, \frac{1}{\epsilon}(\bar{s} - \bar{y}) - 2\rho e^{\alpha(T-\bar{t})}\bar{y}, M_1 \right] + \\ + \mathcal{L} [\underline{v}(\bar{s}, \bar{t})] - \mathcal{L} [\bar{v}(\bar{y}, \bar{t})]. \quad (4.49)$$

The first difference on the right hand side can be approximated as follows

$$A := \mathcal{D}_{p_0} \left[\bar{s}, \bar{t}, \frac{1}{\epsilon}(\bar{s} - \bar{y}) + 2\rho e^{\alpha(T-\bar{t})}\bar{s}, M_0 \right] - \mathcal{D}_{p_0} \left[\bar{y}, \bar{t}, \frac{1}{\epsilon}(\bar{s} - \bar{y}) - 2\rho e^{\alpha(T-\bar{t})}\bar{y}, M_1 \right] = \\ = \frac{1}{\epsilon}(\bar{s} - \bar{y}) [mr(\bar{s}, \bar{t}) - mr(\bar{y}, \bar{t})] + 2\rho e^{\alpha(T-\bar{t})} [mr(\bar{s}, \bar{t})\bar{s} + mr(\bar{y}, \bar{t})\bar{y}] + \\ + \frac{1}{2}\sigma^2(\bar{s}^2 M_0 - \bar{y}^2 M_1). \quad (4.50)$$

Note that from the continuity of $mr(\cdot, t)$ and (4.21), we get that the linear growth condition holds, and thus we get for $c > 0$

$$A \leq c \left[\frac{1}{\epsilon}|\bar{s} - \bar{y}|^2 + 2\rho e^{\alpha(T-\bar{t})}(1 + |\bar{s}|^2 + |\bar{y}|^2) \right] + \frac{1}{2}\sigma^2(\bar{s}^2 M_0 - \bar{y}^2 M_1), \quad (4.51)$$

Moreover, we observe that

$$\text{tr} \left[\begin{pmatrix} M_0 & 0 \\ 0 & -M_1 \end{pmatrix} \begin{pmatrix} \bar{s} \\ \bar{y} \end{pmatrix} \otimes \begin{pmatrix} \bar{s} \\ \bar{y} \end{pmatrix}^\top \right] = \bar{s}^2 M_0 - \bar{y}^2 M_1,$$

where tr is the trace⁴ of a matrix. And together with (4.48), we get in (4.51) that

$$\begin{aligned} A \leq & c \left[\frac{1}{\epsilon} |\bar{s} - \bar{y}|^2 + 2\rho e^{\alpha(T-\bar{t})} (1 + |\bar{s}|^2 + |\bar{y}|^2) \right] + \frac{1}{2} \sigma^2 \left[\left(\frac{3}{\epsilon} + 4\rho e^{\alpha(T-\bar{t})} \right) (\bar{s} - \bar{y})^2 + \right. \\ & \left. + \left(\frac{3}{\epsilon} + 4\rho e^{\alpha(T-\bar{t})} + 4\epsilon \rho^2 e^{2\alpha(T-\bar{t})} \right) (1 + \bar{s}^2 + \bar{y}^2) \right]. \end{aligned} \quad (4.52)$$

Now let

$$\begin{aligned} B & := \mathcal{L} [\underline{v}(\bar{s}, \bar{t})] - \mathcal{L} [\bar{v}(\bar{y}, \bar{t})] = \\ & = \int_{\mathbb{R}_+} [\underline{v}(\bar{s}z, \bar{t}) - \underline{v}(\bar{s}, \bar{t}) - \bar{v}(\bar{y}z, \bar{t}) + \bar{v}(\bar{y}, \bar{t})] f_{e^J}(z) dz. \end{aligned} \quad (4.53)$$

In order to estimate differences of the integro-differential term (4.53) we use the fact that $(\bar{s}, \bar{t}, \bar{y})$ is a global maximum of ϕ and from the inequality $\phi(\bar{s}z, \bar{y}z, \bar{t}) - \phi(\bar{s}, \bar{y}, \bar{t}) \leq 0$ we deduce

$$\begin{aligned} \phi(\bar{s}z, \bar{y}z, \bar{t}) - \phi(\bar{s}, \bar{y}, \bar{t}) & = \underline{v}(\bar{s}z, \bar{t}) - \bar{v}(\bar{y}z, \bar{t}) - \frac{\beta}{\bar{t}} - \frac{1}{2\epsilon} |\bar{s}z - \bar{y}z|^2 - \\ & - \rho e^{\alpha(T-\bar{t})} (|\bar{s}z|^2 + |\bar{y}z|^2) - \underline{v}(\bar{s}, \bar{t}) + \bar{v}(\bar{y}, \bar{t}) + \frac{\beta}{\bar{t}} + \\ & + \frac{1}{2\epsilon} |\bar{s} - \bar{y}|^2 + \rho e^{\alpha(T-\bar{t})} (|\bar{s}|^2 + |\bar{y}|^2). \end{aligned}$$

Performing the computations and rearranging terms we get

$$\begin{aligned} \underline{v}(\bar{s}z, \bar{t}) - \underline{v}(\bar{s}, \bar{t}) - \bar{v}(\bar{y}z, \bar{t}) + \bar{v}(\bar{y}, \bar{t}) & \leq (|z|^2 - 1) \left[\frac{1}{2\epsilon} |\bar{s} - \bar{y}|^2 + \right. \\ & \left. + \rho e^{\alpha(T-\bar{t})} (1 + |\bar{s}|^2 + |\bar{y}|^2) \right] \leq |z|^2 \left[\frac{1}{2\epsilon} |\bar{s} - \bar{y}|^2 + \rho e^{\alpha(T-\bar{t})} (1 + |\bar{s}|^2 + |\bar{y}|^2) \right], \end{aligned}$$

which together with the assumption (4.20) it gives us that

$$B \leq c \left[\frac{1}{2\epsilon} |\bar{s} - \bar{y}|^2 + \rho e^{\alpha(T-\bar{t})} (1 + |\bar{s}|^2 + |\bar{y}|^2) \right]. \quad (4.54)$$

Since $(\bar{s}, \bar{y}, \bar{t})$ is a maximum point of ϕ we have that $\phi(s, s, t) \leq \phi(\bar{s}, \bar{y}, \bar{t})$, i.e.

$$\begin{aligned} \underline{v}(s, t) - \bar{v}(s, t) - \frac{\beta}{t} - 2\rho e^{\alpha(T-t)} |s|^2 & \leq \underline{v}(\bar{s}, \bar{t}) - \bar{v}(\bar{y}, \bar{t}) - \frac{\beta}{\bar{t}} - \\ & - \frac{1}{2\epsilon} |\bar{s} - \bar{y}|^2 - \rho e^{\alpha(T-t)} (|\bar{s}|^2 + |\bar{y}|^2) \end{aligned} \quad (4.55)$$

⁴The trace of a square matrix A is the sum of its diagonal elements.

and from (4.49) we have that

$$\begin{aligned} \underline{v}(\bar{s}, \bar{t}) - \bar{v}(\bar{y}, \bar{t}) - \frac{\beta}{\bar{t}} - \frac{1}{2\epsilon} |\bar{s} - \bar{y}|^2 - \rho e^{\alpha(T-t)} (|\bar{s}|^2 + |\bar{y}|^2) &\leq \\ &\leq \frac{1}{r} [A + B - \alpha \rho e^{\alpha(T-t)} (|\bar{s}|^2 + |\bar{y}|^2)]. \end{aligned} \quad (4.56)$$

Now from (4.55) and (4.56) we get

$$\underline{v}(s, t) - \bar{v}(s, t) - \frac{\beta}{t} - 2\rho e^{\alpha(T-t)} |s|^2 \leq \frac{1}{r} [A + B - \alpha \rho e^{\alpha(T-t)} (|\bar{s}|^2 + |\bar{y}|^2)].$$

Sending $\epsilon \rightarrow 0^+$ and using the estimates from (4.52) and (4.54) we obtain

$$\underline{v}(s, t) - \bar{v}(s, t) - \frac{\beta}{t} - 2\rho e^{\alpha(T-t)} |s|^2 \leq \frac{2\rho e^{\alpha(T-t_0)}}{r} [c(1 + 2|s_0|^2) - \alpha|s_0|^2].$$

Choosing α sufficiently large and sending $\beta, \rho \rightarrow 0^+$ we conclude that

$$\underline{v}(s, t) \leq \bar{v}(s, t).$$

which completes the proof. □

Next we give the main result of this section, the uniqueness theorem.

Theorem 4.4.5. Uniqueness

The value function v is the unique viscosity solution of (4.16)-(4.19).

Proof: let v_1 and v_2 be two viscosity solutions of (4.16). Then applying (4.41) from the comparison theorem we have that $v_1 = v_2$ on $\mathbb{R}_+ \times [0, T]$ and the uniqueness of the solution is proved. □

In this section we have proved the existence and uniqueness of the viscosity solution for the HJB equation associated with the swing pricing problem. In the next part we compute this solution numerically.

4.5 Numerical calculations

We have showed before that in order to price the swing option we have to solve the system of variational inequalities (4.11), computing the following reward payoff function at each time-step

$$\Phi_k(s, t) = \begin{cases} \max \{ \Upsilon_k(s, t); v_k(S_{t+1}^{s,t}, t+1) \} & \text{for } t + \delta_R \leq T \\ \phi_S(t) = \max(K - s, 0) & \text{for } t + \delta_R > T \end{cases}$$

where $\Upsilon_k(s, t) = \phi_S(s, t) + \mathbb{E} [e^{-r\delta_R} v_{k-1}(S_{t+\delta_R}^{s,t}, t + \delta_R)]$.

Thus, we define the following function for $t + \delta_R \leq T$

$$u_{k-1}(s, t) = \mathbb{E} [e^{-r\delta_R} v_{k-1}(S_{t+\delta_R}^{s,t}, t + \delta_R)].$$

This function represents the swing option value with one exercise right less, after the refracting period δ_R . We can compute $u_{k-1}(s, t)$ by letting

$$u_k(S, t) = w(S, 0), \tag{4.57}$$

where $w(S, t)$ is the value of a European option with expiration date δ_R . Then, by Theorem 2.4.3, $w(S, t)$ is the solution to the following PIDE

$$\begin{cases} \frac{\partial w}{\partial t} + \mathcal{D}[w] + \mathcal{L}[w] - rw = 0, & (s, t) \in \mathbb{R}_+ \times [0, \delta_R[\\ w(s, \delta_R) = v_k(s, t + \delta_R). \end{cases} \tag{4.58}$$

In this way, for $t + \delta_R \leq T$ a numerical algorithm calculates inductively the price of the swing option

1. w and u_k are obtained by computing (4.57)-(4.58),
2. then the PIDCP (4.11) is solved for v_k using the following payoff

$$\Upsilon_k(s, t) = \phi_S(s, t) + u_{k-1}(s, t). \tag{4.59}$$

Numerically, in the first step, the problem (4.58) is discretized by finite difference and for the integral term $\mathcal{L}[w]$ we are able to establish a fast recursion formula. This two methodologies are presented in the next sections. Thus, the discretization leads to a system which is solved by an iterative method based on regular splitting of the coefficient matrix. This approach has been studied in d'Halluin, Forsyth and Vetzal [49] for example, so we skip the details here.

In the second step, we apply the finite difference scheme combined with the penalty method, to enforce the early exercise constraint. The resulting system of nonlinear algebraic equations is solved iteratively.

In the next part of the thesis we take a more detailed look on this approach.

4.5.1 The penalty method

A variety of algorithms have been proposed in the literature to solve the linear complementarity problems, including the operator splitting, projective over relaxation or the penalty method. We have to take in account though, that not all of them can be applied to models with jumps.

In this thesis we apply the penalty method to enforce the early exercise constraint in the PIDCP (4.11). This is a very powerful technique which can be used for any type of discretization, in any dimensions, on non-uniform meshes or with nonlinearities. This method has been successfully applied before to price American options in d'Halluin et. al. [48], [49], Forsyth and Vetzal [42] or Reisinger [91]. To our knowledge, the penalty method has never been used in the context of swing option valuation.

The main idea is to replace the system (4.11) by a nonlinear PIDE which includes a penalty term. This term prevents the value of the option from falling below the payoff. Thus, the penalized equation yields

$$\frac{\partial v_k}{\partial t} + \mathcal{D}[v_k] + \lambda \int_{\mathbb{R}_+} v_k(sz, t) f_{e^J}(z) dz - (r + \lambda)v_k + \frac{1}{\epsilon} P(v_k, \Upsilon_k) = 0, \quad (4.60)$$

where P is the penalty term and $\epsilon > 0$ is a positive parameter. P is constructed such that

- If $P \geq 0$ then $\frac{\partial v_k}{\partial t} + \mathcal{D}[v_k] + \lambda \int_{\mathbb{R}_+} v_k(sz, t) f_{e^J}(z) dz - (r + \lambda)v_k \leq 0$.
- If $P(v_k, \Upsilon_k) = 0$ then $\frac{\partial v_k}{\partial t} + \mathcal{D}[v_k] + \lambda \int_{\mathbb{R}_+} v_k(sz, t) f_{e^J}(z) dz - (r + \lambda)v_k = 0$, for $v_k > \Upsilon_k$.
- If $P(v_k, \Upsilon_k) > 0$ for $v_k < \Upsilon_k$ then $\max(\Upsilon_k - v_k, 0) \rightarrow 0$ for $\epsilon \rightarrow 0$.

We discretize the nonlinear problem (4.60) by finite differences and solve the resulting matrix system iteratively.

Next we give a short introduction to the finite difference method. For a detailed presentation of this method we refer to Wilmott [107].

The finite difference method

The finite difference method it is a very powerful and flexible technique, which is able to generate accurate numerical solutions to partial differential equations. As stated at the beginning of this chapter, one main advantage of this technique is its simplicity and thus, it is often used for more complex models and derivatives.

The idea behind this method is to replace the derivatives occurring in the partial differential equations by approximations based on Taylor series expansions. In the following we present the main steps of the finite difference method:

- The infinite domain is reduced to a bounded domain:
 $[0, \infty) \times [0, T] \rightarrow [S_{min}, S_{max}] \times [0, T]$.

- The domain is discretized and an equidistant grid is generated: the S -axis and t -axis are divided into equidistant parts of length ΔS and Δt . In this way the (S, t) plan is split into a dense mesh, with grid points $(i\Delta S, j\Delta t) = (S_i, t_j)$ with $i = 1, \dots, I$, $j = 1, \dots, J$.
- The derivatives are substituted with finite difference approximations at every grid point where the solution is unknown, and thus a system of algebraic equations is formed.
- Convergence analysis is performed to guarantee a good approximation.
- The system of equations (together with the boundary conditions) is solved iteratively and the approximated solution is obtained at each grid point.

The discretized value⁵ of the swing option at time step t_j is denoted by

$$v_i^j = v(S_i, t_j) = v(i\Delta S, j\Delta t). \quad (4.61)$$

There are three commonly used types of finite difference approximations

- $\frac{\partial v}{\partial S}(S, t) = \frac{v_i - v_{i-1}}{\Delta S} + \mathcal{O}(\Delta S)$ forward approximation.
- $\frac{\partial v}{\partial S}(S, t) = \frac{v_{i+1} - v_i}{\Delta S} + \mathcal{O}(\Delta S)$ backward approximation.
- $\frac{\partial v}{\partial S}(S, t) = \frac{v_{i+1} - v_{i-1}}{2(\Delta S)} + \mathcal{O}((\Delta S)^2)$ central approximation.

The second order partial derivatives can be estimated by the symmetric central difference approximation in the following way

$$\frac{\partial^2 v}{\partial S^2}(S, t) = \frac{v_{i+1} - 2v_i + v_{i-1}}{(\Delta S)^2} + \mathcal{O}((\Delta S)^2).$$

The forward, backward and central finite difference approximations lead to *explicit*, *fully implicit* and *Crank-Nicholson* schemes, respectively. For more details about these methods, we refer to Wilmott [107].

Next, we present the discretization of the equation (4.60), neglecting the integral part for the moment. This term is approximated in a later section.

Discretization

We approximate the derivatives in equation (4.60) by using the θ -method, which can be considered as a weighted average of the explicit and implicit schemes. Thus,

⁵From now on we leave away the subscript k from the value function v_k , not to cause any confusion with the discretization indices.

for $i = 2, \dots, I - 1$ and $j = 2, \dots, J - 1$ we get

$$\begin{aligned} & \frac{v_i^{j+1} - v_i^j}{\Delta t} + \theta \left[\frac{\sigma^2 S_i^2}{2} \frac{v_{i+1}^j - 2v_i^j + v_{i-1}^j}{(\Delta S)^2} + \alpha(t_j) S_i \frac{v_{i+1}^j - v_{i-1}^j}{2\Delta S} - (r + \lambda)v_i^j \right] + \\ & + (1 - \theta) \left[\frac{\sigma^2 S_i^2}{2} \frac{v_{i+1}^{j+1} - 2v_i^{j+1} + v_{i-1}^{j+1}}{(\Delta S)^2} + \alpha(t_{j+1}) S_i \frac{v_{i+1}^{j+1} - v_{i-1}^{j+1}}{2\Delta S} - (r + \lambda)v_i^{j+1} \right] + \\ & + \frac{1}{\epsilon} \max(\Upsilon_i - v_i^j, 0) = 0, \end{aligned} \quad (4.62)$$

where Υ_i is the vector of payoffs obtained by the discretization. Moreover, $\theta \in [0, 1]$ is a weight parameter which gives the following approximations: for $\theta = 0$ we recover the explicit scheme, for $\theta = 1$ the fully implicit scheme and for $\theta = \frac{1}{2}$ we get the Crank-Nicolson scheme.

Explicit schemes are typically simple to implement, but suffer from stability issues. Implicit methods are unconditionally stable, but exhibit only linear convergence. Polley et. al. [88] stated that it is advantageous to use Crank-Nicolson time stepping to achieve quadratic convergence. However, this method can lead to oscillations in the numerical solution.

In order to avoid these oscillations, Rannacher [90] introduced a scheme in which the fully implicit time stepping is used for the first four steps and then Crank-Nicolson scheme for the remaining time steps. In this way, high frequency error components are dampened by the implicit steps, leading to smooth convergence. The expected convergence rate remains quadratic since only a finite number of implicit steps are taken. Experimental computations in Polley [87] and Shin [95] also indicate that Rannacher stepping technique improves the stability of the numerical scheme.

In the Rannacher time stepping scheme, θ has the following form

$$\theta = \begin{cases} 1 & \text{for } j = J - 1, J - 2, J - 3 \\ \frac{1}{2} & \text{for } j = J - 4, \dots, 2. \end{cases}$$

Rearranging in (4.62) we get

$$\begin{aligned} & v_i^j [1 + \theta \Delta t (a_i + b_i + r + \lambda)] - v_{i+1}^j \theta \Delta t a_i - v_{i-1}^j \theta \Delta t b_i - \frac{\Delta t}{\epsilon} \max(\Upsilon_i - v_i^j, 0) = \\ & = v_i^{j+1} [1 - (1 - \theta) \Delta t (a_i + b_i + r + \lambda)] + v_{i+1}^{j+1} (1 - \theta) \Delta t a_i + v_{i-1}^{j+1} (1 - \theta) \Delta t b_i. \end{aligned} \quad (4.63)$$

At the missing points $i = \{1, I\}$ we impose Dirichlet boundary conditions. The terms a_i and b_i are determined by choosing between the following discretizations

- Central differences

$$\begin{aligned} a_{i,c} &= \frac{\sigma^2 S_i^2}{2(\Delta S)^2} + \frac{\alpha(\rho^*(t_j) - \ln S_i) S_i}{2\Delta S}, \\ b_{i,c} &= \frac{\sigma^2 S_i^2}{2(\Delta S)^2} - \frac{\alpha(\rho^*(t_j) - \ln S_i) S_i}{2\Delta S}. \end{aligned} \quad (4.64)$$

- Forward differences

$$\begin{aligned} a_{i,f} &= \frac{\sigma^2 S_i^2}{2(\Delta S)^2} + \frac{\alpha(\rho^*(t_j) - \ln S_i)S_i}{\Delta S}, \\ b_{i,f} &= \frac{\sigma^2 S_i^2}{2(\Delta S)^2}. \end{aligned} \quad (4.65)$$

- Backward differences:

$$\begin{aligned} a_{i,b} &= \frac{\sigma^2 S_i^2}{2(\Delta S)^2}, \\ b_{i,b} &= \frac{\sigma^2 S_i^2}{2(\Delta S)^2} - \frac{\alpha(\rho^*(t_j) - \ln S_i)S_i}{\Delta S}. \end{aligned} \quad (4.66)$$

We choose between central, forward, respectively backward differences at each node by using the algorithm given in the Appendix A.1. This algorithm ensures that a_i and b_i satisfy the following *positive coefficient condition*

$$a_i, b_i \geq 0 \quad \text{for all } i = 2, \dots, I-1, \quad j = 2, \dots, J-1. \quad (4.67)$$

This condition avoids oscillatory solutions and is an important tool in proving the convergence of the penalty iteration, as we show later.

The discretization of the diffusion terms in (4.63) leads to a tridiagonal matrix⁶. Solving the resulting discretized system with such a matrix is very efficient and the programming time is extremely low.

On the other hand, in the case of the integral term (4.13) this discretization leads to a full matrix, which is computationally more expensive. However, one main advantage of the electricity spot model proposed in Section 3.2 is that the jumps are modeled by a double exponential distribution. In this case, it is possible to approximate the integral term by a recursion formula, similar to the one presented by Toivanen [99] using the Kou model for pricing American options. Recently, Griebel and Hullmann [45] derived such a formula for the valuation of European basket options, by applying the Galerkin method.

In the next section we approximate the integral term and derive the recursion formula for our model.

4.5.2 The recursion formula for the integral term

We first observe that by using condition (4.15) we can rewrite the integral term in (4.13) as

$$\lambda \int_{\mathbb{R}^+} v(Sz, t) f_{e^J}(z) dz. \quad (4.68)$$

The discretization of this term leads to a $I \times I$ matrix which we denote by Jump.

⁶A tridiagonal matrix is a matrix whose elements are zero except for those on and immediately above and below the leading diagonal.

Using the formulation (3.10) for the density function, we can split the integral (4.68) in two integrals J_1 and J_2 . Then by changing the variable $y = Sz$ we obtain

$$J_1 = \frac{1}{S} \int_S^\infty v(y, t) f_{e^J}(y/S) dy = p\eta_1 S^{\eta_1} \int_S^\infty v(y, t) y^{-\eta_1-1} dy$$

and

$$J_2 = \frac{1}{S} \int_0^S v(y, t) f_{e^J}(y/S) dy = q\eta_2 S^{-\eta_2} \int_0^S v(y, t) y^{\eta_2-1} dy.$$

In the following we consider the approximation of the integral J_1 , while J_2 can be treated in the same way.

At each grid point S_m ($m = 2, \dots, I-1$) we need to approximate

$$(J_1)_m = p\eta_1 S_m^{\eta_1} \int_{S_n}^\infty v(y, t) y^{-\eta_1-1} dy = \sum_{n=0}^{m-1} (J_1)_{m,n},$$

where $(J_1)_{m,n} = p\eta_1 S_m^{\eta_1} \int_{S_n}^{S_{n+1}} v(y, t) y^{-\eta_1-1} dy$.

Linear interpolation for v between the grid points leads to

$$(J_1)_m \approx (M_1)_m = \sum_{n=0}^{m-1} (M_1)_{m,n} \quad (4.69)$$

with

$$(M_1)_{m,n} = p\eta_1 S_m^{\eta_1} \int_{S_n}^{S_{n+1}} \left(\frac{S_{n+1} - y}{S_{n+1} - S_n} v(S_n, t) + \frac{y - S_n}{S_{n+1} - S_n} v(S_{n+1}, t) \right) y^{-\eta_1-1} dy. \quad (4.70)$$

By performing the integration we obtain

$$(M_1)_{m,n} = \frac{pS_m^{\eta_1}}{(\eta_1 - 1)(S_{n+1} - S_n)} \{ [S_{n+1}^{1-\eta_1} - (S_{n+1} - \eta_1(S_{n+1} - S_n))S_n^{-\eta_1}] v(S_n, t) + [S_n^{1-\eta_1} - (S_n + \eta_1(S_{n+1} - S_n))S_{n+1}^{-\eta_1}] v(S_{n+1}, t) \}. \quad (4.71)$$

Following the approach proposed by Toivanen [99] we can now derive a fast and easy to implement recursion formula to determine $(M_1)_m$.

The observation that $\frac{(M_1)_{m-1,n}}{(M_1)_{m,n}} = \left(\frac{S_{m-1}}{S_m}\right)^{\eta_1}$, together with the equation (4.69), leads to the following recursion formula

$$(M_1)_{m-1} = \left(\frac{S_{m-1}}{S_m}\right)^{\eta_1} (M_1)_m + (M_1)_{m-1,m-1} \quad m = 2, \dots, I-1. \quad (4.72)$$

Similarly, we obtain a recursion formula for $(M_2)_m$

$$(M_2)_{m+1} = \left(\frac{S_m}{S_{m+1}}\right)^{\eta_2} (M_2)_m + (M_2)_{m+1,m} \quad m = 2, \dots, I-1 \quad (4.73)$$

where

$$(M_2)_{m,n} = \frac{qS_m^{-\eta_2}}{(\eta_2 + 1)(S_{n+1} - S_n)} \left[(S_{n+1}^{\eta_2+1} - (S_{n+1} + \eta_2(S_{n+1} - S_n))S_n^{\eta_2})v(S_n, t) + [S_n^{\eta_2+1} - (S_n - \eta_2(S_{n+1} - S_n))S_{n+1}^{\eta_2}]v(S_{n+1}, t) \right]. \quad (4.74)$$

Toivanen [99] showed that under suitable regularity assumptions on v , the approximation of the integral is of second-order accuracy and that the accumulating error grows at most linearly with respect to the number of grid points. This means that the formulas are sufficiently stable to be used in the computations.

The following proposition describes an important property which refers to the positivity of the jump matrix terms, and which is useful for the convergence result.

Proposition 4.5.1. *The elements of the jump matrix Jump are positive, i.e.*

$$(\text{Jump})_{m,n} \geq 0,$$

for all $m, n = 1, \dots, I$.

Proof: it is enough to show that the coefficients of $v(S_n, t)$ and $v(S_{n+1}, t)$ appearing in (4.71), respectively (4.74) are positive.

We introduce the following functions

$$\begin{aligned} F_1(S_{n+1}) &= S_{n+1}^{1-\eta_1} - (S_{n+1} - \eta_1(S_{n+1} - S_n))S_n^{-\eta_1} \\ F_2(S_{n+1}) &= S_n^{1-\eta_1} - (S_n + \eta_1(S_{n+1} - S_n))S_{n+1}^{-\eta_1} \\ F_3(S_{n+1}) &= S_{n+1}^{\eta_2+1} - (S_{n+1} + \eta_2(S_{n+1} - S_n))S_n^{\eta_2} \\ F_4(S_{n+1}) &= S_n^{\eta_2+1} - (S_n - \eta_2(S_{n+1} - S_n))S_{n+1}^{\eta_2}, \end{aligned}$$

with $S_n, S_{n+1} \in [S_{min}, S_{max}]$ and $S_{min} = S_1 < S_2 < \dots < S_n = S_{max}$.

We show next that the function F_1 is positive $\forall n = 1, \dots, I$.

For $n = 1$ we have that $F(S_2) = S_2^{1-\eta_1} - (S_2 - \eta_1(S_2 - S_1))S_1^{-\eta_1}$.

We observe that

$$F_1(S_1) = 0. \quad (4.75)$$

Then using the fact that $S_2 - S_1 > 0$ and $\eta_1 > 1$ we have

$$\begin{aligned} F_1'(S_2) &= S_2^{-\eta_1} - \eta_1 S_2^{-\eta_1} - S_1^{-\eta_1} + \eta_1 S_1^{-\eta_1} = \\ &= (S_1^{-\eta_1} - S_2^{-\eta_1})(\eta_1 - 1) > 0, \end{aligned}$$

which together with (4.75) gives us that $F(S_2) > 0$. It is trivial to show then inductively that

$$F_1(S_{n+1}) = S_{n+1}^{1-\eta_1} - (S_{n+1} - \eta_1(S_{n+1} - S_n))S_n^{-\eta_1} > 0, \quad \text{for all } n = 1, \dots, I.$$

The positivity of the functions F_2, F_3 , respectively F_4 can be proved in the same way.

□

For the discretization problem we also need the multiplication of a vector to the jump matrix. That gives us

$$\text{Jump}v_i = \lambda [(M_1)_i + (M_2)_i], \quad i = 2, \dots, I - 1. \quad (4.76)$$

Introducing the jumps approximation into (4.63) we get

$$\begin{aligned} & v_i^j [1 + \theta \Delta t (a_i + b_i + r + \lambda)] - v_{i+1}^j \theta \Delta t a_i - v_{i-1}^j \theta \Delta t b_i - \theta \Delta t \text{Jump} v_i^j - \\ & - v_i^{j+1} [1 - (1 - \theta) \Delta t (a_i + b_i + r + \lambda)] - v_{i+1}^{j+1} (1 - \theta) \Delta t a_i - v_{i-1}^{j+1} (1 - \theta) \Delta t b_i - \\ & - (1 - \theta) \Delta t \text{Jump} v_i^{j+1} - \frac{\Delta t}{\epsilon} \max(\Upsilon_i - v_i^j, 0) = 0. \end{aligned} \quad (4.77)$$

Equation (4.77) represents a system of nonlinear algebraic equations and it is solved iteratively. The iterative solution can be best understand if we rewrite (4.77) in matrix form

$$\begin{aligned} & [\mathcal{I} - (1 - \theta) \Delta t A - \lambda (1 - \theta) \Delta t \text{Jump}] v^{j+1} = \\ & = [\mathcal{I} + \theta \Delta t A - \lambda \theta \Delta t \text{Jump}] v^j - P^j (\Upsilon_i - v^j), \end{aligned} \quad (4.78)$$

where P^j is a diagonal penalty matrix, given by

$$(P^j)_{i,i} = \begin{cases} \text{Large}, & \text{if } v_i^j < \Upsilon_i \\ 0, & \text{otherwise} \end{cases} \quad (4.79)$$

where *Large* is a penalty factor which is related to the desired convergence tolerance, as we show in the next section.

Moreover, in (4.78) \mathcal{I} represents the identity matrix and A is a tridiagonal matrix with the elements $A_{m,n}$

$$A_{m,n} = \begin{cases} -b_i & \text{if } m = n - 1 \\ a_i + b_i + r + \lambda & \text{if } m = n \\ -a_i & \text{if } m = n + 1. \end{cases} \quad (4.80)$$

for $m, n = 2, \dots, I - 1$.

By imposing the positive coefficient condition (4.67), it is trivial to show that A is a M-matrix.⁷

We set Dirichlet boundary conditions at $m = \{1, I\}$ as follows

$$\begin{aligned} A_{m,n} &= 0 \text{ for } m = \{1, I\} \\ P_{m,n} &= 0 \text{ for } m = \{1, I\} \\ \text{Jump}_{m,n} &= 0 \text{ for } m = \{1, I\}. \end{aligned}$$

In the next section we have to evaluate the matrix $D := A - \lambda \text{Jump}$. Thus, we take a closer look on its properties. First we remark that

⁷A M-matrix has a positive diagonal, a non-positive off-diagonal and non-negative row sums, at least one being positive.

Lemma 4.5.1. *The matrix $\lambda\mathcal{I} - \lambda\text{Jump}$ is a diagonally dominant Z-matrix⁸ with the diagonal being positive, i.e.*

$$\sum_j (\lambda\mathcal{I} - \lambda\text{Jump})_{m,m} \geq 0, \quad \text{Jump}_{m,n} \leq 0 \quad \forall m \neq n, \quad i = 1, \dots, I - 1.$$

The proof of this Lemma can be found in Toivanen [99] so we omit the details here.

Proposition 4.5.2. *The matrix $D = A - \lambda\text{Jump}$ is a strictly diagonally dominant M-matrix.*

Proof: it is trivial to show that $A - \lambda\mathcal{I}$ is a M-matrix and together with Lemma 4.5.1 we get that $A - \lambda\mathcal{I} + \lambda\mathcal{I} - \lambda\text{Jump} = A - \lambda\text{Jump}$ is also a M-matrix.

□

We can now rewrite equation (4.78) in the following way

$$[\mathcal{I} - (1 - \theta)\Delta t D] v^{j+1} = [\mathcal{I} + \theta\Delta t D] v^j - P^j(\Upsilon_i - v^j). \quad (4.81)$$

Note that the matrix formulation (4.81) and the discretization (4.77) are equivalent. In the proofs to follow, we use both formulations.

We have showed until now that the problem (4.16)-(4.19) has a unique viscosity solution and then we have introduced the finite difference method to approximate this value. However, the question regarding the convergence of the scheme to the correct solution remains. This issue is treated in the next section.

4.5.3 Convergence analysis

In this section we show that the discretization scheme is monotone, stable and consistent. Then, according to Barles [5], the scheme converges to the unique (cf. Theorem 4.4.5) viscosity solution.

We denote by $\mathcal{S}(\tilde{v}, \theta)$ the finite difference approximation of the penalized matrix formulation

$$\mathcal{S}(\tilde{v}, \theta) = [\mathcal{I} + \theta\Delta t D] v^j - [\mathcal{I} - (1 - \theta)\Delta t D] v^{j+1} - P^j(\Upsilon_i - v^j), \quad (4.82)$$

where let $\tilde{v} = (v_i^j, v_i^{j+1}, v_{\gamma_i}^j, v_{\gamma_i}^{j+1})$ and $\gamma_i \in \{i - 1, i + 1\}$.

d'Halluin et. al. [49] state that usually, consistency follows if any reasonable⁹ discretization method is used. The stability property is usually a consequence of monotonicity of the scheme. Therefore, the most interesting requirement is the monotonicity.

⁸A Z-matrix is a particular case of a M-matrix, in which the off-diagonal entries are less than or equal to zero.

⁹Referring to implicit or Crank-Nicholson schemes.

Definition 4.5.1. Monotone discretization

The discretization (4.82) is monotone if either

$$\begin{aligned} \mathcal{S}(v_i^j, v_i^{j+1} + \varrho_i^{j+1}, v_{\gamma_i}^j + \varrho_{\gamma_i}^j, v_{\gamma_i}^{j+1} + \varrho_{\gamma_i}^{j+1}, \theta) &\geq \mathcal{S}(v_i^j, v_i^{j+1}, v_{\gamma_i}^j, v_{\gamma_i}^{j+1}, \theta) \\ \mathcal{S}(v_i^j + \varrho_i^j, v_i^{j+1}, v_{\gamma_i}^j, v_{\gamma_i}^{j+1}, \theta) &\leq \mathcal{S}(v_i^j, v_i^{j+1}, v_{\gamma_i}^j, v_{\gamma_i}^{j+1}, \theta) \end{aligned} \quad (4.83)$$

or

$$\begin{aligned} \mathcal{S}(v_i^j, v_i^{j+1} + \varrho_i^{j+1}, v_{\gamma_i}^j + \varrho_{\gamma_i}^j, v_{\gamma_i}^{j+1} + \varrho_{\gamma_i}^{j+1}, \theta) &\leq \mathcal{S}(v_i^j, v_i^{j+1}, v_{\gamma_i}^j, v_{\gamma_i}^{j+1}, \theta) \\ \mathcal{S}(v_i^j + \varrho_i^j, v_i^{j+1}, v_{\gamma_i}^j, v_{\gamma_i}^{j+1}, \theta) &\geq \mathcal{S}(v_i^j, v_i^{j+1}, v_{\gamma_i}^j, v_{\gamma_i}^{j+1}, \theta) \end{aligned} \quad (4.84)$$

hold, $\forall \gamma_i \in \{i-1, i+1\}$, $\forall \varrho_i^j, \varrho_{\gamma_i}^j, \varrho_i^{j+1}, \varrho_{\gamma_i}^{j+1} \geq 0$.

Proposition 4.5.3. Monotonicity

1. The fully implicit discretization scheme (4.82) is monotone, independent of the choice of Δt ;
2. The Crank-Nicholson discretization scheme (4.82) is monotone, if the timestep Δt satisfies the following condition

$$\Delta t < \frac{2}{(D)_{i,i}} \quad i = 1, \dots, I. \quad (4.85)$$

Proof: 1. In the case of the fully implicit discretization ($\theta = 1$), (4.82) becomes

$$\mathcal{S}(\tilde{v}, 1) = [\mathcal{I} + \Delta t D] v^j - v^{j+1} - P^j(\Upsilon_i - v^j). \quad (4.86)$$

From Remark (4.5.2) we know that D is a M-matrix, in consequence $\mathcal{I} + \Delta t D$ is a M-matrix as well. That means that $[\mathcal{I} + \theta \Delta t D] v^j$ is an increasing function of v_i^j and decreasing of $v_{\gamma_i}^j$.

From (4.79) we know that the matrix P^j is positive, which means that $-P^j(\Upsilon_i - v^j)$ is an increasing function of v_i^j .

The monotonicity of equation (4.86) follows now directly from Definition 4.5.1.

2. For the Crank-Nicholson discretization ($\theta = \frac{1}{2}$) the scheme (4.82) becomes

$$\mathcal{S}\left(\tilde{v}, \frac{1}{2}\right) = \left[\mathcal{I} + \frac{1}{2}\Delta t D\right] v^j - v^{j+1} + \frac{1}{2}\Delta t D v^{j+1} - P^j(\Upsilon_i - v^j). \quad (4.87)$$

As in the previous case we have that $[\mathcal{I} + \theta \Delta t D] v^j$ is an increasing function of v_i^j and decreasing of $v_{\gamma_i}^j$.

It is known that the Crank-Nicholson discretization is conditionally monotone. In order to achieve monotonicity of (4.87) we need to assure that

$$\mathcal{I} - \frac{1}{2}\Delta t D > 0, \quad (4.88)$$

so that $\left[\mathcal{I} - \frac{1}{2}\Delta t D\right] v^{j+1}$ is a decreasing function of v_i^{j+1} .

Rearranging in (4.88) we get that the Crank-Nicholson scheme is monotone if the timestep is selected such that the following condition holds

$$\Delta t < \frac{2}{(D)_{i,i}}, \quad i = 1, \dots, I.$$

□

The monotonicity of the jump discretization is a very important tool for the finite difference convergence theorem, but as well for the iteration convergence theorem.

Proposition 4.5.4. Monotonicity of the approximated integral

The approximation of the integral term (4.68) is monotone.

Proof: the monotonicity of the jump matrix Jump results from Proposition 4.5.1 and by Definition 4.5.1.

□

Proposition 4.5.5. Stability

1. In the fully implicit case, the scheme (4.82) is stable for any $\Delta S, \Delta t > 0$;
2. If the condition (4.85) is satisfied, then the discretization (4.87) is stable.

Proof: 1. We denote by \mathcal{SD} the scheme without the jump terms

$$\mathcal{SD}(\tilde{v}, 1) = [\mathcal{I} + \Delta t A] v^j - v^{j+1} - P^j(\Upsilon_i - v^j).$$

The stability of \mathcal{SD} follows directly from maximum analysis, we refer to Forsyth et al. [43] for a complete proof. Then the stability of (4.86) is a trivial consequence of the \mathcal{SD} stability and the monotonicity of the integral term (Proposition 4.5.4).

2. If condition (4.85) is satisfied, then by Proposition 4.5.3 the scheme is monotone and so we can apply the same results as in the fully implicit case above.

□

In order to show the last property needed for the convergence, we denote by $\mathcal{G}_i^j(\tilde{v}, \text{Jump}v_i^j, \theta)$ the finite difference discretization of the equation (4.60). Moreover, we introduce the functions $\mathcal{GD}_i^j(\tilde{v}, \theta)$, $\mathcal{GJ}_i^j(\text{Jump}v^j)$, $\mathcal{GP}_i^j(P^j)$, which represent the discretized diffusion terms, the approximation of the integral, respectively the discretized penalty terms. The complete formulation of these functions is given in the Appendix A.1.

Proposition 4.5.6. Consistency

The finite difference scheme $\mathcal{G}_i^j(\tilde{v}, \text{Jump}v_i^j, \theta)$ is consistent if we have

$$\left| \mathcal{G}_i^j(\tilde{v}, \text{Jump}v^j, \theta) - \left(\frac{\partial v}{\partial t} + \mathcal{D}[v] + \mathcal{L}[v] - (r + \lambda)v + \frac{1}{\epsilon} \max(\Upsilon - v, 0) \right) (S, t) \right| \rightarrow 0,$$

when $(\Delta S, \Delta t) \rightarrow 0$, $(S_i, t_j) \rightarrow (S, t)$ and $\theta \in \{1, \frac{1}{2}\}$.

Proof: the idea of the proof is to show that the following properties hold:

$$\left| \frac{v_i^{j+1} - v_i^j}{\Delta t} - \frac{\partial v}{\partial t} \right| \longrightarrow 0 \quad (4.89)$$

$$|\mathcal{GD}_i^j(\tilde{v}, \theta) - \mathcal{D}[v(S, t)]| \longrightarrow 0 \quad (4.90)$$

$$\left| \mathcal{GP}_i^j(P^j) - \frac{1}{\epsilon} \max(\Upsilon - v, 0) \right| \longrightarrow 0 \quad (4.91)$$

$$|\mathcal{GJ}_i^j(\text{Jump}v^j) - \mathcal{L}[v(S, t)]| \longrightarrow 0 \quad (4.92)$$

where \mathcal{D} and \mathcal{L} were defined in (4.12) and (4.13).

The properties (4.89)-(4.91) can be proved using Taylor series analysis, see for example Cont and Voltchkova [32]. Condition (4.92) is proved in Toivanen [99].

□

Theorem 4.5.1. Convergence of the finite difference scheme

If the scheme (4.82) is stable, monotone, consistent and the approximating integral is monotone, then the discretization scheme converges locally uniform to the unique (cf. Theorem 4.4.5) viscosity solution of (4.16) a.s. $(\Delta S, \Delta t) \rightarrow 0$.

Proof: from Propositions 4.5.3, 4.5.5, 4.5.6 and Proposition 4.5.4 we know that the discretization of the penalized equation (4.82) is monotone, stable, consistent and the integral approximation is monotone. Therefore the proof of Theorem 4.5.1 follows directly from the results of Barles [5] and Briani [21].

□

In the next section we present an iterative method to solve the equation (4.81) and we prove that the iteration is convergent.

4.5.4 Matrix iteration

The non-linear discrete equation (4.81) can be solved using generalized Newton iteration. This method has finite termination and it converges to the unique solution of the equation (4.60). This iterative approach was proposed by d'Halluin et. al. [48] for American option pricing.

The algorithm can be summarized as follows

Penalty algorithm

Let $v^0 = v^{j+1}$

$\hat{v}^l = (v^j)^l$

$\hat{P}^l = P((v^j)^l)$

for $l = 0, 1, 2, \dots$ until convergence

Solve

$$\hat{v}^{l+1} = \left(\mathcal{I} + \theta \Delta t A + \hat{P}^l \right)^{-1} \left[[\mathcal{I} - (1 - \theta) \Delta t D] v^{j+1} + \theta \Delta t \lambda \text{Jump} \hat{v}^l + \hat{P}^l \Upsilon \right] \quad (4.93)$$

If $\max_i \frac{|\hat{v}^{l+1} - \hat{v}^l|}{\max(1, |\hat{v}^{l+1}|)} < \text{tol}$ then quit

End for

where tol defines the desired accuracy and the vector \hat{v}^l is suppose to be accurate enough when

$$\max_i \frac{|\hat{v}^{l+1} - \hat{v}^l|}{\max(1, |\hat{v}^{l+1}|)} < \text{tol} = \frac{1}{\text{Large}}.$$

At any time step l , if equation (4.93) is violated, a new penalty matrix is created using equation (4.79) and equation (4.81) is solved again for \hat{v}^{l+1} . When the desired accuracy has been achieved, i.e. equation (4.93) is true, we step backwards in time and solve for the new time step.

Next we introduce some useful lemmas for the convergence theorem of the penalty iteration. These lemmas can be easily proved following d'Halluin, Forsyth and Labahn [48], so we omit the details here.

Lemma 4.5.2. *M-matrices*

Both $[\mathcal{I} + \theta \Delta t A + P^j]$ and $[\mathcal{I} + \theta \Delta t A + P^j - \theta \Delta t \lambda \text{Jump}]$ are *M-matrices*.

Lemma 4.5.3. *Bounded iterates*

Suppose that a_i, b_i satisfy the positive coefficient condition (4.67) for all i . Then for a given timestep, all iterates \hat{v}^{l+1} in the iteration scheme (4.93) are bonded, independent of l .

The iteration (4.93) can be rewritten in the following way

$$\left[\mathcal{I} + \theta \Delta t A + \hat{P}^l \right] (\hat{v}^{l+1} - \hat{v}^l) = (\hat{P}^l - \hat{P}^{l-1})(\Upsilon - \hat{v}^l) + \theta \Delta t \lambda \text{Jump}(\hat{v}^l - \hat{v}^{l-1}). \quad (4.94)$$

The following lemma determines the sign of $(\hat{P}^l - \hat{P}^{l-1})(\Upsilon - \hat{v}^l)$, which is also a convenient result in the proof of the convergence theorem.

Lemma 4.5.4. Positive penalty term

If the penalty matrix is given in (4.79) and the iteration algorithm by (4.93), we have that

$$(\widehat{P}^l - \widehat{P}^{l-1})(\Upsilon - \widehat{v}^l) \geq 0, \quad \text{for all } l \geq 1. \quad (4.95)$$

Lemma 4.5.5. Norm of the iteration matrix

Let A , Jump and \widehat{P}^l be given by (4.80), (4.76) and (4.79) respectively, and the assumptions in Lemma 4.5.3 hold. Then for $U^l := [\mathcal{I} + \theta\Delta t A + \widehat{P}^l]$ we have:

$$\|[U^l]^{-1}\text{Jump}\|_\infty \leq \frac{1}{1 + \theta\Delta t(r + \lambda)}. \quad (4.96)$$

We can now give the main convergence result for the penalty iteration (4.93).

Theorem 4.5.2. Convergence of the penalty iteration

Let A , Jump and \widehat{P}^l be given by (4.80), (4.76), respectively (4.79) and a_i, b_i satisfy the positive coefficient condition (4.67) for all i . Then the iteration (4.93) converges to the unique solution of (4.81), for any initial iterate \widehat{v}^0 .

Proof: we follow the lines of the convergence proof given in d'Halluin, Forsyth and Labahn [48], in the case of American options.

By using the equation (4.96) we can rewrite (4.94) in the following way

$$U^l(\widehat{v}^{l+1} - \widehat{v}^l) = (\widehat{P}^l - \widehat{P}^{l-1})(\Upsilon - \widehat{v}^l) + \theta\Delta t\lambda\text{Jump}(\widehat{v}^l - \widehat{v}^{l-1}).$$

For any $l \geq 1$ we can write

$$\widehat{v}^{l+1} - \widehat{v}^l = Y^l + W^l(\widehat{v}^1 - \widehat{v}^0), \quad (4.97)$$

with

$$\begin{aligned} Y^l &= (U^l)^{-1}(\widehat{P}^l - \widehat{P}^{l-1})(\Upsilon - \widehat{v}^l) + \\ &\quad + \theta\Delta t\lambda(U^l)^{-1}\text{Jump}(U^{l-1})^{-1}(\widehat{P}^{l-1} - \widehat{P}^{l-2})(\Upsilon - \widehat{v}^{l-1}) + \\ &\quad + \dots \\ &\quad + [\theta\Delta t\lambda]^{l-1}(U^l)^{-1}\text{Jump}(U^{l-1})^{-1}\text{Jump}\dots(U^{-1})^{-1}(\widehat{P}^1 - \widehat{P}^0)(\Upsilon - \widehat{v}^1), \end{aligned}$$

$$W^l = [\theta\Delta t\lambda]^l (U^l)^{-1}\text{Jump}(U^{l-1})^{-1}\text{Jump}\dots(U^1)^{-1}\text{Jump}.$$

In order to prove convergence of the iteration, we have to show that both Y^l and W^l tend to zero, as l gets large. We start by proving that $Y^l, W^l \geq 0$.

From Lemma 4.5.4 we have that $(\widehat{P}^l - \widehat{P}^{l-1})(\Upsilon - \widehat{v}^l) \geq 0$ for all $l \geq 0$, while from Remark 4.5.2 we get that $[U^l]^{-1} \geq 0$. Now if we take in account Proposition 4.5.4 it follows that $Y^l \geq 0$. In the same way we can show that $W^l \geq 0$.

Applying Lemma 4.5.5 to the equation (4.98), we get that for each i

$$\|W^i\|_\infty \leq \left[\frac{\theta\lambda\Delta t}{1 + \theta\Delta t(r + \lambda)\Delta t} \right]^i \quad (4.98)$$

and hence

$$\left\| \sum_{i=1}^l W^i \right\|_\infty \leq \sum_{i=1}^l \left[\frac{\theta\lambda\Delta t}{1 + \theta\Delta t(r + \lambda)\Delta t} \right]^i \leq \left[\frac{\theta\lambda\Delta t}{1 + \theta\Delta t(r + \lambda)\Delta t} \right]. \quad (4.99)$$

Thus we have that $\left\{ \sum_{i=1}^l W^i \right\}_{l=1,2,\dots}$ is bounded and consists of non-decreasing elements. Hence, the sequence converges and $W^i \rightarrow 0$ as $l \rightarrow \infty$. Summing up in the equation (4.97) over the index l , we get

$$\hat{v}^{l+1} = \hat{v}^1 + \sum_{l=1}^l Y^i + \sum_{l=1}^k W^i(\hat{v}^1 - \hat{v}^0). \quad (4.100)$$

From equation (4.98) we get that $\sum_{i=1}^l W^i(\hat{v}^1 - \hat{v}^0)$ converges to a finite value. Furthermore from Lemma 4.5.3 we have that the left hand side of equation (4.100) is bounded from above. Consequently, the sequence $\left\{ \sum_{i=1}^l Y^i \right\}_{l=1,2,\dots}$ is bounded and non-decreasing and therefore it is convergent.

In consequence, a convergent limit exists, and from Theorem 4.4.5 this is the unique solution to the equation (4.81).

□

4.6 Numerical results

In this section we present the numerical results for swing option pricing using the approach described before in this chapter. For the computations we used an Intel (R) Core (TM) 2 Quad CPU Q9300 2.50GHz and we implemented the algorithm in MATLAB 7.7.

To verify our implementation, we first consider the Black-Scholes and Kou models and determine the swing option price within these frameworks.

Afterwards we compute the value of swing options with up to 15 exercise rights using the mean-reverting double exponential jump-diffusion model introduced in Chapter 3. We investigate the dependence of swing option value on the model parameters and the number of exercise rights.

Pricing swing option under the Black-Scholes and Kou model

We test our algorithm on the general set of parameters used by Wilhelm [106] in her thesis. She priced swing options under the Black-Scholes model using the finite element method.

We use the following parameters

$$N_s = 5, \quad T = 1, \quad t = 0, \quad K = 100, \quad \sigma = 0.3, \quad r = 0.05, \quad \delta_R = 0.1.$$

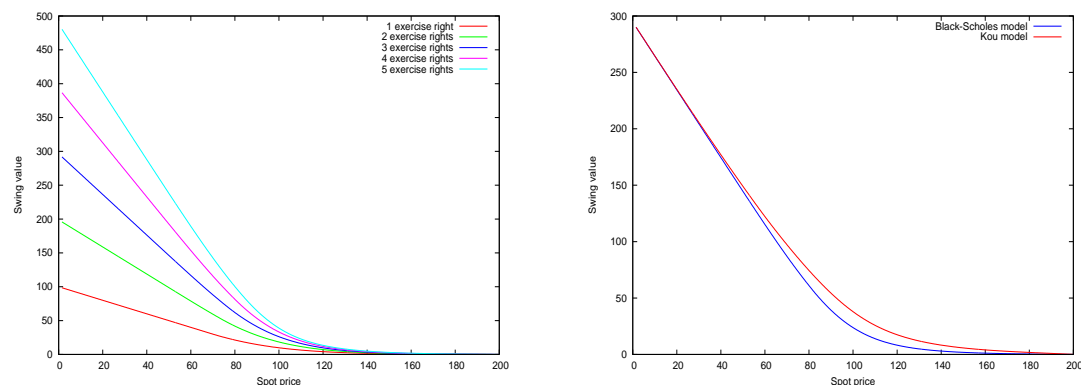


Figure 4.1: **Left:** swing values with up to 5 exercise rights under the Black-Scholes model; **Right:** swing option price with 3 exercise opportunities using the Black-Scholes and Kou model.

Figure 4.1 shows on the left-hand side the swing option put values with up to 5 exercise rights. In this case, the price follows the Black-Scholes model. We applied the method of finite differences combined with the penalty method and we obtained the same results as Wilhelm [106] using the finite element method.

On the right-hand side, Figure 4.1 shows the impact of introducing double exponential jumps into the model, that is the Kou model. The jump parameters for this model are as follows

$$\lambda = 0.5, \quad \eta_1 = 3.0465, \quad \eta_2 = 3.0775, \quad p = 0.3445$$

and have been taken from Toivanen [99]. As we expected, the introduction of jumps increases the swing price, due to the risk introduced by the jumps.

Next we present the numerical results for swing option value using the mean-reverting double exponential jump-diffusion model with seasonality.

Pricing swing options under the mean-reversion double exponential jump-diffusion model

We recall the formulation of the mean-reverting double exponential jump-diffusion model introduced in Chapter 3

$$dS_t = \alpha(\rho(t) - \ln S_t)S_t dt + \sigma S_t dW_t + S_t(e^{J_t} - 1)dN_t,$$

where the time-dependent mean-reverting level is given by

$$\rho(t) = \frac{1}{\alpha} \left(\frac{df(t)}{dt} + \frac{\sigma^2}{2} \right) + f(t)$$

For the numerical valuation we first consider the following set of parameter values

$$\begin{aligned} N_s = 15, S = 3.5, \alpha = 0.5, \sigma = 1.4, T = 1, t = 0, K = 3.5, \\ \lambda = 0.5, \eta_1 = 1.29, \eta_2 = 0.37. \end{aligned} \quad (4.101)$$

For the simulations we used the annual seasonality function (3.13) with the following parameter values:

$$\begin{aligned} a = 2.6, b = -0.018, c(1) = 0.09, c(2) = -0.14, c(3) = 0.02, \\ d(1) = 0.5, d(2) = -0.016, d(3) = -0.02. \end{aligned}$$

To check the validity of our algorithm, we first consider a particular case: if $N_s = 1$, the problem reduces to pricing an American option, which is known to be an upper bound for the corresponding swing option.

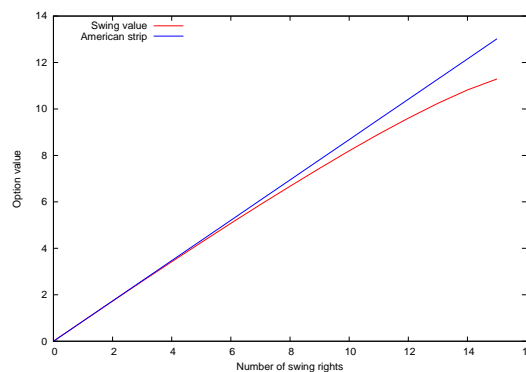


Figure 4.2: Swing option with 15 exercise rights versus a strip of American options.

Figure 4.2 presents the values of one year swing option with up to 15 exercise rights, compared to the value of 15 American options.

Let $k = 1, \dots, 15$ be the number of exercise rights. Then for $k = 1$, the swing option value coincides with the American option value, as we expected. Otherwise, k American options are more expensive than a swing option with k exercise rights, as the rights of a swing option can only be exercised one at a time.

Next we test the sensitivity of the swing option prices to different model parameters. We consider the parameters given in (4.101) as the original parameters.

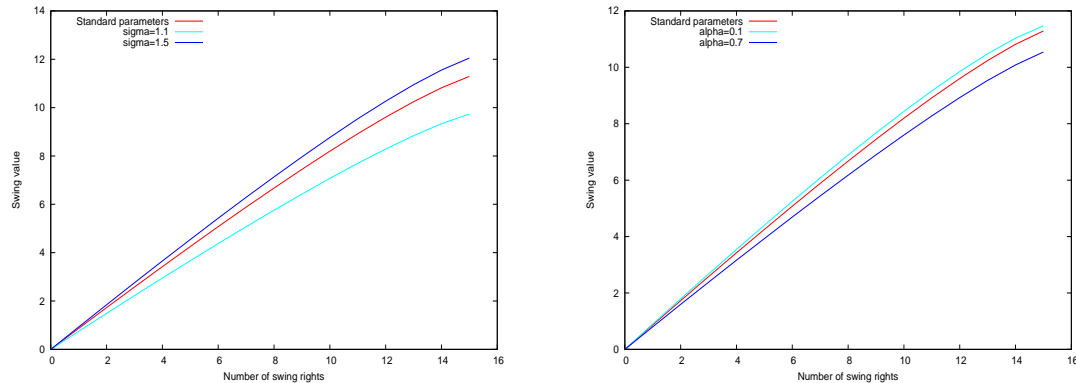


Figure 4.3: Behavior of the swing option price to the model parameters. The volatility σ and the mean-reversion rate α are shifted up and down from the original values.

In Figure 4.3 we deviate the volatility σ and the mean-reversion rate α from their original values, while the rest of the parameters are held constant. Observe that the most significant change is caused by the shifts of the volatility σ . Also, we can see that a change in the mean-reversion parameter α is inversely proportional to the price. That can be explained by looking at the variance of the price process (see Appendix A, (A.9)). According to Kluge [65], there is a direct relationship between the variance and the option price. The variance of the log-price process is given by the multiplication of $\frac{1}{\alpha}$ with another term, and hence the inverse proportionality.

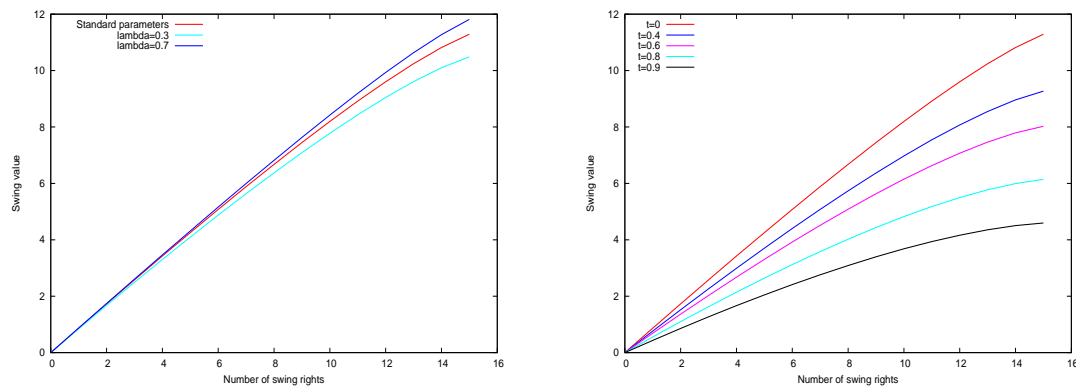


Figure 4.4: Left: the jump intensity parameter λ is shifted up and down from its original value. Right: the swing values are computed against the total number of exercise rights, at different times t .

On the left-hand side, Figure 4.4 shows the behavior of the swing option to changes in the jump intensity parameter λ . As in Figure 4.1, we observe that for a higher jump intensity, the option price will increase, due to the risk introduced by the jumps.

On the right-hand side, Figure 4.4 shows the swing option prices at different times t . For $t = 0$ the holder has the maximum number (N_s) of possibilities to exercise the option, that means that the option value is the largest in this case. As t gets closer to the expiration date ($T = 1$), the holder of the option has less time to exercise all his exercise rights, due to the refraction times. Thus, the option value decreases in time.

Chapter 5

Pricing swing options with power volume constraints

In this chapter we propose a new approach for pricing more general swing options. We show that it is possible to extend the results presented in the previous chapter, to swing options that include the energy volume constraints.

By introducing the volume optionality into the swing formulation, the pricing problem becomes even more difficult: the optimal decision on the quantity of electricity to buy on each date depends not only on the price and on the number of remaining swing rights, but also on the cumulative power consumption.

Designing efficient numerical methods for pricing such general swing derivatives remains a challenging question. The valuation of this type of option is related to a stochastic control problem, which can be formulated as a Hamilton-Jacobi-Bellman (HJB) equation with two state variables, the quantity and the price. Having more than one state variable considerably complicates the problem. However, using the method of characteristics, the problem can be simplified and then solved numerically using an approach similar to the one in the previous chapter.

We apply the penalty method to enforce the early exercise constraint and prove the existence of the viscosity solution for the penalized equation. This nonlinear equation is discretized by finite differences. The integral term can be approximated using the recursion formula presented in Section 4.5.2.

Provided a comparison principle holds, we prove that the finite difference scheme converges to the unique viscosity solution of the penalized equation, by verifying the stability, monotonicity and consistency of the scheme.

The solution can then be found using the generalized Newton iteration, as in the previous chapter.

5.1 Introduction and swing problem formulation

In this section we formulate the optimal multiple stopping time problem associated to the swing option with variable volume. We first start with a short overview on the existing numerical approaches for these specific contracts.

There are only a few theoretical results in the literature that discuss the valuation of general swing options. One of the first attempts was introduced by Lari-Lavassani et. al. [69]. They used the binomial tree method to price swing options with global volume constraints. A similar approach was proposed by Jaillet et. al. [60], using a mean-reverting model. More recently, Wahab and Lee [101] developed a pentanomial lattice method to evaluate swing options in gas markets under a regime switching model for the spot price. Wahab et. al. [102] extended this approach to electricity swing contracts, where the spot price is switching between a mean-reverting and a Geometric Brownian Motion process.

Keppo [63] used a linear programming method to price swing contracts with local and global volume constraints, in a forward market. He provided some upper and lower bounds for the value function and he proved that the optimal exercise strategy is of bang-bang¹ type in the case when no penalties are applied.

An interesting approach was presented by Bardou et. al. [4] who investigated a numerical integration method, so-called optimal quantization, for pricing general swing options. They introduced global volume constraints and showed that the optimal strategy is of bang-bang type in this case. Barrera-Esteve et. al. [10] consider gas swing options with changeable amounts and fixed exercise times. They worked under a forward market and extended the Least Squares Monte-Carlo approach to swing options with variable volume.

Another way to analyze this problem was proposed by Lund and Ollmar [75], who studied flexible load contracts (similar to swing options) by formulating the contract as a stochastic optimization problem under a Ornstein-Uhlenbeck process with local and global constraints. Haarbrücker and Kuhn [47] investigated pricing of electricity swing options using stochastic programming, under a forward price model. They compared this technique with the Least Squares Monte-Carlo method and observed that stochastic programming performs better in the presence of various risk factors and state variables.

Most of the approaches presented above were using a forward price model and did not include refraction periods between swing exercise rights. However, since swing options are influenced by the hourly price behavior, a realistic approach would be to work under a spot-based model.

In this chapter we value general swing options under the double exponential jump-diffusion spot model (3.8). Like in the previous chapter, we impose a refraction time between two successive exercise rights. Moreover, we require that the option holder is allowed to buy a maximum amount of energy Q_{max} MWh during the contract period $[0, T]$ and he has to acquire at least Q_{min} MWh until maturity. On the other hand, at every time t , the owner of the contract is allowed to pur-

¹That means that the swing option is exercised either at the highest or lowest level allowed by the local constraints.

chase a limited amount of energy $q(t)$, at a maximum rate p_{max} MW, respectively at a minimal power rate p_{min} MW. Also denote by $Q(t)$ the amount of electricity purchased up to time t . Then we have that $Q(0) = 0$.

Consequently, the consumption processes $q(t)$ and $Q(t)$ satisfy the following local and global constraints

$$p_{min}(T - t) \leq q(t) \leq p_{max}(T - t), \quad (5.1)$$

$$Q_{min} \leq \sum_{t=0}^T q(t) = Q(T) \leq Q_{max}. \quad (5.2)$$

In the case when the global constraints are allowed to be violated, penalties have to be settled at expiration. We do not consider penalty payments in this thesis, and assume for the rest of the paper that the global constraints are firm. In this case, Keppo [63] proved that it is optimal to always exercise in a bang-bang fashion, that is, either at a minimum power rate p_{min} or at the maximum rate p_{max} .

Without loss of generality we can set $Q_{min} = 0$ and $p = \{p_{min}, p_{max}\}$. This implies that the purchased quantity of energy at time t is given by

$$q(t) = p(T - t). \quad (5.3)$$

We define the new set $\bar{\mathcal{T}}_{t,Q,T}$ of stopping times, which keeps the characteristics of the set $\bar{\mathcal{T}}_{t,T}$ defined in (4.3) and has the following additional properties

- $q(t)$ and $Q(t)$ are \mathcal{F}_t - measurable
- equations (5.1) and (5.2) are satisfied.

Then for $k = 1, \dots, N_s$, the value of the swing option is given by the supremum over the expected discounted payoff at each stopping time

$$v_k(s, Q, t) = \sup_{(Q(\tau_k), \bar{\tau}) \in \bar{\mathcal{T}}_{t,Q,T}} \mathbb{E} \left[\sum_{k=1}^{N_s} e^{-r(\tau_k - t)} \Phi_k(S, Q, \tau_k) | S(t) = s, Q(t) = Q \right] \quad (5.4)$$

with $\bar{\tau} = (\tau_1, \dots, \tau_{N_s})$. As we showed in the previous chapter, this multiple stopping time problem can be reduced to a sequence of N_s optimal single stopping problems

$$v_k(s, Q, t) = \sup_{(Q(\tau), \tau) \in \mathcal{T}_{t,Q,T}} \mathbb{E} [e^{-r(\tau - t)} \Phi_k(S_\tau, Q, \tau) | S(t) = s, Q(t) = Q], \quad (5.5)$$

with the following payoff function

$$\Phi_k(S, Q, t) = \begin{cases} \max \left\{ \Upsilon_k^Q(s, Q, t); v_k(s, Q, t + 1) \right\} & t + \delta_R \leq T \\ q(t)(K - S)^+ & t + \delta_R > T \end{cases} \quad (5.6)$$

$$\text{where } \Upsilon_k^Q = q(t)(K - S)^+ + \mathbb{E} [e^{-r\delta_R} v_{k-1}(S_{t+\delta_R}^{s,t}, Q(t + \delta_R), t + \delta_R)]. \quad (5.7)$$

In this way, the arbitrage free price of a general swing option can be determined by a sequence of single optimal stopping time problems.

5.1.1 Partial integro-differential complementarity formulation

In the previous section we showed that the price of a general swing option can be associated to an optimal multiple stopping time problem. Moreover, in Chapter 4 we saw that it is possible to reduce this problem to a partial integro-differential complementarity formulation.

Due to the new volume variable in the general swing problem formulation, we have to deal with a two state variable partial differentiable equation. In Appendix B we develop the linear complementarity problem for the swing option with volume constraints, in the case when there are no jumps in the price process. This algorithm is based on the delta hedging argument, i.e. a portfolio whose delta is kept to zero as close as possible.

Then combining the results presented in Chapter 4 and Theorem 2.4.3, we get that the value of a general swing option under the double exponential jump-diffusion model (3.8) satisfies the following PIDCP

$$\begin{cases} \left(\frac{\partial v_k}{\partial t} + \mathcal{D}^Q[v_k] + \mathcal{L}^Q[v_k] - rv_k \right) (v_k(s, Q, t) - \Upsilon_k^Q(s, Q, t)) = 0 \\ \frac{\partial v_k}{\partial t} + \mathcal{D}^Q[v_k] + \mathcal{L}^Q[v_k] - rv_k \leq 0 & (s, Q, t) \in \bar{\Omega} \\ v_k(s, Q, t) \geq \Upsilon_k^Q(s, Q, t) \\ v_k(s, Q, T) = q(T)(K - S_T)^+ \end{cases} \quad (5.8)$$

where

$$\begin{aligned} \mathcal{D}^Q[v_k] &= \alpha(\rho(t) - \ln S)S \frac{\partial v_k}{\partial S} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 v_k}{\partial S^2} + p \left(\frac{\partial v_k}{\partial Q} + (K - S)^+ \right) \\ \mathcal{L}^Q[v_k] &= \lambda \int_{\mathbb{R}_+} [v_k(sz, Q, t) - v_k(s, Q, t)] f_{e^J}(z) dz. \end{aligned}$$

We can write (5.8) equivalently, as a Hamilton Jacobi Bellman variational formulation

$$\begin{cases} \min\{rv_k(s, Q, t) - \frac{\partial v_k}{\partial t}(s, Q, t) - \mathcal{D}^Q[v_k] - \mathcal{L}^Q[v_k], v_k(s, Q, t) - \Upsilon_k^Q(s, Q, t)\} = 0 \\ v_k(s, Q, T) = q(T)(K - S)^+. \end{cases} \quad (5.9)$$

Boundary conditions

We specify next the boundary conditions for (5.9). In this case, besides the conditions specified in (4.17)-(4.18) for the problem (4.16), we have to impose boundaries conditions at $Q_{min} = 0$ and Q_{max} . For the sake of completeness, we recall also the

initial and boundary conditions given in (4.17)-(4.19)

$$v_k(s, Q, 0) = 0, \quad (5.10)$$

$$\text{as } S \rightarrow 0 \text{ we set } \mathcal{D}^Q[v_k] = \mathcal{L}^Q[v_k] = 0, \quad (5.11)$$

$$\text{as } S \rightarrow +\infty \text{ we set } v_k(s, Q, t) \rightarrow 0, \quad (5.12)$$

$$\text{as } Q \rightarrow Q_{max} \text{ we set } \frac{\partial v_k}{\partial Q} \rightarrow 0. \quad (5.13)$$

For $Q \rightarrow 0$ no specific boundary condition is needed, since $q(t) \geq 0$ and the PIDE (5.9) is first order hyperbolic in the Q direction.

5.2 Numerical calculations

As in the previous chapter, we denote by $mr(s, t) = \alpha(\rho^*(t) - \ln s)s$ and assume that there exists $c > 0$ such that for all $s, y \in \mathbb{R}_+$, $Q, O \in [0, Q_{max}]$ and $t \in [0, T]$ we have

$$\int_{\mathbb{R}_+} |z|^2 f_{e^J}(z) dz < \infty \quad (5.14)$$

$$|mr(s, t) - mr(y, t)| \leq c|s - y| \quad (5.15)$$

$$|\Upsilon_k^Q(s, Q, t) - \Upsilon_k^Q(y, Q, t)| \leq c|s - y| \quad (5.16)$$

$$|\Upsilon_k^Q(s, Q, t) - \Upsilon_k^Q(s, O, t)| \leq c|Q - O| \quad (5.17)$$

where Υ_k^Q is given in (5.7).

Proposition 5.2.1. *For all $s \in \mathbb{R}_+$, $t \in [0, T]$, $Q, O \in [0, Q_{max}]$ and $k = 1, \dots, N_s$, there exists $c > 0$ such that*

$$|v_k(s, Q, t) - v_k(s, O, t)| \leq c|Q - O|. \quad (5.18)$$

Proof : we prove the Lipschitz inequalities (5.17) and (5.18) concomitant, using mathematical induction on k .

For $k = 1$ we get

$$|\Upsilon_1^Q(s, Q, t) - \Upsilon_1^Q(s, O, t)| \leq |\phi_S(s, t) - \phi_S(s, t)| \leq c|Q - O|$$

and similarly

$$|v_1(s, Q, t) - v_1(s, O, t)| \leq c|Q - O|.$$

Second step: for $k = 2, \dots, N_s$ suppose that exists $c > 0$ such that $\Upsilon_{k-1}^Q(s, Q, t)$ and $v_{k-1}(s, Q, t)$ are Lipschitz in Q . Then

$$|\Upsilon_k^Q(s, Q, t) - \Upsilon_k^Q(s, O, t)| \leq |\phi_S(s, t) - \phi_S(s, t)| + c \mathbb{E}|Q(t + \delta_R) - O(t + \delta_R)|.$$

We know that $Q(t + \delta_R) = Q + q(t + \delta_R) = Q + p\delta_R$, so we get

$$|\Upsilon_k^Q(s, Q, t) - \Upsilon_k^Q(s, O, t)| \leq c|Q - O|.$$

Now for v_k we have

$$\begin{aligned} |v_k(s, Q, t) - v_k(s, O, t)| &\leq \sup_{(Q, \tau) \in \mathcal{T}_{l, q, T}} \mathbb{E} \left[|\Upsilon_k^Q(s, Q, \tau) - \Upsilon_k^Q(s, O, \tau)| \right] \leq \\ &\leq c|Q - O|. \end{aligned}$$

□

For the rest of the paper denote by $\bar{\Omega} = \mathbb{R}_+ \times [0, Q_{max}] \times [0, T]$ the closed domain where our problem is defined and by $\Omega = \mathbb{R}_+ \times [0, Q_{max}] \times [0, T)$.

Proposition 5.2.2. Continuity

Under the assumptions (5.14)-(5.17) the value function v_k is continuous on $\bar{\Omega}$.

Proof: the proof follows from Proposition 4.4.5 and the Lipschitz condition (5.18).

□

Next we give the definition of the viscosity solution associated to the HJB problem (5.9).

Definition 5.2.1. The function $v_k \in C^0(\bar{\Omega})$ is a **viscosity subsolution (super-solution)** of (5.9) if

$$\begin{aligned} \min\{r\varphi(s, Q, t) - \frac{\partial \varphi}{\partial t}(s, Q, t) - \mathcal{D}^Q[\varphi] - \mathcal{L}^Q[\varphi], \varphi(s, Q, t) - \Upsilon_k^Q(s, Q, t)\} &\leq 0 \\ &(\geq 0) \end{aligned}$$

whenever $\varphi \in C^2(\bar{\Omega}) \cap C_2(\bar{\Omega})$ and $v_k - \varphi$ has a global maximum (minimum) at $(s, Q, t) \in \Omega$.

The value function v_k is a **viscosity solution** of (5.9) if it is a sub- and super-solution.

In the definition above $C_2(\bar{\Omega})$ was defined as

$$C_2(\bar{\Omega}) = \left\{ \varphi \in C^0(\bar{\Omega}) \mid \sup_{\bar{\Omega}} \frac{\varphi(s, Q, t)}{1 + |s|^2} < \infty \right\}.$$

We introduce next the modified parabolic superjet and subjet for the problem (5.9)-(5.13). Given $v_k \in C^0(\bar{\Omega})$ and $(s, Q, t) \in \Omega$ we define

$$\begin{aligned} \mathcal{P}_Q^{2,+} v_k(s, Q, t) &= \{(p_0, p_1, p_2, M) \in \mathbb{R} \times \mathbb{R} \times \mathbb{R} \times \mathbb{R} \mid v_k(y, Z, l) - v_k(s, Q, t) \leq \\ &\leq p_0(y - s) + p_1(l - t) + p_2(Z - Q) + \frac{1}{2}M(y - s)^2 + \\ &+ o(|l - t| + |Z - Q| + |y - s|^2) \text{ as } (y, Z, l) \rightarrow (s, Q, t)\}, \end{aligned} \tag{5.19}$$

and its closure

$$\begin{aligned} \overline{\mathcal{P}}_Q^{2,+} v_k(s, Q, t) = & \left\{ (p_0, p_1, p_2, M) = \lim_{n \rightarrow \infty} (p_{0,n}, p_{1,n}, p_{2,n}, M_n), \right. \\ & \text{with } (p_{0,n}, p_{1,n}, p_{2,n}, M_n) \in \mathcal{P}_Q^{2,+} v_k(s, Q, t) \\ & \left. \text{and } \lim_{n \rightarrow \infty} (s_n, Q_n, t_n, v_k(s_n, Q_n, t_n)) = (s, Q, t, v_k(s, Q, t)) \right\}, \end{aligned} \quad (5.20)$$

and the parabolic subset $\mathcal{P}_Q^{2,-} v_k(s, Q, t) = -\mathcal{P}_Q^{2,+} v_k(s, Q, t)$,
and its closure $\overline{\mathcal{P}}_Q^{2,-} v_k(s, Q, t) = -\overline{\mathcal{P}}_Q^{2,+} v_k(s, Q, t)$.

We can give now the characterization of viscosity solutions in terms of superjet and subset.

Lemma 5.2.1. *Let $v_k \in C_2(\overline{\Omega})$ be a viscosity supersolution (resp. subsolution) of (5.9). Then for all $(s, Q, t) \in \Omega$ and $\forall (p_0, p_1, p_2, M) \in \overline{\mathcal{P}}_Q^{2,-} v_k(s, Q, t)$ (resp. $\overline{\mathcal{P}}_Q^{2,+} v_k(s, Q, t)$), there exists $\varphi \in C^2(\Omega)$ such that*

$$\begin{aligned} \min \{ & -p_1 + r v_k(s, Q, t) - \mathcal{D}_{p_0}[s, Q, t, p_0, p_2, M] - \\ & -\mathcal{L}[v_k(s, Q, t)], v_k(s, Q, t) - \Upsilon_k^Q(s) \} \geq 0 \\ & \text{(resp. } \leq 0), \end{aligned}$$

where

$$\begin{aligned} \mathcal{D}_{p_0}[s, Q, t, p_0, p_2, M] &= m r(s, Q, t) p_0 + p p_2 + p(K - s)^+ + \frac{1}{2} \sigma^2 s^2 M \\ \mathcal{L}[v_k(s, Q, t)] &= \lambda \int_{\mathbb{R}_+} [v_k(sz, Q, t) - v_k(s, Q, t)] f_{e^J}(z) dz. \end{aligned}$$

In Section 4.4 we showed that the main result which stands behind the existence and uniqueness of a viscosity solution is the comparison principle.

We assume here that a comparison principle also holds in the case of the swing option problem with volume constrains. A proof of a comparison theorem is behind the scope of our research. However, following Chen [26] we can give enough arguments to show that such a principle holds for our problem.

We can view our two dimensional problem as a three dimensional degenerate elliptic PIDE in the variable $(s, Q, t) \in \overline{\Omega}$. In this way, we are able to apply the results from Barles et. al. [7], [8], who proved that the viscosity solution of degenerate elliptic HJB equation with Dirichlet boundary conditions satisfy the strong comparison result, if several assumptions on the boundary are satisfied. Chen [26] showed that these assumptions hold for the gas storage problem, which has a very similar formulation to the swing problem with volume constrains.

Consequently, we can make the following assumption, which is necessary to ensure that a unique viscosity solution to equation (5.9) exists.

Assumption 5.2.1. *The swing option problem (5.9) and the associated boundary conditions (5.10)-(5.13) satisfy a strong comparison result in $\bar{\Omega}$.*

In the next section we introduce the penalty method and give the existence theorem of a viscosity solution for the penalized equation. This problem is then discretized using finite differences.

5.2.1 Existence of solution for the penalized equation

In order to solve the problem (5.9)-(5.13) we use the penalty method, which we have presented in more detail in the Section 4.5.1.

Thus, instead of solving the system (5.9), we replace it by the following nonlinear PIDE

$$\frac{\partial v_k}{\partial t} + \mathcal{D}^Q[v_k] + \mathcal{L}^Q[v_k] - rv_k + \frac{1}{\epsilon} \max(\Upsilon_k^Q - v_k, 0) = 0, \quad (5.21)$$

where $\epsilon > 0$.

Suppose that for all $\epsilon > 0$ there exists a viscosity solution $v_\epsilon \in \mathcal{C}_I$ for the problem (5.21). We construct upper and lower limits for the approximating v_ϵ , i.e.

$$\bar{v}(s, Q, t) = \limsup_{\substack{(y, O, l) \rightarrow (s, Q, t) \\ \epsilon \rightarrow \infty}} v_\epsilon(s, Q, t), \quad \underline{v}(s, Q, t) = \liminf_{\substack{(y, O, l) \rightarrow (s, Q, t) \\ \epsilon \rightarrow \infty}} v_\epsilon(s, Q, t). \quad (5.22)$$

Theorem 5.2.1. *For all $\epsilon > 0$ there exists a viscosity solution v_ϵ for the problem (5.21) fulfilling*

1. v_ϵ is locally bounded, uniformly with respect to ϵ .
2. The expression under the integral term in $\mathcal{L}^Q[v_k]$ is μ -bounded².

Then \bar{v} and \underline{v} are respectively sub- and supersolutions for the problem (5.9).

Proof: we follow the lines of the proof presented by Amadori [2] in the case of an obstacle problem and apply these results to the general swing option.

From the hypothesis 2 we get that $\bar{v} \in USC_I$ and $\underline{v} \in LSC_I$, where USC_I and LSC_I were defined in (2.38)-(2.39).

We show next that \bar{v} is a subsolution for (5.9). Let $(s, Q, t) \in \Omega$ with $\bar{v}(s, Q, t) > \Upsilon^Q(s)$, $(p_0, p_1, p_2, M) \in \mathcal{P}_Q^{2,+} \bar{v}(s, Q, t)$, and we want to show that

$$\min \{-p_1 + r\bar{v}(s, Q, t) - \mathcal{D}_{p_0}[s, Q, t, p_0, p_2, M] - \mathcal{L}[\bar{v}(s, Q, t)], \bar{v}(s, Q, t) - \Upsilon_k(s)\} \leq 0, \quad (5.23)$$

as in Lemma 5.2.1.

Following Amadori [2] we can find a sequence (s_n, Q_n, t_n) converging to (s, Q, t) , as $\epsilon_n \rightarrow \infty$ and $(p_{0,n}, p_{1,n}, p_{2,n}, M_n) \in \mathcal{P}_Q^{2,+} v_{\epsilon_n}(s_n, Q_n, t_n)$ such that $(v_{\epsilon_n}(s_n, Q_n, t_n), p_{0,n}, p_{1,n}, p_{2,n}, M_n)$ approximates $(\bar{v}(s, Q, t), p_0, p_1, p_2, M)$. In particular we may assume that $v_{\epsilon_n}(s_n, Q_n, t_n) > \Upsilon^Q(s_n)$. For all $t_n > 0$ and since v_{ϵ_n}

²See Definition 2.5.1.

are subsolutions for (5.21) we have that

$$\min \{-p_{1,n} + rv_{\epsilon_n}(s_n, Q_n, t_n) - \mathcal{D}_{p_0}[s_n, Q_n, t_n, p_{0,n}, p_{2,n}, M_n] - \mathcal{L}[v_{\epsilon_n}(s_n, Q_n, t_n)], \\ v_{\epsilon_n}(s_n, Q_n, t_n) - \Upsilon^Q(s_n)\} + \frac{1}{\epsilon_n}(\Upsilon^Q(s_n) - v_{\epsilon_n}(s_n, Q_n, t_n))^+ \leq 0.$$

Passing to limit in the above inequation we get that (5.23) holds.

Next we prove that \underline{v} is a supersolution of (5.9). If $\underline{v}(s, Q, t) \geq \Upsilon^Q(s)$ we may conclude as in the previous step. Otherwise, if $\underline{v}(s, Q, t) < \Upsilon^Q(s)$ we have to show that $\mathcal{P}_Q^{2,-} \underline{v}(s, Q, t) = \emptyset$.

We assume by contradiction that there exists $(p_0, p_1, p_2, M) \in \mathcal{P}_Q^{2,-} \underline{v}(s, Q, t)$ such that we may approximate $(\underline{v}(s, Q, t), p_0, p_1, p_2, M)$ by means of a sequence $(v_{\epsilon_n}(s_n, Q_n, t_n), p_{0,n}, p_{1,n}, p_{2,n}, M_n)$, with $(p_{0,n}, p_{1,n}, p_{2,n}, M_n) \in \mathcal{P}_Q^{2,-} v_{\epsilon_n}(s_n, Q_n, t_n)$ and v_{ϵ_n} solves (5.21). Moreover, since $\underline{v}(s, Q, t) < \Upsilon^Q(s, Q, t)$ we may assume without loss of generality that

$$v_{\epsilon_n}(s_n, Q_n, t_n) \leq \Upsilon^Q(s_n) - \frac{1}{2}(\Upsilon^Q(s) - \underline{v}(s, Q, t)).$$

Thus we get

$$\min \{-p_{1,n} + rv_{\epsilon_n}(s_n, Q_n, t_n) - \mathcal{D}_{p_0}[s_n, Q_n, t_n, p_{0,n}, p_{2,n}, M_n] - \mathcal{L}[v_{\epsilon_n}(s_n, Q_n, t_n)], \\ v_{\epsilon_n}(s_n, Q_n, t_n) - \Upsilon^Q(s_n)\} + \frac{1}{\epsilon_n}(\Upsilon^Q(s_n) - v_{\epsilon_n}(s_n, Q_n, t_n))^+ \geq \\ \geq \frac{1}{\epsilon_n}(\Upsilon^Q(s_n) - v_{\epsilon_n}(s_n, Q_n, t_n))^+ \geq \frac{1}{2\epsilon_n}(\Upsilon^Q(s) - \underline{v}(s, Q, t)).$$

By passing to limit we obtain the contradiction.

□

In the general swing pricing approach, we have two state variables (the spot price S_t and the volume Q) in the partial differential formulation. Using the method of characteristics, the two-factor problem (5.21) can be reduced to a one factor problem at each time. The true option value is obtained by using linear interpolation. Then the diffusion terms are discretized by finite differences. The idea of this approach comes from the optimal tree harvesting decision problem presented in Insley [55], which has a similar formulation to the HJB equation (5.9)-(5.13). The integral term is approximated by the recursion formula presented in Section 4.5.2.

5.2.2 Discretization

We follow the same steps described in Section 4.5.1.

- The infinite domain is reduced to a bounded domain:

$$[0, \infty) \times [0, Q_{max}] \times [0, T] \rightarrow [S_{min}, S_{max}] \times [0, Q_{max}] \times [0, T];$$

- The domain is discretized and an equidistant grid is generated: the S , Q and t -axis are divided into equally spaced nodes at ΔS , ΔQ , respectively Δt distance apart. In this way the (S, Q, t) plan is split into a mesh, with the grid points $(i\Delta S, n\Delta Q, j\Delta t) = (S_i, Q_n, t_j)$ with $i = 1, \dots, I$, $n = 1, \dots, N$, $j = 1, \dots, J$;
- We define³

$$v_{i,n}^j = v(S_i, Q_n, t_j) = v(i\Delta S, n\Delta Q, j\Delta t)$$

the solution of the swing option at the spot price node S_i , for the accumulated volume Q_n and time t_j .

The discretization of the equation (5.21) can be simplified by applying the method of characteristics to $\frac{\partial v}{\partial t} - p \frac{\partial v}{\partial Q}$. This method is discussed in the next section.

The method of characteristics

In general, having more than one state variable considerably complicates the problem and the estimation of v . However, through the method of characteristics we are able to simplify the solution. For more details about this approach we refer to Morton and Mayers [80].

Consider some function $U(X, t)$ with

$$\frac{dU}{dt} = \frac{\partial U}{\partial t} + \frac{\partial U}{\partial X} \frac{dX}{dt}. \quad (5.24)$$

If U satisfies the equation

$$\frac{\partial U}{\partial t} + a(X, t) \frac{\partial U}{\partial X} = 0, \quad (5.25)$$

then from (5.24) we have that $\frac{dU}{dt} = 0$ along the characteristic curves defined by $\frac{dX}{dt} = a(X, t)$. If we consider the simple case where $a(X, t) = ct = a$, then the solution to (5.25) is

$$U(X, t) = U(X - at, 0). \quad (5.26)$$

This can be verified by taking the total derivative of $U(X - at, t)$ and observing that $dU = 0$ when $X_t = a$ and $t = 0$.

In the case when $a(X, t) \neq ct$ we can approximate equation (5.26) in discrete time by

$$\frac{U(X_n, t_{j+1}) - U(X_n - a(X_n, t_{j+1})\Delta t, t_j)}{\Delta t} + \mathcal{O}(\Delta t) = 0. \quad (5.27)$$

³From now on we drop the indices k from the value function v_k , not to cause any confusion with the discretization indices.

Now we rewrite the penalized PIDE for our problem in a more convenient way

$$\begin{aligned} \frac{\partial v}{\partial t} - p \frac{\partial v}{\partial Q} &= (r + \lambda)v - \alpha(\rho^*(t) - \ln S)S \frac{\partial v}{\partial S} - \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 v}{\partial S^2} - p(K - S)^+ - \\ &\quad - \lambda \int_{\mathbb{R}_+} v(sz, Q, t) f_{e^J}(z) dz - \frac{1}{\epsilon} \max(\Upsilon^Q - v, 0). \end{aligned} \quad (5.28)$$

Note that the left hand of the equation (5.28) looks similar to the left hand of the equation (5.27), considering $a(X, t) = -p$, $X = Q$ and $U = v$. Moreover, the right hand side of the equation (5.28) has only derivatives with respect to S . This observation allows us to approximate the two factor nonlinear equation (5.21) by solving a series of one dimensional PIDEs. These equations exchange information at each time step throughout a linear interpolation operation.

Within each timestep, the problem (5.21) is solved using the characteristic approach for a fixed Q_n . Applying as in the previous chapter the θ -scheme discretization, we get

$$\begin{aligned} &\frac{v(S_i, Q_n, t_{j+1}) - v(S_i, Q_n + p\Delta t, t_j)}{\Delta t} + \\ &+ \theta \left[\frac{\sigma^2 S_i^2}{2} \frac{v_{i+1,n}^j - 2v_{i,n}^j + v_{i-1,n}^j}{(\Delta S)^2} + \alpha(t_j) S_i \frac{v_{i+1,n}^j - v_{i-1,n}^j}{2\Delta S} - (r + \lambda)v_{i,n}^j + \right. \\ &\quad \left. + p(K - S_i)^+ + \lambda \text{Jump} v_{i,n}^j \right] + \\ &+ (1 - \theta) \left[\frac{\sigma^2 S_i^2}{2} \frac{v_{i+1,n}^{j+1} - 2v_{i,n}^{j+1} + v_{i-1,n}^{j+1}}{(\Delta S)^2} + \alpha(t_{j+1}) S_i \frac{v_{i+1,n}^{j+1} - v_{i-1,n}^{j+1}}{2\Delta S} - \right. \\ &\quad \left. - (r + \lambda)v_{i,n}^{j+1} + p(K - S_i)^+ + \lambda \text{Jump} v_{i,n}^{j+1} \right] + \frac{1}{\epsilon} (\Upsilon_i^Q - v_i^j) = 0, \end{aligned} \quad (5.29)$$

where $\text{Jump} v_i$ is the approximation of the integral term (see Section 4.5.2).

Rearranging the terms, equation (5.29) can be rewritten as

$$\begin{aligned} &v(S_i, Q_n, t_{j+1}) - v(S_i, Q_n + p\Delta t, t_j) = \\ &-v_{i+1,n}^j \theta \Delta t a_i - v_{i-1,n}^j \theta \Delta t b_i + v_{i,n}^j \theta \Delta t (a_i + b_i + r + \lambda) - \\ &-v_{i+1,n}^{j+1} (1 - \theta) \Delta t a_i - v_{i-1,n}^{j+1} (1 - \theta) \Delta t b_i + v_{i,n}^{j+1} (1 - \theta) \Delta t (a_i + b_i + r + \lambda) - \\ &-p\Delta t (K - S_i)^+ - \theta \Delta t \lambda \text{Jump} v_{i,n}^j - (1 - \theta) \Delta t \lambda \text{Jump} v_{i,n}^{j+1} - \frac{\Delta t}{\epsilon} (\Upsilon_i^Q - v_i^j), \end{aligned} \quad (5.30)$$

where a_i and b_i were defined in (4.64)-(4.66) and they satisfy the positive coefficient condition (4.67). At the missing points $i \in \{1, I\}$ and $n \in \{1, N\}$ we impose Dirichlet boundary conditions as before.

Let Φ be a linear operator of interpolation (of any order) such that

$$(\Phi v^j)_{i,n} = v(S_i, Q_n + p\Delta t, t_j) + \text{interpolation error}. \quad (5.31)$$

Then we can write the discrete equation (5.30) in matrix form

$$\begin{aligned} & [\mathcal{I} - (1 - \theta)\Delta t A] v^{j+1} + (1 - \theta)\Delta t \lambda \text{Jump} v^{j+1} = \\ & = \Phi v^j + \theta \Delta t A v^j - \theta \Delta t \lambda \text{Jump} v^j - p \Delta t g_i - P^j(\Upsilon^Q - v^j), \end{aligned}$$

where A was defined in (4.80), Jump is the jump matrix, g_i is the payoff vector $(K - S)^+$ and P^j is the penalty matrix given by

$$(P^j)_{i,i} = \begin{cases} \frac{1}{\epsilon}, & \text{if } v^j < \Upsilon^Q \\ 0, & \text{otherwise.} \end{cases} \quad (5.32)$$

In order to avoid algebraic complication, we only use the implicit discretization ($\theta = 1$) in this section. In this case, the penalized discretization takes the following form

$$-v^{j+1} + \Phi v^j + \Delta t D v^j - P^j(\Upsilon^Q - v^j) - p \Delta t g_i = 0, \quad (5.33)$$

where $D = A - \lambda \text{Jump}$.

Following Insley et. al. [56] we can give the following important remark.

Remark 5.2.1. *If we use linear interpolation, the interpolation matrix Φ has the property that its entries are non-negative and all row sums are 1.*

5.2.3 Convergence analysis

As the equation (5.33) is nonlinear, the solution of the (5.21) may not be unique. Thus, it is important to ensure that the numerical scheme converges to the unique viscosity solution.

If a strong comparison result holds for the PIDE (5.21), Barles [5] and Briani et. al. [21] proved that, if the numerical scheme is also stable, consistent and monotone then it converges to the unique viscosity solution. We show next that the implicit discretization scheme (5.33) satisfies these properties.

Recall the notations from Chapter 4, $\tilde{v} = (v_i^j, v_i^{j+1}, v_{\gamma_i}^j, v_{\gamma_i}^{j+1})$ and $\gamma_i \in \{i-1, i+1\}$.

Proposition 5.2.3. *Monotonicity*

The fully implicit discretization of the scheme (5.33) is monotone independent of the choice of Δt .

Proof: observing that

- D is a M-matrix (see Remark 4.5.2), therefore $\Delta t D v_{i,n}^j$ is an increasing function of $v_{i,n}^j$, and a decreasing function of $v_{\gamma_i,n}^j$,
- $\Phi v_{i,n}^j$ is an increasing function of $v_{i,n}^j$ (see Remark 5.2.1),
- the penalty matrix P^j is positive, thus $-P^j(\Upsilon^Q - v^j)$ is an increasing function of $v_{i,n}^j$,

we get that the discretization (5.33) is monotone based on Definition 4.5.1.

□

Proposition 5.2.4. Stability

The scheme (5.33) is unconditionally stable for any $\Delta S, \Delta t > 0$.

Proof: for this proof we use the equivalent discrete formulation of (5.33). Denote by \mathcal{SD}^Q the scheme without jumps

$$\begin{aligned} \mathcal{SD}^Q = & -v_{i,n}^{j+1} + \sum_{i,n} w_{i,n} v_{i,n}^j + (a_i + b_i + r + \lambda) \Delta t v_{i,n}^j - a_i \Delta t v_{i+1,n}^j - \\ & -b_i \Delta t v_{i-1,n}^j - p \Delta t (K - S_i)^+ - P^j (\Upsilon_{i,n}^Q - v^j), \end{aligned} \quad (5.34)$$

where $w_{i,n}$ are the linear interpolation weights.

Chen and Insley [27] proved in the case of the tree harvesting problem, that the stability property is equivalent to showing that

$$\|v^j\|_\infty \leq \max \{ \|v^{j+1}\|_\infty, \|\Upsilon^Q\|_\infty \}. \quad (5.35)$$

Let k be an index such that $|v_{k,n}^j| = \|v_n^j\|_\infty$. From (5.34) we have

$$\|v_n^j\|_\infty (1 + (r + \lambda) \Delta t + (P^j)_{k,k}) \leq \|v^{j+1}\|_\infty + (P^j)_{k,k} \|\Upsilon^Q\| \quad (5.36)$$

and rearranging we get

$$\|v_n^j\|_\infty \leq \max \{ \|v^j\|_\infty, \|\Upsilon^Q\| \} \frac{1 + (P^j)_{k,k}}{1 + (r + \lambda) \Delta t + (P^j)_{k,k}}. \quad (5.37)$$

Hence we get the desired inequality

$$\|v^j\|_\infty \leq \max \{ \|v^{j+1}\|_\infty, \|\Upsilon^Q\|_\infty \}. \quad (5.38)$$

Now together with the monotonicity of the integral term (Proposition 4.5.4) we obtain the stability of the scheme (5.33).

□

Next we introduce the following operator

$$\begin{aligned} \mathcal{G}^Q(\tilde{v}, \text{Jump}v^j) = & \frac{v(S_i, Q_n, t_{j+1}) - v(S_i, Q_n + p\Delta t, t_j)}{\Delta t} + \\ & + \left[\frac{\sigma^2 S_i^2}{2} \frac{v_{i+1,n}^j - 2v_{i,n}^j + v_{i-1,n}^j}{(\Delta S)^2} + \alpha(t_j) S_i \frac{v_{i+1,n}^j - v_{i-1,n}^j}{2\Delta S} - (r + \lambda)v_{i,n}^j + \right. \\ & \left. + p(K - S_i)^+ + \lambda \text{Jump}v_{i,n}^j \right] + \frac{1}{\epsilon} (\Upsilon_i^Q - v_i^j). \end{aligned}$$

Proposition 5.2.5. Consistency

The finite difference scheme \mathcal{G}^Q is locally consistent if we have

$$\left| \mathcal{G}^Q(\tilde{v}, \text{Jump}v^j) - \left(\frac{\partial v}{\partial t} + \mathcal{D}^Q[v] + \mathcal{L}^Q[v] - (r + \lambda)v + \frac{1}{\epsilon} \max(\Upsilon^Q - v, 0) \right) (S, Q, t) \right| \rightarrow 0$$

when $(\Delta S, \Delta Q, \Delta t) \rightarrow 0$ and $(S_i, Q_n, t_j) \rightarrow (S, Q, t)$.

Proof: the proof follows from (5.27) and Proposition 4.5.6. □

Theorem 5.2.2. Convergence of the finite difference scheme

If the scheme (5.33) is monotone, stable, consistent, the approximation of the integral is monotone and the equation (5.21) satisfies the strong comparison result, then as $(\Delta S, \Delta Q, \Delta t) \rightarrow 0$, the solution v of the scheme (5.33) converges locally uniformly to the unique continuous viscosity solution of the problem (5.21).

Proof: we give the main ideas of the proof. We follow the lines of the proof presented by Briani [20] for a simple European option under a jump-diffusion process.

The hypothesis of the theorem are satisfied according to Proposition 5.2.3, Proposition 5.2.4, Proposition 5.2.5, respectively Proposition 4.5.4. Moreover from Assumption 5.2.1 the equation (5.21) satisfies the strong comparison result.

Then let \underline{v} and \bar{v} be defined by

$$\underline{v}(s, Q, t) = \liminf_{\substack{(\Delta S, \Delta Q, \Delta t) \rightarrow 0 \\ (i\Delta S, n\Delta Q, j\Delta t) \rightarrow (s, Q, t)}} v_{i,n}^j, \quad (5.39)$$

$$\bar{v}(s, Q, t) = \limsup_{\substack{(\Delta S, \Delta Q, \Delta t) \rightarrow 0 \\ (i\Delta S, n\Delta Q, j\Delta t) \rightarrow (s, Q, t)}} v_{i,n}^j. \quad (5.40)$$

We first prove that \bar{v} and \underline{v} are respectively sub- and supersolutions of the problem (5.33). Thus, we will be able to conclude that $v = \bar{v} = \underline{v}$ and together with the equations (5.39) and (5.40) we will get local uniform convergence.

We only give the proof for \bar{v} , the one for \underline{v} is similar.

We want to prove that $\bar{v} \in USC_I$ and that it is a subsolution to our problem, i.e. for all $\varphi \in C^{2,1,1}(\bar{\Omega}) \cap C_2(\bar{\Omega})$ such that $v_k - \varphi$ has a global maximum at $(\bar{s}, \bar{Q}, \bar{t}) \in \Omega$, we have

$$\begin{aligned} & \frac{\partial \varphi}{\partial t}(\bar{s}, \bar{Q}, \bar{t}) + \mathcal{D}^Q[\varphi](\bar{s}, \bar{Q}, \bar{t}) + \mathcal{L}^Q[\varphi](\bar{s}, \bar{Q}, \bar{t}) - r\varphi(\bar{s}, \bar{Q}, \bar{t}) + \\ & + \frac{1}{\epsilon} \max(\Upsilon_k^Q(\bar{s}) - \varphi(\bar{s}, \bar{Q}, \bar{t}), 0) \leq 0. \end{aligned}$$

In order to show that $\bar{v} \in USC_I$, we have to prove that \bar{v} satisfies the three conditions from (2.38):

1. \bar{v} is upper semicontinuous, i.e.

$$\limsup_{(y, O, l) \rightarrow (s, Q, t)} \bar{v}(y, O, l) \leq \bar{v}(s, Q, t), \quad (5.41)$$

where $(y, O, l) \in \bar{\Omega}$. By definition of \bar{v} we have that

$$\bar{v}(y, O, l) = \limsup_{\substack{(\Delta S, \Delta Q, \Delta t) \rightarrow 0 \\ (i\Delta S, n\Delta Q, j\Delta t) \rightarrow (y, O, l)}} v_{i,n}^j,$$

therefore, it exists $\epsilon > 0$ and (i, n, j) such that

$$\bar{v}(y, O, l) - \epsilon \leq v_{i,n}^j$$

and taking \limsup for $(\Delta S, \Delta Q, \Delta t) \rightarrow 0$ and $(i\Delta S, n\Delta Q, j\Delta t) \rightarrow (s, Q, t)$ we get

$$\bar{v}(y, O, l) - \epsilon \leq \bar{v}(s, Q, t).$$

As ϵ was chosen arbitrary, we obtain (5.41).

2. \bar{v} is locally bounded. This condition is obtained by the solution of the scheme. From (5.2.4) we know that the solution is bounded, independently of $(\Delta S, \Delta Q, \Delta t)$. Thus, let $Y \in \bar{\Omega}$ be a compact set. Then there exists a constant A_Y such that

$$\begin{aligned} |v_{i,n}^j| &\leq A_Y, \quad \forall i, n, j \text{ such that } (i\Delta S, n\Delta Q, j\Delta t) \in Y \Rightarrow \\ &\Rightarrow |\bar{v}(s, Q, t)| \leq A_Y, \quad \forall (s, Q, t) \in Y. \end{aligned}$$

3. $M(\bar{v}(sz, Q, t), \bar{v}(s, Q, t))$ has an upper μ -bound in (s, Q, t) . In our case M is given by the difference under the integral term in $\mathcal{L}^Q[\bar{v}]$ which is clearly Lipschitz. Therefore, according to Definition 2.5.1, we have an upper μ -bound.

We still need to prove that \bar{v} is a viscosity subsolution. Let $(\bar{s}, \bar{Q}, \bar{t})$ be a strict maximum for $\bar{v} - \varphi$ on $\bar{\Omega}$ and $\varphi \in C^{2,1,1}(\bar{\Omega}) \cap C_2(\bar{\Omega})$.

We can assume that $\bar{v}(\bar{s}, \bar{Q}, \bar{t}) = \varphi(\bar{s}, \bar{Q}, \bar{t})$ and that

$$\bar{v}(s, Q, t) - \varphi(s, Q, t) \leq 0 = \bar{v}(\bar{s}, \bar{Q}, \bar{t}) - \varphi(\bar{s}, \bar{Q}, \bar{t}) \quad \text{in } \Omega.$$

Then using the consistency of the scheme (Proposition 5.2.5), the monotonicity (Proposition 5.2.3) and the monotonicity of the approximating integral (Proposition 4.5.4) we get the desired result

$$\begin{aligned} &\frac{\partial \varphi}{\partial t}(\bar{s}, \bar{Q}, \bar{t}) + \mathcal{D}^Q[\varphi](\bar{s}, \bar{Q}, \bar{t}) + \mathcal{L}^Q[\varphi](\bar{s}, \bar{Q}, \bar{t}) - r\varphi(\bar{s}, \bar{Q}, \bar{t}) + \\ &+ \frac{1}{\epsilon} \max(\Upsilon_k^Q(\bar{s}) - \varphi(\bar{s}, \bar{Q}, \bar{t}), 0) \leq 0. \end{aligned}$$

□

The solution to the discrete penalized problem (5.33) can be found using the generalized Newton iteration. We choose as before, the initial guess to be $v^0 = v^{j+1}$. Then the $(l + 1)$ -th iterant is given by

$$\widehat{v}^{l+1} = \left(\Phi + \theta \Delta t A + \widehat{P}^l \right)^{-1} \left[v^{j+1} + p \Delta t g_i + \theta \Delta t \lambda \text{Jump} \widehat{v}^l + \widehat{P}^l \Upsilon \right]. \quad (5.42)$$

Finally, we can give the result which shows that the iteration (5.42) converges to the unique viscosity solution of the equation (5.33).

Theorem 5.2.3. Convergence of the penalty iteration

Let A , Jump and \widehat{P}^j be given by (4.80), (4.76), respectively (5.32) and a_i, b_i satisfy the positive coefficient condition (4.67) for all i . Then the iteration (5.42) converges to the unique solution of (5.33) for any initial iterate \widehat{v}^0 .

Proof: we have showed in Section 5.2.2 that by using the method of characteristics, the problem (5.33) can be approximated by a sequence of one dimensional problems. Thus, it is trivial to show that the convergence of scheme (5.42) reduces to the convergence problem proved in Theorem 4.5.2.

□

Chapter 6

Conclusion

The aim of this thesis was to solve numerically the problem of pricing swing derivatives in the electricity market. We investigated two types of swing options: first, options that included a refraction period in their formulation and in the second part, swing options with variable volume.

The first question was to find a suitable stochastic model which could describe the electricity price dynamics and that could also be used for pricing swing derivatives. We proposed a mean-reverting double exponential jump diffusion model, which is new in the context of swing option valuation. In Chapter 3, we showed that this model is able to reproduce the observed characteristics of the power prices, like mean-reversion, seasonality and jumps. The seasonality function captured the annual and weekly patterns and it was estimated using data from two different energy markets. The jumps from this model followed a double exponential distribution, which reproduced the intensity and size of the positive and also of the negative jumps observed in the market prices.

The model calibration was performed from the deseasonalized data, using the Maximum Likelihood Estimation. The simulation results showed that the proposed model reproduces the observed electricity data accurately.

In Chapter 4, we solved the problem of valuating swing options with refraction times. Pricing of such options is very challenging, because they have no analytical solution, and thus, numerical methods have to be used. We showed that the pricing problem under the double exponential jump diffusion model could be written as a partial integro-differential complementarity problem, which we solved within the framework of viscosity solutions. We applied the penalty method, and transformed the PIDCP into a nonlinear partial integro-differential equation, which we discretized by finite differences. One main difficulty arose due to the jumps in the spot model, which introduced an integral term into the complementarity problem. However, by using the double exponential jump formulation, we were able to approximate this integral term by a recursion formula. These types of formulas were developed for the Kou model in the context of American option pricing. We have showed in this thesis that these recursive procedures can be also applied to our particular electricity price model, and they can be used for pricing swing derivatives. We were thus able to prove that the finite difference scheme is monotone,

stable and consistent. This led us to the convergence of the numerical scheme. Finally, we showed that the discretized penalized equation can be solved iteratively, as long as the convergence to the viscosity solution was guaranteed.

The numerical results showed that the pricing algorithm performs well. We computed the value of the swing option for different parameter values, and got similar results as Kluge [65] in his thesis.

The computation times are quite large, and thus our algorithm is not fast enough to be used in practice. However, as we used a simple MATLAB implementation, we believe that the speed of the algorithm can be improved by using a more performant programming language.

In the last part of the thesis the main question we wanted to answer was if the developed numerical approach presented before, could be also applied to swing options with variable energy quantities. The difficult part arised due to the new volume variable in the integro-differential complementarity formulation. In order to simplify this two dimensional problem, we applied the method of characteristics, which reduced this problem to solving a series of one dimensional partial-integro differential equations. In this way, we were able to apply the penalty method, and to prove the existence of the viscosity solution for the nonlinear penalized equation. Then we discretized the problem by finite differences, and we approximated the integral term by using the recursion formula. We solved the resulting problem iteratively and we showed that the penalty iteration is convergent.

This work could be extended in several ways. Firstly, a natural extension to the spot price model would be to include stochastic volatility, or to model the seasonality as a stochastic process. Thus, the model could be able to fit better the market data. However, it might be more difficult to apply the numerical approach proposed here for swing option pricing.

It would be also interesting to include penalty functions into the pricing problem. These penalties are applied if the overall energy volume purchased until maturity, exceeds the predefined quantity set in the contract. In this case, the optimal exercise strategy is not of a bang-bang type, and it should be determined additionally to the option value.

In practice, swing options have a large number of exercise rights, so it would be desirable to have faster computing times. As we stated before, our algorithm could be optimized, to accelerate the computational time. To our knowledge, at the moment only the Monte Carlo methods are used for valuing swing options in practice. However, these methods exhibit slow convergence, and it would be desirable to use numerical methods with a faster convergence and acceptable computing times.

Appendix A

Solution to the SDE (3.6)

In this section we present the main steps for finding the solution to the log-spot price equation

$$\begin{aligned}\ln S_t &= f(t) + X_t \\ dX_t &= -\alpha X_t dt + \sigma dW_t + J dN_t\end{aligned}\tag{A.1}$$

We first rewrite the log-price process in terms of X_t

$$X_t = \ln S_t - f(t).$$

By taking $Y_t = e^{\alpha t} X_t$, we can rewrite the last equation as

$$e^{\alpha t} dX_t = -\alpha e^{\alpha t} X_t dt + \sigma e^{\alpha t} dW_t + e^{\alpha t} J dN_t.$$

We apply Itô's lemma (2.29) to Y_t and show that

$$dY_t = e^{\alpha t} dX_t + \alpha e^{\alpha t} X_t dt \iff Y_t = Y_0 + \int_0^t e^{\alpha z} dX_z + \alpha \int_0^t e^{\alpha z} X_z dz.$$

Given that

$$Y_t = e^{\alpha t} X_t = Y_0 + \int_0^t e^{\alpha z} dX_z + \alpha \int_0^t e^{\alpha z} X_z dz,$$

we obtain the solution for X_t as

$$\begin{aligned}X_t &= e^{-\alpha t} X_0 + \sigma e^{-\alpha t} \int_0^t e^{\alpha s} dW_s + e^{-\alpha t} \int_0^t J e^{\alpha z} dN_z - \\ &\quad - e^{-\alpha t} \alpha \int_0^t X_z dz + \alpha e^{-\alpha t} \int_0^t e^{\alpha z} X_z dz.\end{aligned}$$

Finally, the solution for the process X_t can be written as

$$X_t = e^{-\alpha t} X_0 + \sigma e^{-\alpha t} \int_0^t e^{\alpha z} dW_z + e^{-\alpha t} \int_0^t J e^{\alpha z} dN_z\tag{A.2}$$

and in terms of log-prices

$$\ln S_t = f(t) + e^{-\alpha t} (\ln S_0 - f(0)) + \sigma e^{-\alpha t} \int_0^t e^{\alpha z} dW_z + e^{-\alpha t} \int_0^t J e^{\alpha z} dN_z.\tag{A.3}$$

Expectation and variance

Applying the expectation operator \mathbb{E} to equation (A.3) we get

$$\mathbb{E}[\ln S_t] = f(t) + e^{-\alpha t}(\ln S_0 - f(0)) + \sigma e^{-\alpha t} \mathbb{E} \left[\int_0^t e^{\alpha z} dW_z \right] + e^{-\alpha t} \mathbb{E} \left[\int_0^t J e^{\alpha z} dN_z \right].$$

From the general properties of a Wiener process (see Section 2.2.1) we have that the first integral disappears. For the second integral we use the results from Section 2.4, and we define

$$M_t = N_t - \lambda t. \quad (\text{A.4})$$

From Theorem 2.4.1 we have that M_t is a martingale. Thus, we get

$$J \int_0^t e^{\alpha z} dM_z = J \int_0^t e^{\alpha z} (dN_z - \lambda dz) \quad (\text{A.5})$$

is a martingale as well. That means

$$J \int_0^t e^{\alpha z} (dN_z - \lambda dz) = 0,$$

and thus,

$$J \int_0^t e^{\alpha z} dN_z = J \int_0^t e^{\alpha z} \lambda dz = 0.$$

Taking in account that J and N_t are independent, we have

$$\mathbb{E} \left[\ln J \int_0^t e^{\alpha z} dN_z \right] = \mathbb{E}[J] \mathbb{E} \left[\int_0^t e^{\alpha z} dN_z \right] = \left(\frac{p}{\eta_1} - \frac{q}{\eta_2} \right) \frac{\lambda}{\alpha} (e^{\alpha t} - 1),$$

where for the computation of $\mathbb{E}[J]$ we used the double exponential density function

$$f_J(z) = p\eta_1 e^{-\eta_1 z} \mathbb{1}_{\{z \geq 0\}} + q\eta_2 e^{\eta_2 z} \mathbb{1}_{\{z < 0\}}. \quad (\text{A.6})$$

Then, we obtain the expectation of the process $\ln S_t$

$$\mathbb{E}[\ln S_t] = f(t) + e^{-\alpha t}(\ln S_0 - f(0)) + \left(\frac{p}{\eta_1} - \frac{q}{\eta_2} \right) \frac{\lambda}{\alpha} (1 - e^{-\alpha t}). \quad (\text{A.7})$$

Next we derive the variance of the process (A.3) in a similar way.

$$\text{Var}[\ln S_t] = e^{-2\alpha t} \sigma^2 \text{Var} \left[\int_0^t e^{\alpha z} dW_z \right] + e^{-2\alpha t} \text{Var} \left[\int_0^t J e^{\alpha z} dN_z \right]. \quad (\text{A.8})$$

Using once again the properties from Section 2.2.1 we have

$$\text{Var} \left[\int_0^t e^{\alpha z} dW_z \right] = \frac{1}{2\alpha} (e^{2\alpha t} - 1).$$

For the second integral in (A.8) we have

$$\text{Var} \left[\int_0^t \ln J e^{\alpha z} dN_z \right] = \mathbb{E} \left[\left(\int_0^t \ln J e^{\alpha z} dN_z \right)^2 \right] - \left(\mathbb{E} \left[\int_0^t \ln J e^{\alpha z} dN_z \right] \right)^2.$$

Using the properties introduced in Section 2.4 we have that $\mathbb{E}[J^2] = \text{Var}[J] + \mathbb{E}^2[J]$ and $(dN_t)^2 = \lambda dt$, and so we get

$$\mathbb{E} \left[\left(\int_0^t J e^{\alpha z} dN_z \right)^2 \right] = \mathbb{E}[(J)^2] \mathbb{E} \left[\int_0^t e^{2\alpha z} \lambda dz \right] = \left(\frac{2p}{\eta_1^2} + \frac{2q}{\eta_2^2} \right) \frac{\lambda}{\alpha} (e^{2\alpha t} - 1),$$

and

$$\left(\mathbb{E} \left[\int_0^t J e^{\alpha z} dN_z \right] \right)^2 = \left(\frac{p}{\eta_1} - \frac{q}{\eta_2} \right)^2 \left(\frac{\lambda}{\alpha} \right)^2 (e^{\alpha t} - 1)^2.$$

Thus, the variance of the process $\ln S_t$ is

$$\begin{aligned} \text{Var}[\ln S_t] &= \frac{\sigma^2}{2\alpha} (1 - e^{-2\alpha t}) + \left(\frac{2p}{\eta_1^2} + \frac{2q}{\eta_2^2} \right) \frac{\lambda}{\alpha} (1 - e^{-2\alpha t}) + \\ &\quad + \left(\frac{p}{\eta_1} - \frac{q}{\eta_2} \right)^2 \left(\frac{\lambda}{\alpha} \right)^2 (1 - e^{-\alpha t})^2. \end{aligned} \tag{A.9}$$

A.1 Positive coefficient condition algorithm

- If $a_{i,c} \geq 0$ and $b_{i,c} \geq 0$,
then $a_i = a_{i,c}$.
- ElseIf $b_{i,f} \geq 0$,
then $a_i = a_{i,f}$.
- Else $a_i = a_{i,b}$.

Discretization functions (consistency)

$$\begin{aligned}
\mathcal{G}_i^j(\tilde{v}, \text{Jump}v_i^j, \theta) &= \frac{v_i^{j+1} - v_i^j}{\Delta t} + \theta \left[\frac{\sigma^2 S_i^2}{2} \frac{v_{i+1}^j - 2v_i^j + v_{i-1}^j}{(\Delta S)^2} + \right. \\
&+ \left. \alpha(t_j) S_i \frac{v_{i+1}^j - v_{i-1}^j}{2\Delta S} - (r + \lambda)v_i^j + \lambda \text{Jump}v_i^j \right] + \\
&+ (1 - \theta) \left[\frac{\sigma^2 S_i^2}{2} \frac{v_{i+1}^{j+1} - 2v_i^{j+1} + v_{i-1}^{j+1}}{(\Delta S)^2} + \alpha(t_{j+1}) S_i \frac{v_{i+1}^{j+1} - v_{i-1}^{j+1}}{2\Delta S} \right. \\
&- \left. (r + \lambda)v_i^{j+1} + \text{Jump}v_i^j \right] + \frac{1}{\epsilon} \max(\Upsilon_i - v_i^j, 0).
\end{aligned}$$

$$\begin{aligned}
\mathcal{GD}_i^j(\tilde{v}, \theta) &= \theta \left[\frac{\sigma^2 S_i^2}{2} \frac{v_{i+1}^j - 2v_i^j + v_{i-1}^j}{(\Delta S)^2} + \alpha(t_j) S_i \frac{v_{i+1}^j - v_{i-1}^j}{2\Delta S} \right. \\
&- \left. (r + \lambda)v_i^j \right] + (1 - \theta) \left[\frac{\sigma^2 S_i^2}{2} \frac{v_{i+1}^{j+1} - 2v_i^{j+1} + v_{i-1}^{j+1}}{(\Delta S)^2} + \right. \\
&+ \left. \alpha(t_{j+1}) S_i \frac{v_{i+1}^{j+1} - v_{i-1}^{j+1}}{2\Delta S} - (r + \lambda)v_i^{j+1} \right].
\end{aligned}$$

$$\begin{aligned}
\mathcal{GJ}_i^j(\text{Jump}v_i^j) &= \theta \lambda \text{Jump}v_i^j + (1 - \theta) \lambda \text{Jump}v_i^{j+1}. \\
\mathcal{GP}_i^j(P^j) &= \frac{1}{\epsilon} \max(\Upsilon_i - v_i^j, 0).
\end{aligned}$$

Appendix B

The determination of the PIDCP with volume constraints

Next we develop the LCP for the general swing option problem. For simplicity, we work under a model which does not include jumps and derive the partial differential pricing equation. Due to the volume variable in the problem formulation, we have to deal with a two state variable partial differential equation.

We use the delta hedging procedure, described in Chapter 2. In order to find the arbitrage-free price of the option we need to introduce another option $V_1(S, t)$, with the same underlying S , and expiration date T_1 . Then we set up a self-financing hedged portfolio containing Δ_1 options with expiration date T_1 . For more details about this procedure we refer to Björk [17].

The portfolio's value at time t is given by

$$\Pi_t = v_k - \Delta_1 V_1.$$

The variation of the value function is given by applying Itô's lemma for the multivariate case (2.5) to $v_k(S, Q, t)$, plus the cash flow acquired in the time interval $(t, t + \delta_R)$

$$dv_k = \frac{\partial v_k}{\partial t} dt + \frac{1}{2} \sigma^2 S^2 \frac{\partial^2 v_k}{\partial S^2} dt + \frac{\partial v_k}{\partial S} dS + \frac{\partial v_k}{\partial Q} dQ + (K - S)^+ dQ. \quad (\text{B.1})$$

In order to determine dQ we first recall that the holder of the option has N_s swing rights, and consequently there are maximum $N_s \delta_R$ refraction periods during $[0, T]$. Thus, during the interval $(t, t + \delta_R)$ we have

$$Q(t + \delta_R) = Q(t) + p \delta_R, \quad p \in \{p_{min}, p_{max}\}$$

Moving to continuous time formulation and taking the limit $\delta_R \rightarrow 0$ we have

$$dQ = p dt. \quad (\text{B.2})$$

Now the change in the portfolio is given by

$$\begin{aligned} d\Pi = & \frac{\partial v_k}{\partial t} dt + \frac{1}{2} \sigma^2 S^2 \frac{\partial^2 v_k}{\partial S^2} dt + \frac{\partial v_k}{\partial S} dS + p \frac{\partial v_k}{\partial Q} dt + p(K - S)^+ dt - \\ & - \Delta_1 \left(\frac{\partial V_1}{\partial t} dt + \frac{1}{2} \sigma^2 S^2 \frac{\partial^2 V_1}{\partial S^2} dt + \frac{\partial V_1}{\partial S} dS \right). \end{aligned}$$

In this equation we have an unpredictable component due to dW_t in dS_t . In order to remove this term, we choose

$$\Delta_1 = \frac{\partial v_k}{\partial S} / \frac{\partial V_1}{\partial S} = \frac{\partial_S v_k}{\partial_S V_1}. \quad (\text{B.3})$$

With this choice of the hedging, the portfolio becomes risk-free and by the absence of arbitrage, it cannot earn more or less than the risk-free rate r

$$d\Pi_t = r\Pi_t dt, \quad (\text{B.4})$$

and substituting the value of Δ_1 from (B.3), we get

$$\begin{aligned} & \frac{\frac{\partial v_k}{\partial t} + \frac{\sigma^2 S^2}{2} \frac{\partial^2 v_k}{\partial S^2} + \alpha(\rho(t) - \ln S) S \frac{\partial v_k}{\partial S} + p \frac{\partial v_k}{\partial Q} + p(K - S)^+ - r v_k}{\partial_S v_k} = \\ & = \frac{\frac{\partial V_1}{\partial t} + \frac{\sigma^2 S^2}{2} \frac{\partial^2 V_1}{\partial S^2} + \alpha(\rho(t) - \ln S) S \frac{\partial V_1}{\partial S} - r V_1}{\partial_S V_1}. \end{aligned}$$

As v_k and V_1 are an arbitrary pair of derivative contracts, both sides of the equation are independent of the maturity date and hence equal a function dependent only on t and S . Thus, we can write

$$\frac{\frac{\partial v_k}{\partial t} + \frac{\sigma^2 S^2}{2} \frac{\partial^2 v_k}{\partial S^2} dt + \alpha(\rho(t) - \ln S) S \frac{\partial v_k}{\partial S} + p \frac{\partial v_k}{\partial Q} + p(K - S)^+ - r v_k}{\partial_S v_k} = -\hat{\alpha}(t)$$

where $\hat{\alpha}(t)$ is the drift of the price process in the risk-neutral world, see (3.12). Therefore the delta-hedging argument shows that the market price of risk is the same for all contingent claims depending on S .

Multiplying by $\partial_S v_k$ and rearranging we get

$$\frac{\partial v_k}{\partial t} + \frac{\sigma^2 S^2}{2} \frac{\partial^2 v_k}{\partial S^2} dt + \hat{\alpha}(t) S \frac{\partial v_k}{\partial S} + p \left(\frac{\partial v_k}{\partial Q} + (K - S)^+ \right) - r v_k = 0. \quad (\text{B.5})$$

Thus the linear complementarity formulation for the swing value under a simple mean-reverting process yields¹

$$\begin{cases} \left(\frac{\partial v_k}{\partial t} + \mathcal{D}^Q[v_k] - r v_k \right) \left(v_k(s, t) - \Upsilon_k^Q(s, Q, t) \right) = 0 \\ \frac{\partial v_k}{\partial t} + \mathcal{D}^Q[v_k] - r v_k \leq 0 & (s, Q, t) \in \bar{\Omega} \\ v_k(s, Q, t) \geq \Upsilon_k^Q(s, Q, t) \\ v_k(s, Q, T) = q(T)(K - S_T)^+ \end{cases} \quad (\text{B.6})$$

where

$$\mathcal{D}^Q[v_k] = \alpha(\rho(t) - \ln S) S \frac{\partial v_k}{\partial S} + \frac{1}{2} \sigma^2 S^2 \frac{\partial^2 v_k}{\partial S^2} + p \left(\frac{\partial v_k}{\partial Q} + (K - S)^+ \right).$$

¹We use the original notations for the parameters, but remark that they might have different values in the risk neutral world, as discussed in Section 3.2.1.

Appendix C

Abbreviations and notations

| | | |
|----------------|---|--|
| DPP | – | Dynamic Programming Principle |
| HJB | – | Hamilton-Jacobi-Bellman |
| OU | – | Ornstein-Uhlenbeck process |
| PDE | – | Partial Differential Equation |
| PIDE | – | Partial Integro-Differential Equation |
| SDE | – | Stochastic Differential Equation |
| a.a. | – | almost all |
| a.s. | – | almost surely |
| e.g. | – | for example |
| et. al. | – | et alia (and others) |
| etc. | – | et cetera (and so on) |
| i.e. | – | that is |
| i.i.d. | – | independent identically distributed |
| ln | – | natural logarithm |
| std. | – | standard deviation |
| resp. | – | respectively |
| f^+ | = | $\max(f, 0)$ |
| $x \vee y$ | = | $\max(x, y)$ |
| $x \wedge y$ | = | $\min(x, y)$ |
| \mathbb{R}_+ | = | $[0, \infty)$ |
| $\mathbb{1}_A$ | – | indicator function, with $\mathbb{1}_A(x) = \begin{cases} 1 & \text{if } x \in A, \\ 0, & \text{if } x \notin A \end{cases}$ |
| $C^0[a, b]$ | – | set of continuous functions on $[a, b]$ |
| $C^1[a, b]$ | – | set of continuously differentiable functions on $[a, b]$ |
| \mathbb{E} | – | expectation value |
| \mathbb{P} | – | probability measure |
| Var | – | variance |

Bibliography

- [1] Alvarez O., Tourin A. - *Viscosity Solutions of Nonlinear Integro-Differential Equations*, Annales de l'institut Henri Poincaré (C) Analyse non linéaire, 3, 293-317, 1996;
- [2] Amadori, A. L. - *The Obstacle Problem for Nonlinear Integro-Differential Equations Arising in Option Pricing*, Working paper, Istituto pre le Applicazione del Calcolo "M. Picone", 2000;
- [3] Amadori, A. L. - *Nonlinear Integro-Differential Evolution Problems Arising in Option Pricing: a Viscosity Solutions Approach*, Differential and Integral Equations, 16, 2003;
- [4] Bardou O., Bouthemy S., Pagés G. - *Optimal Quantization for the Pricing of Swing Options*, Applied Mathematical Finance 16, 183-217, 2009;
- [5] Barles G. - *Convergence of Numerical Schemes for Degenerate Parabolic Equations Arising in Finance*, L.C.G. Rogers and D. Talay, Numerical Methods in Finance, Cambridge University Press, Cambridge, 1997;
- [6] Barles G., Souganidis P. - *Convergence of Approximation Schemes for Fully Nonlinear Second Order Equations*, Asymptotic Analysis 4, 1991;
- [7] Barles G., Burdeau J. - *The Dirichlet Problem for Semilinear Second-Order Degenerate Elliptic Equations and Applications to Stochastic Exit Time Control Problems*, Communications in Partial Differential Equations, 20, 1995;
- [8] Barles G., Rouy E. - *A Strong Comparison Result for the Bellman Equation Arising in Stochastic Exit Time Control Problems and its Applications*, Communications in Partial Differential Equations, 23, 1998;
- [9] Barlow M. - *A Diffusion Model for Electricity Prices*, Mathematical Finance, 12, 287-298, 2002;
- [10] Barrera-Esteve C., Bergeret F., Dossal C., Gobet E., Meziou A., Munos R., Reboul-Salze D. - *Numerical Methods for the Pricing of Swing Options: A Stochastic Control Approach*, Methodology And Computing In Applied Probability, 8, 517-540, 2006;

- [11] Bender C. - *Primal and Dual Pricing of Multiple Exercise Options in Continuous Time*, SIAM J. Finan. Math. 2, pp. 562-586, 2010;
- [12] Bensoussan A., Lions J. - *Contrôle Impulsionnel et Inéquations Quasi Variationnelles*, Dunod, Paris, 1982;
- [13] Bensoussan A. - *On the Theory of Option Pricing*, Acta Applicandae Mathematicae 2, 139-158, 1984;
- [14] Benth E.B., Benth J.B., Koekebakker S. - *Stochastic Modeling of Electricity and Related Markets*, World Scientific, Singapore, 2008;
- [15] Bermaud P. - *Point Processes and Queues: Martingale Dynamics*, Springer Verlag, New York, 1981;
- [16] Bingham N.H, Kiesel R. - *Risk-Neutral Valuation Pricing and Hedging of Financial Derivatives*, Springer Verlag, London, 1998;
- [17] Björk T. - *Arbitrage Theory in Continuous Time*, Oxford University Press, 1998;
- [18] Black F. Scholes M. - *The Pricing of Option and Corporate Liabilities*, J. Political Econom., 72, 1973;
- [19] Bock H. G. - *Randwertproblemmethoden zur Parameteridentifizierung in Systemen nichtlinearer Differentialgleichungen*, Bonner Mathematische Schriften, 183, Universität Bonn, 1987;
- [20] Briani M. - *Numerical Methods for Option Pricing in Jump-Diffusion markets*, PhD thesis, 2003;
- [21] Briani M., La Chioma C., Natalini R. - *Convergence of Numerical Schemes for Viscosity Solutions to Integro-Differential Degenerate Parabolic Problems Arising in Financial Theory*, Numerische Mathematik, 98, 607-646, 2004;
- [22] Burger M., Graeber B., Schindlmayr G. - *Managing Energy Risk*, John Wiley & Sons, Ltd, 2007;
- [23] Carmona R., Touzi N - *Optimal Multiple Stopping and Valuation of Swing Options*, Mathematical Finance, 2006;
- [24] Cartea A., Figueroa M. - *Pricing in Electricity Markets: A Mean Reverting Jump Diffusion Model with Seasonality*, Applied Mathematical Finance 12, 313 - 335, 2005;
- [25] Cartea A., Williams T. - *UK Gas Markets: the Market Price of Risk and Applications to Multiple Interruptible Supply Contracts*, Energy Economics, 30(3), 2008;

- [26] Chen Z. - *Numerical Methods for Optimal Stochastic Control in Finance*, PhD thesis, University of Waterloo, 2008;
- [27] Chen S., Insley M. - *Regime Switching in Stochastic Models of Commodity Prices: An Application to an Optimal Tree Harvesting Problem*, Working Paper, University of Waterloo, 2008;
- [28] Chen Z., Forsyth P. A. - *A Semi-Lagrangian Approach for Natural Gas Storage Valuation and Optimal Operation*, SIAM Journal on Scientific Computing, 30, 2007;
- [29] Clewlow L., Strickland C. - *Energy Derivatives: Pricing and Risk Management*, Lacrima Publications, 2000;
- [30] Cont R., Tankov P. - *Financial Modelling with Jump Process*, Chapman & Hall, London, 2004;
- [31] Cont R., Voltchkova E. - *Integro-Differential Equations for Option Prices in Exponential Lévy Models*, Finance and Stochastics, vol.9, 2005;
- [32] Cont R., Voltchkova E. - *A Finite Difference Scheme for Option Pricing in Jump Diffusion and Exponential Lévy Models*, SIAM Journal on Numerical Analysis, 43, 2005;
- [33] Crandall M.G., Ishii H., Lions P.L. - *User's Guide to Viscosity Solutions of Second Order Partial Differential Equations*, Bull. Amer. Math. Soc., 27, 1-67, 1992;
- [34] Crandall M.G., Lions P.L. - *Viscosity Solutions of Hamilton-Jacobi Equations*, Transactions of the American Mathematical Society, vol. 277, no. 1, 1983;
- [35] Dahlgren M. - *A continuous Time Model to Price Commodity-Based Swing Options*, Review of derivative Research, 8:27-47, 2005;
- [36] Dahlgren M., Korn R. - *The Swing Option on the Stock Market*, International Journal of Theoretical and Applied Finance, 8:123-139, 2005;
- [37] Deng S. - *Stochastic Models of Energy Commodity Prices and their Applications: Mean-Reversion with Jumps and Spikes*, UCEI, PWP-073;
- [38] Dörr U. - *Valuation of Swing Options and Examination of Exercise Strategy by Monte Carlo Techniques*, Master Thesis, University of Oxford, 2003;
- [39] Duffie D. - *Dynamic Asset Pricing Theory*, Princeton University Press, 2001;
- [40] Figueroa M. G. - *Pricing Multiple Interruptible-Swing Contracts*, Technical Reprint BWPEF, University of London, 2006;

- [41] Fleming M., Soner H. M. - *Controlled Markov Processes and Viscosity Solutions*, Springer, 2006;
- [42] Forsyth P. A., Vetzal K. R. - *Quadratic Convergence for Valuing American Options using a Penalty Method*, SIAM Journal on Scientific Computing, 23(6), 2002;
- [43] Forsyth P. A., Labahn G. - *Numerical Methods for Controlled Hamilton-Jacobi-Bellman PDEs in Finance*, Tech. Rep., University of Waterloo, 2006;
- [44] Geman H. - *Commodities and Commodity Derivatives: Modeling and Pricing for Agriculturals, Metals, and Energy*, John Wiley and Sons Ltd, 2005;
- [45] Griebel M., Hullmann A. - *An Efficient Sparse Grid Galerkin Approach for the Numerical Valuation of Basket Options under Kou's Jump-Diffusion Model*, University of Bonn, Preprint, 2012;
- [46] Guo M., Lapenkova M. - *Numerical Methods for Pricing Swing Options in the Electricity Market*, Master thesis, School of Information Science, Computer and Electrical Engineering, Halmstad University, 2010;
- [47] Haarbrücker G., Kuhn D. - *Valuation of Electricity Swing Options by Multistage Stochastic Programming*, Institute for Operations Research and Computational Finance, University of St. Gallen, 2006;
- [48] d'Halluin Y., Forsyth P. A., Labahn G. - *A Penalty Method for American Options with Jump Diffusion Process*, Numer. Math., 97, 321-352, 2004;
- [49] d'Halluin Y., Forsyth P. A., Vetzal K. R. - *Robust Numerical Methods for Contingent Claims under Jump Diffusion Process*, IMA J., Numer. Anal., 25, 2005;
- [50] d'Halluin Y., Forsyth P. A., Labahn G. - *A Semi-Lagrangian Approach for American Asian Options under Jump Diffusion*, SIAM Journal on Scientific Computing, 27, 2005;
- [51] Hamilton J. D. - *Analysis of Time Series Subject to Changes in Regime*, Journal of Econometrics, 45, 39-70, 1990;
- [52] Hambly B. M., Howison S., Kluge T. - *Modelling Spikes and Pricing Swing Options in Electricity Markets*, Quantitative Finance, 9, 937-949, 2009;
- [53] Hull J. - *Options, Futures and other Derivatives*, Prentice Hall International, 2006;
- [54] Ibanez, A. - *Valuation of Contingent Claims with Multiple Early Exercise Opportunities*, Mathematical Finance, 14, 223-248, 2004;

- [55] Insley M. - *A Real Options Approach to the Valuation of a Forestry Investment*, Journal of Environmental Economics and Management 44, 2002;
- [56] Insley M., Rollins K. - *On Solving the Multirotational Timber Harvesting Problem with Stochastic Prices: A Linear Complementarity Formulation*, American Journal of Agricultural Economics, vol. 87 (3), 2005;
- [57] Ishii H. - *Uniqueness of Unbounded Viscosity Solutions of Hamilton-Jacobi Equations*, Ind. Univ. Math., nr. 33, 1984;
- [58] Jacka S. D. - *Optimal Stopping and the American Put*, Math. Finance 1, 1991;
- [59] Jaillet P., Lamberton D., Lapeyre B. - *Variational Inequalities and the Pricing of American options*, Acta Applicandae Mathematicae 21, 263-289, 1990;
- [60] Jaillet P., Ronn E.I., Tompaidis S. - *Valuation of Commodity-Based Swing option*, Management Science 50, 909-921, 2004;
- [61] Kao E. P. C., Wang M - *Finite Element Method for Pricing Swing Options under Stochastic Volatility*, Working paper, 2010;
- [62] Karatzas I., Shreve S.E., - *Methods of Mathematical Finance*, Springer, 1998;
- [63] Keppo J. - *Pricing of Electricity Swing Options*, The Journal of derivatives: a publication of Institutional Investor, 3, 26-43, 2004;
- [64] Kjaer M. - *pricing of Swing Options in a Mean-reverting Model with Jumps*, Appl. Math. Finance, 222, 479-502, 2007;
- [65] Kluge T. - *Pricing Swing options and other Electricity Derivatives*, PhD Thesis, University of Oxford, 2006;
- [66] Kou S.G., Wang H. - *Option Pricing under a Double-Exponential Jump Diffusion Model*, Management Science, 50, 2004;
- [67] Krylov N.V. - *Controlled Diffusion Processes*, Berlin: Springer Verlag, 1980;
- [68] Lamberton D. - *Optimal Stopping and American Options*, Laboratoire d'analyse et de Mathématiques Appliquées, Université Paris-Est, Lecture Notes, 2009;
- [69] Lari-Lavassani A., Simchi M., Ware A. *A Discrete Valuation of Swing Options*, Canadian Applied Mathematics Quarterly, 9, 35-74, 2001;
- [70] Levy G. - *Computational Finance Using C and C#*, Elsevier, 2008;

- [71] Lions P. L. - *Optimal Control of Diffusion Processes and Hamilton-Jacobi-Bellman Equations. Part 1: The Dynamic Programming Principle and Applications; Part 2: Viscosity Solutions and Uniqueness*, Communications in Partial Differential Equations, 8, 1983;
- [72] Løland A., Lindqvist O. - *Valuation of Commodity-Based Swing Options: a Survey*, SAMBA, 38, 2008;
- [73] Longstaff F. A., Schwartz E. S. - *Valuing American Options by Simulation: A Simple Least-Squares Approach*, The Review of Financial Studies, 14, 113-147, 2001;
- [74] Lucia L., Schwartz S. - *Electricity Prices and Power Derivatives: Evidence from the Nordic Power Exchange*, Review of Derivatives Research, 5:5-50, 2002;
- [75] Lund A., Ollmar F. - *Analysing Flexible Load Contracts*, Working paper, 2003;
- [76] Musiela M., Rutkowski M. - *Martingale Methods in Financial Modelling*, Springer-Verlag Berlin Heidelberg, 2009;
- [77] McKean, H. P. Jr. - *Appendix: a Free Boundary Problem for the Heat Equation Arising from a Problem in Mathematical Economics*, Indust. Manage. Rev. 6, 32-39, 1965;
- [78] Meinshausen N., Hambly B. M. - *Monte Carlo Methods for the Valuation of Multiple-Exercise Options*, Mathematical Finance, 14, 557-583, 2004;
- [79] Merton, R.C. - *Option Pricing when Underlying Stock Returns are Siscntinuous*, J. Financial Econ. 125-144, 1976;
- [80] Morton K. W., Mayers D. F. - *Numerical Solution of Partial Differential Equations*, Cambridge University Press, 1994;
- [81] Neftci S. N. - *An Introduction to the Mathematics of Financial Derivatives*, Academic Press, New York, 2000;
- [82] Øksendal, B. K., Sulem. K. - *Applied Stochastic Control of Jump Diffusions*, Springer Verlag, 2005;
- [83] Pham H. - *Optimal Stopping of Controlled Jump Diffusion Process: A Viscosity Solution Approach*, Journal of Mathematical Systems, Estimation and Control, 8(1):1-27, 1998;
- [84] Pham H. - *Optimal Stopping, Free Boundary and American Option in a Jump-Diffusion Model*, Appl. Math. Optim., 35(2):145-164, 1997;
- [85] Pham H - *Continuous-time Stochastic Control and Optimization with Financial Applications*, Springer Verlag, 2009;

- [86] Pilipovic D. - *Energy Risk: Valuing and Managing Energy Derivatives*, McGraw-Hill, 1997;
- [87] Pooley D.M., Forsyth P.A., Vetzal, K.R. - *Numerical Convergence Properties of Option Pricing PDEs with Uncertain Volatility*, Numer. Anal. 23, 2003;
- [88] Pooley D. M., Vetzal K. R., Forsyth P. - *Convergence Remedies For Non-Smooth Payoffs in Option Pricing*, Journal of Computational Finance, 2003;
- [89] Protter P. - *Stochastic Integration and Differential Equations*, Springer-Verlag, 1990;
- [90] Rannacher R. - *Finite Element Solution of Diffusion Problems with Irregular Data*, Numer. Math. 43, 309-327, 1984;
- [91] Reisinger C. - *Finite Differences*, Lecture Notes, University of Oxford, 2008;
- [92] Roch A. F. - *Viscosity Solutions and American Option Pricing in Stochastic Volatility Model of the Ornstein-Uhlenbeck Type*, Journal of Probability and Statistics, 2010;
- [93] Runggaldier W.J. - *Jump-Diffusion Models*, Handbook of Heavy Tailed Distributions in Finance (S.T. Rachev, ed.), Elsevier/North-Holland, 2003;
- [94] Seidel R. - *Tools for Computational Finance*, Springer, Heidelberg, 2002;
- [95] Shin j. - *Convergence Overcomes for Non-Smooth Payoffs in Option Pricing*, working paper, 2010;
- [96] Shiryaev, A. N. - *Optimal Stopping Rules*, Springer Berlin Heidelberg New York, 1978;
- [97] Shreve E.S. - *Stochastic Calculus for Finance II Continuous-Time Models*, Springer, New York, 2004;
- [98] Thompson A. - *Valuation of Path Dependent Contingent Claims with Multiple Exercise Decisions over Time: the Case of Take-or-Pay*, Journal of Financial and Quantitative Analysis, 30(2), 1995;
- [99] Toivanen J. - *Numerical Valuation of European and American Options under Kou's Jump-Diffusion Model*, SIAM J. Sci. Comput. 1949-1970, 2008;
- [100] Touzi N. - *Optimal Stochastic Control, Stochastic Target Problems, and Backward SDEs*, Lecture Notes, Ecole Polytechnique Paris, Departement de Mathematiques Appliquees, 2010;

- [101] Wahab M., Lee C. - *Pricing Swing Options with Regime Switching*, *Annals of Operations Research*, 185, (1), 139-160, 2009,
- [102] Wahab M., Yin Z., Edirisinghe N. - *Pricing Swing Options in the Electricity Markets under Regime-Switching Uncertainty*, *Quantitative Finance* 10, (9), 975-994, 2010;
- [103] Ware T. - *The Valuation of Swing Options in Electricity Markets*, Conference Presentation, University of Calgary, 2007;
- [104] Wegner T. - *Swing Options and Seasonality of Power Prices*, Master Thesis, University of Oxford, 2002;
- [105] Weron R. - *Modelling and forecasting electricity loads and prices: A statistical Approach*, John Wiley & Sons, Ltd, 2006;
- [106] Wilhelm M. - *Modeling, Pricing and Risk Management of Power Derivatives*, PhD Thesis, ETH Zürich, 2007;
- [107] Wilmott P., Howison S., Dewynne J. - *The Mathematics of Financial Derivatives*, Cambridge university Press, 1998;
- [108] Zeghal. A. B., Mnif M. - *Optimal Multiple Stopping and Valuation of Swing Options in Lévy models*, *Int. J. Theor. and Appl. Finance*, 9, 1267-1297, 2006;
- [109] Zhang X. - *Numerical Analysis of American Option Pricing in a Jump-Diffusion Model*, *Mathematics of Operations Research*, 22, (3), 668-690, 1997;