UNIVERSIDAD DE BUENOS AIRES
Facultad de Ciencias Exactas y Naturales
Departamento de Matemática

Problemas parabólicos no lineales provenientes de modelos financieros: existencia y aproximación numérica de las soluciones

Tesis presentada para optar al título de Doctor de la Universidad de Buenos Aires en el área Ciencias Matemáticas

Lic. Andres Mogni

Director de tesis: Dr. Pablo Amster

Consejero de estudios: Dr. Pablo Amster

Buenos Aires, Diciembre 2018

Fecha de defensa: 19/12/2018

Agradecimientos

A Pablo, que sin conocerme me dio la posibilidad de hacer el doctorado. Fueron casi cinco años de estudio, aprendizaje y desarrollo que terminaron en un trabajo del cual estoy sumamente orgulloso. Gracias Pablo por todo el tiempo dedicado a este proyecto y por el tiempo compartido en general.

A Belen, que me acompaño día a día en esta aventura y estuvo a mi lado para verla terminar. A Valen, que se le ocurrió llegar antes de tiempo y tuvo la suerte de ver a su papá doctorarse.

A mis padres y mi hermano, que siempre me apoyaron en la vida y con quienes comparto este logro obtenido.

A Gustavo, que sin su colaboración hubiese sido imposible pensar en combinar el programa de doctorado con las obligaciones laborales diarias.

A Noel, Nico y Sebas, que me dieron un espacio de felicidad y diversión y me permitieron sobrellevar de una mejor manera los pequeños escollos que hubieron que sortear.

Contents

Resumen **vi**

Abstract **vii**

Introducción **viii**

Introduction **xi**

1 Preliminaries **1**
1.1 Finance 1
1.1.1 Financial Instruments 1
1.1.2 Option Pricing via Replication 3
1.1.3 Option Pricing via Expectation 5
1.1.4 Connection between both frameworks: Feynman-Kac Theorem 9
1.1.5 The Greeks 9
1.1.6 The use of financial options 11
1.2 Partial Differential Equations 12
1.2.1 Sobolev spaces and properties 12
1.2.2 Second-order Parabolic Equations 16
1.2.3 Two methods to prove existence 17
1.2.4 Viscosity solutions 20
1.3 Numerical Analysis 22
1.3.1 One-dimensional finite differences methods 23
1.3.2 Two and three-dimensional methods: The ADI method 27
1.4 Resumen del capítulo 30

2 Multi-asset option pricing with general transaction costs **31**
2.1 Introduction 31

2.2 The market model . . . 36
2.3 Existence of solution of the nonlinear PDE . . . 39
2.3.1 Defining the nonlinear equation . . . 39
2.3.2 Degenerate Ellipticity and Leland's condition . . . 41
2.3.3 Perron's method for existence of solution . . . 47
2.4 Numerical . . . 52
2.4.1 Numerical framework . . . 52
2.4.2 Numerical results . . . 54
2.5 Conclusion . . . 61
2.6 Resumen del capítulo . . . 63

3 Counterparty valuation adjustment with transaction costs 65

3.1 Introduction . . . 65
3.1.1 Counterparty Credit Risk . . . 65
3.1.2 CVA calculation by Replication . . . 67
3.1.3 CVA calculation by Expectation . . . 69
3.1.4 More work in CVA modelling . . . 70
3.2 The market model . . . 75
3.3 Existence of solution of the PDE . . . 82
3.3.1 Preliminaries . . . 82
3.3.2 Proof of Theorem 3.3.1 . . . 85
3.4 Numerical . . . 89
3.4.1 Numerical framework . . . 89
3.4.2 Numerical results . . . 91
3.5 Conclusion . . . 96
3.6 Resumen del capítulo . . . 98

Problemas parabólicos no lineales provenientes de modelos financieros: existencia y aproximación numérica de las soluciones

(Resumen)

En la presente tesis estudiamos dos problemas nolineales que se obtienen luego de realizar variaciones en la teoría estandar de valuación de activos financieros. Especificamente, consideramos modelos que incluyan los costos de transacción en su derivación y presentamos las ecuaciones diferenciales en derivadas parciales correspondientes. Una vez determinadas las ecuaciones, estudiamos la existencia de solución utilizando dos metodologías. En la Sección 2 demostramos la existencia de por lo menos una solucion débil viscosa utilizando el método de Perron. En la Sección 3, probamos la existencia de solución usando el Teorema de punto fijo de Schauder. En una segunda instancia nos abocamos a desarrollar esquemas numericos para estudiar el comportamiento de las soluciones en diversos escenarios. Para ello, utilizamos el método de Alternating Direction Implicit (ADI) en la primera sección y desarrollamos un esquema numerico con una cuadricula no uniforme en la segunda. Finalmente, estudiamos el comportamiento de las soluciones ante la presencia de diversos escenarios financieros y analizamos como las modificaciones hechas en la teoría afectan la valuación final de un activo financiero.

Palabras clave: Ecuaciones diferenciales parabólicas no lineales, modelos de valuación de opciones, Modelo de Leland, Teorema de punto fijo de Schauder, Esquema ADI, CVA, Método de Perron.

Nonlinear parabolic problems arising in finance: existence and numerical approximation of solutions

(Abstract)

In this thesis we studied different nonlinear problems that arise from making variations to the standard option pricing theory. Specifically, we considered models which allow the presence of transaction costs and presented the differential equations that explained those dynamics. After determining the corresponding equations, we applied two different methods to prove the existence of solution. In Section 2, we prove the existence of at least on weak viscosity solution using Perron's method. In Section 3, we prove the existence of solution by using the Schauder's Fixed Point theorem. The second part of these works involved developing different numerical schemes to effectively find a solution and analyze the its dynamics under a wide range of scenarios. For this purpose, we used an Alternating Direction Implicit (ADI) scheme in the first Section and an Euler scheme with a non-uniform grid in the second one. Finally, we studied the behaviour of these solutions under the presence of different financial scenarios and analyzed how the variation of the standard theory affects the pricing of a financial instrument.

Keywords: Nonlinear parabolic differential equations, option pricing models, Leland's model, Schauder's fixed point theorem, ADI scheme, CVA, Perron's method.

Introducción

Las matemáticas y las finanzas son dos campos de estudio altamente conectados. En particular, el problema de encontrar el precio óptimo de cualquier instrumento financiero resulta ser una tarea desafiante la cual generalmente es resuelta utilizando distintas técnicas y herramientas matemáticas. Por ejemplo, la dinámica el precio de una acción, la tasa de interés, la volatilidad de una acción y la correlación entre dos activos pueden ser modelados usando movimientos brownianos con drift. Otro ejemplo posible corresponde a los pasos que se aplican para deducir el modelo que explica el precio una opción. En ese contexto, se utilizan diferentes herramientas probabilísticas como por ejemplo el teorema de representación de martingalas y el teorema de Girsanov. Además, la deducción de estos modelos generalmente llevan a diferentes tipos de ecuaciones diferenciales parabólicas.

El objetivo de esta tesis es generalizar dos tipos de modelos financieros muy estudiados en la literatura mediante la relajación de un supuesto clave: la ausencia de costos de transacción al construir el portfolio replicante.

El primer modelo estudiado corresponde al modelo de valuación de opciones estandar el cual, en su forma original, es el modelo de Black-Scholes. Este modelo expresa la dinámica de una opción financiera sobre un activo subyacente. Así mismo, se basa en múltiples supuestos que generalmente no aplican en el mundo real. Nuestra generalización se aplica desde dos lados distintos. Primero, proponemos un modelo multidimensional con N activos subyacentes. Este tipo de generalización ya ha sido estudiada en el pasado y conduce una ecuación diferencial lineal parabólica multidimensional. Nuestra segunda observación se basa en relajación del supuesto de ausencia de costos de transacción. La inclusión de los mencionados costos afecta fuertemente el modelo original dado que aparece un termino no lineal en la ecuación resultante. Dependiendo de la forma que tenga la función de costos de transacción, la no linealidad podra ser de tipo *quasilinear*, *semilinear* or *fully nonlinear*. Dado que nosotros proponemos una función de costos de transacción que cubre todos los casos factibles, deducimos una ecuación no lineal de tipo *fully nonlinear*. En nuestro primer trabajo Ref [7] deducimos este problema general a partir del modelo básico de Black-Scholes en combinación con otras técnicas usadas previamente en distintos trabajos como [4], Ref [20], Ref [33], Ref [39], Ref [45] and Ref [46]. En esta tesis presentamos estos pasos en la Sección 2.2.

Dada la naturaleza no lineal de la ecuación diferencial, decidimos buscar soluciones de tipo viscosas. Estas soluciones corresponden a un tipo de soluciones débiles y suelen encontrarse luego de plantear un par de sub y supersoluciones viscosas. Partiendo de que el correspondiente problema lineal tiene solución, derivamos del mismo el par de sub y supersoluciones.

Luego, utilizamos el método de Perron para garantizar la existencia de una solución viscosa.

Nuestro primer trabajo concluye con el desarrollo de un esquema numérico para encontrar una solución aproximada del problema no lineal y entender como los costos de transacción afectan el precio de una opción. Dado que trabajamos en un problema multidimensional, aplicamos un tipo de *splitting operator* conocido como *Alternating Direction Implicit* (ADI). Esta técnica termina siendo útil para nuestro problema por la existencia de derivadas segundas cruzadas. Esto genera una imposibilidad de resolver problemas tridiagonales y ralentizan el posterior hallazgo de la solución. Finalmente, dado el esquema numérico desarrollado y tres diferentes escenarios de prueba, nos focalizamos en tres distintos análisis: medir el impacto de los costos de transacción en el precio de la opción, analizar la sensibilidad del precio de la opción con respecto a cambios en la frecuencia de rebalanceo del portfolio replicante y analizar la convergencia del método numerico.

El segundo modelo estudiado es una generalización del modelo de *Counterparty Valuation Adjustment* (CVA). CVA es un modelo que surge como consecuencia de la crisis financiera del 2008 dado que la metodología estandar de valuación previa no consideraba las probabilidades de default (PD) de las dos partes de un contrato financiero. Este punto puede reflejarse en clásico modelo de Black-Scholes, el cual entre sus supuestos presenta la ausencia de probabilidad de default de las partes. En nuestro trabajo partimos de el trabajo original de Ref [13] donde la PD de tanto el emisor del contrato como la contraparte son consideradas y a la cual nosotros le agregamos la presencia de costos de transacción en el portfolio replicante. Como consecuencia, obtuvimos una ecuación de tipo *quasilinear* parabólica. Es importante remarcar que determinar el precio de dicha opción mediante la resolución de la ecuación diferencial consume mucho tiempo computacional y es por esto que en la industria se trabaja bajo el modelo de esperanza condicional. Muchos trabajos han sido presentados siguiendo esta linea de trabajo tales como Ref [10], Ref [11] and Ref [12]. Un tercer enfoque es el dearrollado por Ref [14] y expandido por Ref [15] y Ref [16]. En estos trabajos, el autor desarrolla una forma reducida en base a un modelo de *backward stochastic differential equations* para resolver el problema de valuación de CVA permitiendo la existencia de restricciones de fondeo.

Aplicando los pasos de [13], deducimos el modelo de mercado correspondiente y obtuvimos una ecuación diferencial quasilineal. Luego, nos focalizamos en probar la existencia de al menos una solución usando un método de punto fijo. Construimos un operador T tal que el punto fijo del mismo sea al mismo tiempo la solución del problema no lineal. Este operador fue construido siguiendo los parámetros necesarios para la aplicación posterior del teorema de Schauder. Es necesario aclarar que la existencia de solución depende de tres condiciones impuestas a los parámetros del modelo. Estas condiciones son equivalentes a fijar volatilidades ni muy pequeas ni muy grandes y acotar la tasa de crecimiento del activo subyacente bajo la medida libre de riesgo. Bajo estas condiciones impuestas es que deducimos la existencia de la menos una solución convexa.

La segunda parte del trabajo se focaliza en el desarrollo de un esquema numérico para encontrar soluciones aproximadas del problema no lineal original. Desarrollamos un esquema con una grilla no uniforme en la compononente espacial tal que el espaceado resulte fino cerca del strike y grueso lejos del mismo. Siguiendo los trabajos de Ref [43], Ref [9] y Ref [21], obtuvimos la discretización de las primeras y segundas derivadas espaciales y definimos el esquema de diferencias finitas. Dado el esquema numérico, analizamos el comportamiento del precio de una opción *call* de tipo Europea bajo diversos escenarios aplicando un analisis de sensibilidad sobre los diversos parámetros. Además, comparamos nuestros resultados con los obtenidos en el trabajo original Ref [13] y calculamos como los costos de transacción impactan en el el valor final de CVA.

Esquema de la tesis

El Capítulo 1 contiene el marco teórico que va a ser utilizado a lo largo de la tesis. Se encuentra subdividido en las tres grandes areas en las que trabajamos. Primero, presentamos los conceptos básicos relacionados con los instrumentos financieros, requeridos para comprender los problemas propuestos. Segundo, incluimos las definiciones y los resultados más importantes del area de las ecuaciones diferenciales en derivadas parciales. Especificamente, nos focalizamos en resultados de ecuaciones de tipo parabólicas, ecuaciones no lineales y diversas formas de probar la existencia de solución. La tercera sección corresponde al area del cálculo numérico. Definimos el esquema de diferencias finitias explícito y como es generalizado al esquema de Crank-Nicholson. Además, incluimos la definicion del esquema ADI y dos formas distintas de discretizar las ecuaciones: los esquemas de Peaceman-Rachford y Douglas-Rachford.

El Capítulo 2 contiene todos los resultados correspondientes al trabajo Ref [7]. La sección 2.1 incluye una introducción en valuación de opciones considerando la presencia de costos de transacción y diversos trabajos y modelos existentes en la literatura. La sección 2.2 está dedicada a la construcción del modelo de mercado y la consecuente derivación de la ecuación diferencial no lineal. La sección 2.3 comprende todos los pasos aplicados para probar la existencia de al menos una solución viscosa, incluyendo la construcción de ambas sub y super soluciones y la aplicación del método de Perron. La sección 2.4 incluye el desarrollo del esquema numérico y los resultados correspondientes al impacto de los costos de transacción en el precio de la opción como así también la convergencia del esquema numérico.

El Capítulo 3 contiene todos los resultados correspondientes al trabajo Ref [6]. La sección 3.1 presenta una introducción del modelo CVA y diferentes formas de derivar el esquema de valuación. La sección 3.2 está dedicada a la construcción del modelo de mercado incluyendo la presencia de costos de transacción en cada paso del rebalanceo del portfolio replicante. La sección 3.3 comprende todos los pasos aplicados para probar la existencia de al menos una solución débil utilizando el teorema de punto fijo de Schauder. La sección 3.4 incluye el desarrollo del esquema numérico utilizado para encontrar una solución aproximada del problem no lineal y provee conclusiones relacionadas con el impacto de los costos de transacción en el valor de CVA.

Introduction

Mathematics and finance are two fields that are highly connected. In particular, the problem of finding the optimal price of any financial instrument is a challenging task which is generally solved using different mathematical tools and techniques. For example, the dynamics of a stock price, an interest rate, the volatility of a stock and the correlation between two assets can be modelled using a brownian motion with drift. Another example correspond to the steps that are applied to find the model of the price of a financial option. In that context, probabilistic tools such as the martingale representation theorem and the Girsanov theorem are required to deduce the model. Moreover, the deduction of these models usually leads to different types of parabolic partial differential equations.

The aim of this thesis is to generalize two types of financial models widely known in the literature by relaxing one key assumption: the absence of transaction costs when constructing the replicant portfolio.

The first model corresponds to the standard option pricing model which, in its original form, correspond to the Black-Scholes model. This model expresses the dynamic of a financial option over one underlying asset. Moreover, it relies on multiple assumptions that do not generally apply in real life. Our generalization comes from two different sides. First, we propose to model a multi-asset option with N underlying assets. This generalization has been studied before and leads to a multidimensional linear parabolic partial differential equation. Our second observation comes from the relaxation of one of the assumptions: the absence of transaction costs in the construction of the replicant portfolio. This inclusion strongly affects the original model as a nonlinear term appears in final equation. From a PDE perspective, the model is still parabolic but not any more linear. Depending the shape of the transaction costs, the nonlinearity can be quasilinear, semilinear or fully nonlinear. As we propose a general transaction costs function that covers all the feasible functions, we derive a fully nonlinear problem. In our first work Ref [7] we deduce this general problem following the standard Black-Scholes steps in a combination of techniques used previously in different works such as Ref [4], Ref [20], Ref [33], Ref [39], Ref [45] and Ref [46]. In this thesis we present those steps in Section 2.2.

Given the nonlinear nature of the differential equation, we decide to look for a viscosity solution. This solutions belong to a type of weak solutions and are usually found by proposing a pair of sub and supersolutions. Using the fact that the linear problem is solvable, we take advantage of this to derive the pair of sub and supersolutions. One important tool within this technique is Perron's method. This methodology was originally proposed to find a solution of the Dirichlet problem for the Laplace equation. Under the viscosity framework, Perron's method guarantees the existence of a solution of the correspondent Dirichlet problem.

Our first work ends with a development of a numerical framework to find an approximate solution of the original nonlinear problem and understand how do the transaction costs affects the price of a financial option. As we work with a multidimensional equation, we apply one type of splitting operator method, known as Alternating Direction Implicit (ADI) method. This technique results to be useful due to the existence of crossed derivatives in the numerical scheme and the impossibility to solve one tridiagonal system. Finally, given the numerical scheme and different testing scenarios, we focus on three different analysis: measure the impact of transaction costs in the option price, analyze the sensitivity of the option price to changes in the frequency of rebalancing of the replicant portfolio and observe the convergence of the numerical scheme.

The second model in which we worked with is a generalization of the Counterparty Valuation Adjustment (CVA) model. CVA is a model that arose as a consequence of the 2008 financial crisis as the standard valuation methodology of that time did not consider the probability of default (PD) of both parties that participated in a contract. This situation can be seen in the widely used Black-Scholes model as one of its assumptions is the absence of probability of default. In our work, we start from the seminal paper of Ref [13] where the PD of both the issuer and the counterparty are considered and include the presence of transaction costs in the replicant portfolio of the option. As a consequence, we obtain a quasilinear parabolic partial differential equation. It is important to remark that finding a numerical solution of this problem is time consuming and it is a reason of why practitioners prefer to work under the conditional expectation approach. Many works have been done following this line of work as of Ref [10], Ref [11] and Ref [12]. A third approach is taken by Ref [14] and further on with Ref [15] and Ref [16]. On its works, the author develop a reduced-form backward stochastic differential equations (BSDE) approach to the problem of pricing and hedging of the CVA by allowing the presence of multiple funding constraints.

After deducing the market model by adapting the steps of Ref [13] and getting a quasilinear differential equation, we focus on proving the existence of at least one solution. For this purpose, we choose to use a fixed point approach. We have to construct an operator T such that its fixed point is at the same time a solution of our nonlinear problem. This operator has to be constructed in a way so that the Schauder fixed point theorem can be applied. One important consideration for this problem is that three different conditions on the parameters have to be applied to assure the existence of solution. These conditions are equivalent to setting the volatility either not to high or not to low and that the stock growth rate under the risk neutral measure has to be bounded. As a consequence, the existence of at least one convex solution is deduced.

The second part of the work relates to the development of a numerical framework to find an approximate solution of the original nonlinear problem. We develop a scheme with a non-uniform grid in the spatial component such that the spacing is fine near the strike value and coarse away from the strike. Following Ref [43], Ref [9] and Ref [21] we obtain the discretization of the first and second spatial derivatives and define the finite difference framework. Given the numerical scheme we analyze the behavior of the option price for an European call under different scenarios by performing a sensitivity analysis on different parameters. We also compare our results with the ones obtained by the original model Ref [13] and calculate how the transaction costs impact the final CVA value.

Thesis outline

Chapter 1 contains the theoretical framework that will be used along the thesis. It is subdivided in the three main areas in which we worked. First, we provide all the basic concepts related to financial instruments that will be required to understand the problems proposed. Second, we include important definitions and results from the partial differential equation area. Specifically, we focus on different results about parabolic equations, nonlinear equations and different ways to prove the existence of solution. The third Section correspond to the numerical analysis area. We define an explicit difference scheme and how is generalized into the Crank-Nicholson scheme. Moreover, we include the definition of the ADI scheme and two different ways of discretizing the differential equation: the Peaceman-Rachford and Douglas-Rachford scheme.

Chapter 2 contains all the results related to Ref [7]. Section 2.1 provides an introduction on option pricing with transaction costs and different works and models available in the literature. Section 2.2 is devoted to the construction of the correspondent market model and the derivation of the nonlinear differential equation. Section 2.3 comprises all the steps applied to prove the existence of at least one viscosity solution, including the construction of both sub and supersolutions and the application of Perron's method. Section 2.4 includes the development of the numerical scheme and the results regarding the impact of the transaction costs in the option price as well as the convergence of the numerical scheme.

Chapter 3 contains all the results related to Ref [6]. Section 3.1 presents an introduction on the CVA model and different ways to derive the pricing framework. Section 3.2 is devoted to the construction of the market model by including the presence of transaction costs on each rebalancing step. Section 3.3 comprises all the steps applied to prove the existence of at least one weak solution following Schauder fixed point theorem. Section 3.4 includes the development of a numerical framework to find an approximate solution of the nonlinear problem and provides conclusions regarding the impact of transaction costs in the CVA value.

Chapter 1

Preliminaries

In this chapter we present different results that are used along the entire thesis and are needed to derive the main results presented in Chapters 2 and 3. Moreover, we split this chapter into three Sections to cover the main theory that arises on each working topic.

In the first Section we introduce the basic financial language and discuss the fundamental results of the option pricing theory. This will help us to understand the financial problems that we want to model and solve using different techniques. The second Section covers all the definitions and results that we require to prove the existence of solution of nonlinear parabolic partial differential equations. These include the introduction of the spaces in which we will work, different fixed-point theorems, iterative methods, super-sub solution methods and the main results of parabolic PDE theory. The third and last Section is related to the theoretical background of the numerical methods that we employ to solve the partial differential equations following a numerical approach. We will also discuss about both Euler method and splitting operators methodology.

1.1 Finance

In this Section we are going to present basic definitions and concepts needed to understand the problems that we discuss in Chapters 2 and 3. Specifically, we are going to introduce the theory of option pricing that leads to the parabolic equations that we will study afterwards. The bibliography used for this purpose are Ref [8], Ref [27] and Ref [44].

1.1.1 Financial Instruments

We are going to work with two basic financial instruments: a discount bond and a stock. The first instrument correspond to an agreement to pay some money now in exchange for receiving a larger sum later. It is defined based on two values: the length or *maturity* of the contract and the extra amount that will be paid in the future (that is measured via the *interest rate*). If we define $P(t,T)$ as the price at time t of this bond that pays 1 dollar at time T, we have that

$$P(t, T) = e^{-r(T-t)} \tag{1.1.1}$$

where r is a constant interest rate. The second instrument, the share of a stock, represents a fractional ownership of a company. The stock price is modelled following a log-normal distribution such that, if we define S_t as the price of the stock S at time t, it can be modelled as

$$\log S_T = \log S_0 + X \tag{1.1.2}$$

where X is a normally distributed random variable. One important difference between the price of a bond and the price of a stock is that the first one is deterministic; given the value of the interest rate, the price is known for every t. On the other side, the price of the stock is a random variable with a pre-defined mean and variance. This implies that the price is stochastic and can vary with certain probability. In fact, the stochastic component of the stock price is modelled following a brownian motion.

Definition 1.1.1. The process $W = \{W_t\}_{t\geq 0}$ is a brownian motion if and only if

i. W_t is continuous and $W_0 = 0$,

ii. W_t follows a normal distribution with zero mean and variance equal to t,

iii. the increments $W_{s+t} - W_s$ are independents.

Given the definition of a brownian motion, the following equation shows the dynamics of the stock price by using the notation of an stochastic differential equation

$$dS = \mu S dt + \sigma S dW_t \tag{1.1.3}$$

where μ and σ are the respective mean and volatility of the stock price. The first term of the equation contains the deterministic change of the stock price. Indeed, the price moves with a trend given by the value of μ. The second term involves the stochastic component which is expressed with a brownian motion.

One important tool from the stochastic calculus is the Itô's formula, which is an analogous of Taylor's formula for stochastic processes.

Definition 1.1.2 (Itô's formula)**.** If X is an stochastic process satisfying the stochastic differential equation $dX_t = \mu dt + \sigma dW_t$ and f is a deterministic twice continuously differentiable function, then $Y_t = f(X_t)$ is also a stochastic process and is given by

$$dY_t = \left(\mu f'(X_t) + \frac{1}{2}\sigma^2 f''(X_t)\right) dt + \left(\sigma f'(X_t)\right) dW_t \tag{1.1.4}$$

The third and last financial instrument that we present is the option. An option is a contract which gives the buyer the right, but not the obligation, to buy or sell an underlying asset at a specific price (strike price) on a specified date. Every option has a payoff, which is its value at the time of exercise. For example, the most common option, the European call, correspond to a contract which gives the right to buy the stock S at time T at price K. Hence, the payoff of this contract is given by

$$\max\left(S_T - K, 0\right) \tag{1.1.5}$$

If the price at maturity is higher than the strike price, the holder of the option will exercise the contract and buy the stock. As a consequence, he will earn $S_T - K$. However, if the strike price is higher, the holder will not exercise the option as it will loose money. Hence, he will earn nothing from this contract.

Up to this point, we will focus on the valuation of different options. We need to answer the question of which is the optimal price of an option given the underlying instrument and its own characteristics. In the following two Sections we will show the answer of this question using two different frameworks. The first one will lead to a PDE of the option price. The second framework shows that the option price is given by a conditional expectation of the discounted payoff under an equivalent measure.

1.1.2 Option Pricing via Replication

In this Section we present the Black-Scholes methodology, which is a procedure that leads to the price of an option over one asset (and to the well-known Black-Scholes equation). Let us first define the main seven assumptions that we will follow on this procedure:

- The asset price follows a lognormal random walk.
- The risk-free interest rate r and the volatility of the asset σ are known functions over the life of the option.
- There are no transaction costs when buying or selling any asset.
- The underlying asset pays no dividends during the life of the option.
- Every risk-free portfolio earns the same return (r, which is the risk-free interest rate).
- The trading of the underlying asset can take place continuously and
- Assets are divisible and can be sold without owning it (short-selling).

Let us define $V\left(S, t\right)$ as the price of an option that depends only on the price of the asset S and the time t. Hence, using Ito's lemma from (1.1.4), and considering the dynamics of the asset S as in equation (1.1.3) we know that

$$dV = \sigma S \frac{\partial V}{\partial S} dW + \left(\mu S \frac{\partial V}{\partial S} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + \frac{\partial V}{\partial t} \right) dt \tag{1.1.6}$$

The next step is to construct a portfolio Π which has one option V and $-\Delta$ amounts of the stock S. Its value can be represented as

$$\Pi = V - \Delta S \tag{1.1.7}$$

The one-step change of the value of this portfolio is given by

$$d\Pi = dV - \Delta dS \tag{1.1.8}$$

as Δ is held fixed during each time step. If we apply equation (1.1.6) in equation (1.1.8), we obtain

$$d\Pi = \sigma S \left(\frac{\partial V}{\partial S} - \Delta \right) dW + \left(\mu S \frac{\partial V}{\partial S} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + \frac{\partial V}{\partial t} - \mu \Delta S \right) dt \tag{1.1.9}$$

It can be seen that if we choose $\Delta = \frac{\partial V}{\partial S}$ on each time step, we will be able to remove the randomness of the problem. Following this step, we define a portfolio whose increments are deterministic

$$d\Pi = \left(\frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + \frac{\partial V}{\partial t} \right) dt \tag{1.1.10}$$

It is important to note that $d\Pi = r\Pi dt$ as if the right-hand side of equation (1.1.10) were greater or lower than $r\Pi$, an strategy can be created to make an instantaneous profit. This is not accepted as per the assumptions of the model. Then, we have

$$r\Pi dt = \left(\frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + \frac{\partial V}{\partial t} \right) dt \tag{1.1.11}$$

Finally, we obtain the following deterministic equation, also known as the Black-Scholes equation.

$$\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + rS \frac{\partial V}{\partial S} - rV = 0 \tag{1.1.12}$$

Equation (1.1.12) gives us the dynamics of every option over one underlying stock S. Boundary and terminal conditions will be the ones that represent each existing option. For example, for an European option we have the following conditions

$$\begin{aligned} V(S,T) &= \max(S_T - K, 0) \\ V(0,t) &= 0 \quad \text{for all} \quad 0 \le t \le T \\ V(S,t) &\to S \quad \text{as} \quad S \to +\infty \end{aligned} \tag{1.1.13}$$

Moreover, the equation (1.1.12) with conditions given by (1.1.13) has the following solution

$$\begin{aligned} V(S,t) &= N(d_1) S_t - N(d_2) K e^{-r(T-t)} \\ d_1 &= \frac{1}{\sigma\sqrt{T-t}}\left[\log\left(\frac{S_t}{K}\right) + \left(r + \frac{\sigma^2}{2}(T-t)\right)\right] \\ d_2 &= d_1 - \sigma\sqrt{T-t} \end{aligned} \tag{1.1.14}$$

where N is the cumulative distribution function of the standard normal distribution, T is the maturity, σ is the stock's volatility, K is the strike price and r is the risk-free interest rate.

Example 1.1.3. Let us apply formula (1.1.14) at $t = 0$ with the following parameters: $S_0 = 10$, $T = 1$, $r = 0.05$, $\sigma = 0.25$. The results are given in the following table.

Strike	Price	Strike	Price
5	5.2446	11	0.8026
7	3.3856	13	0.3046
9	1.8141	15	0.1038

The results show the expected behaviour of the option price. When the option is in-the-money ($S > K$), there is a positive intrinsic value. Hence, the probability of the option maturing with a positive payoff is indeed positive. On the other hand, when the option is out-of-the-money ($S < K$), the option has no intrinsic value. As the probability of the option maturing with a positive payoff decrease as the strike is higher, it is expected that option price will tend to zero.

1.1.3 Option Pricing via Expectation

In this Section we derive the pricing framework that is known as *risk neutral pricing.* We follow the steps explained in Ref [40] in order to find the optimal price of a financial option with the underlying stock S.

Let us start by defining the discount process $D(t)$ as of

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

where $R(s)$ is an interest rate process. Let us note that, by using Itô's formula (1.1.4), the variation of the discount process is equal to

$$dD(t) = -R(t)\, D(t)\, dt. \tag{1.1.15}$$

Let us create a portfolio $X(t)$ which starts with an initial capital $X(0)$ and at each time t holds Δ shares of stock $S(t)$ by borrowing or investing at interest rate $R(t)$ as necessary to finance this operation. Then,

$$dX(t) = \Delta(t)\, dS(t) + R(t)\,(X(t) - \Delta(t)\, S(t))\, dt \tag{1.1.16}$$

We will again model the stock process as in equation (1.1.3) so that equation (1.1.16) becomes

$$\begin{aligned} dX(t) &= \Delta(t)\,(\mu(t)\, S(t)\, dt + \sigma(t)\, S(t)\, dW(t)) + R(t)\,(X(t) - \Delta(t)\, S(t))\, dt \\ &= R(t)\, X(t)\, dt + \Delta(t)\,(\mu(t) - R(t))\, S(t)\, dt + \Delta(t)\,\sigma(t)\, S(t)\, dW(t) \\ &= R(t)\, X(t)\, dt + \Delta(t)\,\sigma(t)\, S(t)\,[\Theta(t)\, dt + dW(t)] \end{aligned} \tag{1.1.17}$$

where $\Theta(t)$ is known as the market price of risk and is denoted as

$$\Theta(t) = \frac{\mu(t) - R(t)}{\sigma(t)} \tag{1.1.18}$$

If we apply the Itô's product rule to the process of the discounted value of the portfolio $D(t)\, X(t)$, we get that

$$d(D(t)\, X(t)) = \Delta(t)\,\sigma(t)\, D(t)\, S(t)\,[\Theta(t)\, dt + dW(t)] \tag{1.1.19}$$

The next step requires to present the Girsanov's Theorem.

Theorem 1.1.4 (Girsanov's Theorem)**.** *Let $W(t)$ be a brownian motion on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and let $\mathcal{F}(t)$ be a filtration for this brownian motion. Let $\Omega(t)$ be an adapted process. Let us define*

$$Z(t) = \exp\left(-\int_0^t \Theta(u)\, dW(u)\, du - \frac{1}{2}\int_0^t \Theta^2(u)\, du\right),$$
$$\tilde{W}(t) = W(t) + \int_0^t \Theta(u)\, du, \tag{1.1.20}$$

and assume that

$$\mathbb{E}\int_0^T \Theta^2(u)\, Z^2(u)\, du < \infty \tag{1.1.21}$$

Then $\mathbb{E}Z(T) = 1$ *and under the probability measure* $\tilde{\mathbb{P}}$ *the process* $\tilde{W}(t)$ *is a brownian motion where the new probability measure* $\tilde{\mathbb{P}}$ *is given by*

$$\tilde{\mathbb{P}}(A) = \int_A Z(w)\, dP(w) \tag{1.1.22}$$

Hence, we can use the Girsanov's theorem in equation (1.1.19) so that

$$d(D(t)\, X(t)) = \Delta(t)\, \sigma(t)\, D(t)\, S(t)\, d\tilde{W}(t) \tag{1.1.23}$$

In equation (1.1.23) we deduce that the stochastic process of discounting the portfolio value is a martingale.

Definition 1.1.5. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and $\mathcal{F}(t)$ be a filtration of sub-σ-algebras of $\mathcal{F}$. Consider an adapted stochastic process $M(t)$. Then, $M(t)$ is a martingale if

$$\mathbb{E}\left[M(t)\,|\mathcal{F}(s)\right] = M(s) \quad \text{for all} \quad 0 \le s \le t \le T. \tag{1.1.24}$$

Remark 1.1.6. The brownian motion process is a martingale

Using the fact that the discounted portfolio's value is a martingale, we can use the definition and note that

$$D(t)\, X(t) = \tilde{\mathbb{E}}\left[D(T)\, X(T)\,|\mathcal{F}(t)\right], \quad 0 \le t \le T \tag{1.1.25}$$

Let us go back and recall the problem of pricing a financial derivative. Let $V(T)$ be a measurable $\mathcal{F}(T)$ random variable which represents the payoff of the option at maturity time

T. So, we want to know how many $\Delta(t)$ shares of $S(t)$ such that $X(t) = V(t)$ almost surely. Hence, once it is done, we will have that

$$D(t)\,V(t) = \tilde{\mathbb{E}}\left[D(T)\,V(T)\,|\mathcal{F}(t)\right] \tag{1.1.26}$$

In order to determine the value of $\Delta(t)$ we need to define the Martingale Representation's Theorem.

Theorem 1.1.7 (Martingale Representation's Theorem)**.** *Let $W(t)$ be a brownian motion on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and let $\mathcal{F}(t)$ be the filtration generated by the brownian motion. Let $M(t)$ be a martingale with respect to this filtration. Then, there exists an adapted process $\Gamma(t)$ such that*

$$M(t) = M(0) + \int_0^t \Gamma(u)\,dW(u) \tag{1.1.27}$$

By using this theorem we can define the process $\Delta(t)$ and find the value of $V(t)$. So, given that $D(t)\,V(t)$ is a martingale, there exists a process $\tilde{\Gamma}(u)$ such that

$$D(t)\,V(t) = V(0) + \int_0^t \Gamma(u)\,d\tilde{W}(u) \tag{1.1.28}$$

Moreover, if we recall equation (1.1.23), we know that

$$D(t)\,X(t) = X(0) + \int_0^t \Delta(u)\,\sigma(u)\,D(u)\,S(u)\,d\tilde{W}(u) \tag{1.1.29}$$

In order to have $X(t) = V(t)$ for all t, we will shall clear out $Delta(t)$ from $X(0) = V(0)$ so that

$$\Delta(t)\,\sigma(t)\,D(t)\,S(t) = \tilde{\Gamma}(t) \tag{1.1.30}$$

which is equivalent to

$$\Delta(t) = \frac{\tilde{\Gamma}(t)}{\sigma(t)\,D(t)\,S(t)} \tag{1.1.31}$$

Hence, with this choice of $\Delta(t)$ we could define the price of the option $V(t)$ as of

$$V(t) = D(t)^{-1}\,\tilde{\mathbb{E}}\left[D(T)\,V(T)\,|\mathcal{F}(t)\right] \tag{1.1.32}$$

In particular, we can note that the price of a financial derivative at time zero is the expected value of the discounted payoff.

1.1.4 Connection between both frameworks: Feynman-Kac Theorem

In the previous two Sections we derived the price of an option following two different frameworks. In Section 1.1.2 we found that the price is the solution of a parabolic PDE with certain final and boundary conditions. On the other side, in Section 1.1.3 we showed that the price of an option is equal to the conditional expectation of the payoff discounted. In order to show that both frameworks generate the same results, we need to define the Feynman-Kac Theorem. In fact, the relationship between geometric brownian motions and the Black-Scholes PDE equation is a particular case of the relationship between stochastic differential equations and PDE's.

Theorem 1.1.8 (Feynman-Kac Theorem)**.** *Consider the stochastic differential equation*

$$dX(u) = \beta(u, X(u))\, du + \gamma(u, X(u))\, dW(u). \tag{1.1.33}$$

Let $h(y)$ be a Borel-measurable function. Fix $T > 0$ and let $t \in [0, T]$ be given. Define the function

$$g(t, x) = \mathbb{E}\left[e^{-r(T-t)}\left(h(X(T)) + f(t, x)\right) | \mathcal{F}_t\right]. \tag{1.1.34}$$

Then, $g(t, x)$ satisfies the partial differential equation

$$\frac{\partial g(t, x)}{\partial t} + \beta(t, x) \frac{\partial g(t, x)}{\partial x} + \frac{1}{2}\gamma^2(t, x) \frac{\partial^2 g(t, x)}{\partial x^2} - rg(t, x) = f(t, x) \tag{1.1.35}$$

with terminal condition

$$g(T, x) = h(x) \quad \text{for all } x \tag{1.1.36}$$

Hence, recalling that under the risk-neutral measure the dynamics of the stock S is given by

$$dS(t) = rS(t)\, dt + \sigma S(t)\, dW(t) \tag{1.1.37}$$

if we set $\beta = rS$, $\gamma = \sigma S$ and $h(T)$ the payoff of the option $g(t, x)$, the theorem confirms that both frameworks produce the same solution.

1.1.5 The Greeks

The next step after learning how to price an option is to understand and measure its inherent risks. The derivatives of the option with respect to the different variables are known as *Greeks*.

If we recall the Black-Scholes formula (1.1.12), we can see that the option's riskiness is derived from movements in the underlying asset S, the volatility σ, the risk-free interest rate r and the time to maturity $\tau = T - t$.

The first risk that every option has is due to the unpredictable movements on the underlying (for example, in the example (1.1.3), the underlying would be S). The delta (Δ) is defined as the rate of change of the option price with respect to the price of the underlying asset. Then,

$$\Delta = \frac{\partial V}{\partial S} \tag{1.1.38}$$

The Delta of an European call can be derived from (1.1.12) so that

$$\Delta = N\left(d_1\right). \tag{1.1.39}$$

Given that N is the cumulative distribution function of the standard normal distribution, Delta is always in the interval $[0, 1]$. Moreover, if we analyse the components of d_1 we can observe that Delta is smaller when the option is out-of-the-money ($S_t <<< K$) and higher when the option is in-the-money ($S_t >>> K$).

The rate of change of the option's Delta with respect of the price of the underlying asset is known as Gamma (Γ). This is the second derivative of the option with respect to the asset price

$$\Gamma = \frac{\partial^2 V}{\partial S^2}. \tag{1.1.40}$$

The Gamma for an European call option is given by

$$\Gamma = \frac{N'\left(d_1\right)}{S\sigma\sqrt{T-t}}. \tag{1.1.41}$$

where N' is the probability density function for a standard normal distribution. We can observe that the maximum is reached when the option is near at-the-money ($S \sim K$) and decreases in both out-of-the-money and in-the-money scenarios.

The Theta (Θ) of an option is the rate of change of the the value of the option with respect to the passage of time

$$\Theta = -\frac{\partial V}{\partial \tau}. \tag{1.1.42}$$

The Theta of an European call is given by

$$\Theta = -\frac{SN'(d_1)\,\sigma}{2\sqrt{T-t}} - rKe^{-r(T-t)}N(d_2) \tag{1.1.43}$$

As the option tends to be less valuable as the time passes, Theta is then negative. Moreover, when the stock price is low, Theta tends to zero and as the price becomes larger, Theta tends to $-rKe^{-r(T-t)}$.

Vega ($\mathcal{V}$) represents the rate of change of the value of a portfolio with respect to the volatility of the underlying asset. Hence, it is defined as

$$\mathcal{V} = \frac{\partial V}{\partial \sigma}. \tag{1.1.44}$$

For an European call, Vega is given by

$$\mathcal{V} = S_t\sqrt{T-t}\,N'(d_1). \tag{1.1.45}$$

Finally, we have Rho (ρ) which represents the rate of change of the value of an option with respect to changes in the interest rate. Then.

$$\rho = \frac{\partial V}{\partial r}. \tag{1.1.46}$$

For an European call, Rho is given by

$$\rho = K(T-t)\,e^{-r(T-t)}\,N(d_2) \tag{1.1.47}$$

1.1.6 The use of financial options

Options can be used for many purposes but we will list the two most important ones. The first one is for speculation and the second one is for hedging.

Let us pick example (1.1.3) and suppose that an investor thinks that the stock will rise its value over the next year. Then, he has two possibilities. The first one is to actually buy the stock at time $t = 0$. If his capital worth 1000 \$, he will be able to buy 100 stocks. The second possibility is to buy an option with maturity of 1 year. Let suppose that he bought the option with strike $K = 13$ so that he acquired 3283 options of stock S. Now, let us analyse the two possible scenarios. The first case occur when at $t = T$ the price of the stock is higher than the strike value which we can suppose is $S_T = K + C = 13 + C$, where C is a positive value. Then, if he invested in the stock market, he would have earned $(100 \times C)$ \$. However, if he had invested in the option market the earning would rise up to $(3283 \times C)$ \$: almost 33 times higher!! The second case is when the price of the stock is lower than the strike value which we can suppose as $S_T = K - C = 13 - C$, where C is a positive value. Then, in the stock market he would have lost $(100 \times C)$ \$ but in the option market he would have lost everything (as the payoff of an European option is given by $\max(S_T - K, 0)$. This analysis also show why options are riskier than stocks and why earnings can be larger but always by assuming more risk.

The same example can be used for a different analysis. We have seen that we need 1000\$ to earn $(100 \times C)$ \$ in the stock market. By doing an inverse calculation we can note that we need just 30.46 \$ to buy the 100 options and earn the same money by speculating in the options market.

However, hedging is the original purpose of financial options. They were meant to be used to reduce the risk of a portfolio at a reasonable cost, just like a standard insurance policy. Let us see an example of how hedging can be done with an option. Suppose we have a portfolio Π consisting of N_S amount of stocks of S and we want to sell an unknown amount N_V of options V. Then our portfolio's value is given by

$$\Pi = N_S S - N_V V. \tag{1.1.48}$$

Our aim is to define N_V so that $\partial\Pi/\partial S$ is equal to zero at each time step. Then, we have that

$$0 \equiv \frac{\partial \Pi}{\partial S} = N_S \frac{\partial S}{\partial S} - N_V \frac{\partial V}{\partial S}. \tag{1.1.49}$$

We can clear out N_V and obtain that, at each time step, $\Delta = N_S/N_V$. This implies that given a stock, an option and its delta at time t we can create a portfolio that eliminates the risk of the stock's price movements.

1.2 Partial Differential Equations

As we mentioned before, we recall the main definitions and results which are needed to solve the financial problems presented in Chapters 2 and 3. For this purpose we use the following books Ref [1, 32, 5, 34, 18]. as they comprise all the theoretical background that we need.

1.2.1 Sobolev spaces and properties

In this Section we are going to present the definition and main properties of Sobolev spaces. These spaces are of great importance as they allow to naturally find solutions to different partial differential equations without requiring to assume, for example, that the solutions are differentiable in the classical sense.

To define the Sobolev space we must first present the concept of *weak derivative* and, then, the Sobolev norm.

Definition 1.2.1. Let α be a multi-index and let v and h be locally integrable functions on Ω such that

$$\int_\Omega \phi\, h\, dx \;=\; (-1)^{|\alpha|} \int_\Omega v\, D^\alpha \phi\, dx \tag{1.2.1}$$

for all ϕ in $C_0^\infty(\Omega)$. Then, h is the D^α Sobolev derivative of v and write $h = D^\alpha v$.

Definition 1.2.2. Let Ω be an open set such that $\Omega \subset \mathbb{R}^n$, $k \in \{0, 1, 2, ...\}$, $1 \le p \le \infty$ and α a multi-index. We define the Sobolev norm of a function u as

$$\|u\|_{W_p^k(\Omega)} = \sum_{0\le|\alpha|\le k} \|D^\alpha u\|_p \tag{1.2.2}$$

$$\|u\|_{W_\infty^k(\Omega)} = \max_{0\le|\alpha|\le k} \|D^\alpha u\|_\infty \tag{1.2.3}$$

where $\|\cdot\|_p$ is the classical norm in $L^p(\Omega)$.

Given the the definition of weak derivative in Definition 1.2.1 and the norm defined in Definition 1.2.2, we can now construct Sobolev spaces.

Definition 1.2.3. Given k a positive integer and $1 \le p \le \infty$, we define the Sobolev space $W_p^k(\Omega)$ as

$$W_p^k(\Omega) \equiv \{u \in L^p(\Omega) : D^\alpha u \in L^p(\Omega) \text{ for } 0 \le |\alpha| \le k\}. \tag{1.2.4}$$

Further, we denote the space $W_{p,\text{loc}}^k(\Omega)$ as

$$W_{p,\text{loc}}^k \equiv \{u \in L^p(\Omega) : u\phi \in W_p^k \text{ for any } \phi \in C_0^\infty\}. \tag{1.2.5}$$

One of the main characteristics of the Sobolev spaces is that the vector space endowed with the norm defined above form a Banach space.

Theorem 1.2.4. *For each $k \in \{1, 2, 3, ...\}$ and $1 \le p \le \infty$, the Sobolev space $W_p^k(\Omega)$ is a Banach space. Moreover, $C_0^k(\bar{\Omega})$ is dense in $W_p^k(\Omega)$.*

Another important characteristics that will be used in Chapter 2 are the embeddings of various Sobolev spaces into others. Essentially, if having certain $u \in W_p^k$, we want to know: *does this function also belong to other spaces?* For this purpose, we present the Sobolev embedding theorem which comprises all possible embeddings.

Theorem 1.2.5 (The Sobolev Embedding Theorem). *Let Ω be an open set in $\mathbb{R}^n$ and, for $1 \le j \le n$, let Ω_j be the intersection of Ω with a plane of dimension j in $\mathbb{R}^n$. (If $j = n$, then $\Omega_j = \Omega$.) Let $k \ge 0$ and $m \ge 1$ be integers and let $1 \le p \le \infty$ and suppose Ω satisfies the cone condition. Then,*

Case A *If either $mp > n$ or $m = n$ and $p = 1$, then*

$$W_p^{k+m}(\Omega) \to C^k(\Omega) \tag{1.2.6}$$

Moreover, if $1 \le j \le n$, then

$$W_p^{k+m}(\Omega) \to W_q^k(\Omega_j) \quad \textit{for} \quad p \le q \le \infty \tag{1.2.7}$$

Case B *If $1 \le j \le n$ and $mp = n$, then*

$$W_p^{k+m}(\Omega) \to W_q^k(\Omega_j) \quad \textit{for} \quad p \le q \le \infty \tag{1.2.8}$$

Case C *If $mp < n$ and either $n - mp < j$ or $p = 1$ and $n - m \le j \le n$, then*

$$W_p^{k+m}(\Omega) \to W_q^k(\Omega_j) \quad \textit{for} \quad p \le q \le jp/(n - mp). \tag{1.2.9}$$

Another important result to add is that most of these embeddings are in fact compact. Let us first recall the definition of a compact embedding:

Definition 1.2.6. Let X and Y be Banach spaces, $X \subset Y$. We say that X is compactly embedded in Y if

i) $\|x\|_Y \le C\|x\|_X$ for some constant C,

ii) each bounded sequence in X is precompact in Y.

Considering the hypothesis of Theorem 1.2.5 and a bounded subset Ω_0 of Ω, the Rellich-Kondrachov Theorem provide the compactness embedding results.

Theorem 1.2.7 (The Rellich-Kondrachov Theorem). *Let Ω be an open set in $\mathbb{R}^n$, Ω_0 a bounded subdomain of Ω and let Ω_0^j the intersection of Ω_0 with a plane of j dimension in $\mathbb{R}^n$. Let $k \ge 0$ and $m \ge 1$ be integers and $1 \le p \le \infty$.*

PART I: *Suppose Ω satisfies the cone condition and $mp \le n$, then the following embedding is compact*

$$W_p^{k+m}(\Omega) \to W_q^k\left(\Omega_0^j\right) \tag{1.2.10}$$

for $0 < n - mp < k \leq n$ and $1 \leq q < kp/(n - mp)$.

PART II: *Suppose Ω satisfies the cone condition and $mp > n$, then the following embedding is compact*

$$W_p^{k+m}(\Omega) \to W_q^k\left(\Omega_0^j\right) \tag{1.2.11}$$

for $1 \leq q < \infty$.

We have listed the most important definitions and theorems of Sobolev spaces that we need to look for solutions of certain partial differential equations. However, our equations are actually *parabolic* instead of *elliptic*. This implies that we will need to work in spaces of the form $\mathbb{R}^n \times \mathbb{R}$ to consider the temporal variable. Previous definitions and theorems will help us to define the spaces and results that will be in our scope.

Under this new scope we will recall Ω as an open set of $\mathbb{R}^n$ and $\Omega_T = \Omega \times (0, T)$ the parabolic domain for some $T > 0$. Let again α be a multi-index and δ a positive constant such that $0 < \delta < 1$. We first extend the definition of weak derivative presented in Definition 1.2.1 to consider the presence of the temporal variable.

Definition 1.2.8. Let v and h be locally integrable functions on Ω_T and such that

$$\int_{\Omega_T} \phi\, h\, dx \;=\; (-1)^{|\alpha|+\rho} \int_{\Omega_T} v\, D^\alpha \partial_t^\rho \phi\, dx \tag{1.2.12}$$

for all ϕ in $C_0^\infty(\Omega)$. Then, h is the $D^\alpha \partial_t^\rho$ Sobolev derivative of v and write $h = D^\alpha \partial_t^\rho v$.

The next step is to define the Sobolev space and the respective norm.

Definition 1.2.9. Given k a positive integer and $1 \leq p \leq \infty$, we define the Sobolev space $W_p^{2k,k}(\Omega_T)$ as

$$W_p^{2k,k}(\Omega_T) \equiv \{u \in L^p(\Omega_T) : D^\alpha \partial_t^\rho u \in L^p(\Omega_T) \text{ for } 0 \leq |\alpha| + 2\rho \leq 2k\}. \tag{1.2.13}$$

This space is actually a Banach space when we endow it with the following norm

Definition 1.2.10. We define the Sobolev norm of a function u as

$$\|u\|_{W_p^{2k,k}(\Omega_T)} = \sum_{0\leq|\alpha|+2\rho\leq 2k} \|D^\alpha \partial_t^\rho u\|_p \tag{1.2.14}$$

where $\|\cdot\|_p$ is the classical norm in $L^p(\Omega_T)$.

The inclusions of $W_\infty^{2,1}$ spaces will be similar as the ones of W_∞^2 but considering the temporal variable.

Theorem 1.2.11. *Let Ω be an open set in $\mathbb{R}^n$ and $\Omega^T = \Omega\times(0,T)$. The, the following embeddings are compact*

$$W_\infty^{2,1}\left(\Omega^T\right) \to C^{1,0}\left(\Omega_T\right) \tag{1.2.15}$$
$$W_\infty^{2,1}\left(\Omega^T\right) \to W_2^{2,1}\left(\Omega_T\right) \tag{1.2.16}$$

1.2.2 Second-order Parabolic Equations

In this Section we present the most important results that arise from the study of second-order parabolic equations. For this purpose, let us first denote the following initial/boundary-value problem

$$\begin{aligned} u_t + Lu = f \quad &\text{in} \quad \Omega_T \\ u = 0 \quad &\text{in} \quad \partial\Omega\times(0,T) \\ u = g \quad &\text{in} \quad \Omega\times\{t=0\} \end{aligned} \tag{1.2.17}$$

where f and g are given functions and u is the unknown function $u(x,t)$. The letter L corresponds to the elliptic operator of the form

$$Lu = -\sum_{i,j=1}^{n} a^{ij}(x,t)\, u_{x_i,x_j} + \sum_{i=1}^{n} b^i(x,t)\, u_{x_i} + c(x,t)\, u \tag{1.2.18}$$

for given coefficients a^{ij},b^i and c.

Remark 1.2.12. For the purpose of deriving existence and uniqueness of solution for problem (1.2.17), we will let the coefficients of the operator L to follow this assumptions:

$$|b^i(X)| \leq \frac{B}{d(X)}, \qquad |c(X)| \leq \frac{c_1}{d(X)^2} \tag{1.2.19}$$
$$|a^{ij}(X) - a^{ij}(Y)| \leq w\left(\frac{|X-Y|}{d(X,Y)}\right) \tag{1.2.20}$$

where $d(\cdot)$ is the distance function, w is a positive, continuous and increasing function with $w(0) = 0$ and B and c_1 positive constants.

Definition 1.2.13. The partial differential operator $\frac{\partial}{\partial t} + L$ is parabolic if there exists a constant $\theta > 0$ such that

$$\sum_{i,j=1}^{n} a^{ij}(x,t)\,\eta_i\eta_j \geq |\eta|^2 \tag{1.2.21}$$

for all $(x,t) \in \Omega_T$ and $\eta \in \mathbb{R}^n$.

One important property of solutions of parabolic equations is that they follow a *strong maximum principle.* Hence, the maximum of this solutions is achieved in the boundary of the domain.

Theorem 1.2.14 (Strong Maximum Principle)**.** *Assume $u \in C^{1,2} \in (\Omega_T)$ and $c \geq 0$ in Ω_T. Moreover, suppose that Ω is connected. Then, if*

$$u_t + Lu \leq (\geq)\, 0 \quad in \quad \Omega_T$$

and u attains a nonnegative maximum (minimum) over $\overline{\Omega}_T$ at a point $(x_0, t_0) \in \Omega_T$, the u is constant in $\Omega \times (0, t_0)$.

The following theorem is crucial as it shows that under certain conditions, there exist a unique solution for the problem (1.2.17).

Theorem 1.2.15. *Suppose $\Omega_T \subset \mathbb{R}^{n+1}$, $p > 1$ and the coefficients of the elliptic operator L follow the conditions* (1.2.19) *and* (1.2.20)*. Then, for any $\phi \in W_p^{2,1}$ and any $f \in L^p$, there is a unique solution of*

$$\begin{aligned} u_t + Lu &= f \quad in \quad \Omega_T \\ u &= \phi \quad in \quad \mathcal{P}\Omega \end{aligned} \tag{1.2.22}$$

where $\mathcal{P}\Omega$ is the parabolic boundary of Ω_T. Moreover, u satisfies the estimate

$$\|u\|_p + \|Du\|_p + \|D^2u\|_p + \|u_t\|_p \leq C\left(\|f\|_p + \|\phi\|_p + \|D\phi\|_p + \|D^2\phi\|_p + \|\phi_t\|_p\right) \tag{1.2.23}$$

1.2.3 Two methods to prove existence

In this Section we present the main techniques that can be used to solve non-linear problems. These include different fixed point theorems as well as the super and sub solution method. We will use this techniques to prove the existence of solution of our parabolic partial differential equations.

Fixed Point Theorems

Fixed point theorems are one of the main techniques of the family of the Topological methods. There are at least two distinct classes of theorems that will be useful for us. First we have the fixed point theorems for strict contractions. Secondly, the ones for compact mappings. Let us first present these both definitions of contractions and compact mappings.

Definition 1.2.16. Let X and Y be two metric spaces, we say that $T : X \to Y$ is a contraction if there exists $\alpha < 1$ such that

$$\forall\, x, y \in X, \quad d_Y\left(T_x, T_y\right) \leq \alpha d_X\left(x, y\right) \tag{1.2.24}$$

where d_X and d_Y are the correspondent distance functions.

Definition 1.2.17. Let X and Y be two metric spaces, we say that a continuous $T : X \to Y$ is compact if $T\,\bar{(}B)$ is compact for every bounded set $B \subset X$.

The first theorem that we present is the well-known Banach theorem.

Theorem 1.2.18 (Banach)**.** *Let X be a complete metric space and let $T : X \to X$ a contraction. Then, T has a unique fixed point $\hat{x}$. Moreover, $\hat{x}$ can be calculated in an iterative way from the sequence $x_{n+1} = T\left(x_n\right)$, starting from any $x_0 \in X$.*

The second important fixed point theorem is due to Brouwer.

Theorem 1.2.19 (Brouwer)**.** *Let $B = B_1\left(0\right) \subset \mathbb{R}^n$ and $f : \bar{B} \to \bar{B}$. Then, there exists $x \in \bar{B}$ such that $f\left(x\right) = x$.*

Brouwer's theorem can be extended to Banach spaces by working on compact subsets.

Theorem 1.2.20 (Schauder)**.** *Let X be a Banach space and suppose $K \subset X$ is compact and convex. Assume also that $A : K \to K$ is continuous. Then, A has a fixed point in K.*

Last theorem is due to Schaefer. The advantage of Schaefer's theorem is that it is not necessary to identify and explicit convex, compact set.

Definition 1.2.21 (Schaefer)**.** Let X be a Banach space and $A : X \to X$ a continuous and compact mapping. Moreover, assume that the set

$$\left\{u \in X | u = \lambda A\left[u\right] \text{ for some } 0 \leq \lambda \leq 1\right\} \tag{1.2.25}$$

is bounded. Then A has a fixed point.

Method of subsolutions and supersolutions

Let us first present the definition of sub (super) solution of a parabolic partial differential equation.

Definition 1.2.22. A function u is call a sub (super) solution of a parabolic problem if

$$\begin{aligned} Lu &\leq (\geq) f(x,t) && \text{in} \quad \Omega_T \\ u(x,0) &\leq (\geq) u_0(x) && \text{in} \quad \Omega \\ u(x,t) &\leq (\geq) g(x,t) && \text{in} \quad \partial\Omega \times (0,T). \end{aligned} \tag{1.2.26}$$

where L is a parabolic operator. If α is a subsolution and β is a supersolution, we say that the pair α, β is ordered if

$$\alpha(x,t) \leq \beta(x,t)$$

The main idea behind the method of subsolutions and supersolutions is to exploit the ordering properties for solutions of partial differential equations. Essentially, if we can find one subsolution $\bar{u}$ and a supersolution $\underline{u}$ of a particular boundary-value problem, and if also $\underline{u} \leq \bar{u}$, then there exists a solution satisfying

$$\underline{u} \leq u \leq \bar{u}. \tag{1.2.27}$$

Despite given the sub and supersolutions the existence of the main solution can be granted, the main difficulty is actually how to find it. The first step relies into converting the problem into a fixed point problem as of $x = Tx$ where T will be an operator related to the partial differential equation (for example, if we recall Equation (1.2.26), we can define $T = L^{-1}f$ and work with the fixed point problem).

Suppose now we have X a Banach space equipped with an order $\leq$ induced by some cone K and let $T : X \to X$ a continuous monotone nondecreasing operator. As before, we say that α and β are sub and supersolutions of a fixed point problem $x = Tx$ if

$$\alpha \leq T\alpha \quad \text{and} \quad \beta \geq T\beta \tag{1.2.28}$$

respectively. Moreover, lets define the sequences $u_{n+1} = Tu_n$ and $v_{n+1} = Tv_n$ such that $\alpha = u_0$ and $\beta = v_0$. Without including any other assumption, this sequences might not converge. Hence, we have to add two more conditions.

Definition 1.2.23. Let X be a Banach space and $K \subset X$ a closed cone. The order $\leq$ induced by K is normal if there exists a constant $c > 0$ such that

$$0 \leq x \leq y \Rightarrow \|x\| \leq c\|y\| \tag{1.2.29}$$

We first need that the normality is fulfilled. This will imply that the sequences of sub and supersolutions are actually bounded. Further, we will need that the operator T is compact. The inclusion of these two assumptions lead to the following theorem

Theorem 1.2.24. *Let X be a Banach space and $K \subset X$ a closed cone. Assume that the order induced by K is normal and that $T : X \to X$ is compact and nondecreasing. If (α, β) is a well-ordered couple of sub and supersolutions, then the sequences defined by*

$$\begin{aligned} u_0 &= \alpha, \quad u_{n+1} = Tu_n, \\ v_0 &= \beta, \quad v_{n+1} = Tv_n \end{aligned}$$

converge respectively to fixed points u, v of T such that $\alpha \leq u < v \leq \beta$.

1.2.4 Viscosity solutions

Another type of solutions of nonlinear parabolic equations are the ones known as "viscosity solutions". They correspond to a different notion of weak solution as the ones of the usual Sobolev framework. The term "viscosity" follows from historical reasons and refers the existence of a method of obtaining a solution by adding an artificial viscosity term to the equation and obtaining a solution by passing to a vanishing viscosity limit. We are going to list some definitions that are important to then define a viscosity solution. First of all, given an operator F, Definition 1.2.25 presents the definition of a degenerate elliptic operator.

Definition 1.2.25. A nonlinear operator F is degenerate elliptic if

$$A \leq B \implies F(t, x, p, s, A) \geq F(t, x, p, s, B). \tag{1.2.30}$$

The degeneracy of the nonlinear operator is a necessary condition to assure the existence of a viscosity solution. Next, in order to define a viscosity solution, we shall start by defining the notion of upper and lower semi-continuity. During this Section we are going to follow the notes that correspond Ref [28].

Given an open set $\Omega_T \subset \mathbb{R}^{N+1}$, we recall that V is *lower semi-continuous* (LSC) or *upper semi-continuous* (USC) at (t, x) if for all sequences $(s_n, y_n) \to (t, x)$,

$$\begin{aligned} V(t, x) &\leq \liminf_{n\to\infty} V(s_n, y_n) \quad \text{(LSC)} \\ V(t, x) &\geq \limsup_{n\to\infty} V(s_n, y_n) \quad \text{(USC)}. \end{aligned}$$

Moreover, we define V_* the *lower semi-continuous envelope of V* as the largest lower semi-continuous function lying below V and V^* the correspondent *upper semi-continuous envelope of V* as the smallest upper semi-continuous function lying above V. Let us now define our Dirichlet problem as of

$$\frac{\partial V}{\partial \tau} + F\left(\tau, x, V, DV, D^2V\right) = 0 \quad \text{in} \quad \Omega \times [0, T]$$
$$V\left(0, x_1, ..., x_N\right) = V_0\left(x_1, ..., x_N\right) \quad \text{in} \quad \Omega \tag{1.2.31}$$

where F is a degenerate elliptic nonlinear operator and $V_0\left(x_1, ..., x_N\right)$ is the initial condition.

Let us continue by presenting the definition of viscosity solutions, which are the type of solutions that we will look for. Let us consider an open set $\Omega_T \subset \mathbb{R}^{N+1}$ and a function $V \in C^{1,2}\left(\Omega_T\right)$. Then, we have the following definitions.

Definition 1.2.26. U is a subsolution of (1.2.31) if U is upper semi-continuous and if, for all $(t, x) \in \Omega_T$ and all the test functions ϕ such that $U \leq \phi$ in a neighbourhood of (t, x) and $U\left(t, x\right) = \phi\left(t, x\right)$, we have that

$$\frac{\partial \phi}{\partial \tau} + F\left(\tau, x, \phi, D\phi, D^2\phi\right) \leq 0. \tag{1.2.32}$$

U is a supersolution of (1.2.31) if U is lower semi-continuous and if, for all $(t, x) \in \Omega_T$ and all the test functions ϕ such that $U \geq \phi$ in a neighbourhood of (t, x) and $U\left(t, x\right) = \phi\left(t, x\right)$, we have that

$$\frac{\partial \phi}{\partial \tau} + F\left(\tau, x, \phi, D\phi, D^2\phi\right) \geq 0. \tag{1.2.33}$$

Finally, U is a solution of (1.2.31) if it is both a sub and supersolution.

Proposition 1.2.27. *Two important properties of sub and super-solutions are the following ones.*

- *Let* $\left(V_\alpha\right)_\alpha$ *be a family of sub-solutions of problem* (1.2.31) *in* Ω_T *such that the upper semi-continuous envelope* V *of* $\sup_\alpha V_\alpha$ *is finite in* Ω_T*. Then* V *is also a sub-solution of problem* (1.2.31) *in* Ω_T*.*
- *If* $\left(V_n\right)$ *is a sequence of sub-solutions of problem* (1.2.31)*, then the upper relaxed-limit* V *of the sequence defined as follows*

  $$\overline{V}\left(\tau, x\right) = \limsup_{n \to \infty, (s,y) \to (\tau, x)} V_n\left(s, y\right)$$

  is everywhere finite in Ω_T*, then it is a subsolution of problem* (1.2.31) *in* Ω_T*.*

Definition 1.2.26 presents not only the actual definition of a viscosity solution but also introduces the first step to find it: It is crucial to identify a pair of sub and super-solutions of problem (1.2.31). The method that helps us to derive the existence of a viscosity solution is the Perron process. The general idea of this methodology is to construct a sub-solution V^- and a super-solution V^+ of the nonlinear parabolic equation (1.2.31) such that $V^- \leq V^+$. Using Proposition 1.2.27, we can construct a maximal sub-solution V lying between V^- and V^+. Following a general argument, we can prove that the lower semi-continuous envelope of the maximal subsolution V is in fact a supersolution.

Now we can present the Perron method to find a solution of problem (1.2.31). We are going to require first that the nonlinear operator F is degenerate elliptic. Then, Perron method is defined as follows.

Theorem 1.2.28. *Suppose w is a subsolution of problem* (1.2.31) *and v is a supersolution of problem* (2.3.35) *such that $w \leq v$. Suppose also that there is a subsolution $\underline{u}$ and a supersolution $\overline{u}$ of problem* (1.2.31) *that satisfy the boundary condition $\underline{u}_* (t,x) = \overline{u}^* (t,x) = g(t,x)$. Then,*

$$W(t,x) = \sup\{w(t,x) : \underline{u} \leq w \leq \overline{u} \, and w \; is \; a \; subsolution \; of \; (2.3.35)\}. \tag{1.2.34}$$

is a solution of problem (1.2.31).

The method of Perron guarantees that the supreme of the set of sub-solutions of problem (1.2.31) that lies between the original sub and super-solutions is indeed a viscosity solution. However, we need some extra result to assure that boundary conditions are being fulfilled. For this purpose, we need some type of comparison principle. Following the notes of Ref [28], we recall the next principle.

Proposition 1.2.29 (Comparison Principle)**.** *If u is a sub-solution of problem* (1.2.31) *and v is a super-solution of problem* (1.2.31) *in Ω_T and $u \leq v$ on the parabolic boundary $\partial_p \Omega_T$, then $u \leq v$ in Ω_T.*

Finally, the combination of both Theorem 1.2.28 and Proposition 1.2.29 let us assure the existence of solution of problem (1.2.31).

1.3 Numerical Analysis

Our work presented in Chapters 2 and 3 involves finding the solutions of the partial differential equations via a numerical framework. In both problems we use a standard explicit finite differences approach with forward differences in the temporal variable and central differences in the spatial variable. The results of convergence, consistency and stability for this method are considered valid during the work's course. In this Section we recall this results from the standard bibliography (Ref [35] and Ref [41]) and isolate the main properties that this framework has. Moreover, given that in Chapter 3 we solve a three-dimensional problem (one temporal dimension and two spatial dimensions), we will also present and explain the Alternating Direction Implicit (ADI) method following Ref [3], Ref [37] and Ref [42]. This methodology belongs to the family of the operating splitting methods and is useful when dealing with crossed derivatives inside the partial differential equation.

1.3.1 One-dimensional finite differences methods

In this Section we present the results that arise from the study and use of the finite difference methods on parabolic problems. Moreover, we recall the main steps that are needed to develop a numerical framework on parabolic equations. For this purpose, and to show the main properties of this methods, we will reduce our parabolic problems to the following: find $u(x,t)$ with $x \in [0,1]$ and $t \geq 0$ such that

$$\begin{aligned} \frac{\partial u}{\partial t} &= \frac{\partial^2 u}{\partial x^2} \quad \text{for} \quad t > 0, 0 < x < 1 \\ u(0,t) = u(1,t) &= 0 \quad \text{for} \quad t > 0 \\ u(x,0) &= u_0(x) \quad \text{for} \quad 0 \leq x \leq 1. \end{aligned} \tag{1.3.1}$$

We first need to set the scheme in which we will be defining our solution u. We form an equally spaced grid on the closed domain $\overline{\Omega} = [0,1] \times [0, t_C]$ where t_C can be as large as desirable. If Δx and Δt are the line spacings, we define each of the points of the grid as

$$\begin{aligned} x_j &= j\Delta x \quad \text{for} \quad 1 \leq j \leq J \\ t_n &= n\Delta t \quad \text{for} \quad 1 \leq n \leq N \end{aligned}$$

Under this framework, our aim is to approximate the solution u by its values in the grid points as of $u_j^n \simeq u(x_j, t_n)$. We then discretize both temporal and spatial derivatives and apply them in equation (1.3.1). Hence, we find that

$$\frac{\partial u}{\partial t}(x_j, t_n) \simeq \frac{u(x_j, t_{n+1}) - u(x_j, t_n)}{\Delta t} \tag{1.3.2}$$

and

$$\frac{\partial^2 u}{\partial x^2}(x_j, t_n) \simeq \frac{u(x_{j+1}, t_n) - 2u(x_j, t_n) + u(x_{j-1}, t_n)}{\Delta x^2} \tag{1.3.3}$$

and Equation (1.3.1) can be approximated by the following *explicit difference scheme*

$$u_j^{n+1} = u_j^n + \mu\left(u_{j+1}^n - 2u_j^n + u_{j-1}^n\right) \tag{1.3.4}$$

where $\mu = \Delta t / (\Delta x)^2$.

The first question that arises from this scheme corresponds to the error that is get from the approximation done.

Definition 1.3.1 (Truncation Error)**.** We define the truncation error as the difference between the two sides of Equation (1.3.4). Moreover, if we calculate the Taylor series expansions of both Equations (1.3.2) and (1.3.3), we get that

$$T(x,t) := \frac{1}{2}\frac{\partial^2 u}{\partial t^2}(x,\eta)\,\Delta t - \frac{1}{12}\frac{\partial^4 u}{\partial x^4}(\nu,t)\,(\Delta x)^2\,. \tag{1.3.5}$$

One desired property of a finite-difference scheme is to have a solution that actually converge to the solution of the differential equation as the mesh lengths tend to zero. This can be traduced in terms of the convergence of the truncation error, i.e.

Definition 1.3.2 (Consistency)**.** A finite-difference scheme is said to be unconditionally consistent if

$$T(x,t) \to 0 \quad \text{as} \quad \Delta x, \Delta t \to 0\, \forall\,(x,t) \in (0,1) \times [\tau, t_C]\,. \tag{1.3.6}$$

The second and third conditions that have to be satisfied if the solution of the finite-difference equation is to be an accurate approximation of the corresponding parabolic partial differential equation are (1) the convergence of the solution of the approximated difference equation to the actual solution and (2) the controlled decay or boundedness of the rounding errors that are introduced during the computation.

Definition 1.3.3 (Convergence)**.** We say that the scheme is convergent if for any fixed point (x^*, t^*) in the domain $\overline{\Omega}$ we have that

$$x_j \to x^* \quad \text{and} \quad t_n \to t^* \quad \text{implies} \quad u_n^j \to u(x^*, t^*) \tag{1.3.7}$$

The following convergence theorem guarantees that an arbitrarily high accuracy can be attained by using a sufficiently fine mesh. We generalize the definition of μ to consider different mesh sizes. Then, we define μ_i as

$$\mu_i = \frac{(\Delta t)_i}{(\Delta x)_i^2}$$

Theorem 1.3.4. *Suppose $\mu \leq 1/2$ for all sufficiently large values of i, the positive numbers n_i,j_i are such that*

$$n_i\,(\Delta t)_i \to t > 0, \quad j_i\,(\Delta x)_i \to x \in [0,1] \tag{1.3.8}$$

and $|\frac{\partial^4 u}{\partial x^4}| \leq M$ uniformly in $\overline{\Omega}$. Then, the approximations $u_{n_i}^{j_i}$ generated by the explicit difference scheme (1.3.4) *converge uniformly to the solution $u(x,t)$ of the differential equation in the region.*

Last condition desired for a numerical scheme is to be stable. The idea behind the stability is that the numerical process should limit the amplification of all components of the initial conditions. This analysis is usually done by expressing the solution as a Fourier series and observing the characteristics of the Fourier modes. Suppose we substitute

$$u_j^n = \lambda^n e^{ik(j\Delta x)} \tag{1.3.9}$$

in the equation (1.3.4). If we set $u_j^{n+1} = \lambda u_j^n$ and divide the whole equation by u_j^n, we get

$$\lambda = 1 - 4\mu \sin^2 \frac{1}{2} k\Delta x \tag{1.3.10}$$

where λ is known as the amplification factor for the mode. This amplification factor will help us to measure the speed at which the solution grows. Hence, a desirable stable scheme would require that the difference between two solutions of the difference equations to be bounded. This idea leads to the following definition of stability.

Definition 1.3.5 (Stability - Von-Neumann)**.** A difference scheme is said to be stable when there exists a positive number M independent of Δx and Δt such that

$$|\lambda(k)| \leq 1 + M\Delta t \quad \text{for all} \quad k. \tag{1.3.11}$$

This condition applied to problem (1.3.4) leads to requiring $\mu \leq 1/2$.

The scheme proposed in Equation (1.3.4) can be generalized to use not only the last three temporal steps (forward scheme) but also the three actual temporal steps (implicit scheme). This idea can be applied in terms of a weighted average of both schemes. Effectively, we can propose a six-point scheme

$$u_j^{n+1} = u_j^n + \mu \left[\theta \left(u_{j+1}^{n+1} - 2u_j^{n+1} + u_{j-1}^{n+1}\right) + (1-\theta)\left(u_{j+1}^n - 2u_j^n + u_{j-1}^n\right)\right] \tag{1.3.12}$$

The parameter θ lies in the interval $[0, 1]$ where $\theta = 0$ corresponds to an explicit scheme and $\theta = 1$ to an implicit scheme. When the value of θ is set as $1/2$, the scheme is known as the Crank-Nicolson scheme. This choice has good properties with respect to the stability, consistency and convergence of the solution.

To analyze the stability of the Crank-Nicolson scheme we apply a Fourier analysis. If we substitute the mode (1.3.9) into equation (1.3.4), we obtain

$$\lambda - 1 = \mu \frac{\lambda + 1}{2} \left(-4\sin^2 \frac{k\Delta x}{2}\right) \tag{1.3.13}$$

which is equivalent to

$$\lambda = \frac{1 - 2\mu \sin^2 \frac{k\Delta x}{2}}{1 + 2\mu \sin^2 \frac{k\Delta x}{2}} \tag{1.3.14}$$

Given that $\lambda < 1$, instability can arise from the case when $\lambda < -1$. However, as we have chosen $\theta = 1/2$, we see that λ always lies in $[-1, 1]$ for every μ.

The truncation error can be calculate by performing a Taylor expansion in (1.3.4). This expansion is done over the point $\left(x_j, t_{n+1/2}\right)$. After substituting all the expansions in the scheme and ripping out all the terms that are cancelled out, it can be showed that the truncation error for the Crank-Nicolson scheme is given by

$$T_j^{n+1/2} = -\frac{1}{12}\left[(\Delta x)^2 \frac{\partial^4 u}{\partial x^4} + (\Delta t)^2 \frac{\partial^3 u}{\partial t^3}\right]_j^{n+1/2} + O\left(\Delta x\right)^4 \tag{1.3.15}$$

Therefore, the Crank-Nicolson scheme is consistent and always second-order accurate in both Δt and Δx. Moreover, the following theorem shows that the scheme also converge when $\mu \leq 1$.

Theorem 1.3.6. *Suppose* $\mu\left(1-\theta\right) \leq 1/2$, $0 \leq \theta \leq 1$ *and* u_j^n *satisfying*

$$u_{min} \leq u_j^n \leq u_{max} \tag{1.3.16}$$

where

$$u_{min} = \min\{u_j^n\} \tag{1.3.17}$$

and

$$u_{max} = \max\{u_j^n\}. \tag{1.3.18}$$

Then, for any refinement path which satisfies the stability condition 1.3.5, the approximations given by (1.3.12) *with consistent initial and Dirichlet boundary conditions converge uniformly on* $[0,1]\times[0,t_C]$ *if the initial data are smooth enough for the truncation error* $T_j^{n+1/2}$ *to tend to zero along the refinement path uniformly in this domain.*

In particular, if we consider this theorem, the truncation error shown in Equation (1.3.15) and the value of λ in equation (1.3.14), we can assert that the Crank-Nicolson scheme converge.

1.3.2 Two and three-dimensional methods: The ADI method

The previous Section comprised the analysis of the basic numerical framework to be applied on a parabolic unidimensional equation. However, the computational cost of treating the equation similarly increases when more dimensions are added. Indeed, an important characteristic of Equation (1.3.4) is that the system matrix A is tridiagonal. This implies that fast algorithms can be applied to solve the iterative problem. As the number of dimensions increases, crossed-derivatives become real and the resultant system matrix is no longer tridiagonal. In this framework is when the *Alternating Direction Implicit* (ADI) method takes a crucial role.

Let us first define the forward difference operators in the temporal variable and the central difference operators in the spatial variable.

$$\Delta_t u\left(x,t\right) = u\left(x,t+\Delta t\right) - u\left(x,t\right), \quad \delta_t u\left(x,t\right) = u\left(x,t+\frac{1}{2}\Delta t\right) - u\left(x,t-\frac{1}{2}\Delta t\right)$$
$$\Delta_x u\left(x,t\right) = u\left(x+\Delta x,t\right) - u\left(x,t\right), \quad \delta_x u\left(x,t\right) = u\left(x+\frac{1}{2}\Delta x,t\right) - u\left(x-\frac{1}{2}\Delta x,t\right)$$

Suppose we have a bidimensional heat equation given by

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2}. \tag{1.3.19}$$

The first approach would be to apply the standard Crank-Nicolson framework in two dimensions. This would give us

$$u_{j,l}^{n+1} = u_{j,l}^{n} + \frac{1}{2}\frac{\Delta t}{\Delta x}\left(\delta_x^2 u_{j,l}^{n+1} + \delta_x^2 u_{j,l}^{n} + \delta_y^2 u_{j,l}^{n+1} + \delta_y^2 u_{j,l}^{n}\right). \tag{1.3.20}$$

As we mentioned before, this scheme has one issue. When trying to solve the coupled linear equations, the system is no longer tridiagonal. One way to solve this problem is to generalize the Crank-Nicolson scheme differently. The main idea is to split each time step into two steps of size $\delta/2$. In each substep, a different dimension is treated implicitly. Then, the scheme proposed is of the form

$$u_{j,l}^{n+1/2} = u_{j,l}^{n} + \frac{1}{2}\frac{\Delta t}{\Delta x}\left(\delta_x^2 u_{j,l}^{n+1/2} + \delta_y^2 u_{j,l}^{n}\right) \tag{1.3.21}$$
$$u_{j,l}^{n+1} = u_{j,l}^{n+1/2} + \frac{1}{2}\frac{\Delta t}{\Delta x}\left(\delta_x^2 u_{j,l}^{n+1/2} + \delta_y^2 u_{j,l}^{n+1}\right) \tag{1.3.22}$$

As it can be observed, the main gain of this method is that, in each substep, the solution is obtained via solving a tridiagonal system. This method can indeed be generalize in terms of splitting operators. If we rewrite Equation (1.3.19) as of

$$\frac{\partial u}{\partial t} = \mathcal{L}u \tag{1.3.23}$$

where $\mathcal{L}$ is the parabolic operator. Suppose we can write $\mathcal{L}$ as a linear sum of m pieces as of

$$\mathcal{L}u = \mathcal{L}_1 u + ... + \mathcal{L}_m u \tag{1.3.24}$$

Moreover, we construct $\mathcal{L}_i$ such that we know how to solve each $\mathcal{L}_i u$ from time step n to $n+1$ if that piece of the operator where the only one on the right-hand side of the equation (for example, via a tridiagonal system). These updates can be noted as $u^{n+1} = \mathcal{T}_i\left(u^n, \Delta t\right)$. Hence, the steps to update u^m to u^{m+1} are the following:

$$\begin{aligned} u^{n+1/m} &= \mathcal{T}_1\left(u^n, \Delta t/m\right) \\ u^{n+2/m} &= \mathcal{T}_2\left(u^{n+1/m}, \Delta t/m\right) \\ &\cdots \\ u^{n+1} &= \mathcal{T}_m\left(u^{n+(m-1)/m}, \Delta t/m\right) \end{aligned} \tag{1.3.25}$$

Based on this procedure, different ADI methods can be proposed to solve parabolic equations. One of the first methods is due to Peaceman and Rachford Ref [36].

Peaceman and Rachford scheme

If we continue working with the bidimensional equation (1.3.19), we want to find $\mathcal{T}_1$ and $\mathcal{T}_2$ to solve the equation

$$\frac{\partial u}{\partial t} = \mathcal{T}_1 u + \mathcal{T}_2 u. \tag{1.3.26}$$

Following the idea of the Crank-Nicolson scheme, and without discretizing in the spatial variable, by Taylor series we have that

$$\frac{u^{n+1} - u^n}{\Delta t} = \frac{1}{2}\left(\mathcal{T}_1 u^{n+1} + \mathcal{T}_1 u^n\right) + \frac{1}{2}\left(\mathcal{T}_2 u^{n+1} + \mathcal{T}_2 u^n\right) + O\left(\Delta t^2\right) \tag{1.3.27}$$

which is equivalent to

$$\left(I - \frac{\Delta t}{2}\mathcal{T}_1 - \frac{\Delta t}{2}\mathcal{T}_2\right) u^{n+1} = \left(I + \frac{\Delta t}{2}\mathcal{T}_1 + \frac{\Delta t}{2}\mathcal{T}_2\right) u^n + O\left(\Delta t^3\right). \tag{1.3.28}$$

By adding $\Delta t^2 \mathcal{T}_1 \mathcal{T}_2 u^{n+1}/4$ on both sides of the equation and the factorizing the terms that multiply u^{n+1} and u^n we obtain

$$\left(I - \frac{\Delta t}{2}\mathcal{T}_1\right)\left(I - \frac{\Delta t}{2}\mathcal{T}_2\right)u^{n+1} = \left(I + \frac{\Delta t}{2}\mathcal{T}_1\right)\left(I + \frac{\Delta t}{2}\mathcal{T}_2\right)u^n + \frac{\Delta t^2}{4}\mathcal{T}_1\mathcal{T}_2 u^{n+1} + \left(\Delta t^3\right). \tag{1.3.29}$$

Given that $\Delta t^2 \mathcal{T}_1 \mathcal{T}_2 u^{n+1}/4$ is of order Δt^3 we have

$$\left(I - \frac{\Delta t}{2}\mathcal{T}_1\right)\left(I - \frac{\Delta t}{2}\mathcal{T}_2\right)u^{n+1} = \left(I + \frac{\Delta t}{2}\mathcal{T}_1\right)\left(I + \frac{\Delta t}{2}\mathcal{T}_2\right)u^n + \left(\Delta t^3\right). \tag{1.3.30}$$

Peaceman and Rachford solve equation (1.3.30) by proposing the following two steps

$$\begin{aligned}\left(I - \frac{\Delta t}{2}\mathcal{T}_1\right)u^{n+1/2} &= \left(I + \frac{\Delta t}{2}\mathcal{T}_2\right)u^n \\ \left(I - \frac{\Delta t}{2}\mathcal{T}_2\right)u^{n+1} &= \left(I + \frac{\Delta t}{2}\mathcal{T}_1\right)u^{n+1/2}\end{aligned}$$

Douglas and Rachford scheme

The second important scheme is due to Douglas and Rachford Ref [17]. If we start with the backward-time-central-space scheme for equation (1.3.26), we have

$$\left(I - \Delta t\mathcal{T}_1 - \Delta t\mathcal{T}_2\right)u^{n+1} = u^n + O\left(\Delta t^2\right) \tag{1.3.31}$$

which is equivalent to

$$\left(I - \Delta t\mathcal{T}_1 - \Delta t\mathcal{T}_2 + \Delta t^2\mathcal{T}_1\mathcal{T}_2\right)u^{n+1} = u^n + \Delta t^2\mathcal{T}_1\mathcal{T}_2 u^n + \Delta t^2\mathcal{T}_1\mathcal{T}_2\left(u^{n+1} - u^n\right) + O\left(\Delta t^2\right). \tag{1.3.32}$$

If we don't consider the terms of order higher that Δt^2, we have

$$\left(I - \Delta t\mathcal{T}_1 -\right)\left(I - \Delta t\mathcal{T}_2\right)u^{n+1} = \left(I + \Delta t^2\mathcal{T}_1\mathcal{T}_2\right)u^n. \tag{1.3.33}$$

The Douglas and Rachford scheme is then given by

$$\begin{aligned}\left(I - \Delta t\mathcal{T}_1 -\right)u^{n+1/2} &= \left(I + \Delta t\mathcal{T}_2\right)u^n. \\ \left(I - \Delta t\mathcal{T}_2\right)u^{n+1} &= u^{n+1/2} - \Delta t\mathcal{T}_2 u^n\end{aligned}$$

Further results regarding the convergence, stability and consistency of the numerical method can be found in Ref [29] and Ref [30].

1.4 Resumen del capítulo

El objetivo del Capítulo 1 es introducir el marco teórico a utilizar a lo largo del trabajo de tesis. El mismo está dividido en tres secciones. Esencialmente, presentamos los conceptos básicos financieros que servirán para entender el trasfondo de los problemas de valuación, la teoría escencial de los espacios de Sobolev y las soluciones débiles en ecuaciones diferenciales parabólicas, la teoría de soluciones viscosas y finalmente los diversos esquemas numéricos utilizados para encontrar soluciones aproximadas a los problemas planteados.

En la Sección 1.1 introducimos el concepto de opción financiera y la forma de modelar su activo subyacente utilizando el movimiento browniano. Asimismo, presentamos la formula de Itô la cual resulta ser una herramienta de vital importancia a la hora de deducir la ecuación diferencial correspondiente. A continuación mostramos dos formas de arribar al valor de una opción financiera. La Sección 1.1.2 presenta la primera forma, llamada valuación via replicación, la cual consiste en construir un portfolio compuesto por una opción y una cantidad no fija del activo subyacente. Utilizando la fórmula de Itô y diversos pasos algebraicos obtenemos finalmente la ecuación diferencial que modela el comportamiento de la opción en cuestión. La Sección 1.1.3 presenta la segunda forma, llamada valuación via expectación, la cual se basa en la construcción de un portfolio con cantidades de el activo subyacente y un bono libre riesgo. El objetivo es notar que, bajo un cambio de medida, el valor del portfolio descontado resulta ser una martingala. Tanto la definición de martingala como el teorema de representación de martingalas nos permiten notar que el valor de la opción no es otra cosa que el valor esperado condicional del payoff descontado. En la Sección 1.1.4 incluimos el teorema de Feynman-Kac dado que permite conectar ambos resultados: la solución de la ecuación diferencial con el valor de la expectación condicional. La Sección 1.1.5 introduce las griegas de una opción financiera, las cuales permiten estudiar el comportamiento de la misma ante cambios en diferentes variables del modelo. Finalmente, el objetivo de la Sección 1.1.6 consiste en explicar por que el problema de valuación de opciones es relevante y cuales son los posibles usos que se les puede dar a la misma.

La Sección 1.2 introduce la definición de los espacios de Sobolev y sus principales propiedades. Además, incluimos dos métodos para probar la existencia de solución de distintas ecuaciones no lineales: los métodos de punto fijo y el método de super y sub soluciones. Finalmente, presentamos las nociones básicas correspondientes a las soluciones viscosas, incluyendo el método de Perron para demostrar la existencia de solución.

El objetivo de la Sección 1.3 es recordar los métodos mas comunes para resolver numericamente ecuaciones diferenciales parabólicas. Presentamos el esquema de Crank-Nicholson, con sus propiedades principales, resultados de estabilidad y convergencia. Dado que nosotros trabajaremos con ecuaciones de mas de una dimension espacial, presentamos el método ADI y dos posibles esquemas numericos: el Peaceman-Rachford y el Douglas-Rachford.

Chapter 2

Multi-asset option pricing with general transaction costs

This chapter is devoted to the presentation and explanation of our first work Ref [7].

2.1 Introduction

In the previous Section 1.1 we showed the steps that have to be applied in order to construct the equation that governs the price of an option. Specifically, the Black-Scholes equation (1.1.12), is of the parabolic type and depends on the values of the volatility of the underlying asset σ, the risk-free interest rate r and the price of the asset S. However, this equation remains valid under the assumptions listed in Section 1.1.2. If we relax those assumptions, we may find that the option's price follow different dynamics. We will now present some examples of 'adjusted' Black-Scholes formulas for multiple changes of the original settings.

If we consider that the stock can pay dividends over the life of the option, a Black-Scholes equation can be constructed following the same steps. If we define q as the dividend payment rate of stock S (i.e. the dividend payment is equal to $q\,S_t\,dt$), the new Black-Scholes formula is given by

$$\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + (r-q)\,S\frac{\partial V}{\partial S} - rV = 0 \tag{2.1.1}$$

One parameter that can be modelled with an stochastic component is the volatility. The well-known Heston model Ref [26] propose a stochastic volatility of the form

$$dS_t = \mu S_t dt + \sqrt{v_t} S_t dW_t^s \tag{2.1.2}$$
$$dv_t = k\,(\theta - v_t)\,dt + \sigma\sqrt{v_t} dW_t^v \tag{2.1.3}$$

where the deterministic component correspond to a mean-reversion. Under this volatility model, the price of an European call has the following dynamics

$$\frac{\partial V}{\partial t}+\frac{1}{2}v^2S^2\frac{\partial^2 V}{\partial S^2}+\rho\sigma vS\frac{\partial^2 V}{\partial S\partial v}+\frac{1}{2}\sigma^2v^2\frac{\partial^2 V}{\partial v^2}+rS\frac{\partial V}{\partial S}+\left[k\left(\theta-v\right)-\lambda\right]\frac{\partial V}{\partial v}-rV=0 \tag{2.1.4}$$

where the λ parameter is known as the market price of the volatility. According to Ref [19], the premium of volatility risk is set as $\lambda\left(S,v,t\right)=\lambda v$ with $\lambda\in\mathbb{R}$. We can see that the problem now becomes bidimensional but the linearity of the equation is preserved.

Nonlinear equations arise when we include the existence of transaction cost in the construction of the portfolio Π in (1.1.7). On each step, we will have to own Δ shares of stock S. However, the cost of buying or selling a certain amount those stocks is not considered. The seminal paper on this topic is due to Leland Ref [33]. In his work, he proposed that, on each time step, the transaction costs are equal to $k|v|S$, where v is the number of shares bought or sold and k is a proportional constant characteristic to the individual investor. Hence, the one-step change of the value of the portfolio is given by

$$d\Pi=dV-\Delta dS-k|v|S \tag{2.1.5}$$

After applying the same steps as the ones in the standard Black-Scholes model, the equation obtained is of the form

$$\frac{\partial V}{\partial t}+\frac{1}{2}\sigma^2S^2\frac{\partial^2 V}{\partial S^2}+rS\frac{\partial V}{\partial S}-rV-k\sigma S^2\sqrt{\frac{2}{\pi dt}}\left|\frac{\partial^2 V}{\partial S^2}\right|=0 \tag{2.1.6}$$

In this model, dt is a non-infinitesimal fixed time-stop not to be taken $dt\to 0$. This first model lead to different variations and generalizations when considering transaction costs in option pricing. In Ref [4], the authors proposed that transaction costs follow a linear decreasing function as of $\left(a-b|v|\right)S|v|$. Under this scenario, Leland's equation (2.1.6) is modified to include a new term

$$\frac{\partial V}{\partial t}+\frac{1}{2}\sigma^2S^2\frac{\partial^2 V}{\partial S^2}+rS\frac{\partial V}{\partial S}-rV-a\sigma S^2\sqrt{\frac{2}{\pi dt}}\left|\frac{\partial^2 V}{\partial S^2}\right|+bS^3\sigma^2\frac{\partial^2 V}{\partial S^2}^2=0 \tag{2.1.7}$$

Moreover, multiple assumptions can be relaxed at the same type. In Ref [20] the author combine the existence of transaction costs with the presence of an stochastic volatility. The stochastic representation of both the asset price and the volatility are of the form

$$dS_t=\mu\left(S_t\right)dt+\sqrt{v_t}S_tdW_t^s \tag{2.1.8}$$

$$dv_t=\alpha\left(v_t\right)dt+\sigma\sqrt{v_t}dW_t^v \tag{2.1.9}$$

Under the new assumption that the volatility can be considered as a traded asset, the adjusted Black-Scholes equation is

$$\begin{aligned}
&\frac{\partial V}{\partial t}+\frac{1}{2}v^2S^2\frac{\partial^2 V}{\partial S^2}+\rho\sigma vS\frac{\partial^2 V}{\partial S\partial v}+\frac{1}{2}\sigma^2v^2\frac{\partial^2 V}{\partial v^2}+rS\frac{\partial V}{\partial S}+rv\frac{\partial V}{\partial v}-rV\\
&-kS\sqrt{\frac{2}{\pi dt}}\sqrt{v^2S^2\left(\frac{\partial^2 V}{\partial S^2}\right)^2+2\rho\sigma v^2S\frac{\partial^2 V}{\partial S^2}\frac{\partial^2 V}{\partial S\partial v}+\sigma^2v^2\left(\frac{\partial^2 V}{\partial S\partial v}\right)^2}\\
&-k_1v\sqrt{\frac{2}{\pi dt}}\sqrt{v^2S^2\left(\frac{\partial^2 V}{\partial S\partial v}\right)^2+2\rho\sigma v^2S\frac{\partial^2 V}{\partial v^2}\frac{\partial^2 V}{\partial S\partial v}+\sigma^2v^2\left(\frac{\partial^2 V}{\partial v^2}\right)^2}=0
\end{aligned} \tag{2.1.10}$$

where ρ is the correlation between dW_t^v and dW_t^s.

A similar equation is obtained in Ref [38] when both transaction costs and stochastic interest rates are considered. Under this scenario, the stock's price and the interest rate process are modelled following these equations

$$dS_t=\mu S_tdt+\sigma S_tdW_t^s \tag{2.1.11}$$

$$dr_t=u\left(r,t\right)dt+w\left(r,t\right)dW_t^r \tag{2.1.12}$$

The nonlinear equation that arises from this setting is of the form

$$\begin{aligned}
&\frac{\partial V}{\partial t}+\frac{1}{2}\sigma^2S^2\frac{\partial^2 V}{\partial S^2}+\rho\sigma wS\frac{\partial^2 V}{\partial S\partial r}+\frac{1}{2}w^2\frac{\partial^2 V}{\partial r^2}+rS\frac{\partial V}{\partial S}+\left(u-\lambda w\right)\frac{\partial V}{\partial r}-rV\\
&-kS\sqrt{\frac{2}{\pi dt}}\sqrt{\sigma^2S^2\left(\frac{\partial^2 V}{\partial S^2}\right)^2+2\rho\sigma wS\frac{\partial^2 V}{\partial S^2}\frac{\partial^2 V}{\partial S\partial r}+w^2\left(\frac{\partial^2 V}{\partial S\partial r}\right)^2}\\
&-k_1\frac{Z}{\eta}\sqrt{\frac{2}{\pi dt}}\sqrt{\sigma^2S^2\left(\frac{\partial^2 V}{\partial S\partial r}\right)^2+2\rho\sigma wS\left(\frac{\partial^2 V}{\partial r^2}-\frac{\partial V}{\partial r}\right)\frac{\partial^2 V}{\partial S\partial r}+w^2\left(\frac{\partial^2 V}{\partial r^2}-\frac{\partial V}{\partial r}\right)^2}=0
\end{aligned} \tag{2.1.13}$$

From equations (2.1.6), (2.1.7), (2.1.10) and (2.1.13), we can observe that the transaction costs function is always known and pre-defined. Moreover, the market models presented correspond to options over one single asset. In Ref [46] and Ref [45], the author generalize this framework and apply on different financial instruments. The first extension of Leland's method is to price and hedge a portfolio of strongly path-dependent European options on a stock. The payoff of this option will be given by $V\left(T,S_T,Y_T\right)$ where Y_T is a path-dependant quantity. If we represent Y_T in an integral form

$$Y\left(T\right)=\int_t^T f\left(s,S\left(s\right)\right)ds \tag{2.1.14}$$

it can be seen that the price of the aforementioned option is given by

$$\frac{\partial V}{\partial t} + rS\frac{\partial V}{\partial S} + \frac{1}{2}\sigma^2 S^2 \left[1 - A\,\mathrm{sgn}\frac{\partial^2 V}{\partial S^2}\right]\frac{\partial^2 V}{\partial S^2} + f\left(t, S\right)\frac{\partial V}{\partial Y} - rV = 0 \tag{2.1.15}$$

where

$$A = \frac{k}{\sigma}\sqrt{\frac{8}{\pi dt}} \tag{2.1.16}$$

and k is the parameter of the transaction costs function.

The second extension is developed to derive the price of basket options. These are options that are applied over multiple underlyings. We can suppose that each of the N underlyings are modelled following a lognormal distribution as of equation (1.1.3). Under this scenario, the N brownian motions are correlated between each of them where we denote ρ_{ij} as the correlation parameter between W_{S_i} and W_{S_j}. If on each time step the transaction costs that arises of trading asset S_i are equal to $k_i|v|S_i$, the PDE that governs the dynamics of the basket option is given by

$$\frac{\partial V}{\partial t} + r\sum_{i=1}^{N} S_i\frac{\partial V}{\partial S_i} + \frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N}\sigma_i\sigma_j S_i S_j\frac{\partial^2 V}{\partial S_i \partial S_j} - \sqrt{\frac{2}{\pi dt}}\sum_{i=1}^{N} k_i\Theta_i S_i - rV = 0 \tag{2.1.17}$$

where

$$\Theta_i = \sqrt{\left|\sum_{j=1}^{N}\sum_{k=1}^{N}\frac{\partial^2 V}{\partial S_i \partial S_j}\frac{\partial^2 V}{\partial S_i \partial S_k}\sigma_j\sigma_k\rho_{jk}S_j S_k\right|} \tag{2.1.18}$$

The second type of generalization has been done in Ref [39] to consider different forms of transaction costs functions. In their work, they approximate the one time-step change in the transaction costs function by its expected value. Hence, the equation (2.1.5) can be rewritten as

$$d\Pi = dV - \Delta dS - \mathbb{E}\left[\Delta TC\right] \tag{2.1.19}$$

Moreover, they assume that the cost C per one transaction is a nonincreasing function of the amount of transactions $|\Delta\delta|$ per unit of time Δt. Then, the one-step change in the transactions costs is given by

$$\Delta TC = S\,C\left(|\Delta\delta|\right)|\Delta\delta| \tag{2.1.20}$$

If we define the mean value modification of the transaction costs function as

$$\tilde{C}\left(\alpha\right) = \sqrt{\frac{\pi}{2}}\mathbb{E}\left[C\left(\alpha|\phi|\right)|\phi|\right] \tag{2.1.21}$$

where ϕ is random variable that follows a standard normal distribution, the original problem of Leland (2.1.6) can be rewritten as

$$\frac{\partial V}{\partial t} + \frac{1}{2}\hat{\sigma}^2 S^2 \frac{\partial^2 V}{\partial S^2} + rS\frac{\partial V}{\partial S} - rV = 0 \tag{2.1.22}$$

where

$$\hat{\sigma}^2 = \sigma^2\left(1 - \sqrt{\frac{2}{\pi}}\tilde{C}\left(\sigma S\left|\frac{\partial^2 V}{\partial S^2}\right|\sqrt{\Delta t}\right)\frac{\operatorname{sgn}\left(S\frac{\partial^2 V}{\partial S^2}\right)}{\sigma\sqrt{\Delta t}}\right). \tag{2.1.23}$$

Hence, the dynamics of multiple type of options can be modelled by proposing different transaction costs functions. Also, the existence of solution of these family of problems is considered in Ref [39]. The authors transform the fully nonlinear Black-Scholes equation into a quasilinear Gamma equation. If we denote the function $\beta\left(H\right) = 1/2\hat{\sigma}\left(H\right)^2 H$ and apply the change of variables $x = \ln\left(S/K\right)$, $\tau = T - t$ and $H\left(x,\tau\right) = S\partial^2 V/\partial S^2$, it is showed that if H is a solution of problem

$$\frac{\partial H}{\partial \tau} = \frac{\partial^2 \beta\left(H\right)}{\partial x^2} + \frac{\partial \beta\left(H\right)}{\partial x} + r\frac{\partial H}{\partial x} \tag{2.1.24}$$

then

$$V\left(S,t\right) = aS + be^{-r(T-t)} + \int_{-\infty}^{+\infty} \max\left(S - Ke^x\right) H\left(x, T-t\right) dx \tag{2.1.25}$$

is a solution of (2.1.22) with $a, b \in \mathbb{R}$.

We can observe that the unidimensional problem proposed in the Black-Scholes equations is then transformed by relaxing different assumption in two ways. If another component is deemed stochastic, the hedging portfolio will have to include another instrument to hedge the risk that arises from this new stochastic source. Then, the pricing equation will increase its dimensionality as many times as stochastic components are included in the model. The other way is by including the presence of transaction costs in the hedging strategy. Under this situation, different nonlinear terms are included into the equation affecting the linearity of the original Black-Scholes equation.

With respect to our work, we propose a model to price a multidimensional option when a general transaction costs function is considered. We start from the results developed by Ref [39], Ref [45] and Ref [46] and from there we develop the market model and derive the nonlinear PDE that governs the dynamics of the option price. We prove that there exists at least one weak solution by linearising the equation and applying an iterative approach. The second part of the work involves developing an ADI framework (presented in Section 1.3.2) to price a best-cash-or-noting option. Moreover, we analyze different aspects of the ADI model and how the transaction costs affect the final option price.

2.2 The market model

We start this Section by presenting the type of option that we want to price and the financial instruments needed for this purpose. Let's suppose we have N different assets which its price is denoted by S_i for $1 \leq i \leq N$. Each of this prices are modelled as lognormal processes. Then, for $1 \leq i \leq N$, we note these stochastic processes as of

$$dS_t^i = \mu_i \, S_t^i \, dt \, + \, \sigma_i \, S_t^i \, dW_t^i \tag{2.2.1}$$

where μ_i and σ_i are the correspondent mean and variance and $W = \left(W_t^1, ..., W_t^N\right)$ a multivariate normal vector with zero mean and correlations ρ_{ij}. To avoid the excess of sub-indices, we will extract the temporal index and just note $S_i := S_t^i$.

Given these n assets we define an option $V := V\left(t, S_1, ..., S_n\right)$ whose dynamic depends on the price of the n underlyings. An example of this type of option would be the basket option. In this option, the underlying is the weighted sum or average of the n instruments. Then, given a strike price K, the payoff is equal to $\left(n^{-1} \sum_i^n S_T^i - K\right)^+$. Just as the basket option, different type of multidimensional options can be defined. In Section 2.4 we will be working with a best-cash-or-nothing option which pays out a predefined cash amount K if assets S_1 or S_2 (bidimensional problem) are above or equal to a strike price X.

To get the price of option V we will follow the Black-Scholes steps presented in 1.1.2. Moreover, we will assume the existence of transaction costs when buying or selling any financial instruments present in the hedging portfolio.

Let's start by defining Π to be the portfolio that contains δ_i of asset S_i and an option V over those assets at time t. This portfolio can be represented by

$$\Pi = V + \sum_{i=1}^{N} \delta_i S_i. \tag{2.2.2}$$

If we calculate the variation of the portfolio Π at each time t, the transaction costs that arise from buying or selling any asset S_i appear in the formula. Specifically, we can observe that

$$\Delta\Pi = \Delta\left(V + \sum_{i=1}^{N} \delta_i S_i\right) + \sum_{i=1}^{N} \Delta TC_i, \tag{2.2.3}$$

where $\delta_i\, \Delta TC_i$ is the amount of transaction costs when buying or selling δ_i assets of S_i. Hence, we can observe that the change in the value of the portfolio Π will occur due to changes in the price of the option V, changes in the price of each asset S_i and the lost of value due to the presence of transaction costs. Following the Black-Scholes approach, we take $\delta_i = -V_{S_i}$, so that on each time step, the amount of the shares of assets hold are equal to the Delta of the asset. Then, we have that

$$\Delta\Pi = \Delta V - \sum_{i=1}^{N} \frac{\partial V}{\partial S_i} \Delta S_i - \sum_{i=1}^{N} \Delta TC_i. \tag{2.2.4}$$

We recall from equation (2.1.20) that the variation of the transaction costs can be modelled as of

$$\Delta TC_i = S_i\, C\left(|\Delta\delta_i|\right)|\Delta\delta_i| \tag{2.2.5}$$

where C is a nonincreasing function that models the transaction costs per unit of time. Following the rational proposed in Ref [39], we define r^i_{TC} to be the expected value of the change of the transaction costs per unit time interval Δt and price S_i. Hence, r^i_{TC} is equal to

$$r^i_{TC} = \frac{E\left[\Delta TC_i\right]}{S_i \Delta t} = \frac{E\left[C\left(|\Delta\delta_i|\right)|\Delta\delta_i|\right]}{\Delta t}. \tag{2.2.6}$$

Thus, we are approximating the transaction costs by the expected value of the transaction costs function applied to the amount of assets bought or sold and multiplied by these amount again. This value is then multiplied by the price of asset S_i in order to get a transaction cost in dollar terms. We can use this approximation of the transaction costs and apply it in equation (2.2.4).

Then, we have that

$$\Delta\Pi = \Delta V - \sum_{i=1}^{N} \frac{\partial V}{\partial S_i} \Delta S_i - \sum_{i=1}^{N} S_i\, r^i_{TC}\, \Delta t. \tag{2.2.7}$$

The one-step variation of the option price V can be modelled via the Itô's formula. If we recall equation (1.1.4), we get that

$$\Delta V = \frac{\partial V}{\partial t}\Delta t + \sum_{i=1}^{N}\frac{\partial V}{\partial S_i}\Delta S_i + \frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N}\sigma_i\sigma_j\rho_{ij}S_iS_j\frac{\partial^2 V}{\partial S_i \partial S_j}\Delta t. \tag{2.2.8}$$

Now, we can use equation the formula of ΔV given by (2.2.8) in equation (2.2.7) to obtain

$$\Delta \Pi = \left(\frac{\partial V}{\partial t} + \frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N}\sigma_i\sigma_j\rho_{ij}S_iS_j\frac{\partial^2 V}{\partial S_i \partial S_j}\right)\Delta t - \sum_{i=1}^{N} S_i\, r_{TC}^i\, \Delta t. \tag{2.2.9}$$

Moreover, we know from the Black-Scholes model that the growth rate of the portfolio Π is equal to r. Hence, we saw in Section 1.1.2 that

$$\Delta \Pi = r\Pi\Delta t. \tag{2.2.10}$$

If we replace the value of $\Delta\Pi$ in equation (2.2.9) we obtain the PDE that represents the dynamics of the option price

$$rV + \sum_{i=1}^{N} r_{TC}^i\, S_i = \frac{\partial V}{\partial t} + \frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N}\sigma_i\sigma_j\rho_{ij}S_iS_j\frac{\partial^2 V}{\partial S_i \partial S_j} + r\sum_{i=1}^{N}\frac{\partial V}{\partial S_i}S_i \tag{2.2.11}$$

where

$$r_{TC}^i S_i = \frac{E\left[\Delta TC_i\right]}{\Delta t} = \frac{E\left[C\left(|\Delta\delta_i|\right)|\Delta\delta_i|S_i\right]}{\Delta t}. \tag{2.2.12}$$

However, we need to find the value of $\Delta\delta_i$ and calculate the explicit form of the term $r_{TC}^i S_i$ to obtain the complete and final expression of the PDE.

If we recall some previous steps, we remember that we set $\delta_i = -\partial V/\partial S_i$. Then, we need to apply Itô's formula to δ_i which is a function that depends on $S_1, ..., S_N$ variables. The multidimensional version of Itô's formula (1.1.4) tell us that given $dX_t^i = \mu_i dt + \sigma_i dW_t^i$ and f a deterministic twice continuously differentiable function, then $Y_t = f\left(X_t^i\right)$ is also a stochastic process and is given by

$$dY_t = \sum_{i=1}^{N}\frac{\partial f}{\partial x_i}dX_t^i + \frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N}\frac{\partial^2 f}{\partial x_i \partial x_j}dX_t^i dX_t^j \tag{2.2.13}$$

Therefore, if we set $f = -\partial V/\partial S_i$ and consider only the terms of order $\mathcal{O}\left(\sqrt{\Delta t}\right)$, we obtain that

$$\Delta\delta_i = -\Delta\frac{\partial V}{\partial S_i} \sim \sum_{j=1}^{N}\frac{\partial^2 V}{\partial S_i \partial S_j}\Delta S_j \sim \sum_{j=1}^{N}\frac{\partial^2 V}{\partial S_i \partial S_j}\sigma_j S_j \phi_j \sqrt{\Delta t} \tag{2.2.14}$$

with ϕ_j being a standard normal variable. Then, following the calculation in (2.2.12), we find that

$$|\Delta\delta_i| = \left|\sum_{j=1}^{N}\frac{\partial^2 V}{\partial S_i \partial S_j}\Delta S_j\right| = \left|\sum_{j=1}^{N}\frac{\partial^2 V}{\partial S_i \partial S_j}\sqrt{\Delta t}\,\sigma_j\, S_j\, \phi_j\right| = \sqrt{\Delta t}\left|\sum_{j=1}^{N}\frac{\partial^2 V}{\partial S_i \partial S_j}\,\sigma_j\, S_j\, \phi_j\right|. \tag{2.2.15}$$

We can set $\Phi_i = \sum_{j=1}^{N} \frac{\partial^2 V}{\partial S_i \partial S_j}\,\sigma_j\, S_j\, \phi_j$ and obtain that $\Phi_i \sim N\left(0, \Theta_i\right)$ with

$$\Theta_i = \sum_{j=1}^{N}\sum_{k=1}^{N}\frac{\partial^2 V}{\partial S_i \partial S_j}\frac{\partial^2 V}{\partial S_i \partial S_k}\sigma_j\sigma_k\rho_{jk}S_jS_k \tag{2.2.16}$$

where ρ_{jk} is the correlation parameter between ϕ_j and ϕ_k. Hence, we find that the expected value of the change of the transaction costs per unit time interval Δt is approximate to

$$\begin{aligned} r_{TC}^{i}S_i &= \frac{E\left[\Delta TC\right]}{\Delta t} = \frac{E\left[C\left(|\Delta\delta_i|\right)|\Delta\delta_i|S_i\right]}{\Delta t} \\ &= \frac{\sqrt{\Delta t}\,E\left[C\left(\sqrt{\Delta t}\ |\Phi_i|\right)|\Phi_i|\ \ S_i\right]}{\Delta t} \\ &= \frac{S_i}{\sqrt{\Delta t}}E\left[C\left(\sqrt{\Delta t}\ |\Phi_i|\right)|\Phi_i|\right]. \end{aligned} \tag{2.2.17}$$

The last step is to use the result of equation (2.2.17) in equation (2.2.11) to obtain the PDE that models the dynamics of a multi-asset option with the presence of any type of transaction costs:

$$rV + \sum_{i=1}^{N}\frac{S_i}{\sqrt{\Delta t}}E\left[C\left(\sqrt{\Delta t}\ |\Phi_i|\right)|\Phi_i|\right] = \frac{\partial V}{\partial t} + \frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N}\sigma_i\sigma_j\rho_{ij}S_iS_j\frac{\partial^2 V}{\partial S_i \partial S_j} + r\sum_{i=1}^{N}\frac{\partial V}{\partial S_i}S_i \tag{2.2.18}$$

2.3 Existence of solution of the nonlinear PDE

2.3.1 Defining the nonlinear equation

In Section 2.2 we derived the nonlinear PDE which becomes our object of study. Our aim is to prove that, under certain 'feasible' conditions, there exists at least one weak viscosity solution of equation (2.2.18). We are going first to rewrite equation (2.2.18) to match with the notation used in Section 1.2.4 and Ref [28].

Let C be a measurable bounded transaction costs function such that $C : \mathbb{R}_0^+ \to \mathbb{R}_0^+$, $C \in L^2\left(\mathbb{R}_0^+\right)$ and let $\overline{C}, \underline{C} > 0$ be such that $\underline{C} < C(x) < \overline{C}$ for every $x \in \mathbb{R}_0^+$. Let us denote G to be the nonlinear operator

$$G\left(S, D^2V\right) = \sum_{i=1}^{N} \frac{S_i}{\sqrt{\Delta t}} E\left[C\left(\sqrt{\Delta t}\,|\Phi_i|\right)|\Phi_i|\right] \tag{2.3.1}$$

$$= \sum_{i=1}^{N} \frac{S_i}{\sqrt{\Delta t}} \sqrt{\frac{2}{\pi}}\, 2\sqrt{\Theta_i} \int_0^{+\infty} C\left(\sqrt{\Delta t\, 2\,\Theta_i} y\right) y\, e^{-y^2}\, dy \tag{2.3.2}$$

where Θ_i is given by

$$\Theta_i = \sum_{j=1}^{N}\sum_{k=1}^{N} \frac{\partial^2 V}{\partial S_i \partial S_j} \frac{\partial^2 V}{\partial S_i \partial S_k} \sigma_j \sigma_k \rho_{jk} S_j S_k. \tag{2.3.3}$$

where ρ_{jk} is the correlation parameter between ϕ_j and ϕ_k, both standard normal variables. Moreover, let us denote F to be the following nonlinear elliptic operator

$$F\left(\tau, S, V, DV, D^2V\right) = -\frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N} \sigma_i\sigma_j\rho_{ij} S_i S_j \frac{\partial^2 V}{\partial S_i \partial S_j} - r\sum_{i=1}^{N} \frac{\partial V}{\partial S_i} S_i + rV + G\left(S, D^2V\right). \tag{2.3.4}$$

Then, we define the nonlinear PDE for the problem of pricing a multi-asset option with general transaction costs as of

$$\begin{aligned} \frac{\partial V}{\partial \tau}(\tau, S) + F\left(\tau, S,\ V, DV, D^2V\right) &= 0 \quad \text{in} \quad \Omega \times [0, T] \\ V(0, S_1, ..., S_N) &= V_0(S_1, ..., S_N) \quad \text{in} \quad \Omega \end{aligned} \tag{2.3.5}$$

Hence, our objective is to find a viscosity solution of problem (2.3.5).

Remark 2.3.1. Equation (2.3.5) can be rewritten following a matricial form. If we denote the matrix A as of

$$(A)_{ij} = \sigma_i \sigma_j \rho_{ij} S_i S_j \tag{2.3.6}$$

then the function F can be set as

$$F\left(\tau, S, V, DV, D^2V\right) = -\frac{1}{2}\, tr\left(A\, D^2V\right) - rDV \cdot S + rV + G\left(S, D^2V\right) \tag{2.3.7}$$

For the nonlinear term that correspond to the function G we first note that the value of Θ_i is equivalent to the i-th term of the diagonal of the product $D^2V\, A\, D^2V$, i.e.

$$\Theta_i = \left(D^2V\, A\, D^2V\right)_{ii} \tag{2.3.8}$$

Then, the function G noted in a matricial form as of

$$G\left(S, D^2V\right) = \sum_{i=1}^{N} \frac{S_i}{\sqrt{\Delta t}} \sqrt{\frac{2}{\pi}}\, 2\, \sqrt{(D^2V\, A\, D^2V)_{ii}} \int_0^{+\infty} C\left(\sqrt{\Delta t\, 2\, \left(D^2V\, A\, D^2V\right)_{ii}}\, y\right) y\, e^{-y^2}\, dy \tag{2.3.9}$$

2.3.2 Degenerate Ellipticity and Leland's condition

Deriving the conditions

We are going to prove the existence of a viscosity solution of problem (2.3.5) using Perron's process. The main idea of the method is to construct a subsolution u^- and a supersolution u^+ of the nonlinear parabolic equation such that $u^- \leq u^+$. Moreover, a subsolution u lying between u^- and u^+ can be constructed and show that the lower semi-continuous envelope of the subsolution u is a supersolution. Before applying the Perron's process, we need to set different conditions on the nonlinear operator F. Let us start by recalling the definition of degenerate ellipticity stated in Section 1.2.26. For this purpose, we will denote by $\mathbb{S}_N$ the space of N-dimensional square symmetric matrices. As usual, for $X, Y \in \mathbb{S}_N$ we shall say that $X \leq Y$ if and only if $Y - X$ is (nonstrictly) positive definite.

Definition 2.3.2. A nonlinear function $F : [0, T] \times \Omega^+ \times \mathbb{R} \times \mathbb{R}^N \times \mathbb{S}_N \to \mathbb{R}$ is degenerate elliptic if

$$X \leq Y \implies F\left(t, x, p, s, X\right) \geq F\left(t, x, p, s, Y\right). \tag{2.3.10}$$

Given the definition of degenerate ellipticity we have to set the correspondent conditions such that the nonlinear operator F satisfies (2.3.10). Let us start by denoting the differential of F with respect to the second derivative component Y as

$$D_Y F\left(t, x, p, s, B\right) = \left.\frac{\partial F\left(t, x, p, s, Y\right)}{\partial Y}\right|_{Y=B} \tag{2.3.11}$$

that is, $D_Y F\,(t,x,p,s,B) : \mathbb{R}^{N\times N} \to \mathbb{R}$. As usual, this operator shall be identified with the matrix D given by $D_{ij} := \frac{\partial F}{\partial Y_{ij}}$, namely the gradient of F with respect to the variable Y, through the so-called Frobenius inner product:

$$D(U) = D \cdot U := \sum_{i,j=1}^{N} D_{ij}U_{ij}.$$

Following Definition 2.3.2, given a positive definite matrix U, we want to see that, for all B:

$$D_Y F\,(t,x,p,s,B)\,(U) \leq 0$$

Indeed, if the latter condition is fulfilled, then for arbitrary symmetric matrices $X \leq Y$ it is seen that, for some $B \in (X,Y)$,

$$\begin{aligned} F\,(t,x,p,s,Y) - F\,(t,x,p,s,X) &= D_Y F\,(t,x,p,s,B)\,(Y-X) \\ &\leq 0 \end{aligned} \tag{2.3.12}$$

and hence F is degenerate elliptic.

Let us recall the Leland condition which is present in the unidimensional problem with a constant transaction costs function. The aim of the this condition is in fact to define a degenerate elliptic operator such that the matrix of coefficients that correspond to the second derivatives is definite positive. In our work, the generalized Leland condition will act as the same and will be deduced from the following two Lemmas.

Lemma 2.3.3. *Let $D, U \in \mathbb{S}_N$. Then $D\,(U) = Tr\,(D\,U)$*

Proof. It follows by simple computation that

$$\begin{aligned} \mathrm{Tr}\,(D\,U) &= \sum_{i=1}^{N}\sum_{j=1}^{N} D_{ij}U_{ji} \\ &= \sum_{j=1}^{N}\sum_{i=1}^{N} D_{ij}U_{ij} = D \cdot U = D(U). \end{aligned}$$

□

Lemma 2.3.4. *Assume that $D_Y F$ is symmetric. Then $D_Y F$ is negative definite if and only if $Tr\,(D_Y F\,U) \leq 0$ for all $U \geq 0$.*

Proof. Let us firstly observe that, as $D_Y F$ is symmetric, there exists a diagonal matrix $\tilde{D}$ and a change of basis matrix C such that $D = C^{-1}\tilde{D}C$. Then, we have:

$$\mathrm{Tr}\left(D_Y F\, U\right) = \mathrm{Tr}\left(C^{-1}\,\tilde{D}\,C\,U\right) = \mathrm{Tr}\left(C^{-1}\,\tilde{D}\,C\,U\,C^{-1}\,C\right). \tag{2.3.13}$$

Denote $W = C\,U\,C^{-1}$, then the previous equality can be rewritten as

$$\mathrm{Tr}\left(D_Y F\, U\right) = \mathrm{Tr}\left(C^{-1}\,\tilde{D}\,W\,C\right) = \mathrm{Tr}\left(\tilde{D}\,W\right). \tag{2.3.14}$$

Moreover, U is positive definite if and only if W is positive definite. Assume that $Tr\left(D_Y F\, U\right) \leq 0$ for all $U \geq 0$, then $Tr(\tilde{D}W) \leq 0$ for all $W \geq 0$. Fix a sparse matrix W such that $W_{ij} = 1$ if $i = j = k$ and 0 otherwise, then $[\tilde{D}W]_{ii} = \tilde{D}_{kk}$ if $i = k$ and 0 otherwise. Thus we deduce that all the eigenvalues of $D_Y F$ are nonpositive.

Conversely, let us now assume that $D_Y F$ is negative definite and W is positive definite, then

$$\mathrm{Tr}\left(\tilde{D}\,W\right) = \sum_{i=1}^{N} \tilde{D}_{ii} W_{ii} \leq 0. \tag{2.3.15}$$

□

Both Lemmas 2.3.3 and 2.3.4 can be summarized in the following line: If the differential matrix $D_Y F$ is symmetric, then for any $U \geq 0$ the following equivalences hold:

$$D_Y F \leq 0 \iff \mathrm{Tr}\left(D_Y F\, U\right) \leq 0 \iff D_Y F\left(U\right) \leq 0$$

Hence, in view of (2.3.12) the nonlinear operator F is degenerate elliptic if the differential matrix $D_Y F$ is symmetric definite negative. In the following section we will see that the condition of being symmetric definite negative is the generalization of the Leland condition defined for the unidimensional problem with constant transaction costs.

Differential Matrix calculation

In this section we perform the computation of the differential matrix $D_Y F$. Let us recall Equation (2.3.7) such that

$$F\left(t, x, p, s, Y\right) = -\frac{1}{2}\, tr\left(A\,Y\right) - rs \cdot S + rp + G\left(S, Y\right) \tag{2.3.16}$$

Then, by applying standard calculations and discarding function dependencies, we have that

$$D_Y F\,(t,x,p,s,Y) = -\frac{\partial}{\partial Y}\, tr\left(\frac{1}{2} A\, Y\right) + \frac{\partial}{\partial Y} G\,(S,Y) \tag{2.3.17}$$

The first derivative follows trivially by linearity, that is

$$\begin{aligned}\frac{\partial}{\partial Y_{kl}}\, tr\left(\frac{1}{2} A\, Y\right) &= \frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N} A_{ij}\frac{\partial Y_{ji}}{\partial Y_{kl}},\\ &= \frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N} A_{ij}\delta_{jk}\delta_{il},\\ &= \frac{1}{2}A_{lk} = \frac{1}{2}A_{kl}.\end{aligned} \tag{2.3.18}$$

Then, it follows that

$$\frac{\partial}{\partial Y}\, tr\left(\frac{1}{2} A\, Y\right) = \frac{1}{2}A \tag{2.3.19}$$

The derivative of the second term involves the product rule, namely

$$\begin{aligned}\frac{\partial}{\partial Y} G\,(S,Y) &= \frac{\partial}{\partial Y}\left[\sum_{i=1}^{N}\frac{S_i}{\sqrt{\Delta t}}\sqrt{\frac{2}{\pi}}\, 2\sqrt{\sum_{j=1}^{N}\sum_{k=1}^{N} Y_{ij}\, A_{jk}\, Y_{ki}}\int_0^{+\infty} C\left(\sqrt{2\,\Delta t\sum_{j=1}^{N}\sum_{k=1}^{N} Y_{ij}\, A_{jk}\, Y_{ki}\, y}\right)\, y\, e^{-y^2}\, dy\right]\\ &= \sum_{i=1}^{N}\frac{S_i}{\sqrt{\Delta t}}\sqrt{\frac{2}{\pi}}\, 2\left[\frac{\partial}{\partial Y}\sqrt{\sum_{j=1}^{N}\sum_{k=1}^{N} Y_{ij}\, A_{jk}\, Y_{ki}}\int_0^{+\infty} C\left(\sqrt{2\,\Delta t\sum_{j=1}^{N}\sum_{k=1}^{N} Y_{ij}\, A_{jk}\, Y_{ki}\, y}\right)\, y\, e^{-y^2}\, dy\right.\\ &\quad\left.+\sqrt{\sum_{j=1}^{N}\sum_{k=1}^{N} Y_{ij}\, A_{jk}\, Y_{ki}}\int_0^{+\infty}\frac{\partial}{\partial Y} C\left(\sqrt{2\,\Delta t\sum_{j=1}^{N}\sum_{k=1}^{N} Y_{ij}\, A_{jk}\, Y_{ki}\, y}\right)\, y\, e^{-y^2}\, dy\right]\end{aligned} \tag{2.3.20}$$

The calculation above can be verified by analysing two derivatives. The first one correspond to the Θ_i function defined in (2.3.3).

$$\begin{aligned}
\frac{\partial}{\partial Y_{lm}}\sqrt{\sum_{j=1}^{N}\sum_{k=1}^{N}Y_{ij}\,A_{jk}\,Y_{ki}} &= \frac{1}{2}\left(\sum_{j=1}^{N}\sum_{k=1}^{N}Y_{ij}\,A_{jk}\,Y_{ki}\right)^{-1/2}\frac{\partial}{\partial Y}\sum_{j=1}^{N}\sum_{k=1}^{N}Y_{ij}\,A_{jk}\,Y_{ki}\\
&= \frac{1}{2}\left(\sum_{j=1}^{N}\sum_{k=1}^{N}Y_{ij}\,A_{jk}\,Y_{ki}\right)^{-1/2}\sum_{j=1}^{N}\sum_{k=1}^{N}\left(\frac{\partial Y_{ij}}{\partial Y_{lm}}A_{jk}Y_{ki}+Y_{ij}A_{jk}\frac{\partial Y_{ki}}{\partial Y_{lm}}\right)\\
&= \frac{1}{2}\left(\sum_{j=1}^{N}\sum_{k=1}^{N}Y_{ij}\,A_{jk}\,Y_{ki}\right)^{-1/2}\sum_{j=1}^{N}\sum_{k=1}^{N}\left(\delta_{il}\delta_{jm}A_{jk}Y_{ki}+Y_{ij}A_{jk}\delta_{kl}\delta_{im}\right)\\
&= \frac{1}{2}\left(\sum_{j=1}^{N}\sum_{k=1}^{N}Y_{ij}\,A_{jk}\,Y_{ki}\right)^{-1/2}\left[\sum_{k=1}^{N}\delta_{il}A_{mk}Y_{ki}+\sum_{j=1}^{N}\delta_{im}Y_{ij}A_{jl}\right]\\
&= \frac{1}{2}\left(\sum_{j=1}^{N}\sum_{k=1}^{N}Y_{ij}\,A_{jk}\,Y_{ki}\right)^{-1/2}\left[\delta_{il}\left(AY\right)_{mi}+\delta_{im}\left(YA\right)_{il}\right] \quad (2.3.21)
\end{aligned}$$

If we denote the matrix P_i as

$$\left(P_i\right)_{ml} = \left[\delta_{il}\left(AY\right)_{mi}+\delta_{im}\left(YA\right)_{il}\right] \tag{2.3.22}$$

the derivative with respect to matrix Y is equal to

$$\frac{\partial}{\partial Y}\sqrt{\Theta_i} = \frac{1}{2}\Theta_i^{-1/2}P_i. \tag{2.3.23}$$

The second derivative corresponds to the derivative of the transaction costs function C with respect to matrix Y. If we denote

$$H_i\left(y\right) = y\sqrt{2\,\Delta t\,\sum_{j=1}^{N}\sum_{k=1}^{N}B_{ij}\,A_{jk}\,B_{ki}} \tag{2.3.24}$$

then, it follows that

$$\begin{aligned}
\frac{\partial}{\partial Y_{lm}}C\left(H_i\left(y\right)\right) &= C'\left(H_i\left(y\right)\right)\frac{\partial}{\partial Y_{lm}}H_i\left(y\right)\\
&= C'\left(H_i\left(y\right)\right)\,y\,\frac{1}{2}\left(2\Delta t\Theta_i\right)^{-1/2}\,2\,\Delta t\frac{\partial}{\partial Y_{lm}}\sum_{j=1}^{N}\sum_{k=1}^{N}Y_{ij}A_{jk}Y_{ki}\\
&= C'\left(H_i\left(y\right)\right)\,y\,\frac{1}{2}\left(2\Delta t\Theta_i\right)^{-1/2}\,2\,\Delta t\left(P_i\right)_{lm} \quad (2.3.25)
\end{aligned}$$

Then, the derivative with respect to matrix Y is equal to

$$\frac{\partial}{\partial Y} C\left(\sqrt{2\,\Delta t\,(YAY)_{ii}}\,y\right) = C'\left(H_i\left(y\right)\right)\,y\,P_i\sqrt{\frac{\Delta t}{2}}\,\Theta_i^{-1/2} \tag{2.3.26}$$

Then, we can write Equation (2.3.17) as of

$$\begin{aligned} DY_F\left(t,x,p,s,Y\right) = -\frac{1}{2}A + P_i\frac{2}{\sqrt{\Delta t}}\sqrt{\frac{2}{\pi}}\sum_{i=1}^{N} S_i\left[\frac{1}{2}\Theta_i^{-1/2}\int_0^{+\infty} C\left(\sqrt{2\,\Delta t\,(YAY)_{ii}}\,y\right)\,y\,e^{-y^2}\,dy\right. \\ \left. + \sqrt{\frac{\Delta t}{2}}\int_0^{+\infty} C'\left(\sqrt{2\,\Delta t\,(YAY)_{ii}}\,y\right)\,y^2\,e^{-y^2}\,dy\right]. \end{aligned} \tag{2.3.27}$$

Thus, $D_Y F\left(t,x,p,s,Y\right)$ is strictly negative definite if

$$\begin{aligned} \frac{1}{2}A > P_i\frac{2}{\sqrt{\Delta t}}\sqrt{\frac{2}{\pi}}\sum_{i=1}^{N} x_i\left[\frac{1}{2}\Theta_i^{-1/2}\int_0^{+\infty} C\left(\sqrt{2\,\Delta t\,(YAY)_{ii}}\,y\right)\,y\,e^{-y^2}\,dy\right. \\ \left. + \sqrt{\frac{\Delta t}{2}}\int_0^{+\infty} C'\left(\sqrt{2\,\Delta t\,(YAY)_{ii}}\,y\right)\,y^2\,e^{-y^2}\,dy\right]. \end{aligned} \tag{2.3.28}$$

As a consequence, the generalized Leland condition follows from Lemmas 2.3.3 and 2.3.4 and can be summarized under the following definition:

Definition 2.3.5. The nonlinear operator F satisfies the *generalized Leland condition* if (2.3.28) is satisfied.

Remark 2.3.6. Let us show that effectively our condition reduces to Leland's condition in the unidimensional case with constant transaction costs. For this purpose, let us rewrite all the equation components under these assumptions. Then,

$$A = S^2\sigma^2, \quad \Theta = \frac{\partial^2 V}{\partial S^2}\sigma^2 S^2, \quad C\left(\sqrt{2\,\Delta t\,(D^2VAD^2V)}\,y\right) = \frac{\tilde{C}}{2} \tag{2.3.29}$$

If we apply this definitions on Equation (2.3.27) and we consider the convexity of solution V, we get that

$$\begin{aligned} D_Y F\left(t,x,p,s,\frac{\partial^2 V}{\partial S^2}\right) &= -\frac{1}{2}S^2\sigma^2 + 2\frac{\partial^2 V}{\partial S^2}S^2\sigma^2\frac{2\,S}{\sqrt{\Delta t}}\sqrt{\frac{2}{\pi}}\frac{1}{2}\frac{\tilde{C}}{4}\left(\frac{\partial^2 V}{\partial S^2}S^2\sigma^2\right)^{-1/2} \\ &= -\frac{1}{2}S^2\sigma^2 + S^2\sigma^2\,\mathrm{sgn}\left(\frac{\partial^2 V}{\partial S^2}\right)\frac{S}{\sqrt{\Delta t}}\sqrt{\frac{2}{\pi}}\frac{\tilde{C}}{2\sigma S} \\ &= \frac{1}{2}S^2\sigma^2\left[-1 + \frac{\tilde{C}}{\sqrt{\Delta t}}\sqrt{\frac{2}{\pi}}\frac{1}{\sigma}\,\mathrm{sgn}\left(\frac{\partial^2 V}{\partial S^2}\right)\right] \end{aligned} \tag{2.3.30}$$

Then, $D_Y F$ is negative if and only if

$$\frac{\tilde{C}}{\sqrt{\Delta t}}\sqrt{\frac{2}{\pi}}\frac{1}{\sigma} < 1 \tag{2.3.31}$$

2.3.3 Perron's method for existence of solution

Let us start this section by setting the framework in which the well-known Perron method shall be employed to derive the existence of a viscosity solution. Let us firstly apply a change of variables in order to obtain a nonlinear operator F with constant coefficients. More precisely, define

$$x_i = \log\left(S_i\right)$$

so we obtain:

$$\bar{F}\left(\tau, x, V, DV, D^2V\right) = -\frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N}\sigma_i\sigma_j\rho_{ij}\frac{\partial^2 V}{\partial x_i \partial x_j} - \sum_{i=1}^{N}\frac{\partial V}{\partial x_i}\left(r - \frac{\sigma_i^2}{2}\right) + rV + \bar{G}\left(x, D^2V\right), \tag{2.3.32}$$

and the nonlinear function G becomes

$$\bar{G}\left(x, D^2V\right) = \sum_{i=1}^{N}\frac{e^{x_i}}{\sqrt{\Delta t}}\sqrt{\frac{2}{\pi}}\, 2\,\sqrt{\Theta_i}\int_0^{+\infty} C\left(\sqrt{\Delta t\, 2\,\Theta_i} y\right) y\, e^{-y^2}\, dy, \tag{2.3.33}$$

with

$$\Theta_i = e^{-2x_i}\left[\sum_{j\neq i}^{N}\sum_{k\neq i}^{N}\frac{\partial^2 V}{\partial x_i \partial x_j}\frac{\partial^2 V}{\partial x_i \partial x_k}\sigma_j\sigma_k\rho_{jk} + 2\sum_{j\neq i}^{N}\frac{\partial^2 V}{\partial x_i \partial x_j}\left(\frac{\partial^2 V}{\partial x_i^2} - \frac{\partial V}{\partial x_i}\right)\sigma_i\sigma_j + \left(\frac{\partial^2 V}{\partial x_i^2} - \frac{\partial V}{\partial x_i}\right)^2\sigma_i^2\right]. \tag{2.3.34}$$

Given Equations (2.3.32) and (2.3.33), our Dirichlet problem becomes

$$\begin{aligned} &\frac{\partial V}{\partial \tau} + \bar{F}\left(\tau, x, V, DV, D^2V\right) = 0 \quad \text{in} \quad \Omega\times[0,T] \\ &V\left(0, x_1, ..., x_N\right) = V_0\left(x_1, ..., x_N\right) \quad \text{in} \quad \Omega \end{aligned} \tag{2.3.35}$$

where $V_0\left(x_1, ..., x_N\right)$ is the initial condition.

Lemma 2.3.7. *Let F be the nonlinear operator defined in Equation* (2.3.4) *and $\bar{F}$ be the transformed operator defined in Equation* (2.3.32)*. If F is degenerate elliptic, then $\bar{F}$ is degenerate elliptic.*

Proof. Let $\bar{A}$ and $\bar{B}$ be symmetric matrices such that $\bar{B} - \bar{A} > 0$. Observe that the change of variables leads to

$$\frac{\partial V}{\partial x_j} = \frac{\partial V}{\partial S_j} S_j$$
$$\frac{\partial^2 V}{\partial x_i x_j} = \frac{\partial^2 V}{\partial S_i S_j} S_i S_j + \delta_{ij} \frac{\partial V}{\partial S_j},$$

so we may denote

$$\bar{F}\left(t, x, \bar{p}, \bar{q}, \bar{A}\right) = F\left(t, S, p, q, A\right)$$

where

$$S_j = e^{x_j}, \quad p = \bar{p}, \quad q_j = \frac{\bar{q}_j}{S_j}$$
$$A_{ij} = \frac{\bar{A}_{ij} - \delta_{ij} q_j}{S_i S_j}.$$

Hence, because

$$\bar{F}\left(t, x, \bar{p}, \bar{q}, \bar{B}\right) - \bar{F}\left(t, x, \bar{p}, \bar{q}, \bar{A}\right) = F\left(t, S, p, q, B\right) - F\left(t, S, p, q, A\right),$$

it suffices to prove that $B > A$. To this end, observe that

$$(B - A)_{ij} = \frac{\left(\bar{B} - \bar{A}\right)_{ij}}{S_i S_j}$$

and, for an arbitrary square matrix M denote by $m_k(M)$ the leading principal minor of order k. A simple computation shows that

$$m_k\left(B - A\right) = \frac{m_k(\bar{B} - \bar{A})}{\prod_{i \le k} S_i^2}$$

and the result follows from Sylvester's criterion.

□

Hence, the main theorem of this work can be stated as follows:

Theorem 2.3.8. *Assume that the original nonlinear operator F satisfies Definition 2.3.5. Then, problem* (2.3.35) *has at least one viscosity solution.*

Before passing to the proof of the theorem, we are going to recall some definitions stated in Section 1.2.4 that will be used afterwards. Given an open set $\Omega_T \subset \mathbb{R}^{N+1}$, we recall that V is *lower semi-continuous* (LSC) or *upper semi-continuous* (USC) at (t,x) if for all sequences $(s_n, y_n) \to (t,x)$,

$$
\begin{aligned}
V(t,x) &\leq \liminf_{n\to\infty} V(s_n, y_n) \quad \text{(LSC)} \\
V(t,x) &\geq \limsup_{n\to\infty} V(s_n, y_n) \quad \text{(USC)}.
\end{aligned}
$$

Moreover, we define V_* the *lower semi-continuous envelope of V* as the largest lower semi-continuous function lying below V and V^* the correspondent *upper semi-continuous envelope of V* as the smallest upper semi-continuous function lying above V.

Now we can present Perron's method to find a solution of problem (2.3.35). We are going to require first that the nonlinear operator F is degenerate elliptic. Then, Perron's method is defined as follows.

Moreover, we recall Perron's method from Section 1.2.26.

Theorem 2.3.9. *Suppose w is a subsolution of problem* (2.3.35) *and v is a supersolution of problem* (2.3.35) *such that $w \leq v$. Suppose also that there is a subsolution $\underline{u}$ and a supersolution $\overline{u}$ of problem* (2.3.35) *that satisfy the boundary condition $\underline{u}_*(t,x) = \overline{u}^*(t,x) = g(t,x)$. Then,*

$$
W(t,x) = \sup\{w(t,x) : \underline{u} \leq w \leq \overline{u} \; and \, w \; is \; a \; subsolution \; of \; (2.3.35)\}. \tag{2.3.36}
$$

is a solution of problem (2.3.35).

In order to apply the Perron's method we first have to set a subsolution and supersolution of problem (2.3.35). Then, we will have to construct a maximal subsolution such that it lies between both sub and supersolutions. Finally, we will have to define the proper comparison principle such that the boundary condition defined in Theorem 2.3.9 holds.

Hence, let us start by recalling the equivalent "Black-Scholes" linear problem. If we denote the linear elliptic operator as of

$$
\tilde{F}\left(\tau, x, V, DV, D^2V\right) = -\frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N}\sigma_i\sigma_j\rho_{ij}\frac{\partial^2 V}{\partial x_i \partial x_j} - \sum_{i=1}^{N}\frac{\partial V}{\partial x_i}\left(r - \frac{\sigma_i^2}{2}\right) + rV, \tag{2.3.37}
$$

then there exists a unique solution Λ of the problem

$$\begin{aligned} \frac{\partial V}{\partial \tau} + \tilde{F}\left(\tau, x, V, DV, D^2 V\right) = 0 \quad & \text{in} \quad \Omega \times [0,T] \\ V\left(0, x_1, ..., x_N\right) = V_0\left(x_1, ..., x_N\right) \quad & \text{in} \quad \Omega \end{aligned} \tag{2.3.38}$$

Based on the existence of this unique solution Λ, we will construct our sub and supersolutions. Moreover, the existence of Λ helps us to set an upper bound for the nonlinear term $\bar{G}$ as its second derivatives are bounded. Following the replicant portfolio strategy it is observed that transaction costs are proportional to the size of the second derivatives of the option price. From the already known Λ, we know that its second derivatives reach a maximum near the strike value and tend to zero when prices are either too high or too low. Then, based on the dynamics of the replicant portfolio, little amount of stocks are traded on each time step under the scenarios of low or high prices. As a result, transaction costs tend to zero for these stock prices and have an upper bound when the stock price tends to the strike value. Based on the existence of this upper bound for the nonlinear term, the following Lemma presents both sub and supersolutions of problem (2.3.35).

Lemma 2.3.10. *Let F be the nonlinear elliptic operator defined in Equation* (2.3.32)*. Then the following functions are sub and supersolutions of problem* (2.3.35)*.*

$$\begin{aligned} \overline{V} &= \Lambda + C\tau \\ \underline{V} &= \Lambda - C\tau \end{aligned}$$

where Λ is the unique solution of problem (2.3.38) *and C is a positive constant such*

$$C \geq \sup_{x \in \Omega} \left|\bar{G}\left(x, D^2\Lambda\right)\right|$$

Proof. Let us see that the $\underline{V}$ is a subsolution of (2.3.35). Firstly, the upper semi-continuity of $\underline{V}$ follow from the continuity of the solution $\Lambda(\tau, x)$. Let us see that for all test functions ϕ such that $\underline{V} \leq \phi$ in a neighbourhood of (τ, x) and $\underline{V}(\tau, x) = \phi(\tau, x)$, it follows that $\frac{\partial \phi}{\partial \tau} + \bar{F}\left(\tau, x, \phi, D\phi, D^2\phi\right)$ is negative.

Let ϕ be a test function such that $\underline{V} \leq \phi$. Then, we have that

$$\begin{aligned} \frac{\partial \phi}{\partial \tau}(\tau, x) &= \frac{\partial \underline{V}}{\partial \tau}(\tau, x) \\ D\phi(\tau, x) &= D\underline{V}(\tau, x) \\ D^2\phi(\tau, x) &\geq D^2\underline{V}(\tau, x) \end{aligned}$$

Now we use the condition of degenerate ellipticity required on F. This condition implies that

$$\begin{aligned}\frac{\partial \phi}{\partial \tau}+\bar{F}\left(\tau, x, \phi, D\phi, D^2\phi\right) &\leq \frac{\partial \underline{V}}{\partial \tau}+\bar{F}\left(\tau, x, \underline{V}, D\underline{V}, D^2\underline{V}\right)\\ &\leq \bar{G}\left(\tau, D^2\Lambda\right)-C\\ &\leq 0\end{aligned}$$

where the last inequality holds using the lower bound of the constant C.

Let us now confirm that $\overline{V}$ is in fact a supersolution. In this case, the lower semi-continuity follows from the continuity of the solution Λ. Let us see that for all test functions ϕ such that $\overline{V} \geq \phi$ in a neighbourhood of (τ, x) and $\overline{V}(\tau, x) = \phi(\tau, x)$, it follows that $\frac{\partial \phi}{\partial \tau}+\bar{F}\left(\tau, x, \phi, D\phi, D^2\phi\right)$ is positive.

Let ϕ be a test function such that $\underline{\Lambda} \geq \phi$. Then, we have that

$$\begin{aligned}\frac{\partial \phi}{\partial \tau}(\tau, x) &= \frac{\partial \overline{V}}{\partial \tau}(\tau, x)\\ D\phi(\tau, x) &= D\overline{V}(\tau, x)\\ D^2\phi(\tau, x) &\leq D^2\overline{V}(\tau, x)\end{aligned}$$

Now we use the condition of degenerate ellipticity required on $\bar{F}$ and the lower bound of the positive constant C. Both conditions imply that

$$\begin{aligned}\frac{\partial \phi}{\partial \tau}+\bar{F}\left(\tau, x, \phi, D\phi, D^2\phi\right) &\geq \frac{\partial \overline{V}}{\partial \tau}+\bar{F}\left(\tau, x, \overline{V}, D\overline{V}, D^2\overline{V}\right)\\ &\geq \bar{G}\left(x, D^2\Lambda\right)+C\\ &\geq 0.\end{aligned}$$

Then, $\overline{V}$ is a supersolution of problem (2.3.35). □

Remark 2.3.11. By definition, it remains valid that $\underline{V} \leq \overline{V}$.

Following Lemma 2.3.15 from Ref [28], there exists a function U such that $\underline{V} \leq U \leq \overline{V}$ and U^* is a subsolution of (2.3.35) and U_* is a supersolution of (2.3.35). Then, to finally prove Theorem 2.3.8, we need to confirm that $U^*(\tau, S) = U_*(\tau, S)$. For this purpose, we will need to use the comparison principle stated in Ref [28].

Proposition 2.3.12 (Comparison Principle)**.** *If u is a subsolution of problem* (2.3.35) *and v is a supersolution of problem* (2.3.35) *in Ω_T and $u \leq v$ on the parabolic boundary $\partial_p\Omega_T$, then $u \leq v$ in Ω_T.*

Hence, our last Lemma is stated below:

Lemma 2.3.13. *Let $\underline{V}$ and $\overline{V}$ be the sub and supersolutions of problem* (2.3.35) *and U the function obtained by Lemma 2.3.15 from Ref [28] such that $\underline{V} \leq U \leq \overline{V}$. Then, $U^*(\tau, S) = U_*(\tau, S)$.*

Proof. Let us first observe that the inequality $U_* \leq U^*$ holds by definition of the semi-continuous envelopes. For the other inequality let us recall $\underline{V}$ and $\overline{V}$ and, using the continuity of the linear solution Λ and the distance function $f(S)$, we have that

$$\begin{gathered}(\underline{V})_* = \underline{V} = (\underline{V})^* \\ \left(\overline{V}\right)_* = \overline{V} = \left(\overline{V}\right)^*\end{gathered}$$

In particular, in the parabolic boundary, we find that both sub and supersolutions are equal to Λ. Then, it is valid that

$$\left(\overline{V}\right)^* \leq (\underline{V})_* \text{ in } \partial_p\Omega_T \tag{2.3.39}$$

Moreover, as for Lemma 2.3.15 from Ref [28], $\underline{V} \leq U \leq \overline{V}$. Using this result and the previous inequality, it follows that

$$U^* \leq U_* \text{ in } \partial_p\Omega_T \tag{2.3.40}$$

Finally, the expected inequality is obtained following the comparison's principle result. □

2.4 Numerical

2.4.1 Numerical framework

This Section will include the development of a numerical framework to solve the nonlinear problem presented in Equation (2.3.35). To this purpose, we will use the concepts showed in Section 1.3 with the regards to the recommended schemes usually used for multidimensional PDE's with the presence of crossed derivatives. As we discussed before, we consider an ADI scheme (see Section 1.3.2) to find an approximate solution of problem (2.3.35). This ADI scheme is applied on an iterative framework which results of linearizing our original nonlinear problem resulting on the following multi-step equation.

$$\begin{aligned} -U^n_\tau + \mathcal{L}U^n &= G\left(U^{n-1}\right) && \text{in} \quad \Omega \times [0,T] \\ U^n\left(0, x_1, ..., x_N\right) &= U_0\left(0, x_1, ..., x_N, 0\right) && \text{in} \quad \Omega \\ U^0\left(\tau, x_1, ..., x_N\right) &= 0 && \text{in} \quad \Omega \times (0,T) \end{aligned} \tag{2.4.1}$$

with $\dim \Omega = 2$. For numerical convenience, we approximate the original smooth domain by a discrete one $\hat{\Omega}_T \subset [a,b] \times [a,b] \times [0,T]$, setting a and b in order to cover a set of feasible logarithmic stock prices. The step of the spatial variables is uniformly set as $\Delta x = (b-a)/S_x$, being S_x the number of grid points in the x- direction. The step of the temporal variable is also uniformly set as $\Delta\tau = T/T_x$ being T_x the number of grid points in the τ- direction. We define n to be the step of the iterative problem and, on each n, m to be each of the temporal steps. Hence, we define the solution to the n-step iterative problem as $U_{ij}^m = U\left(x_i, y_j, m\Delta\tau\right)$ where $0 \leq i, j \leq S_x$ and $0 \leq m \leq T_x$.

Given that we have to work on an iterative problem, in each step n, we will have to solve a linear problem involving both second and mixed derivatives of U. Considering these settings, we follow the work of Ref [31] to develop the two-step procedure of our ADI scheme. In Section 1.3.2 we discussed about the intermediate step $m + 1/2$ that is created to eliminate the crossed derivative term. In fact, we will take the first half step implicitly in the x-direction and explicitly in the y-direction. The other half step will be taken implicitly in the y-direction and explicitly in the x-direction. With this schema, we will be always solving tridiagonal problems on every step.

Hence, let us start by showing how the PDE equation (2.4.1) is treated. In the first place, we split the temporal derivative as shown on (2.4.2)

$$U_\tau \simeq \frac{U_{ij}^{m+1} - U_{ij}^m}{\Delta t} = \frac{U_{ij}^{m+1} - U_{ij}^{m+\frac{1}{2}}}{\Delta t} + \frac{U_{ij}^{m+\frac{1}{2}} - U_{ij}^m}{\Delta t}. \tag{2.4.2}$$

In a second step, we discretize the lineal operator

$$\mathcal{L}U = \frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N}\sigma_i\sigma_j\rho_{ij}\frac{\partial^2 U}{\partial x_i \partial x_j} + \sum_{i=1}^{N}\frac{\partial U}{\partial x_i}\left(r - \frac{\sigma_i^2}{2}\right) - rU, \tag{2.4.3}$$

by applying forward differences in the first order derivatives and central differences in the second order derivatives

$$\begin{aligned}
\frac{\partial U}{\partial x_1} &\simeq \frac{U_{i+1,j}^m - U_{i,j}^m}{\Delta x}, \\
\frac{\partial U}{\partial x_2} &\simeq \frac{U_{i,j+1}^m - U_{i,j}^m}{\Delta x}, \\
\frac{\partial^2 U}{\partial x_1^2} &\simeq \frac{U_{i+1,j}^m - 2U_{i,j}^m + U_{i-1,j}^m}{\Delta x^2}, \\
\frac{\partial^2 U}{\partial x_2^2} &\simeq \frac{U_{i,j+1}^m - 2U_{i,j}^m + U_{i,j-1}^m}{\Delta x^2}, \\
\frac{\partial^2 U}{\partial x_1 x_2} &\simeq \frac{U_{i+1,j+1}^m + U_{i-1,j-1}^m - U_{i-1,j}^m - U_{i,j-1}^m}{4\Delta x^2}.
\end{aligned}$$

Following Section 2.1 of Ref [31] we split the discretization of the operator $\mathcal{L}$ between

$$
\begin{aligned}
\mathcal{L}^x = & \frac{\sigma_1^2}{4}\frac{U_{i+1,j}^{m+\frac{1}{2}} - 2U_{i,j}^{m+\frac{1}{2}} + U_{i-1,j}^{m+\frac{1}{2}}}{\Delta x^2} + \frac{\sigma_2^2}{4}\frac{U_{i,j+1}^{m} - 2U_{i,j}^{m} + U_{i,j-1}^{m}}{\Delta x^2} \\
& + \frac{1}{2}\sigma_1\sigma_2\rho\frac{U_{i+1,j+1}^{m} + U_{i-1,j-1}^{m} - U_{i-1,j}^{m} - U_{i,j-1}^{m}}{4\Delta x^2} \\
& + \frac{1}{2}\left(r - \frac{\sigma_1^2}{2}\right)\frac{U_{i+1,j}^{m+\frac{1}{2}} - U_{i,j}^{m+\frac{1}{2}}}{\Delta x} + \frac{1}{2}\left(r - \frac{\sigma_2^2}{2}\right)\frac{U_{i,j+1}^{m} - U_{i,j}^{m}}{\Delta x} - \frac{1}{2}rU_{ij}^{m+\frac{1}{2}}
\end{aligned}
$$

and

$$
\begin{aligned}
\mathcal{L}^y = & \frac{\sigma_1^2}{4}\frac{U_{i+1,j}^{m+\frac{1}{2}} - 2U_{i,j}^{m+\frac{1}{2}} + U_{i-1,j}^{m+\frac{1}{2}}}{\Delta x^2} + \frac{\sigma_2^2}{4}\frac{U_{i,j+1}^{m+1} - 2U_{i,j}^{m+1} + U_{i,j-1}^{m+1}}{\Delta x^2} \\
& + \frac{1}{2}\sigma_1\sigma_2\rho\frac{U_{i+1,j+1}^{m+\frac{1}{2}} + U_{i-1,j-1}^{m+\frac{1}{2}} - U_{i-1,j}^{m+\frac{1}{2}} - U_{i,j-1}^{m+\frac{1}{2}}}{4\Delta x^2} \\
& + \frac{1}{2}\left(r - \frac{\sigma_1^2}{2}\right)\frac{U_{i+1,j}^{m+\frac{1}{2}} - U_{i,j}^{m+\frac{1}{2}}}{\Delta x} + \frac{1}{2}\left(r - \frac{\sigma_2^2}{2}\right)\frac{U_{i,j+1}^{m+1} - U_{i,j}^{m+1}}{\Delta x} - \frac{1}{2}rU_{ij}^{m+1}
\end{aligned}
$$

obtaining a two-stage full scheme

$$
\begin{aligned}
\frac{U_{ij}^{m+\frac{1}{2}} - U_{ij}^{m}}{\Delta t} &= \mathcal{L}^x U_{ij}^{m+\frac{1}{2}}, \\
\frac{U_{ij}^{m+1} - U_{ij}^{m+\frac{1}{2}}}{\Delta t} &= \mathcal{L}^y U_{ij}^{m+1}.
\end{aligned}
$$

Up to this moment we have not considered the nonlinear operator G. However, we must remember that after the linearization, this term only depends of the value of x and t and no longer involves the different derivatives of U. Then, we decide to add this term on the second stage of the procedure by redefining $\tilde{\mathcal{L}}^y = \mathcal{L}^y - G$ such that the whole framework becomes

$$
\frac{U_{ij}^{m+1} - U_{ij}^{m}}{\Delta t} = \mathcal{L}^x U_{ij}^{m+\frac{1}{2}} + \mathcal{L}^y U_{ij}^{m+1} - G(\cdot) = \mathcal{L}^x U_{ij}^{m+\frac{1}{2}} + \tilde{\mathcal{L}}^y U_{ij}^{m+1}. \tag{2.4.4}
$$

2.4.2 Numerical results

In order to implement the framework proposed in Section 2.4.1, we select an specific type of multi-asset option and a transaction costs function. First, we choose to price a best cash-or-nothing option call on two assets. This option pays out a predefined cash amount K if assets

S_1 or S_2 are above or equal to the strike price X. The closed-form formula is presented on Ref [25] as

$$\begin{aligned}
c_{best} &= Ke^{-rT}\left[M\ (y, z_1; -\rho_1) + M\,(-y, z_2; -\rho_2)\right] && (2.4.5)\\
y &= \frac{\ln\left(S_1/S_2\right) + \frac{\sigma^2}{2}T}{\sigma\sqrt{T}}, \qquad \sigma = \sqrt{\sigma_1^2 + \sigma_2^2 - 2\sigma_1\sigma_2\rho}\\
z_1 &= \frac{\ln\left(S_1/X\right) + \frac{\sigma_1^2}{2}T}{\sigma_1\sqrt{T}}, \qquad z_2 = \frac{\ln\left(S_2/X\right) + \frac{\sigma_2^2}{2}T}{\sigma_2\sqrt{T}}\\
\rho_1 &= \frac{\sigma_1 - \rho}{\sigma}, \qquad \rho_2 = \frac{\sigma_2 - \rho}{\sigma}
\end{aligned}$$

where S_1 and S_2 are the stock prices, σ_1 and σ_2 are the volatilities, ρ is the correlation between both assets, T is the maturity and $M\,(a, b; \rho)$ is

$$M\,(a, b; \rho) = \frac{1}{2\pi\sqrt{1-\rho^2}} \int_{-\infty}^{a}\int_{-\infty}^{b} \exp\left[-\frac{x^2 + y^2 - 2\rho xy}{2\left(1-\rho^2\right)}\right] dx\, dy. \tag{2.4.6}$$

Second, given the nonlinear function G defined on (2.3.33), we choose an exponential decreasing transaction costs function defined as

$$C\,(x) = C_0\, e^{-\tilde{k}\,x} \tag{2.4.7}$$

for each asset x. Hence, by recalling (2.2.17), we can see that

$$\begin{aligned}
E\left[C\left(\sqrt{\Delta t}\ |\Phi_i|\right)|\Phi_i|\right] &= \int_0^{+\infty} C_0\, e^{-\tilde{k}\sqrt{\Delta t}x}\, \frac{2\,x}{\sqrt{2\pi\Theta_i}} e^{-x^2/2\Theta_i}\, dx\\
&= C_0\sqrt{\frac{2}{\pi}} \int_0^{+\infty} e^{-\tilde{k}\sqrt{\Delta t}x}\, \frac{x}{\sqrt{\Theta_i}}\, e^{-x^2/2\Theta_i}\, dx\\
&= C_0\sqrt{\frac{2}{\pi}} \int_0^{+\infty} e^{-\tilde{k}\sqrt{\Delta t\Theta_i}y}\, \sqrt{\Theta_i}y\, e^{-y^2/2}\, dy\\
&= C_0\sqrt{\Theta_i}\sqrt{\frac{2}{\pi}}\left[1 - e^{\tilde{k}^2\Delta t\Theta_i/2}\,\tilde{k}\sqrt{\Delta t\Theta_i}\,\mathrm{ERFC}\left(\tilde{k}\sqrt{\frac{\Delta t\Theta_i}{2}}\right)\right].
\end{aligned}$$

Then,

$$G\,(x) = C_0\sqrt{\frac{2}{\pi}}\sum_{i=1}^{2}\frac{e^{x_i}}{\Delta t}\sqrt{\Theta_i}\left[1 - e^{\tilde{k}^2\Delta t\Theta_i/2}\,\tilde{k}\sqrt{\Delta t\Theta_i}\,\mathrm{ERFC}\left(\tilde{k}\sqrt{\frac{\Delta t\Theta_i}{2}}\right)\right]. \tag{2.4.8}$$

In this Section we are going to analyze different aspects of the ADI algorithm implemented and the dynamics of the general transaction cost model proposed. We focus on the following three points:

1. Measure the impact of transaction costs in the option price.
2. Given an optimal number of iterations such that convergence is achieved, analyze the sensitivity of the final output to the choice of Δt_{TC}.
3. Given the iteration procedure proposed in (2.4.1), determine the optimal number of n such that the convergence is achieved and how the error diminishes as more steps are added.

In Table 2.1 we present the parameters chosen for the numerical implementation. Three different tests are then applied by varying the values of the stocks price, volatility, interest rate and strike among others.

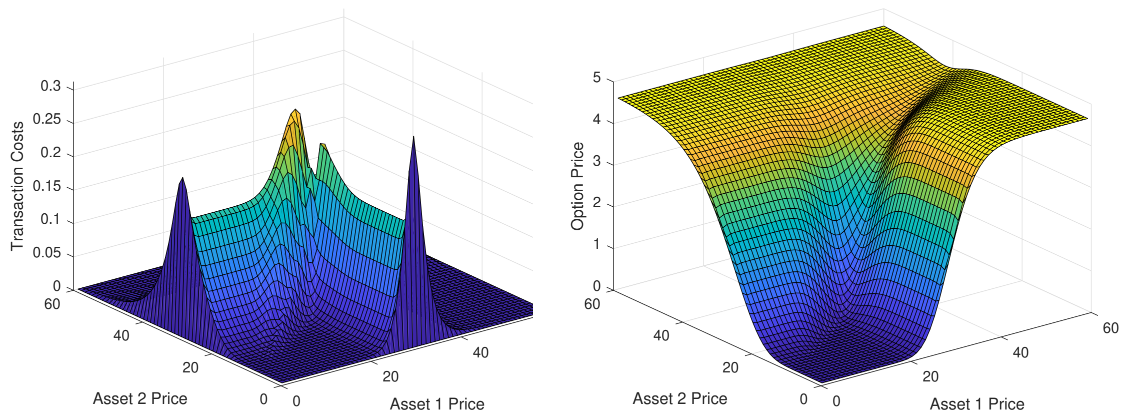
Parameters	Testing 1		Testing 2		Testing 3	
	Asset 1	Asset 2	Asset 1	Asset 2	Asset 1	Asset 2
σ	0.30	0.15	0.05	0.1	0.2	0.2
ρ	0.5		-0.3		0.2	
r	0.08		0.02		0.1	
T	1 year		1 year		1 year	
K	5		8		6	
X	30		40		15	
Δx	1		1		1	
Δt_{TC}	1/261		1/261		1/261	
C_0	0.005		0.001		0.003	
$\tilde{k}$	1		0.5		0.7	

Table 2.1: Numerical implementation parameters

Transaction Costs impact

Figure (2.1) to Figure (2.3) present the results for both transaction costs function and option price with transaction costs at $t = 0$. By recalling the transaction costs function G in (2.4.2) it can be noted that the costs are proportional to the assets spot price, the size of the second derivatives (i.e. Gamma) of the option price and the volatilities of each asset.

Testing 1 is defined based on a strike at $X = 30$ with a premium paid at $K = 5$. Figure (2.1a) shows an exponential transaction costs function where the maximum is reached around the strike value. This behavior is expected as the maximum of Gamma is found near the at-the-money price. As these derivatives converge to zero when deep out-of-the-money or in-the-money, the transactions costs function vanishes. Figure (2.1b) describes the dynamics of the option price when considering the transaction costs function.



(a) Transaction costs at time $t = 0$. (b) Option price at time $t = 0$.

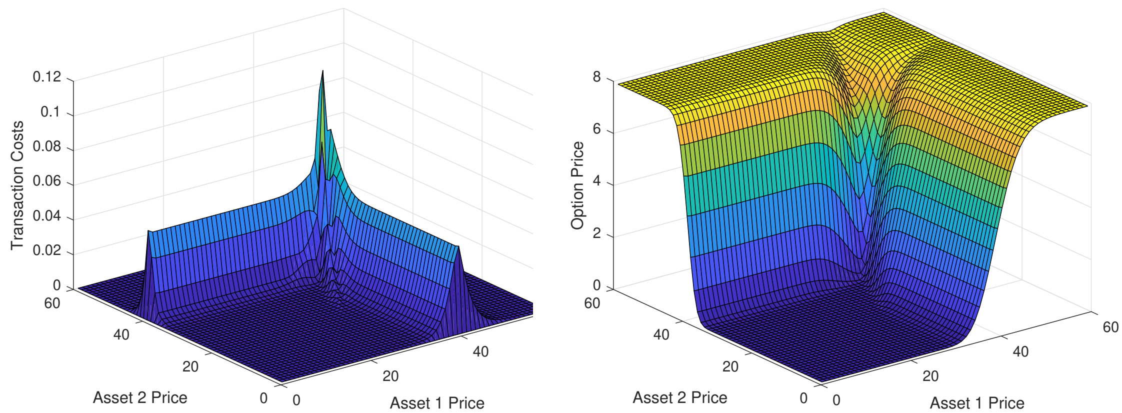
Figure 2.1: Testing 1.

The results for Testing 2 framework are presented in Figure (2.2). It is defined a strike value at $X = 40$, a premium paid at $K = 8$ and with two low volatile assets. The transaction costs function presented in Figure (2.2a) shows a similar increasing pattern on its value up to the at-the-money region. Moreover, the higher volatilty of Asset 2 is observed by noting that transaction costs are higher when fixing a price for Asset 2 in comparison with Asset 1. As the option gets out-of-the-money, the shape of the transaction costs function becomes more symmetric and smoother.

Testing framework 3 is presented on Figure (2.3). Both assets are defined to have the same volatility but almost uncorrelated. The strike price is fixed at $X = 15$ and the premium paid is equal to $K = 6$. The symmetry observed in both Figures (2.3a) and (2.3b) are expected due to the design of the testing. Again, the maximum of the transaction costs function is reached when the prices are near the strike value and the converge to zero is seen when the option is deeper out-of-the-money or in-the-money. The option prices reflect the complementary pattern by showing a decrease in its value when the option is near the strike price.

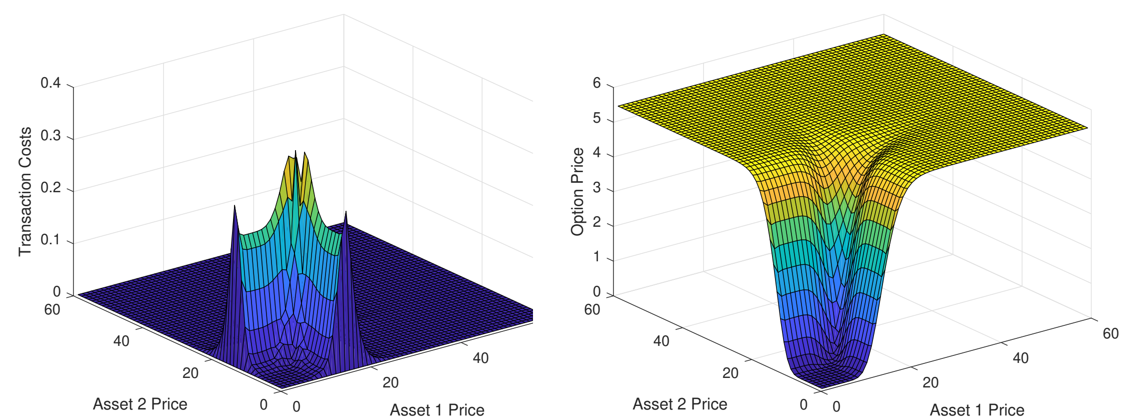
Sensitivity of the option to changes in Δt_{TC}

In this Section we study the sensitivity of the option price to changes in the size of the time-step Δt_{TC} for rebalancing the replicant portfolio. By observing Equation (2.4.2), it can be



(a) Transaction costs at time $t = 0$. (b) Option price at time $t = 0$.

Figure 2.2: Testing 2.



(a) Transaction costs at time $t = 0$. (b) Option price at time $t = 0$.

Figure 2.3: Testing 3.

seen that the transaction costs function tends to infinity if Δ_{TC} tends to zero. Hence, we expect to see this results in the numerical testing. For this purpose, we ran Testing 2 framework under 100 possible values of Δ_{TC} ranging from $7.6E-05$ (approximately rebalancing every 29 minutes) to 0.007 (approximately rebalancing every 2 years). The results can be observed in Figures 2.4 and 2.5.

Figure 2.4 presents two different plots which show two states of the option price. In the figure of the left side it can be observed how the transaction costs behave when the price of Asset 1 is equal to $S_1 = 15$ and the parameter $t = 0$. It can be noted that the maximum value is reached at-the-money with a transaction cost of almost 2. This maximum is also reached when Δ_{TC} is minimum. When the option becomes deeper in-the-money and out-of-the-money and Δ_{TC} increases, transaction costs tend to zero. A similar pattern is observed in the figure on the right side. The main difference reside on how the transaction costs highly increase as the the Asset 1 price is set as of $S_1 = 40$. As Gamma is maximum near the at-the-money moneyness of the option, transaction costs explode near this pricing area. As it can be seen, the costs are of 34 when both prices are set as 40. This will be the case in which rebalancing is done too often so that the option price becomes negative due to the high amount of transaction costs payed.

The plot of the left side of Figure 2.5 shows the same dynamics for the case when Asset 1 price is equal to $S_1 = 55$. These dynamics are similar to the one observed in the plot of the left side of Figure 2.4. As Gamma decreases when prices are to low or to high, transaction costs present the usual spike near the strike value. Moreover, this costs tend to zero as Δ_{TC} becomes larger and rebalancing is done less periodically. The plot of the right side helps us to understand how far can the transaction costs increase. For this purpose, we fixed the prices of both assets as of $S_1 = 40$ and $S_2 = 40$. Then, we plotted the value of the transaction cost with respect to its time to maturity and the size of Δ_{TC}. It can be observed that when $t = T$, transaction costs are almost zero as the option price is the predefined payoff. As time passes and the payoff is discounted, transaction costs increases as Gamma increases. When we reach the actual date $t = 0$, transaction costs grow up to 34. This result help us to confirm the following expected conclusion: As the frequency of rebalancing of the replicant portfolio increases, transaction costs increase such that after a certain point of time, the option price turns into negative and the model becomes ill-posed.

Convergence Analysis

The third item of the previous list involves measuring the convergence of the iterative framework in terms of the differences between consecutive solutions. Our approach will follow from the observation that the result of each iteration correspond to a square matrix. Hence, fixing the last time step $\tau = T$, we calculate the distance between two consecutive final results. For this purpose, we use three different p-norms matrix as of: the 1 norm, the 2 norm and the ∞ norm. In summary, for each step n and solution U^n, we calculate

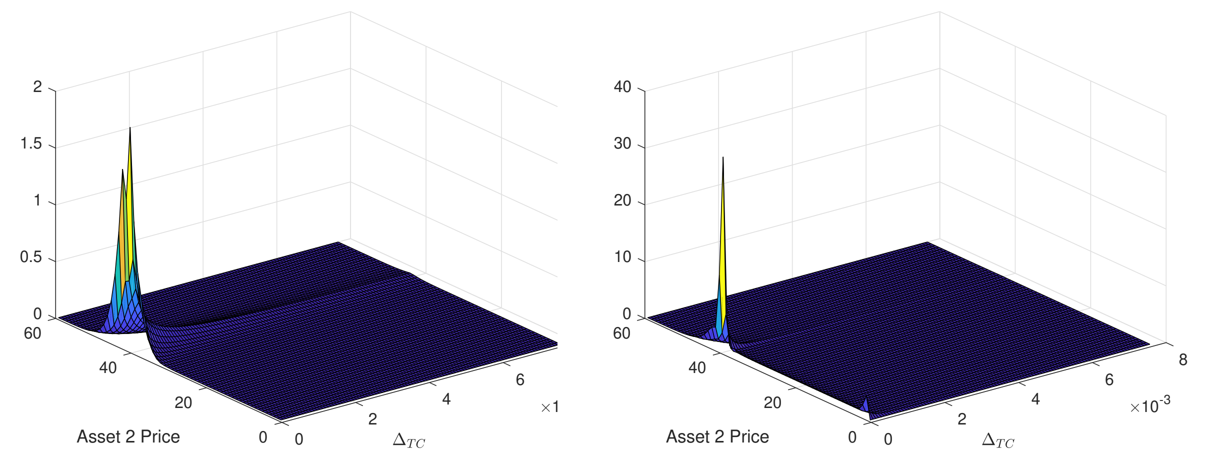


Figure 2.4: Testing 2 - The figure on the left shows the transaction costs function at time $t = 0$ when Asset 1 price is equal to 15. The figure on the right shows the transaction costs function at time $t = 0$ when Asset 1 price is equal to 40.

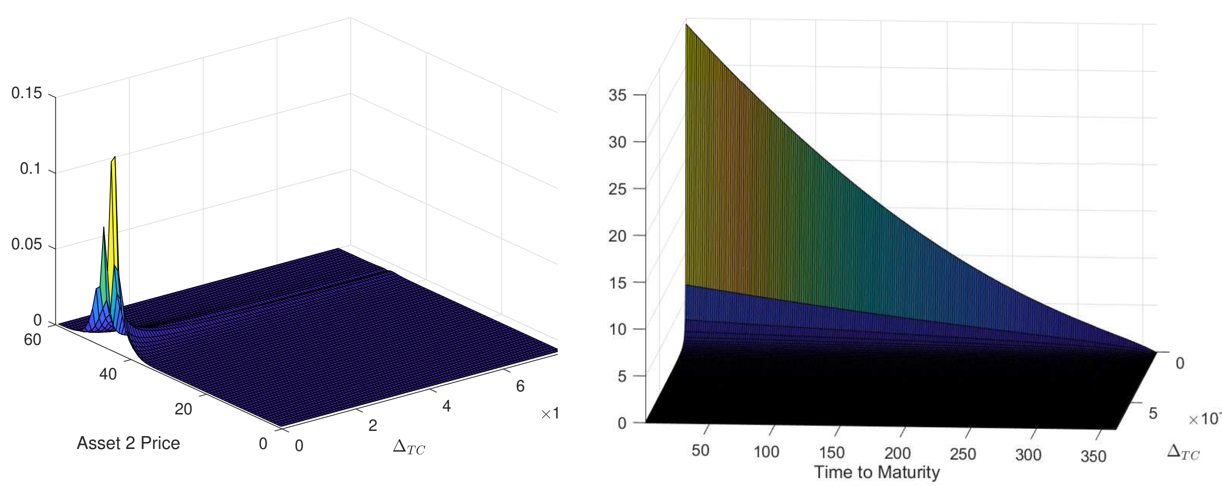
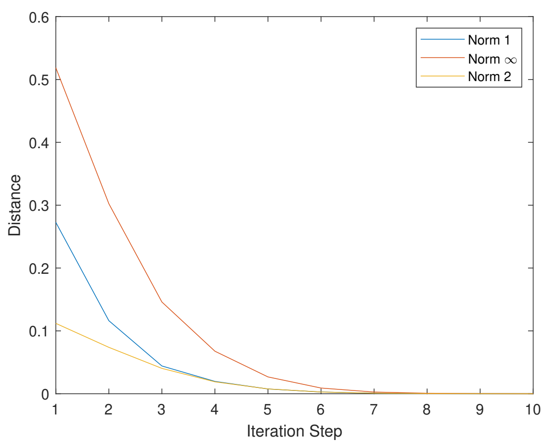


Figure 2.5: Testing 2 - The figure on the left shows the transaction costs function at time $t = 0$ when Asset 1 price is equal to 55. The figure on the right shows how the transaction costs function explodes when the option is at-the-money.

$$d\left(U^n, U^{n-1}\right) = \|U^n - U^{n-1}\|_p. \tag{2.4.9}$$

Our objective is to see that this distance tends to zero as n increases. Figure (2.6) to (2.8) present the plots of the results for the three scenarios. The figures plot the distance between two consecutive solutions against the iteration step $n+1$. In the right side, we provide a table with all the numerical results up to iteration $n = 11$. In the first case, it can be seen that the three norms exponentially decrease to zero and, between iteration 7 and 8, convergence is achieved. In the second case, convergence is achieved even faster as by step 2, the distance between both consecutive results is of order $E-05$. The third case is similar as the first scenario by noting that convergence is achieved at iteration 5 with a distance between consecutive solutions of order $E-04$.

‘



Iteration	Norm 1	Norm 2	Norm ∞
1	0.2727	0.1120	0.5187
2	0.1162	0.0738	0.3024
3	0.0442	0.0403	0.1460
4	0.0196	0.0189	0.0676
5	0.0076	0.0076	0.0267
6	0.0028	0.0027	0.0090
7	8.8E-4	8.6E-4	0.0027
8	2.5E-4	2.5E-4	7.5E-4
9	8.1E-5	6.7E-5	1.9E-4
10	2.3E-5	1.6E-5	4.5E-5

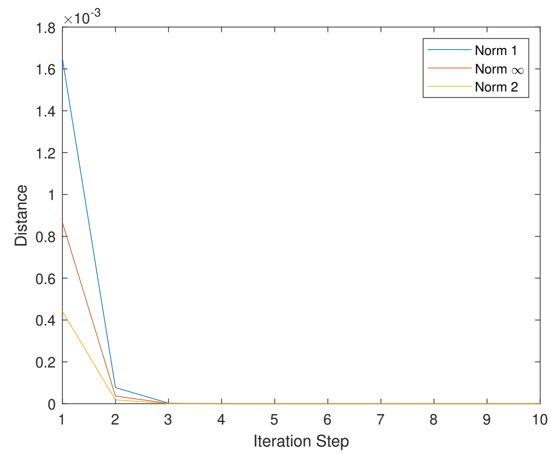
Figure 2.6: Convergence Analysis - Testing 1

2.5 Conclusion

Chapter 2 was dedicated to the presentation and deduction of an option pricing model for a multiasset option considering a general transaction costs function.

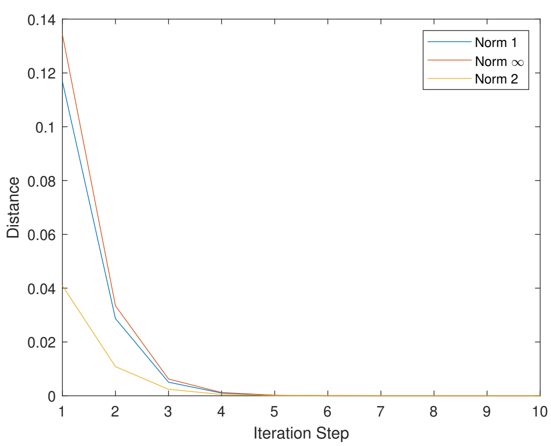
Section 2.1 was devoted to provide an introduction of the past works that have been done over the last years in relaxing multiple assumptions of the Black-Scholes model, including specific transaction costs functions and modelling multi-asset financial instruments.

Section 2.2 was devoted to adapt the standard steps applied in the Black-Scholes model by including the transaction costs on each time step. As a consequence, the dynamics of the pricing model can be explained by a nonlinear parabolic PDE. One important feature of this equation is that the "type" of nonlinearity depends of the shape of the transaction costs function. In fact, the nonlinear term is found to be proportional to the product of the asset price by the expected value of the transaction costs value.



Iteration	Norm 1	Norm 2	Norm ∞
1	0.0017	4.4E-4	8.6E-4
2	7.7E-5	1.8E-5	3.6E-5
3	2.6E-5	6.9E-7	1.2E-6
4	8.3E-8	2.1E-8	3.6E-8
5	2.4E-9	6E-10	9E-10
6	6E-11	1E-11	2E-11
7	1E-12	3E-13	5E-13
8	2E-14	8E-15	1E-14
9	7E-16	2E-16	2E-16
10	3E-16	1E-16	1E-16

Figure 2.7: Convergence Analysis - Testing 2



Iteration	Norm 1	Norm 2	Norm ∞
1	0.1169	0.0410	0.1344
2	0.0287	0.0108	0.0335
3	0.0050	0.0024	0.0062
4	0.0010	4.8E-4	0.0012
5	1.7E-4	8.6E-5	2.0E-4
6	2.6E-5	1.4E-5	3.0E-5
7	3.7E-6	2.1E-6	4.2E-6
8	5.0E-7	3.0E-7	5.5E-7
9	6.7E-8	4.0E-8	6.9E-8
10	8.5E-9	5.0E-9	8.3E-9

Figure 2.8: Convergence Analysis - Testing 3

Section 2.3 was devoted to the generalization of Leland's condition in a multi-asset framework with a general transaction costs function and proving the existence of a weak viscosity solution of the corresponding nonlinear problem. By transforming the nonlinear operator into a matricial form and forcing the nonlinear operator to be degenerate elliptic, it let us identify the shape of the generalized Leland's condition and showed that the original Leland's condition is in fact a particular case of our result. The second result involved proving the existence of a weak viscosity solution. For this purpose, we found a pair of sub and supersolutions using the already known result of existence of the Black-Scholes problem. After determining these solutions, we proved the existence of viscosity solution by using Perron's method and the correspondent comparison principle.

Section 2.4 was devoted to present a numerical framework to find a solution of the pricing problem using numerical techniques. As our PDE had crossed-derivatives terms, we decided to apply an ADI methodology. This choice let us split the numerical framework in two where both steps involved inverting two tridiagonal matrices. Moreover, we applied an iterative framework to solve the linearized equation. After defining the numerical scheme, we proposed three different pricing problems to analyse and understand different aspects of the option dynamics and the numerical framework. Moreover, we chose to price an specific multi-asset option from which an analytical solution is known for the model without transaction costs. The first testing involved analysing the transaction costs impact in the option price. We showed that the results are in line with the factors that form the transaction costs. When Gamma reaches its maximum correspond to the spot price in which transaction costs are higher. Moreover, as the option involves two different assets, the transaction costs pattern is similar to the Gamma of each asset. The second testing helped us to understand the sensitivity of the option price to changes in the frequency of rebalancing of the replicant portfolio. We observed results that are in line with the shape of the transaction costs function. As the frequency of rebalancing increases, transaction costs increase. Hence, the size of the frequency of rebalancing becomes crucial to define a well-posed problem and obtain positive option prices. The last testing was a convergence analysis. By using a metric that measures the matrix distance between consecutive solutions, we wanted to see the how the iterative framework converged after a several amount of steps. By considering the three testing frameworks proposed, the iterative method is seen to converge after not more than seven iterations.

2.6 Resumen del capítulo

En el Capítulo 2 presentamos y deducimos un modelo de valuación de opciones financieras de múltiples activos considerando una función de costos transacción general.

En la Sección 2.1 incluimos una introducción presentando diversos resultados de trabajos relacionados con el nuestro. Incluimost trabajos que muestran como modelar la valuación de opciones financieras al relajar los supuestos de Black-Scholes, al incluir diversos tipos de funciones de costos de transacción y al modelar instrumentos financieros con multiples activos como subyacentes.

En la Sección 2.2 nos abocamos a adaptar la metodología estandar de Black-Scholes al incluir costos de transacción en cada paso temporal. De esta forma, el modelo obtenido resulta ser una ecuación diferencial parabólica no lineal. Es importante notar que la forma que tenga la función de costos de transacción influirá en el tipo de no linealidad de la ecuación. Además, notamos que el término no lineal será siempre proporcional al producto del valor del activo por el valor esperado de los costos de transacción.

En la Sección 2.3 mostramos como la generalización de la condición de Leland permite definir un operador no lineal elíptico degenerado y así demostrar la existencia de al menos una solución viscosa. Al reescribir el operador no lineal en forma matricial, llegamos a la condición de Leland generalizada y observamos como la condición original no es más que un caso particular de nuestra condición. El segundo resultado consistió en demostrar la existencia de al menos una solución viscosa. Para ello, considerando la existencia de solución del problema lineal de Black-Scholes, construimos un par de sub y supersoluciones. Finalmente, utilizamos el método de Perron y un correspondiente principio de comparación para demostrar la existencia de solución.

En la Sección 2.4 presentamos el esquema numérico utilizado para encontrar una solución aproximada de nuestro problema de valuación. Dado que nuestra ecuación diferencial presenta segundas derivadas cruzadas, decidimos utilizar el algoritmo ADI. Esta elección nos permitió dividir en dos los pasos del algoritmo tal que en cada paso se inviertan matrices tridiagonales. Este algoritmo fue utilizado para resolver un problema iterativo que resultó de linealizar la ecuación original. El algoritmo fue probado bajo tres distintos escenarios económicos para entender la naturaleza del problema de valuación y la convergencia del método numerico. Elegimos valuar una opción financiera determinada para la cual conocíamos su fórmula cerrada para el modelo sin costos de transacción. El primer testeo realizado consistió en entender como los costos de transacción impactan en la valuación de la opción financiera. Observamos que los resultados se corresponden con el modelo teórico: los costos son máximos cuando Gamma es máxima. El segundo testeo nos permitió analizar la sensibilidad del precio de la opción ante cambios en la frecuencia de rebalanceo del portfolio replicante. Nuevamente, los resultados se condicen con el modelo teórico. A medida que la frecuencia de rebalanceo aumenta, los costos de transacción aumentan. Así, la frecuencia de rebalanceo resulta ser un factor crucial a la hora de definir una ecuación bien definida y así tener siempre precios positivos. El último testeo resultó ser una análisis de convergencia. Para ello, medimos la distancia entre dos resultados consecutivos del proceso iterativo. Observamos que en los tres escenarios planteados, el esquema iterativo converge con menos de siete iteraciones.

Chapter 3

Counterparty valuation adjustment with transaction costs

This chapter is devoted to the presentation and explanation of our second work Ref [6].

3.1 Introduction

3.1.1 Counterparty Credit Risk

In Section 1.1 we explained the standard steps followed to price a financial derivative under a broad set of assumptions. In this chapter we are going to work under a framework in which the credit riskiness of both the issuer of the contract and the counterparty are consider. So let us go one step back and provide an example of the type of pricing problem that we may want to study. Let us suppose that there exists two parts in a contract. Party A will buy a call option of certain stock S with strike K to party B. At maturity, two possible situations might occur: If the value of the stock S lies below the strike value, the option worth nothing and investor A will lose just the money used to buy the calls. However, if the asset price surpasses the strike value, investor B must have the stock (or the equivalent money) to close the contract with A. So, at this point is when we modify the standard Black-Scholes framework and ask ourselves: what happens if B doesn't have the stock or the money to close the contract? The answer to this equation depends on where the contract has been set. If the contract has been traded in an exchange (i.e. NYSE, NASDAQ, etc), each party is required to have a margin account so that the margin can be used when the party cannot close the contract. However, if the contract has been traded over-the-counter (OTC) there are no disclosure requirements and any possible contract and condition can be defined. In this type of scenarios is when the *counterparty credit risk* becomes important on any pricing framework.

As it is explained in Ref [23], the counterparty credit risk is the risk that a counterparty defaults prior to the expiration of the contract and fails to make future payments. This type of risk is only present when one counterparty has an exposure to the other. We will define the *exposure at default* (EAD) as the total amount owed by the defaulting party to the non-defaulting party

$$EAD = \max(V, 0) \tag{3.1.1}$$

where V is the total value. Given the exposure at the time of default, the amount recovered is measured as a percentage R, also known as the *recovery rate*. Then, the recovered amount is given by

$$R \times EAD = R \times \max(V, 0). \tag{3.1.2}$$

Another component that we need to consider is some measure that help use to define the probability of default. We define the hazard rate $\lambda(t)$ as the instantaneous probability of default of a security. Then, we will define the *probability of default* at time τ as of

$$P(\tau \leq T) = 1 - \exp\left(-\int_{\tau}^{T} \lambda(s)\, ds\right) \tag{3.1.3}$$

The *expected loss* in a financial contract can be obtained by multiplying the exposure by the loss given default by the probability of default, i.e.

$$EL = PD \times LGD \times EAD = PD \times (1 - R) \times EAD. \tag{3.1.4}$$

During this chapter, we are going to study the credit risk that is present during the existence of a financial contract. We define the *Credit Valuation Adjustment* or CVA to the market price of credit risk on a financial instrument that is marked-to-market, usually an OTC derivative. The CVA can be defined as the difference between the price of the instrument including credit risk and the price where both counterparties are free of credit risk. We will denote CVA as a positive value such that it represents a costs and a lost of value to the financial contract. Hence, the value of the option V can be denoted as

$$V = V_{BS} - CVA \tag{3.1.5}$$

where V_{BS} is the price of the option under the Black-Scholes framework. An additional term can be added to consider the credit risk of the issuer. This value is known as *Debit Valuation Adjustment* or DVA. The DVA acts oppositely as the CVA by adding value to the option when the issuer risk increases. Then, the adjusted option price is given by

$$V = V_{BS} - CVA + DVA$$

where CVA is a cost and DVA is a benefit. We will call unilateral CVA when only the the counterparty credit risk is considered and bilateral CVA when the issuer credit risk is also considered.

In the following two Sections we will study two possible ways to calculate the CVA. The first one is following the steps applied in the Black-Scholes model in Section 1.1.2 by creating the correspondent replicant portfolio. The second way of deriving CVA is by calculating the expectation of the future payoffs of the contract.

3.1.2 CVA calculation by Replication

Given the definitions of CVA and DVA provided above, we are going to show one methodology that can be used to find both values. This methodology will be similar to the Black-Scholes replication framework used previously in Section 1.1.2 using the work of Ref [22] and Ref [13]. Let us recall (3.1.5) and rename the functions as of

$$\hat{V} = V + U \tag{3.1.6}$$

where $\hat{V}$ is the value of a contract including the CVA, V the value without considering the counterparty risk and U the CVA.

Let us create a portfolio Π that replicates the price of the contract including the underlying asset S, a zero-coupon counterparty bond P_C and β amounts of money. Moreover, the market contains a zero-coupon risk-free bond P_R. The dynamics of these three instruments are given by

$$\begin{aligned} dS &= \mu_S S dt + \sigma_S S dW \\ dP_C &= r_C P_C dt - P_C dJ_C \\ dP_R &= r P_R dt \end{aligned} \tag{3.1.7}$$

where μ_S and σ_S are the mean and volatility of the underlying asset S, r is the risk-free interest rate, r_C is the yield of the counterparty bond of C and J_C is a point process that jumps from 0 to 1 on the default of C.

We define the value of the replicant portfolio Π as of

$$-\hat{V}(t) = \Pi(t) = \delta(t) S(t) + \alpha_C(t) P_C(t) + \beta(t) \tag{3.1.8}$$

where $\delta(t)$ is the amount of units held of $S(t)$, $\alpha_C(t)$ is the amount of units held of $P_C(t)$ and $\beta(t)$ are units of cash.

Given that we want the portfolio to be self-financing, we required that

$$-d\hat{V}(t) = \delta(t) dS(t) + \alpha_C(t) dP_C(t) + d\beta(t) \tag{3.1.9}$$

where $d\beta(t)$ is decomposed into $d\beta(t) = d\beta_S(t) + d\beta_C(t)$ with $d\beta_S$ being difference between the dividend income γ_S and the financing cost q_S of purchasing $\delta(t)$ units of S and $d\beta_C$ being the financing cost of short-selling the counterparty bond. Hence, $d\beta(t)$ can be decomposed into

$$d\beta(t) = \delta(\gamma_S - q_S) S dt - r\alpha_C P_C dt. \tag{3.1.10}$$

Then, we can rewrite equation (3.1.9) in terms of the components of $d\beta$

$$-d\hat{V} = \delta dS + \alpha_C dP_C + \delta(\gamma_S - q_S) S dt - r\alpha_C P_C dt. \tag{3.1.11}$$

Moreover, we can use the dynamics model for each financial instrument presented in (3.1.7) to get

$$-d\hat{V} = [(\gamma_S - q_S)\delta S + (r_C - r)\alpha_C P_C] dt - \alpha_C P_C dJ_C + \delta dS \tag{3.1.12}$$

Now we can recall equation (1.1.4) and apply Itô's formula to the process $\hat{V}$ in order to obtain

$$d\hat{V} = \frac{\partial \hat{V}}{\partial t} dt + \frac{\partial \hat{V}}{\partial S} dS + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 \hat{V}}{\partial S^2} dt + \triangle \hat{V}_C \, dJ_C \tag{3.1.13}$$

where $\triangle \hat{V}_C$ is calculated based on default conditions. In Ref [13] it is showed that

$$\triangle \hat{V}_C = \hat{V}(t, S, 0, 1) - \hat{V}(t, S, 0, 0) = -\hat{V} + R_C M^+ + M^- \tag{3.1.14}$$

Finally, we can match the results found in (3.1.12) and (3.1.13). We can eliminate all the stochastic components of the portfolio by taking

$$\begin{aligned} \delta &= -\frac{\partial \hat{V}}{\partial S} \\ \alpha_C &= \frac{\triangle \hat{V}_C}{P_C} = -\frac{\hat{V} - (M^- + R_C M^+)}{P_C} \end{aligned} \tag{3.1.15}$$

where M is the value of the financial contract at default. Hence, we obtain the PDE that governs the price of a financial contract under counterparty credit risk as

$$\begin{aligned} &\frac{\partial \hat{V}}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 \hat{V}}{\partial S^2} + (q_S - \gamma_S) S \frac{\partial \hat{V}}{\partial S} - r\hat{V} + (r_C - r) \triangle \hat{V}_C = 0 \\ &\hat{V}(T, S) = H(S) \end{aligned} \tag{3.1.16}$$

If we assume that the value at default M is equal to $\hat{V}$, we can use the Feynman-Kac theorem to derive the value of CVA. In fact, if we decompose $\hat{V} = V + U$, we get that

$$\frac{\partial U}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 U}{\partial S^2} + (q_S - \gamma_S)\, S \frac{\partial U}{\partial S} - rU - (1 - R_C)(r_C - r)(V + U)^+ = 0$$
$$U(T, S) = 0 \tag{3.1.17}$$

so that by Feynman-Kac theorem (Theorem 1.1.8) we can deduce the following non-linear integral equation

$$U(t, S) = -(1 - R_C) \int_t^T (r_C - r)\, D_r(t, u)\, \mathbb{E}\left[(V(u, S(u)) + U(u, S(u)))^+\right] du \tag{3.1.18}$$

where $D_r(t, u)$ is a discount factor between times t and u given by

$$D_r(t, u) = \exp\left(-\int_t^u r(v)\, dv\right). \tag{3.1.19}$$

3.1.3 CVA calculation by Expectation

The idea here is to follow similar steps as the ones applied in Section 1.1.3 using the work of Ref [22] and Ref [24]. For this purpose we may assume that parties A and B enter into a contract with discounted value $V(t, T)$ with maturity T in the absence of default, that only the counterparty credit risk of party B is considered and that the event of default occurs at time τ. If the event of default happens after the maturity of the contract the payoff of the contract is essentially

$$I_{\tau > T} V(t, T) \tag{3.1.20}$$

On the other side, if the default of counterparty B occurs before the maturity of the contract, the payoff of the contract is equal to the value of the position before the default time plus the recoverable part of the defaulted position. Hence, we can write this payoff as of

$$I_{\tau \leq T} V(t, \tau) + I_{\tau \leq T} \left(RV(\tau, T)^+ + V(\tau, T)^-\right) \tag{3.1.21}$$

where $x^+ = \max(x, 0)$ and $x^- = \min(x, 0)$. Then, if we put together the results from equations (3.1.20) and (3.1.21) and use the pricing framework by expectation explained in Section 1.1.3, we have that the price of the risky contract under the risk-neutral measure is

$$\begin{aligned}\hat{V}(t,T) &= \mathbb{E}\left[V(t,\tau)\, I_{\tau<T} + V(t,T)\, I_{\tau\geq T}|\mathcal{F}_t\right] + \mathbb{E}\left[\left(RV(\tau,T)^+ + V(\tau,T)^-\right) I_{\tau<T}|\mathcal{F}_t\right] \\ &= \mathbb{E}\left[V(t,\tau)\, I_{\tau<T} + V(t,T)\, I_{\tau\geq T}|\mathcal{F}_t\right] + \mathbb{E}\left[\left(RV(\tau,T)^+ + V(\tau,T) - V(\tau,T)^+\right) I_{\tau<T}|\mathcal{F}_t\right] \\ &= \mathbb{E}\left[V(t,\tau)\, I_{\tau<T} + V(t,T)\, I_{\tau\geq T}|\mathcal{F}_t\right] + \mathbb{E}\left[\left([R-1]\, V(\tau,T)^+ + V(\tau,T)\right) I_{\tau<T}|\mathcal{F}_t\right]\end{aligned}$$

Then, if we note that $V(\tau,T) = V(t,T)\, I_{\tau<T} + V(t,T)\, I_{\tau\geq T}$, we get the complete formula for the risky option price

$$\begin{aligned}\hat{V}(t,T) &= \mathbb{E}\left[V(t,T)\, I_{\tau<T} + V(t,T)\, I_{\tau\geq T} + (R-1)\, V(\tau,T)^+\, I_{\tau\geq T}|\mathcal{F}_t\right] \\ &= V(t,T) - \mathbb{E}\left[(1-R)\, V(\tau,T)^+\, I_{\tau\geq T}|\mathcal{F}_t\right]\end{aligned} \tag{3.1.22}$$

Hence, if we recall equation (3.1.5) and discard the assumption that $V(t,T)$ is already discounted, we define the value of the unilateral CVA as of

$$CVA = \mathbb{E}\left[D(t,\tau)\,(1-R)\, V(\tau,T)^+\, I_{\tau\geq T}|\mathcal{F}_t\right] \tag{3.1.23}$$

3.1.4 More work in CVA modelling

Although the use of CVA is recent, multiple works and frameworks have been developed over the years. In this Section we will recall different papers in which the authors show how to define and solve the CVA calculations under different scenarios and for different OTC derivatives.

In Ref [2], the authors derive the bilateral counterparty valuation adjustment terms by decomposing the un-defaultable portfolio into a set of binary states. Given defined the price of a default free portfolio as of $V(t)$ and τ, τ' the first-time-to-default of both parties, we can extend the price of that portfolio as of

$$\begin{aligned}V(t) &= V\left(t|\tau' > t, \tau > t\right) + V\left(t|\tau' < t, \tau < t\right) + V\left(t|\tau' < t, \tau > t\right) + V\left(t|\tau' > t, \tau < t\right) \\ &= V_s\left(t,\tau,\tau'\right) + V_d\left(t,\tau,\tau'\right)\end{aligned} \tag{3.1.24}$$

where V_s and V_d are the correspondent surviving and defaulted portolio's value which are defined as of

$$\begin{aligned}&V_s\left(t,\tau,\tau'\right) = V(t)\, I_{\tau'>t,\tau>t} \\ &V_d\left(t,\tau,\tau'\right) = V(t)\left[I_{\tau'<t,\tau<t} + I_{\tau'<t,\tau>t} + I_{\tau'>t,\tau<t}\right]\end{aligned} \tag{3.1.25}$$

Two more decompositions on the surviving portfolio can be uses to expand the formula for $V(t)$. The first expands in two binary states regarding the incoming or outgoing cash flows as of

$$V_s\left(t,\tau,\tau'\right) = V_s^+\left(t,\tau,\tau'\right) - V_s^-\left(t,\tau,\tau'\right) \tag{3.1.26}$$

Given a time horizon η to be δt further away of t such that $\eta = t + \delta t$, default times can fall before or after η. Then

$$\begin{aligned} 1^+ &= I_{\tau'<\eta,\tau>\tau'} + I_{\tau'<\eta,\tau<\tau'} + I_{\tau'>\eta} \\ 1^- &= I_{\tau<\eta,\tau'>\tau} + I_{\tau<\eta,\tau'<\tau} + I_{\tau>\eta}. \end{aligned} \tag{3.1.27}$$

Hence, the defaulted portfolio can be written as of

$$V_s\left(t,\tau,\tau'\right) = V_s^+\left(t,\tau,\tau'\right)1^+ - V_s^-\left(t,\tau,\tau'\right)1^- \tag{3.1.28}$$

The last expansion is done to include the recovery rates of both counterparties. If R and R' are both recovery rates, formula (3.1.28) can be expanded as of

$$\begin{aligned} V_s\left(t,\tau,\tau'\right) =& V_s^+\left(t,\tau,\tau'\right)\left[I_{\tau'<\eta,\tau>\tau'}\left(1-R'\right) + I_{\tau'<\eta,\tau>\tau'}R' + I_{\tau'<\eta,\tau<\tau'} + I_{\tau'>\eta}\right] \\ & - V_s^-\left(t,\tau,\tau'\right)\left[I_{\tau<\eta,\tau'>\tau}\left(1-R\right) + I_{\tau<\eta,\tau'>\tau}R + I_{\tau<\eta,\tau'<\tau} + I_{\tau>\eta}\right] \end{aligned} \tag{3.1.29}$$

Then, if we use the results of equation (3.1.25) into (3.1.29), we get that the value of the survival portfolio is given by

$$\begin{aligned} V_s\left(t,\tau,\tau'\right) =& V^+(t)\, I_{\tau'>t,\tau>t} I_{\tau'<\eta,\tau>\tau'}\left(1-R'\right) && (3.1.30) \\ & - V^-(t)\, I_{\tau'>t,\tau>t} I_{\tau<\eta,\tau'>\tau}\left(1-R\right) && (3.1.31) \\ & + V^+(t)\, I_{\tau'>t,\tau>t}\left[I_{\tau'<\eta,\tau>\tau'}R' + I_{\tau'<\eta,\tau<\tau'} + I_{\tau'>\eta}\right] && (3.1.32) \\ & - V^-(t)\, I_{\tau'>t,\tau>t}\left[I_{\tau<\eta,\tau'>\tau}R + I_{\tau<\eta,\tau'<\tau} + I_{\tau>\eta}\right] && (3.1.33) \end{aligned}$$

The authors interpret this equation so that the contributions to the value of the surviving portfolio come from two sources of loss (3.1.30) and (3.1.31) and two sources of recovery (3.1.32) and (3.1.33) of both counterparties. Then, by defining $V_s^*(t,\tau,\tau')$ to be the net total of recovery values for all cash flows between t and η, conditional on both counterparties surviving until t, the authors define CVA as the difference between $V_s^*(t,\tau,\tau')$ and $V_s(t,\tau,\tau')$ so that

$$\begin{aligned} CVA\left(t,\tau,\tau'\right) =& -V^+(t)\, I_{\tau'>t,\tau>t}\left[I_{\tau'<\eta,\tau>\tau'}R' + I_{\tau'<\eta,\tau<\tau'} + I_{\tau'>\eta}\right] && (3.1.34) \\ & + V^-(t)\, I_{\tau'>t,\tau>t}\left[I_{\tau<\eta,\tau'>\tau}R + I_{\tau<\eta,\tau'<\tau} + I_{\tau>\eta}\right] && (3.1.35) \end{aligned}$$

Other important work in CVA modelling and OTC pricing has been done by Ref [11] and Ref [10].

In the first article Ref [11] they consider the counterparty risk for a credit-default-swap (CDS) with correlation defaults of the counterparty and CDS reference credit. A CDS is a contract that works as an insurance. The buyer makes quarterly payments to the seller so that it will compensate the buyer with certain payoff in an event of credit default.

We denote τ_1 the default time of the CDS, by τ_2 the default time of the counterparty and the investor who considers the transaction with the counterparty to be default-free. Let us call T the maturity of the contract so that if $\tau_2 \leq T$, the counterparty cannot fulfil its obligations and the following could happen: if the net present value (NPV, which is the value of the contract discounted at time t) is negative for the investor, it is completely paid by the investor itself. If it is positive, only a recovery fraction R of the NPV is exchanged. If we denote $\Pi^D(t,T)$ to be the sum of all payoff terms between t and T discounted back at t and subject to counterparty default risk and $\Pi(t,T)$ the analogous without considering the same risk, the following formula is found valid

$$\begin{aligned}\mathbb{E}\left[\Pi^D(t,T)\,|\mathcal{F}_t\right] &= \mathbb{E}\left[\Pi(t,T)\,|\mathcal{F}_t\right] - (1-R)\,\mathbb{E}\left[I_{t\leq\tau_2\leq T} D(t,\tau_2)\,NPV(\tau_2)^+\,|\mathcal{F}_t\right] \\ &= \mathbb{E}\left[\Pi(t,T)\,|\mathcal{F}_t\right] - CVA \end{aligned} \tag{3.1.36}$$

We can approximate the formula defined in (3.1.36) by generating a time grid $T_0, T_1, ..., T_b = T$ and set for convenience $t = 0$ so that

$$\mathbb{E}\left[\Pi^D(0,T_b)\right] = \mathbb{E}\left[\Pi(0,T_b)\right] - (1-R)\sum_{j=1}^{b}\mathbb{E}\left[I_{T_{j-1}\leq\tau_2\leq T_j} D(0,T_j)\left(\mathbb{E}\left[\Pi(T_j,T_b)\,|\mathcal{F}_{T_j}\right]\right)^+\right] \tag{3.1.37}$$

where the approximation consists in postponing the default time to the first T_i following τ_2. In order to apply the formula defined in (3.1.36) to price a CDS contract, the following assumption is done. The stochastic intensity model is set as a *CIR++* model so that

$$\lambda_j(t) = y_j(t) + \Phi_j(t,\beta) \tag{3.1.38}$$

so that Φ is a free deterministic function and $y_j(t)$ is given by

$$dy_j(t) = k(\mu - y_j(t))\,dt + \nu\sqrt{y_j(t)}dW_j(t) \tag{3.1.39}$$

with W_j correlated brownian motions. From equation (3.1.37) it can be seen that the only non trivial term to compute is

$$\mathbb{E}\left[I_{T_{j-1}\leq\tau_2\leq T_j}\left(\mathbb{E}\left[\Pi\left(T_j,T_b\right)|\mathcal{F}_{T_j}\right]\right)^+\right] \tag{3.1.40}$$

Under these assumption, the authors show that this term which is at the same time the T_j-credit adjustment for counterparty risk equals to

$$\mathbb{E}\left[I_{T_{j-1}\leq\tau_2\leq T_j}\left(\mathbb{E}\left[\Pi\left(T_j,T_b\right)|\mathcal{F}_{T_j}\right]\right)^+\right] = \mathbb{E}\left[\overline{CDS}_{a,b}\left(T_j,S\right)^+\mathbb{E}\{I_{T_j-1\leq\tau_2\leq T_j,\tau_1>T_j}|\mathcal{F}_{T_j}\}\right]$$

where $\overline{CDS}_{a,b}\left(T_j,S\right)$ is given by

$$\begin{aligned}\overline{CDS}_{a,b}\left(T_j,S\right) =&S\left[-\int_{\max(T_a,T_j)}^{T_b}D\left(T_j,t\right)\left(t-T_{\gamma(t)-1}\right)\,dt\mathbb{Q}\left(\tau_1\geq t|\mathcal{F}_{T_j}\right)\right.\\ &\left.+\sum_{\max(a,j)+1}^{b}D\left(T_j,T_i\right)\alpha_i\mathbb{Q}\left(\tau_1\geq T_i|\mathcal{F}_{T_j}\right)\right] &(3.1.41)\\ &+\left(1-R\right)\left[\int_{\max(T_a,T_j)}^{T_b}D\left(T_j,t\right)\,dt\mathbb{Q}\left(\tau_1\geq t|\mathcal{F}_{T_j}\right)\right] &(3.1.42)\end{aligned}$$

with $D\left(T_j,t\right)$ is the discounting factor, $T_{\gamma(t)}$ is the first T_j following t, $\mathbb{Q}$ is the risk neutral measure, S is the CDS premium rate and α_i is the intensity of the jumps of the hazard rate (can be taken equal to zero if no jumps are assumed).

This framework is generalized in the second article Ref [10] by modelling a bilateral counterparty risk scenario. Equation (3.1.36) is generalized so that the bilateral CVA is given by

$$\begin{aligned}BR\text{-}CVA\left(t,T\right) =& \left(1-R_2\right)\mathbb{E}\left[I_{\tau_2\leq\tau_0\leq T,\tau_2\leq T\leq\tau_0}D\left(t,\tau_2\right)NPV\left(\tau_2\right)^+|\mathcal{F}_t\right]\\ &-\left(1-R_0\right)\mathbb{E}\left[I_{\tau_0\leq\tau_2\leq T,\tau_0\leq T\leq\tau_2}D\left(t,\tau_0\right)\left[-NPV\left(\tau_0\right)^+\right]|\mathcal{F}_t\right] &(3.1.43)\end{aligned}$$

where τ_0 is the investor's default time and τ_2 is the counterparty's default time. Hence, the bilateral CVA for a CDS contract with stochastic hazard rates is

$$\begin{aligned}BR\text{-}CVA\text{-}CDS_{a,b}\left(t,S\right) =& \left(1-R_2\right)\mathbb{E}\left[I_{\tau_2\leq\tau_0\leq T,\tau_2\leq T\leq\tau_0}D\left(t,\tau_2\right)\left[I_{\tau_1>\tau_2}\overline{CDS}_{a,b}\left(\tau_2,S\right)\right]^+|\mathcal{F}_t\right]\\ &-\left(1-R_0\right)\mathbb{E}\left[I_{\tau_0\leq\tau_2\leq T,\tau_0\leq T\leq\tau_2}D\left(t,\tau_0\right)\left[-I_{\tau_1>\tau_0}\overline{CDS}_{a,b}\left(\tau_0,S\right)\right]^+|\mathcal{F}_t\right]\end{aligned} \tag{3.1.44}$$

In Ref [12] different methods and formulas for modelling bilateral CVA on interest-rate portfolios are introduced to expand the coverage of the results obtained in Ref [10]. Interest rates are modelled using a $G2++$ model so that the dynamics of the short-rate process under the risk-neutral measure are given by

$$r(t) = x(t) + z(t) + \Phi(t, \alpha) \tag{3.1.45}$$

where α is a set of parameters, Φ is a deterministic function with $\Phi(0, \alpha) = r_0$ and the processes x and z are $\mathcal{F}_t$ adapted and satisfy

$$dx(t) = -ax(t)\,dt + \sigma dZ_1(t) \tag{3.1.46}$$
$$dy(t) = -bz(t)\,dt + \eta dZ_2(t) \tag{3.1.47}$$

with Z_1 and Z_2 two brownian motions with instantaneous correlation ρ_{12}. Hazard rates are again modelled following equation (3.1.38) where the correspondent brownian motions is correlated with Z_1 and Z_2. In their paper, the authors analyze the impact of correlations, interest-rate curve, credit spread levels and volatility scenarios on the bilateral CVA calculation for an interest rate swap (IRS).

A different approach is taken by Ref [14] and further on with Ref [15] and Ref [16]. In their works, the author develop a reduced-form backward stochastic differential equations (BSDE) approach to the problem of pricing and hedging of the CVA by allowing the presence of multiple funding constraints.

The third approach taken and the one that we will consider as a starting point of our work is the one developed by Ref [13] in which they derive a market model similar to the one presented in Section 3.1.2. In their work they find the PDE that governs the dynamics of a derivative contract $\tilde{V}$ on an asset S between a seller B and a counterparty C that may both default as of

$$\frac{\partial \hat{V}}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 \hat{V}}{\partial S^2} + (q_S - \gamma_S) S \frac{\partial \hat{V}}{\partial S} - r\hat{V} = (1 - R_B)\lambda_B \tilde{V}^- + (1 - R_C)\lambda_C \tilde{V}^+ + s_F \tilde{V}^+ \tag{3.1.48}$$

where R_B and R_C are the correspondent recovery rates, λ_B and λ_C the correspondent hazard rates and s_F is the difference between the seller funding rate and the risk-free-rate. Hence, by applying the Feynman-Kac theorem (1.1.8) the found that the value of the CVA is given by

$$\begin{aligned} U(t, S) = &- (1 - R_B) \int_t^T (r_B - r) D_r(t, u) \mathbb{E}\left[(V(u, S(u)) + U(u, S(u)))^-\right] du \\ &- (1 - R_C) \int_t^T (r_C - r) D_r(t, u) \mathbb{E}\left[(V(u, S(u)) + U(u, S(u)))^+\right] du \\ &- \int_t^T s_F(u) D_r(t, u) \mathbb{E}\left[(V(u, S(u)) + U(u, S(u)))^+\right] du \end{aligned}$$

where $D_r(t,u)$ is the same as of in Section 3.1.2.

Our work expands the framework of Burgard and Kjaer in Ref [13] to include the transaction costs that arise of trading the underlying assets and both the issuer and the counterparty bonds. We propose an initial constant transaction cost function proportional to the amount of assets traded. As a consequence, we derive a nonlinear PDE that extends the results found in Ref [13] and prove the existence of solution by applying the Schauder Fixed-Point theorem. In the second part of the work, we develop a numerical approach to solve the PDE by considering a non-uniform grid on the spatial variable. The main greeks of the option (Delta, Gamma, Rho and Vega) are calculated and analyzed to understand how both the value adjustments and transaction costs affect the behavior of the option price. Nonetheless, we perform a sensitivity analysis on the remaining parameters (hazard rate, recovery rates, etc) to complete the study of the option price dynamics.

3.2 The market model

As we discussed in the previous Section, the seminal paper of Ref [13] derives the PDE for the value of a financial derivative considering both bilateral counterparty risk and funding costs. In order to build their market model they propose a economy consisting of a risk-free zero-coupon bond, two default risky zero-coupon bond of party B and C and a spot asset with no default risk. B will refer to the seller and C to the counterparty. Notation will be followed from the original work Ref [13] .

The dynamics of the four tradable assets under the historical probability measure are defined as follows:

$$\begin{cases} dP_R &= P_R\, r\, dt, \\ dP_B &= P_B\, r_B\, dt \,-\, P_B\, dJ_B, \\ dP_C &= P_C\, r_C\, dt \,-\, P_C\, dJ_C, \\ dS &= \mu\, S\, dt \,+\, \sigma\, S\, dW_t. \end{cases}$$

The default risky zero-coupon bonds are modeled by considering both r_B and r_C interest rates and J_B and J_C the two independent point processes that jump from 0 to 1 on default of B and C respectively. The default risk-free zero-coupon bond is a deterministic process with drift equal to r and the spot asset is modeled following a geometric brownian motion with drift μ and volatility σ. Among this Section the parameters r, r_B, r_C, μ and σ are held positive and constant. The value of the derivative at time t will be denoted $\hat{V}(t, S, J_B, J_C)$ and depends on the spot S and the default states J_B and J_C. In our calculations we will drop the dependencies of S, J_B and J_C. Also we will recall the following notation which will be useful in further steps:

$$\begin{aligned} x^+ &= \max(x,0) \\ x^- &= \min(x,0)\,. \end{aligned}$$

In order to derive the price of a financial option $\hat{V}$, we are going to adapt the standard Black-Scholes framework discussed in Section 1.1.2 and further applied in Ref [13] to consider the Leland's approach already used in Chapter 2. We will have to define three transaction costs functions which will be needed to calculate the costs that arises from trading both risky zero-coupon bonds and the underlying instrument.

Let us start by creating the self-financing portfolio which covers all the underlying risk factors that hedges the option. Let $\Pi(t)$ be the sellers portfolio which consists of $\delta(t)$ units of $S(t)$, $\alpha_B(t)$ units of $P_B(t)$, $\alpha_C(t)$ units of $P_C(t)$ and $\beta(t)$ units of cash. For hedging purposes we set $\Pi(t) + \hat{V}(t) = 0$ and

$$-\hat{V}(t) = \Pi(t) = \delta(t)\ S(t) + \alpha_B(t)\ P_B(t) + \alpha_C(t)\ P_C(t) + \beta(t). \tag{3.2.1}$$

We define the transaction costs function for both default risky bonds P_B and P_C and the spot asset S as follows:

$$\begin{cases} TC_B(t, P_B) & = C_B\,|\alpha_B(t)|\,P_B(t)\,, \\ TC_C(t, P_C) & = C_C\,|\alpha_C(t)|\,P_C(t)\,, \\ TC_S(t, S) & = C_S\,|\delta(t)|\,S(t) \end{cases}$$

where C_B, C_C and C_S are positive constants. This definition of transaction costs is the standard approach applied initially in Ref [33] and is the initial step before creating more complex dynamics. In this case, the costs are defined to be proportional to the amount of assets traded multiplied by the price of each asset. For the purpose of enhancing clarity, we drop the dependencies on every function.

By forcing the portfolio to be self-financing, we find that

$$-d\hat{V} = \delta\, dS + \alpha_B\, dP_B + \alpha_C\, dP_C + d\beta. \tag{3.2.2}$$

where $d\beta$ is decomposed as $d\beta = d\beta_S + d\beta_F + d\beta_C$ corresponding to the variations in the cash position due to each of the three assets. In this step we consider the effect of the transaction cost in the hedging strategy. On each time step, there would be a decrease in the cash account because of the cost of buying or selling a different amount of assets. Hence, the original calculations of Ref [13] are modified as follows:

- The share position provides a dividend income, a financing cost and a transaction cost. The variation in the position is found to be

$$d\beta_S = \delta\,\gamma_S\,S\,dt - \delta\,q_S\,S\,dt - dTC_S \tag{3.2.3}$$

- After the own bonds are purchased, if any surplus in cash is available, it must earn the free-risk-rate r. If borrowing money, the seller needs to pay the rate r_F. In this case, transaction costs appear when calculating the surplus after the own bonds purchasing. The variation in this position is determined by

$$\begin{aligned} d\beta_F &= r\left(-\hat{V}-\alpha_B\,P_B-TC_B\right)^{+}\,dt + r_F\left(-\hat{V}-\alpha_B\,P_B-TC_B\right)^{-}\,dt \\ &= r\left(-\hat{V}-\alpha_B\,P_B-TC_B\right)\,dt + s_F\left(-\hat{V}-\alpha_B\,P_B-TC_B\right)^{-}\,dt \end{aligned} \tag{3.2.4}$$

where $s_F = r_F - r$ is the funding spread.

- Finally, a financing cost due to short-selling the counterparty bond and its related transaction costs are considered for calculating the variation in the cash counterparty position as follows:

$$d\beta_C = -\alpha_C\,r\,P_C\,dt - dTC_C. \tag{3.2.5}$$

By applying equations (3.2.3), (3.2.4) and (3.2.5) in (3.2.2), we obtain

$$\begin{aligned} -d\hat{V} =& \delta\,dS + \alpha_B\,P_B\,(r_B\,dt - dJ_B) + \alpha_C\,P_C\,(r_C\,dt - dJ_C) - dTC_C - dTC_S + \\ & \left[r\left(-\hat{V}-\alpha_B\,P_B-TC_B\right)^{+} + r_F\left(-\hat{V}-\alpha_B\,P_B-TC_B\right)^{-} + \delta\,(\gamma_S - q_S)\,S\right. \\ & \left. -\ \alpha_C\,r\,P_C\ \right]dt \end{aligned} \tag{3.2.6}$$

$$\begin{aligned} -d\hat{V} =& \left[-r\hat{V} + s_F\left(-\hat{V}-\alpha_B\,P_B-TC_B\right)^{-} + (r_B - r)\,\alpha_B\,P_B + (r_C - r)\,\alpha_C\,P_C\right. \\ & \left. -\ r\,TC_B + \delta\,(\gamma_S - q_S)\right]dt - dTC_S - dTC_C + \delta\,dS - \alpha_B\,P_B\,dJ_B - \alpha_C\,P_C\,dJ_C. \end{aligned} \tag{3.2.7}$$

In this step we will need to apply the Itô's formula to the process $\hat{V}$. However, we can observe that the process $\hat{V}$ contains a jump process within. Hence, we will need to define first the Itô's formula adapted to jump processes.

Definition 3.2.1 (Itô's formula for single jumps)**.** If X is an stochastic process satisfying the stochastic differential equation $dX_t = \mu dt + \sigma dW_t - JdN$ where N is a Poisson process and J is the size of the jump and f is a deterministic twice continuously differentiable function, then $Y_t = f\left(X_t\right)$ is also a stochastic process and is given by

$$dY_t = \left(\mu f'\left(X_t\right) + \frac{1}{2}\sigma^2 f''\left(X_t\right)\right)dt + \left(\sigma f'\left(X_t\right)\right)dW_t + \left[f\left(X_t + J\right) - f\left(X_t\right)\right]dN \tag{3.2.8}$$

We can use this adapted Itô's formula on our function $\hat{V}$ that depends on both possible jumps J_B and J_C. We denote $\triangle\hat{V}_B$ and $\triangle\hat{V}_C$ to the difference in the price before and after the jumps. In Ref [13] it is showed that

$$\triangle \hat{V}_B = \hat{V}\,(t,S,1,0) - \hat{V}\,(t,S,0,0) = -\hat{V} + \hat{V}^{+} + \; R_B\,\hat{V}^{-}$$
$$\triangle \hat{V}_C = \hat{V}\,(t,S,0,1) - \hat{V}\,(t,S,0,0) = -\hat{V} + R_C\,\hat{V}^{+} + \; \hat{V}^{-}$$

Then, we observe that

$$d\hat{V} = \frac{\partial \hat{V}}{\partial t}\,dt + \frac{\partial \hat{V}}{\partial S}\,dS + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 \hat{V}}{\partial S^2}\,dt + \triangle \hat{V}_B\,dJ_B + \triangle \hat{V}_C\,dJ_C \tag{3.2.9}$$

Hence, if we add together (3.2.7) and (3.2.9) we can clear out the values of δ, α_B and α_C in order to hedge the risks related to the corporate bonds and the spot asset. Then, we obtain that

$$\delta = -\frac{\partial \hat{V}}{\partial S} \tag{3.2.10}$$
$$\alpha_B = \frac{-\hat{V} + \hat{V}^{+} + \; R_B\,\hat{V}^{-}}{P_B} \tag{3.2.11}$$
$$\alpha_C = \frac{-\hat{V} + R_C\,\hat{V}^{+} + \; \hat{V}^{-}}{P_C} \tag{3.2.12}$$

so that the formula becomes

$$\begin{aligned} 0 = \Bigg[& -r\hat{V} + s_F\left(-\hat{V} - \alpha_B\,P_B - TC_B\right)^{-} + (r_B - r)\,\alpha_B\,P_B + (r_C - r)\,\alpha_C\,P_C - r\,TC_B \\ & + \delta\,S\,\left(\gamma_S - q_S\right) + \frac{\partial \hat{V}}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 \hat{V}}{\partial S^2}\Bigg]\,dt - dTC_S - dTC_C \end{aligned} \tag{3.2.13}$$

By recalling the definition of the transaction costs, we can compute dTC_S and dTC_C. Then, for the calculation of the transaction costs of the spot asset, we recall the value of δ and note that

$$dTC_S = C_S\,|d\delta|\,S \;\sim C_S\,\left|\frac{\partial^2 \hat{V}}{\partial S^2}\right|\,\sigma\,S^2\sqrt{dt}\,\sqrt{\frac{2}{\pi}} \tag{3.2.14}$$

where the approximation is made by taking the expected value of $|d\delta|$ and taking the lowest order $\mathcal{O}\left(\sqrt{\Delta t}\right)$ as it follows

$$E\left(|d\delta|\right) = E\left(\left|\frac{\partial^2 \hat{V}}{\partial S^2}\right|\,dS\right) = \left|\frac{\partial^2 \hat{V}}{\partial S^2}\right|\,\sigma\,S\,\sqrt{dt}\,E\left[\Phi\right] \tag{3.2.15}$$

and setting Φ as a standard normal random variable.

The variation of the transaction costs of the counterparty bond position is computed by applying the same rationale as before but over $|d\alpha_C|$ in this case. Then,

$$dTC_C = C_C\,|d\alpha_C|\,P_C \sim C_C \left| R_C\,\frac{\partial \hat{V}}{\partial S}^{+} + \frac{\partial \hat{V}}{\partial S}^{-} - \frac{\partial \hat{V}}{\partial S} \right| \sigma\,S\sqrt{dt}\,\sqrt{\frac{2}{\pi}} \tag{3.2.16}$$

where on this occasion the approximation is obtained by taking the expected value of $|d\alpha_C|$ as it follows

$$E\left(|d\alpha_C|\right) = E\left(\left| R_C\,\frac{\partial \hat{V}}{\partial S}^{+} + \frac{\partial \hat{V}}{\partial S}^{-} - \frac{\partial \hat{V}}{\partial S} \right| dS \right) = \left| R_C\,\frac{\partial \hat{V}}{\partial S}^{+} + \frac{\partial \hat{V}}{\partial S}^{-} - \frac{\partial \hat{V}}{\partial S} \right| \sigma\,S\,\sqrt{dt}\,E\left[\Phi\right]$$

and setting again Φ as a standard normal random variable.

By recalling (3.2.14) and (3.2.16) and applying those computations in (3.2.13), we obtain the following nonlinear parabolic partial derivative equation

$$\begin{aligned}
0 = \Bigg[& -r\hat{V} + s_F\left(-\hat{V} - \alpha_B\,P_B - TC_B\right)^{-} + (r_B - r)\,\alpha_B\,P_B + (r_C - r)\,\alpha_C\,P_C - r\,TC_B \\
& + \delta\,S\,\left(\gamma_S - q_S\right) + \frac{\partial \hat{V}}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 \hat{V}}{\partial S^2}\Bigg]\,dt - C_S \left|\frac{\partial^2 \hat{V}}{\partial S^2}\right| \sigma\,S^2\sqrt{dt}\,\sqrt{\frac{2}{\pi}} \\
& - C_C \left| R_C\,\frac{\partial \hat{V}}{\partial S}^{+} + \frac{\partial \hat{V}}{\partial S}^{-} - \frac{\partial \hat{V}}{\partial S} \right| \sigma\,S\sqrt{dt}\,\sqrt{\frac{2}{\pi}} \\
0 = \Bigg[& -r\hat{V} + s_F\left(-\hat{V} - \alpha_B\,P_B - C_B\,\left|-\hat{V} + \hat{V}^{+} + R_B\,\hat{V}^{-}\right|\right)^{-} + (r_B - r)\,\alpha_B\,P_B \\
& + (r_C - r)\,\alpha_C\,P_C - r\,\left(C_B\,\left|-\hat{V} + \hat{V}^{+} + R_B\,\hat{V}^{-}\right|\right) - \frac{\partial \hat{V}}{\partial S}\,S\,\left(\gamma_S - q_S\right) + \frac{\partial \hat{V}}{\partial t} \\
& + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 \hat{V}}{\partial S^2}\Bigg]\,dt - C_S \left|\frac{\partial^2 \hat{V}}{\partial S^2}\right| \sigma\,S^2\sqrt{dt}\,\sqrt{\frac{2}{\pi}} \\
& - C_C \left| R_C\,\frac{\partial \hat{V}}{\partial S}^{+} + \frac{\partial \hat{V}}{\partial S}^{-} - \frac{\partial \hat{V}}{\partial S} \right| \sigma\,S\sqrt{dt}\,\sqrt{\frac{2}{\pi}}
\end{aligned} \tag{3.2.17}$$

If we set $\lambda_B = r_B - r$, $\lambda_C = r_C - r$ and apply the definitions of α_B and α_C, (3.2.17) becomes

$$\begin{aligned}0 =& \Big[-r\hat{V} + s_F \left(-\hat{V}^+ - R_B \hat{V}^- - C_B \left| -\hat{V} + \hat{V}^+ + R_B \hat{V}^- \right| \right)^- + \lambda_B \left(-\hat{V} + \hat{V}^+ + R_B \hat{V}^- \right) \\ &+ \lambda_C \left(-\hat{V} + R_C \hat{V}^+ + \hat{V}^- \right) - r \left(C_B \left| -\hat{V} + \hat{V}^+ + R_B \hat{V}^- \right| \right) - \frac{\partial \hat{V}}{\partial S} S \left(\gamma_S - q_S \right) \\ &+ \frac{\partial \hat{V}}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 \hat{V}}{\partial S^2} \Big] \, dt - C_S \left| \frac{\partial^2 \hat{V}}{\partial S^2} \right| \sigma S^2 \sqrt{dt} \sqrt{\frac{2}{\pi}} \\ &- C_C \left| R_C \frac{\partial \hat{V}}{\partial S}^+ + \frac{\partial \hat{V}}{\partial S}^- - \frac{\partial \hat{V}}{\partial S} \right| \sigma S \sqrt{dt} \sqrt{\frac{2}{\pi}} \end{aligned} \tag{3.2.18}$$

The absolute value that involves the transaction costs due to the own bonds purchase can be reduced by noting that when $\hat{V} \geq 0$, its value is 0 and when $\hat{V} < 0$ is equal to $(R_B - 1)\,\hat{V}$. Hence,

$$\left| -\hat{V} + \hat{V}^+ + R_B \hat{V}^- \right| = (R_B - 1)\, \hat{V}^- \tag{3.2.19}$$

Using this reduction in the minimum function that is multiplied by s_F, we get that

$$\begin{aligned}\left(-\hat{V}^+ - R_B \hat{V}^- - C_B \left| -\hat{V} + \hat{V}^+ + R_B \hat{V}^- \right| \right)^- &= \left(-\hat{V}^+ - R_B \hat{V}^- - C_B \,(R_B - 1)\, \hat{V}^- \right)^- \\ &= \left(-\hat{V}^+ - \hat{V}^- \left[R_B - C_B \,(R_B - 1) \right] \right)^- \\ &= \begin{cases} -\hat{V} & \text{if} \quad \hat{V} \geq 0 \\ 0 & \text{if} \quad \hat{V} < 0 \end{cases} \\ &= -\hat{V}^+. \end{aligned} \tag{3.2.20}$$

Thus, by implementing (3.2.20) in (3.2.18), we get

$$\begin{aligned}0 =& - r\hat{V} - s_F \hat{V}^+ + \lambda_B \left(-\hat{V} + \hat{V}^+ + R_B \hat{V}^- \right) + \lambda_C \left(-\hat{V} + R_C \hat{V}^+ + \hat{V}^- \right) \\ &- r\, C_B \,(R_B - 1)\, \hat{V}^- - \frac{\partial \hat{V}}{\partial S} S \left(\gamma_S - q_S \right) + \frac{\partial \hat{V}}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 \hat{V}}{\partial S^2} \\ &- \sigma S^2 \sqrt{\frac{2}{\pi\, dt}} \left(C_S \left| \frac{\partial^2 \hat{V}}{\partial S^2} \right| + S^{-1} C_C \left| R_C \frac{\partial \hat{V}}{\partial S}^+ + \frac{\partial \hat{V}}{\partial S}^- - \frac{\partial \hat{V}}{\partial S} \right| \right) \end{aligned} \tag{3.2.21}$$

If we introduce the parabolic operator $\mathcal{A}_t$ as

$$\mathcal{A}_t \equiv \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 \hat{V}}{\partial S^2} + \frac{\partial \hat{V}}{\partial S} S \left(q_S - \gamma_S \right) \tag{3.2.22}$$

then it follows that $\hat{V}$ is the solution of

$$
\begin{aligned}
0 = & \frac{\partial \hat{V}}{\partial t} + \mathcal{A}_t \hat{V} - (\lambda_B + \lambda_C + r)\,\hat{V} + (\lambda_B + \lambda_C\, R_C - s_F)\,\hat{V}^+ \\
& + (\lambda_B\, R_B + \lambda_C - r\,(R_B - 1)\,C_B)\,\hat{V}^- \\
& - \sigma\, S^2 \sqrt{\frac{2}{\pi\, dt}} \left(C_S \left| \frac{\partial^2 \hat{V}}{\partial S^2} \right| + S^{-1}\, C_C \left| R_C\, \frac{\partial \hat{V}}{\partial S}^+ + \frac{\partial \hat{V}}{\partial S}^- - \frac{\partial \hat{V}}{\partial S} \right| \right)
\end{aligned} \tag{3.2.23}
$$

Looking forward to compare (3.2.23) with equation 26 from Ref [13], we rearrange the terms that involve $\hat{V}$, $\hat{V}^+$ and $\hat{V}^-$ and obtain the following nonlinear parabolic PDE

$$
\begin{aligned}
\frac{\partial \hat{V}}{\partial t} + \mathcal{A}_t \hat{V} - r\,\hat{V} = & s_F\,\hat{V}^+ + \lambda_C\,(1 - R_C)\,\hat{V}^+ + \lambda_B\,(1 - R_B)\,\hat{V}^- - r\,(1 - R_B)\,C_B\,\hat{V}^- \\
& + \sigma\, S^2 \sqrt{\frac{2}{\pi\, dt}} C_S \left| \frac{\partial^2 \hat{V}}{\partial S^2} \right| + \sigma\, S \sqrt{\frac{2}{\pi\, dt}} C_C \left| R_C\, \frac{\partial \hat{V}}{\partial S}^+ + \frac{\partial \hat{V}}{\partial S}^- - \frac{\partial \hat{V}}{\partial S} \right|.
\end{aligned} \tag{3.2.24}
$$

The first three terms on the right hand side of (3.2.24) are equal to the nonlinear terms of the original model. The inclusion of the transaction costs in the hedging strategy brings to the model three new terms:

- The fourth term in the right hand side of (3.2.24) corresponds to the amount of cash that is not invested at r rate when considering the surplus held by the seller after the purchase of its own bonds as it is shown in (3.2.4).

- The fifth term is the effect of the transaction costs due to buying or selling δ assets of S. It shall be noted that the term is equal to one find in Leland's standard approach.

- The sixth term is the effect of the transaction costs due to shorting the counterparty bond.

If we want to compare (3.2.23) with Leland's notation, we can define the *modified volatility* as

$$
\hat{\sigma}^2 = \sigma^2 \left(1 - \sqrt{\frac{2}{\pi\, dt}}\, \frac{C_S}{\sigma}\, \mathrm{sgn}\left(\frac{\partial^2 \hat{V}}{\partial S^2} \right) \right) \tag{3.2.25}
$$

and noting that

$$\left| R_C \, \frac{\partial \hat{V}}{\partial S}^{+} + \frac{\partial \hat{V}}{\partial S}^{-} - \frac{\partial \hat{V}}{\partial S} \right| = \begin{cases} \left|(1 - R_C)\, \frac{\partial \hat{V}}{\partial S}\right| & \text{if} \quad \hat{V} \geq 0 \\ 0 & \text{if} \quad \hat{V} < 0 \end{cases}$$
$$= (1 - R_C) \left| \frac{\partial \hat{V}}{\partial S}^{+} \right| \tag{3.2.26}$$

we obtain the following differential equation

$$\frac{\partial \hat{V}}{\partial t} + \frac{1}{2}\, \hat{\sigma}^2 \, S^2 \, \frac{\partial^2 \hat{V}}{\partial S^2} + \frac{\partial \hat{V}}{\partial S}\, S \, (q_S - \gamma_S) - r\, \hat{V} = \, s_F \, \hat{V}^{+} + \lambda_C \, (1 - R_C)\, \hat{V}^{+}$$
$$+ \lambda_B \, (1 - R_B)\, \hat{V}^{-} - r \, (1 - R_B)\, C_B \, \hat{V}^{-} + \sigma \, S \sqrt{\frac{2}{\pi \, dt}} C_C \, (1 - R_C) \left| \frac{\partial \hat{V}}{\partial S}^{+} \right| \tag{3.2.27}$$

Remark 3.2.2. The left-hand side of equation (3.2.27) is effectively a Black-Scholes operator with a volatility parameter $\hat{\sigma}$, a dividend yield γ_S and a financing cost (different to the risk-free interest rate) q_S. The right-hand side of the equation contains the nonlinear terms that arises from considering the existence of transaction costs and default risk. The inclusion of these 'extra' costs can be thought as a perturbation to the original model. By assessing the magnitude of the parameters of each term, it can be noted that they are indeed small. Hazard rates, recovery rates and interest rates are always below one and the transaction costs per unit of asset can be modeled between 0.025 to 0.04 as it is done in Ref [33].

3.3 Existence of solution of the PDE

3.3.1 Preliminaries

We finish Section 3.2 by presenting the PDE that models the price of a financial contract between two parties B and C with an underlying S allowing probability of default and considering the correspondent transaction costs in the replication strategy. Equation (3.2.27) shows that the PDE has nonlinear terms involving different minimum and maximum functions. Then, we are going to present the Sobolev spaces in which we are going to search for a weak solution of the PDE. Let us recall the definitions presented in Section 1.2.1 to define the Sobolev space and the correspondent norm. Let Ω be a bounded open set, $\Omega \subset \mathbb{R}$ and $\Omega_T = \Omega \times (0, T)$. Let $1 \leq p \leq \infty$ and $k \in \mathbb{N}$. We define the following Sobolev space

$$W_p^{2k,k} (\Omega_T) \, = \, \left\{ u \, \in \, L^p \, (\Omega_T) \, | D^\alpha \partial_t^\beta u \, \in \, L^p \, (\Omega_T) \, , 1 \leq |\alpha| + 2\beta \leq 2k \right\} \tag{3.3.1}$$

where $D^\alpha \partial_t^\beta u$ is the weak partial derivative of u. These spaces are actually Banach spaces when assigning the following norms

$$\|u\|_{W_p^{2k,k}(\Omega_T)} = \sum_{0\leq|\alpha|+2\beta\leq 2k} \left\| D^\alpha \partial_t^\beta u \right\|_{L^p(\Omega)}. \tag{3.3.2}$$

One important consideration regarding the solution of equation (3.2.27) is that we are going to look specifically for convex solutions. This type of problems refer to any derivative whose payoff correspond to a convex function as it could be an European option call. Hence, the modified volatility defined in (3.2.25) is changed to

$$\hat{\sigma}^2 = \sigma^2 \left(1 - \sqrt{\frac{2}{\pi\, dt}}\, \frac{C_S}{\sigma}\right). \tag{3.3.3}$$

Also, we are going to apply the change of variables $x = \log(S)$ and $\tau = T - t$. We define the parabolic operator $\mathcal{L}$ and the nonlinear operator $\mathcal{N}$ as

$$\begin{aligned}
\mathcal{L}\, V \;&= -\frac{\partial V}{\partial \tau} + \frac{1}{2}\,\hat{\sigma}^2\, \frac{\partial^2 V}{\partial x^2} + \frac{\partial V}{\partial x}\, \left(q_S - \gamma_S - \frac{1}{2}\hat{\sigma}^2\right) - r\, V \\
\mathcal{N}\, V \;&= V^+ \left[s_F + \lambda_C\, (1 - R_C)\right] + V^-\, (\lambda_B - r\, C_B)\, (1 - R_B) + \sigma \sqrt{\frac{2}{\pi\, dt}}\, C_C\, (1 - R_C) \left| \frac{\partial V^+}{\partial x} \right|
\end{aligned}$$

such as the problem reads as

$$\begin{aligned}
\mathcal{L}\, \hat{V}(\tau, x) &= \mathcal{N}\, \hat{V}(\tau, x) \quad \text{in} \quad \Omega \times [0, T] \\
\hat{V}(0, x) &= g(x) \quad \text{in} \quad \Omega \\
\hat{V}(\tau, x) &= f(x) \quad \text{in} \quad \partial\Omega \times (0, T).
\end{aligned} \tag{3.3.4}$$

where $g(x)$ is the initial condition (i.e. the payoff of the derivative) and $f(x)$ is the boundary condition. For example, we define the conditions for an European call option as

$$g(x) = (\exp(x) - K)^+,$$

$$f(x) = \begin{cases} 0 & \text{if} \quad x \to 0 \\ \exp(x) & \text{if} \quad x \to \infty. \end{cases}$$

In order to find a solution of problem (3.3.4), we define an operator $T : C^{1,0}\left(\bar{\Omega}\right) \to C^{1,0}\left(\bar{\Omega}\right)$ such that $T(u) = v$, where $v \in W_p^{2,1}$ is the unique solution of the problem $\mathcal{L}v = \mathcal{N}u$. Our objective is to find a fixed point of the operator T which at the same time will be the solution of problem (3.3.4). We will set three conditions that the parameters of the model must fulfill to assess the existence of a convex solution.

The first condition is required to define a well-posed equation. As it is explained in Ref [33], the modified volatility shown in equation (3.3.3) must be positive. This can be addressed by setting a lower bound for the volatility parameter as it is seen below

$$\sigma > \sqrt{\frac{2}{\pi\, dt}}\, C_S \tag{3.3.5}$$

The second one is a sufficient condition which is required to find a fixed point of the operator T. We first recall c as a positive constant which depends only of the domain which will be defined below. Then, the following inequality must remain valid

$$c\,|\Omega|^{1/p} \left([s_F + \lambda_C\ (1 - R_C)] + 2\ (\lambda_B - r\, C_B)\ (1 - R_B) + \sigma \sqrt{\frac{2}{\pi\, dt}}\, C_C\ (1 - R_C) \right) < 1. \tag{3.3.6}$$

Given all the parameters of the model set, this assumption can be rewritten in terms of an upper bound for the volatility parameter

$$\sigma < \frac{1 - c\,|\Omega|^{1/p} \left([s_F + \lambda_C\ (1 - R_C)] + 2\ (\lambda_B - r\, C_B)\ (1 - R_B) \right)}{c\,|\Omega|^{1/p} \sqrt{\frac{2}{\pi\, dt}}\, C_C\ (1 - R_C)} \tag{3.3.7}$$

The third and last condition is required to prove that the solution found is indeed convex. For this reason, the stock growth rate under the risk neutral measure has to be bounded. This condition reads as

$$q_S - \gamma_S < M \tag{3.3.8}$$

with

$$M = \max \left(r_1 + \sigma\, S \sqrt{\frac{2}{\pi\, dt}} C_C\ (1 - R_C)\, , r_2 \right),$$

where $r_1 = r - [s_F + \lambda_C\ (1 - R_C)]$ and $r_2 = r - (\lambda_B - rC_B)\ (1 - R_B)$.

Hence, the main theorem of this Section reads as follows.

Theorem 3.3.1. *Suppose that assumptions* (3.3.5), (3.3.7) *and* (3.3.8) *are valid, that both the initial and boundary conditions belong to the* $W_p^{2,1}$ *space and the initial condition is a convex function. Then, the problem* (3.3.4) *admits at least one solution.*

3.3.2 Proof of Theorem 3.3.1

This Section covers all the steps required to prove Theorem 3.3.1. The main idea of the proof of Theorem 3.3.1 is to apply the Schauder fixed point theorem, presented previously in Section 1.2.3, on the operator T. Given K a nonempty convex subset of Ω we have to check that the operator T is a compact continuous mapping of K into itself such that $T(K) \subset K$.

To begin with the proof of Theorem 3.3.1, we first recall Theorem 7.32 from Ref [34] which not only assures that the operator T is well defined but also provides a lower estimate for $\mathcal{L}u$.

Theorem 3.3.2. *Suppose $\Omega \subset \mathbb{R}^{n+1}$ and let $p > 1$, $p(1-\alpha) < 1$ and L a parabolic operator with coefficients satisfying*

$$|b^i| \leq B \quad , \quad |c| \leq c_1$$
$$|a^{ij}(X) - a^{ij}(Y)| \leq w(|X - Y|).$$

Then, for any $\phi \in W_p^{2,1}$ and any $f \in L^p(\Omega)$, there is a unique solution of $Lu = f$ in Ω, $u = \phi$ in $\mathcal{P}\Omega$. Moreover, u satisfies the estimate

$$\|u\|_p + \|Du\|_p + \|D^2u\|_p + \|u_t\|_p \leq C\left(\|f\|_p + \|\phi\|_p + \|D\phi\|_p + \|D^2\phi\|_p + \|\phi_t\|_p\right) \tag{3.3.9}$$

In this Theorem the parabolic operator L is defined as

$$Lu = -u_t + a^{ij}D_{ij}u + b^iD_iu + cu.$$

Moreover, $\alpha \in (0,1)$ such that $\mathcal{P}\Omega \in H_{1+\alpha}$ where the parabolic boundary $\mathcal{P}\Omega$ is defined to be the set of all points $x_0 \in \partial\Omega$ such that for any $\epsilon > 0$, the cylinder $\mathcal{Q}(\S_\prime, \epsilon)$ contains points not in Ω. In the case that $\Omega = D \times (0,T)$ for some $D \subset \mathbb{R}^N$ and $T > 0$, the parabolic boundary is the union of the bottom of Ω, $B\Omega = D \times \{0\}$, the side of Ω, $S\Omega = \partial D \times (0,T)$ and the corner of Ω, $C\Omega = \partial D \times \{0\}$. Finally, w is defined as a positive, continuous and increasing function such that $w(0) = 0$.

By adapting Theorem (3.3.2) to our problem, we get the following lemma

Lemma 3.3.3. *Let $u \in C^{1,0}\left(\bar{\Omega}\right)$, $\mathcal{L}$ be as in Theorem 7.32 from Ref [34] and $f := \mathcal{N}u$. Then there exists a unique solution of problem $\mathcal{L}v = f$ in Ω, $v = g$ in $\{0\} \times \Omega$ and $v = f$ in $\partial\Omega \times (0,T)$. Moreover, there exists $C > 0$ independent of f such that v satisfies the estimate*

$$\|v\|_{W_p^{2,1}} \leq C\left(\|\mathcal{L}v\|_p + \|f\|_{W_p^{2,1}} + \|g\|_{W_p^{2,1}}\right). \tag{3.3.10}$$

Proof: The result is obtained by noting the definition of $\|v\|_{W_p^{2,1}}$ and applying Theorem 3.3.2 with $f = \mathcal{L}v$. $\square$

This lemma gives us an a priori estimate that let us derive the existence of a fixed point of T. The following lemma will be useful to address the continuity of the operator T.

Lemma 3.3.4. *Let $p > N$ and $u_n \in W_p^{2,1}$ a bounded sequence such that $u_n \to u$ in $W_p^{1,0}$. Given $G(x) = \max(x, 0)$ it follows that*

$$\begin{aligned} G(u_n) &\to G(u) \quad in \quad L^p \\ \frac{\partial G(u_n)}{\partial x} &\to \frac{\partial G(u)}{\partial x} \quad in \quad L^p \end{aligned} \tag{3.3.11}$$

Proof: The proof of the first statement follows from noting that $|G'(x)| \leq 1$ so then $\|G(u_n) - G(u)\|_p \leq \|u_n - u\|_p \to 0$. To address the second statement, we first note that

$$\frac{\partial}{\partial x}(G \circ u) = (G' \circ u)\frac{\partial u}{\partial x}.$$

Then, we can rewrite

$$(G' \circ u)\frac{\partial u}{\partial x} - (G' \circ u_n)\frac{\partial u_n}{\partial x} = (G' \circ u - G' \circ u_n)\frac{\partial u}{\partial x} + (G' \circ u_n)\left(\frac{\partial u}{\partial x} - \frac{\partial u_n}{\partial x}\right)$$

For the first term, since $|G'(u) - G'(u_n)| \leq 1$ it follows that

$$\left|(G'(u) - G'(u_n))\frac{\partial u}{\partial x}\right|^p \leq \left|\frac{\partial u}{\partial x}\right|^p$$

Then, by dominated convergence theorem

$$\left\|(G'(u) - G'(u_n))\frac{\partial u}{\partial x}\right\|_p^p = \int \left|(G'(u) - G'(u_n))\frac{\partial u}{\partial x}\right|^p \to 0.$$

To assess the second term, we consider the inclusion $W_p^{2,1} \hookrightarrow W_p^{1,0}$ to obtain a convergent subsequence in $W_p^{1,0}$. Nonetheless, given that $|G'(u_n)| \leq 1$, it follows that

$$\left|G'(u_n)\left(\frac{\partial u}{\partial x} - \frac{\partial u_n}{\partial x}\right)\right|^p \leq \left|\frac{\partial u}{\partial x} - \frac{\partial u_n}{\partial x}\right|^p \to 0.$$

and hence

$$\left\| G'(u_n)\left(\frac{\partial u}{\partial x} - \frac{\partial u_n}{\partial x}\right)\right\|^p \leq \left\| \frac{\partial u}{\partial x} - \frac{\partial u_n}{\partial x}\right\|^p \to 0$$

□

Given Lemma 3.3.3 and Lemma 3.3.4, we can now address the proof of Theorem 3.3.1. The first step is to define the constant c of Equation (3.3.7). The constant of imbedding $W_p^{2,1} \hookrightarrow C^{1,0}$ will be set as c_1 and the constant obtained from Lemma (3.3.3) with $p > 1$ will be set as c_2 and we will define $c = c_1 c_2$. The second step is to apply Schauder Fixed Point theorem on the operator T. We define our invariant subset as $K = \overline{B_R}(u_0)$. To apply Schauder theorem, we have to prove that T is a compact continuous mapping such that $T\left(\overline{B_R}(u_0)\right) \subset \overline{B_R}(u_0)$. Within our proof we will set $u_0 = 0$

Let first see the continuity of the operator T. Let $u_n, u \in W_p^{2,1}$ such that $u_n \to u$ in $W_p^{2,1}$. By recalling Lemma 3.3.3, there exists a constant $C > 0$ such that

$$\begin{aligned} \|Tu_n - T_u\|_{W_p^{2,1}} &\leq C\ \|\mathcal{L}\,Tu_n - \mathcal{L}\,Tu\|_p \\ &\leq C\ \|\mathcal{N}\,u_n - \mathcal{N}\,u\|_p \\ &\leq C\ [s_F + \lambda_C\ (1 - R_C)]\ \left\|u_n^+ - u^+\right\|_p + C\ (\lambda_B - r\,C_B)\ (1 - R_B)\ \left\|u_n^- - u^-\right\|_p \\ &+ C\,\sigma\sqrt{\frac{2}{\pi\,dt}}\,C_C\ (1 - R_C)\ \left\|\frac{\partial u_n}{\partial x}^+ - \frac{\partial u}{\partial x}^+\right\|_p \end{aligned} \tag{3.3.12}$$

By applying Lemma 3.3.4, we know that $\|u_n^+ - u^+\|_p \to 0$ and $\left\|\frac{\partial u_n}{\partial x}^+ - \frac{\partial u}{\partial x}^+\right\|_p \to 0$. The same lemma is valid by changing the function G into $G(x) = \min(x, 0)$ so that $\|u_n^- - u^-\|_p \to 0$.

The next step is to address the compactness of the operator T in the whole space. Let S a bounded subset of $C^{1,0}$. By the definition of T, $T(S) \subset W_p^{2,1}$ and, as $p > 1$, the compact inclusion $W_p^{2,1} \hookrightarrow C^{1,0}$ guarantees that $T(S) \subset C^{1,0}$. Hence, it suffices to prove that $T(S)$ is bounded for the $W_p^{2,1}$ norm. Let $v \in S$ and let us use Lemma 3.3.3 to see that

$$\begin{aligned} \|Tv\|_{W_p^{2,1}} = \|u\|_{W_p^{2,1}} &\leq C\ \left(\|\mathcal{L}u\|_p + \|f\|_{W_p^{2,1}} + \|g\|_{W_p^{2,1}}\right). \\ &\leq C\ \left(\|\mathcal{N}v\|_p + \|f\|_{W_p^{2,1}} + \|g\|_{W_p^{2,1}}\right). \end{aligned} \tag{3.3.13}$$

Hence, the result follows from the bounds of both initial and boundary conditions and the bound of v.

Further, let R be a positive number such that

$$R > \frac{\|f\|_{W_p^{2,1}} + \|g\|_{W_p^{2,1}}}{1 - c\,|\Omega|^{1/p}\left([s_F\; + \lambda_C\;(1 - R_C)]\; +\; 2\;(\lambda_B - r\,C_B)\;(1 - R_B)\; +\; \sigma\,\sqrt{\frac{2}{\pi\,dt}}\,C_C\;(1 - R_C)\right)} \tag{3.3.14}$$

and u such that $\|u\|_{C^{1,0}} \leq R$. Then, there exists a constant $c_1 > 0$ given by the embedding $W_p^{2,1} \hookrightarrow C^{1,0}$ so that

$$\|Tu\|_{C^{1,0}} \leq c_1\,\|Tu\|_{W_p^{2,1}}\,. \tag{3.3.15}$$

Given the inequality presented in equation (3.3.15), we can use the result of Lemma 3.3.3. Hence, there exists a constant $c_2 > 0$ such that

$$\|Tu\|_{C^{1,0}} \leq c_1\,c_2\,\left(\|\mathcal{N}u\|_p + \|f\|_{W_p^{2,1}} + \|g\|_{W_p^{2,1}}\right). \tag{3.3.16}$$

By recalling (3.3.4), the nonlinear term $\mathcal{N}$ is bounded by

$$\begin{aligned}
\|\mathcal{N}u\|_p \;\leq\; & [s_F\; + \lambda_C\;(1 - R_C)]\;\left\|u^+\right\|_p + (\lambda_B - r\,C_B)\;(1 - R_B)\;\left\|u^-\right\|_p \\
& + \sigma\,\sqrt{\frac{2}{\pi\,dt}}\,C_C\;(1 - R_C)\;\left\|\frac{\partial u^+}{\partial x}\right\|_p \\
& < [s_F\; + \lambda_C\;(1 - R_C)]\;|\Omega|^{1/p}\,R\; +\; (\lambda_B - r\,C_B)\;(1 - R_B)\;|\Omega|^{1/p}\,2R \\
& + \sigma\,\sqrt{\frac{2}{\pi\,dt}}\,C_C\;(1 - R_C)\;|\Omega|^{1/p}\,R \\
& < k_1\,R
\end{aligned} \tag{3.3.17}$$

using that $\|u\|_{C^{1,0}} \leq R$ where

$$k_1 = |\Omega|^{1/p}\left([s_F\; + \lambda_C\;(1 - R_C)]\; +\; (\lambda_B - r\,C_B)\;(1 - R_B)\;2\; +\; \sigma\,\sqrt{\frac{2}{\pi\,dt}}\,C_C\;(1 - R_C)\right).$$

By applying (3.3.17) in (3.3.16), we get that

$$\begin{aligned}
\|Tu\|_{C^{1,0}} & \leq c_1\,c_2\,k_1\,R + c_1\,c_2\,\left(\|f\|_{W_p^{2,1}} + \|g\|_{W_p^{2,1}}\right). \\
& <\; R
\end{aligned} \tag{3.3.18}$$

which follows from the assumption (3.3.7) by setting $c = c_1\, c_2$ and the lower bound of R.

The last step of the proof is to show that the solution is indeed convex. In order to confirm this assumption, we can analyze the similarity between equation (3.2.27) and a Black-Scholes equation with dividends and notice that given a convex initial condition, the solution would remain convex. We analyze separately when $\hat{V}$ is positive or negative. When the solution is positive, equation (3.2.27) reduces to

$$\frac{\partial \hat{V}}{\partial t} + \frac{1}{2}\,\hat{\sigma}^2\,S^2\,\frac{\partial^2 \hat{V}}{\partial S^2} + \frac{\partial \hat{V}}{\partial S}\,S\,(q_S - \gamma_S) - r\,\hat{V} = [s_F + \lambda_C\,(1 - R_C)]\,\hat{V}\, + \sigma\,S\,\sqrt{\frac{2}{\pi\,dt}}C_C\,(1 - R_C)\,\left|\frac{\partial \hat{V}}{\partial S}\right|.$$

By rearranging terms and defining $r_1 = r - [s_F + \lambda_C\,(1 - R_C)]$ the equation above becomes

$$\frac{\partial \hat{V}}{\partial t} + \frac{1}{2}\,\hat{\sigma}^2\,S^2\,\frac{\partial^2 \hat{V}}{\partial S^2} + \frac{\partial \hat{V}}{\partial S}\,S\,\left(q_S - \gamma_S - sgn\left(\frac{\partial \hat{V}}{\partial S}\right)\sigma\,S\,\sqrt{\frac{2}{\pi\,dt}}C_C\,(1 - R_C)\right) - r_1\,\hat{V} = 0 \tag{3.3.19}$$

When the solution is negative, equation (3.2.27) reduces to

$$\frac{\partial \hat{V}}{\partial t} + \frac{1}{2}\,\hat{\sigma}^2\,S^2\,\frac{\partial^2 \hat{V}}{\partial S^2} + \frac{\partial \hat{V}}{\partial S}\,S\,(q_S - \gamma_S) - r\,\hat{V} = (\lambda_B - rC_B)\,(1 - R_B)\,\hat{V}.$$

By rearranging terms and defining $r_2 = r - (\lambda_B - rC_B)\,(1 - R_B)$ the equation above becomes

$$\frac{\partial \hat{V}}{\partial t} + \frac{1}{2}\,\hat{\sigma}^2\,S^2\,\frac{\partial^2 \hat{V}}{\partial S^2} + \frac{\partial \hat{V}}{\partial S}\,S\,(q_S - \gamma_S) - r_2\,\hat{V} = 0 \tag{3.3.20}$$

Equation (3.3.19) and (3.3.20) can be thought as a Black-Scholes equation with dividend yield γ_S and free-risk interest rate r_1 and r_2 respectively. Moreover, the condition stated in (3.3.8) can be used to derive an upper bound for $q_S - \gamma_S$. If $q_S - \gamma_S < M$, we see that the growth rate of the stock under the risk-free measure is lower than the free-risk interest rate. This dynamic is the one expected for a Black-Scholes model with dividends. equation (3.2.27). As the initial condition of the problem is indeed convex the solution $\hat{V}$ is also convex.

3.4 Numerical

3.4.1 Numerical framework

In this Section we are going to develop a numerical framework to solve the problem defined in (3.3.4) by applying a forward Euler method. Hence, we recall the nonlinear problem

$$\begin{aligned} \mathcal{L}\,\hat{V}\left(\tau,x\right) &= \mathcal{N}\,\hat{V}\left(\tau,x\right) \quad \text{in} \quad \Omega\times[0,T] \\ \hat{V}\left(0,x\right) &= g\left(x\right) \quad \text{in} \quad \Omega \\ \hat{V}\left(\tau,x\right) &= f\left(x\right) \quad \text{in} \quad \partial\Omega\times(0,T)\,. \end{aligned} \tag{3.4.1}$$

with $\mathcal{L}$ and $\mathcal{N}$ defined in (3.3.4). For numerical convenience, we approximate the original smooth domain by a discrete one $\hat{\Omega}_T \subset [a,b]\times[0,T]$, setting a and b in order to cover a set of feasible logarithmic stock prices. The step of the temporal variable is uniformly set as $\Delta\tau = T/T_x$ being T_x the number of grid points in the τ- direction. For the spatial variable, we are going to apply a non-uniform grid where the spacing is fine near the strike and coarse away from the strike. In Ref [43], the following grid is proposed

$$x_i = x^* + \alpha\,\sinh\left(c_2\frac{i}{N} + c_1\left(1-\frac{i}{N}\right)\right) \tag{3.4.2}$$

where

$$\begin{aligned} c_1 &= \sinh^{-1}\left(\frac{x^- - x^*}{\alpha}\right) \\ c_2 &= \sinh^{-1}\left(\frac{x^+ - x^*}{\alpha}\right). \end{aligned}$$

This is a transformation that maps the interval $[0,1]$ into $[x^-,x^+]$ by concentrating the points near x^*. The value of α sets how non-uniform the grid will be and N to be the amount of points within the grid. In our problem we set $x^*=K$ and $[x^-,x^+]$ accordingly to cover all the possible logarithmic prices. Hence, we define the solution to the m-temporal step as $\hat{V}_i^m = \hat{V}\left(x_i, m\Delta\tau\right)$ where $1\le i\le N$ and $1\le m\le T_x$. We also define $\hat{U}=\max\left(\hat{V},0\right)$ for numerical notation convenience.

To derive the expression of the numerical framework we follow Ref [9] and Ref [21] in which this grid had been applied. By following the same steps, we obtain that the discretization of the first and second spatial derivatives are given by

$$\begin{aligned} \frac{\partial\hat{V}}{\partial x} &= \frac{\hat{V}_{i+1}^m - \hat{V}_i^m}{x_{i+1}-x_i}, \\ \frac{\partial^2\hat{V}}{\partial x^2} &= h_i^+\,\frac{\hat{V}_{i+1}^m - \hat{V}_i^m}{x_{i+1}-x_i} - h_i^-\,\frac{\hat{V}_i^m - \hat{V}_{i-1}^m}{x_i - x_{i-1}} \end{aligned}$$

where $h_i = x_i - x_{i-1}$ and

$$h_i^+ = \frac{2}{h_{i+1}\,(h_{i+1}+h_i)},$$
$$h_i^- = \frac{2}{h_i\,(h_{i+1}+h_i)}.$$

Given that the temporal step is set uniformly, the finite difference framework is defined as below

$$\begin{aligned}\mathcal{L}\,\hat{V} &= -\left(\frac{\hat{V}_i^{m+1}-\hat{V}_i^m}{\Delta\tau}\right) + \frac{1}{2}\,\hat{\sigma}^2\left[h_i^+\left(\hat{V}_{i+1}^m-\hat{V}_i^m\right) - h_i^-\left(\hat{V}_i^m-\hat{V}_{i-1}^m\right)\right] + \frac{\hat{V}_{i+1}^m-\hat{V}_i^m}{x_{i+1}-x_i} \\ &\quad \left(q_S-\gamma_S-\frac{1}{2}\,\hat{\sigma}^2\right) - r\,\hat{V}_i^m. \qquad (3.4.3)\\ \mathcal{N}\,\hat{V} &= \max\left(\hat{V}_i^m,0\right)\left[s_F+\lambda_C\,(1-R_C)\right] + \min\left(\hat{V}_i^m,0\right)\,(\lambda_B-r\,C_B)\,(1-R_B) \\ &\quad + \sigma\sqrt{\frac{2}{\pi\,dt}}\,C_C\,(1-R_C)\left|\frac{\hat{U}_{i+1}^m-\hat{U}_i^m}{x_{i+1}-x_i}\right|. \qquad (3.4.4)\end{aligned}$$

By rearranging and combining terms we obtain the following iterative process

$$\begin{aligned}\hat{V}_i^{m+1} &= \hat{V}_i^m\left(1-\frac{\hat{\sigma}^2\,\Delta\tau}{2}\left(h_i^+ + h_i^-\right) - \frac{\Delta\tau}{x_{i+1}-x_i}\left(q_S-\gamma_S-\frac{\hat{\sigma}^2}{2}\right) - r\,\Delta\tau\right) + \hat{V}_{i-1}^m\left(\frac{\hat{\sigma}^2\,\Delta\tau}{2}h_i^-\right) \\ &\quad + \hat{V}_{i+1}^m\left(\frac{\hat{\sigma}^2\,\Delta\tau}{2}h_i^+ + \frac{\Delta\tau}{x_{i+1}-x_i}\left(q_S-\gamma_S-\frac{\hat{\sigma}^2}{2}\right)\right) - \Delta\tau\,\max\left(\hat{V}_i^m,0\right)\left[s_F+\lambda_C\,(1-R_C)\right] \\ &\quad - \Delta\tau\,\min\left(\hat{V}_i^m,0\right)\,(\lambda_B-r\,C_B)\,(1-R_B) - \Delta\tau\,\sigma\sqrt{\frac{2}{\pi\,dt}}\,C_C\,(1-R_C)\left|\frac{\hat{U}_{i+1}^m-\hat{U}_i^m}{x_{i+1}-x_i}\right|. \qquad (3.4.5)\end{aligned}$$

If we let $\hat{V}_i^0 = g\,(x_i)$, this framework can be used to find the solution of the problem (3.4.1) at each time step m.

3.4.2 Numerical results

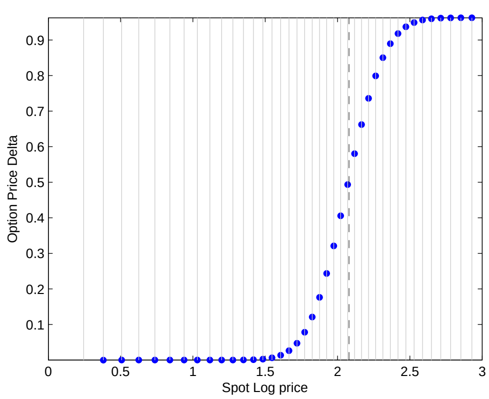
In this Section we analyze the behavior of the option price for an European call under different scenarios. We perform a sensitivity analysis on the volatility, free-risk interest rate, transaction costs, recovery rates and hazard rates by stressing its values. Nonetheless, we compare our results with the ones obtained by the original model proposed in Ref [13] and calculate how the transaction costs impact the final CVA value. To further analyze the behavior of the option price, we calculate its derivatives with respect to certain parameters. This derivatives are known as Greeks and consist of Delta (derivative with respect to the option price), Gamma (second derivative with respect to the option price), Vega (derivative with respect to the volatility) and Rho (derivative with respect to the interest rate).

For notation purposes we recall BK to the original model and BK_{TC} the model with transaction costs. Also, for each scenario, the parameters set for both models are defined in the caption of each figure and results are obtained at time $\tau = T$. Within each figure, two types of vertical lines are included. The grey-shaded lines correspond to the non-uniform x_i grid defined in Section 3.4.1 and the black dashed-line represents the strike value.

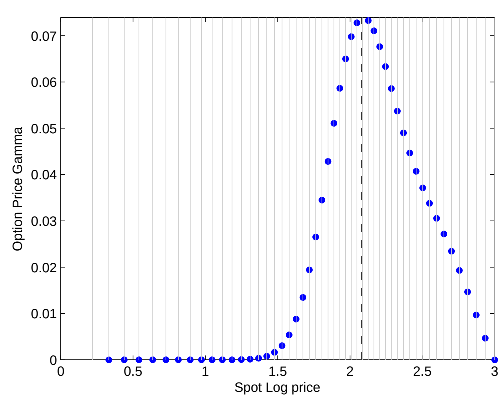
Delta and Gamma

Figure (3.1) presents the two derivatives with respect to the option price, which are Delta and Gamma. Delta shows a similar behavior to an European call. For that vanilla option, Delta's formula correspond to a normal cumulative function. When including CVA and transaction costs, it can be seen that when the option is deep out-of-the-money, Delta is near zero which implies that the portfolio defined in Equation (3.2.1) needs no shares of S to hedge the option. As the option gets at-the-money, Delta grows approximately up to 0.5. The option is more sensitive to changes in the spot price so then almost 50% of the hedging portfolio has to be covered with shares of S. This trend continues to converge to a Delta equal to 1 when the option gets deeper in-the-money. At this point, the option price changes at the same rate with respect to the spot price and the hedging portfolio has to be only long on shares of S to cover its hedging purpose.

As Gamma represents the second derivative of the option price with respect to the spot price, its maximum is actually reached when the option is at-the-money and diminishes when the option go either in-the-money or out-of-the-money. This behavior is again similar to the one seen on a vanilla European call and shows to us how sensitive is Delta to movements in the spot price.



(a) Option Delta

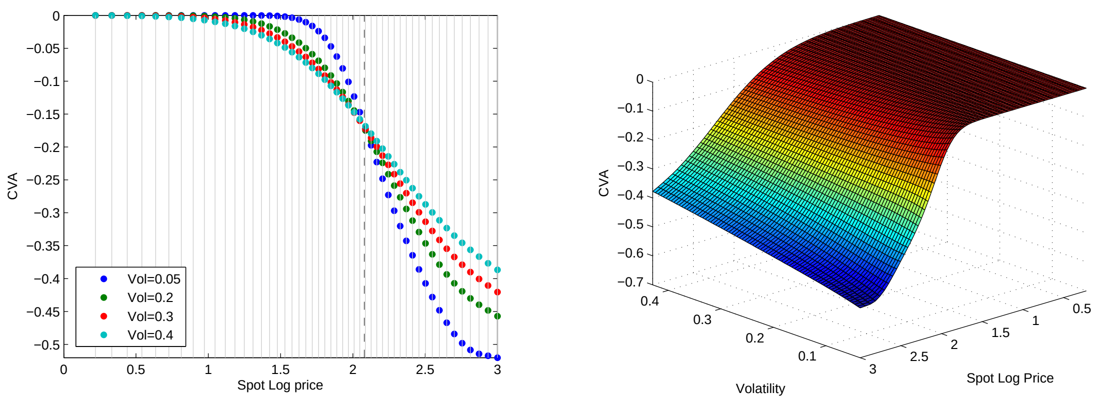


(b) Option Gamma

Figure 3.1: $C_S = 0.002$, $r = 0.05$, $q_S = 0.05$, $\gamma_S = 0.03$, $\sigma = 0.1$, $S_f = 0$, $\lambda_B = 0.05$, $\lambda_C = 0.01$, $R_B = 0.4$, $R_C = 0.4$, $C_B = 0.001$, $C_C = 0.001$, $dt = 1/261$, $\Delta\tau = 1/261$, $K = 8$.

Volatility

Figure (3.2a) presents the sensitivity of the CVA to changes in the volatility parameter. The figure shows that the strike price serves as threshold where the behavior of the CVA changes. When the option is out-of-the money ($S < K$), higher volatility produces higher CVA (more negative). However, when the option is in-the-money ($S > K$), the convexity changes leading to higher CVA as the volatility decreases. Figure (3.2b) expands these results over the entire set of possible volatilities.



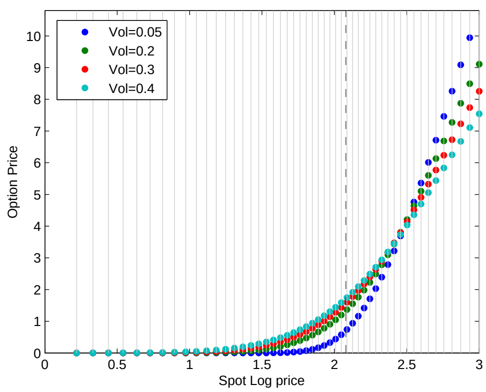
(a) BK_{TC} CVA for different volatilities (b) BK_{TC} CVA by Spot Price and Volatility

Figure 3.2: $C_S = 0.002$, $q_S = 0.05$, $\gamma_S = 0.03$, $r = 0.05$, $S_f = 0$, $\lambda_C = 0.01$, $R_C = 0.4$, $\lambda_B = 0.05$, $R_B = 0.4$, $C_B = 0.001$, $C_C = 0.001$, $dt = 1/261$, $\Delta\tau = 1/261$, $K = 8$.

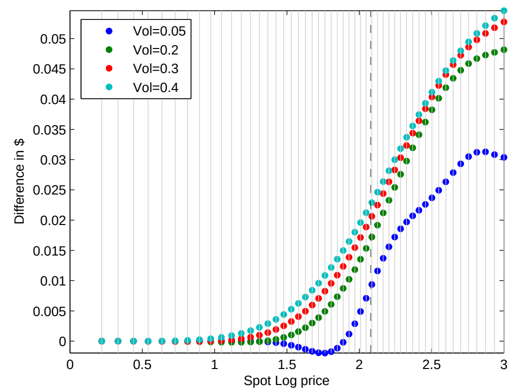
In Figure (3.3a), the sensitivity of the option price with respect to different volatility parameters is presented. Under the usual Black-Scholes framework, it is expected to get higher option prices as volatility increases. This pattern is confirmed up to a certain spot price. Under our framework, as the option gets deeper in-the-money, Delta (Figure (3.1a)), which represents the amount of shares to buy in the replicant strategy, tends to one and the impact of the transaction costs increase by generating a decrease in the option price. This behavior can be confirmed by assessing the first derivative of the option price with respect to the volatility (usually known as Vega). In Figure (3.4), Vega is split with respect to the moneyness of the option. Figure (3.4a) shows that, when the option is out-of-the-money, Vega is positive as it is under the Black-Scholes model. Further, Figure (3.4b) demonstrate that not only Vega becomes negative as the option gets in-the-money but also that its sign changes in the same spot price as seen in Figure (3.3a). Hence, if we consider the impact of the volatility not only in the parabolic side of the PDE but also in the nonlinear term , it is expected to find these relationship between the volatility parameter and the option price.

Interest Rate

Figure (3.5a) presents the sensitivity of the CVA to changes in the interest rate. Figure (3.5b) expand this results to the entire interval of Spot Log prices. Both figures show that

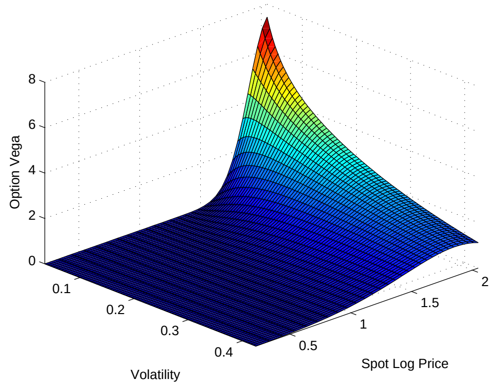


(a) BK_{TC} Option Price for different volatilities

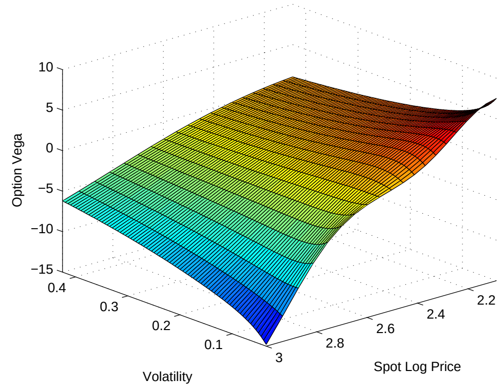


(b) Difference between BK and BK_{TC}

Figure 3.3: $C_S = 0.002$, $q_S = 0.05$, $\gamma_S = 0.03$, $r = 0.05$, $S_f = 0$, $\lambda_C = 0.01$, $R_C = 0.4$, $\lambda_B = 0.05$, $R_B = 0.4$, $C_B = 0.001$, $C_C = 0.001$, $dt = 1/261$, $\Delta\tau = 1/261$, $K = 8$.



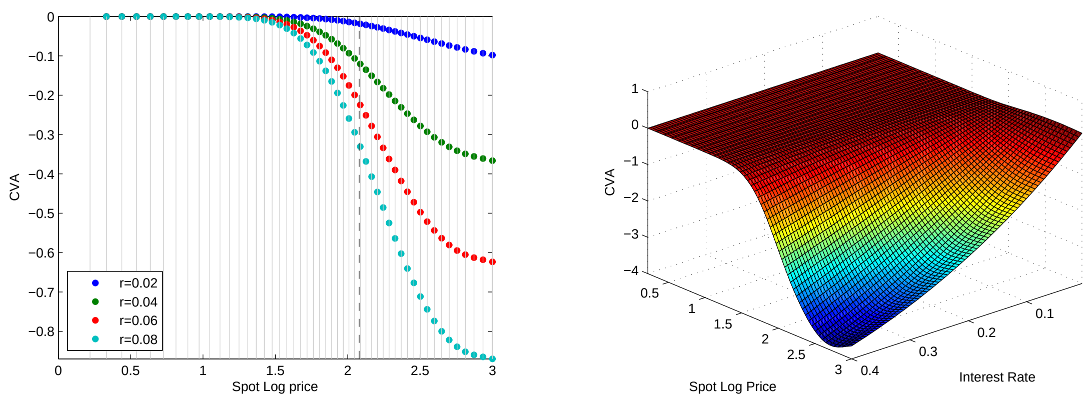
(a) Option Vega out-of-the-money



(b) Option Vega in-the-money

Figure 3.4: $C_S = 0.002$, $r = 0.05$, $q_S = 0.05$, $\gamma_S = 0.03$, $\sigma = 0.1$, $S_f = 0$, $\lambda_B = 0.05$, $\lambda_C = 0.01$, $R_B = 0.4$, $R_C = 0.4$, $C_B = 0.001$, $C_C = 0.001$, $dt = 1/261$, $\Delta\tau = 1/261$, $K = 8$.

the CVA decreases as the interest rate increases and the its size is larger when the option is deep in-the-money.



(a) BK_{TC} CVA for different interest rates. (b) BK_{TC} CVA by Spot price and interest rate.

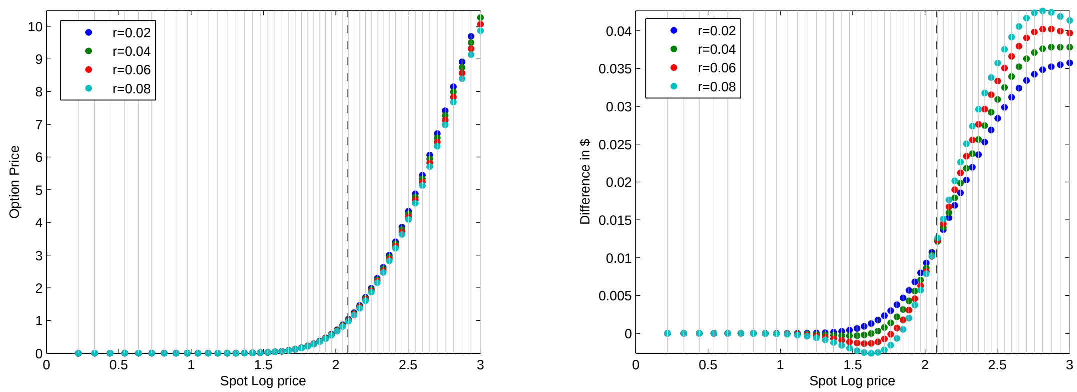
Figure 3.5: $C_S = 0.002$, $q_S = 0.05$, $\gamma_S = 0.03$, $\sigma = 0.1$, $S_f = 0$, $\lambda_C = 0.01$, $R_C = 0.4$, $\lambda_B = 0.05$, $R_B = 0.4$, $C_B = 0.001$, $C_C = 0.001$, $dt = 1/261$, $\Delta\tau = 1/261$, $K = 8$.

Figure (3.6a) shows that the option price as decreasing monotonic function with respect to the interest rate. This results is confirmed by analyzing the first derivative of the option price with respect to the interest rate presented in figure (3.7), also known as Rho. When the option is out-of-the-money, Rho is approximately zero. But as the option gets in-the-money, it is observed a negative slope. This result is counter-intuitive by considering that, under Black-Scholes model, the derivative is always positive. This discrepancy can be assessed by noting that, in equation (3.2.27), the terms that go with $\frac{\partial \hat{V}}{\partial S} S$ are equal to $(q_S - \gamma_S)$ instead of $(r - \gamma_S)$. Given that under the BK_{TC} model, q_S is being modeled as a constant function, the positive sensitivity of the option to the interest rate is not observed. In order to match the expected behavior, an improvement of the modeling approach of the financing cost and its relationship with the interest rate has to be done.

Transaction Costs

Figure (3.8) presents the variation on the CVA due to changes in the transaction costs that arise of trading δ amount of shares S and α_C amounts of bond P_C. By recalling equation (3.2.27) it can be noted that an increase in C_S leads to a decrease in the modified volatility. We actually can assume that the modified volatility behaves similarly to the actual volatility, so that the analysis done in Section 3.4.2 can be applied. By considering the pattern showed in Figure (3.8) it can be seen that it is in line with the behavior of the CVA when varying the volatility in Figure (3.8a). In both cases, the convexity changes near the strike value due to the same issues presented in the aforementioned Section.

On the other side, the presence of C_C in Equation (3.2.27) actually shows that larger costs generates a lower option price. Also, as transaction costs are multiplied by Delta, the gap



(a) BK_{TC} Option Price for different interest rates (b) Difference between BK and BK_{TC}

Figure 3.6: $C_S = 0.002$, $q_S = 0.05$, $\gamma_S = 0.03$, $\sigma = 0.1$, $S_f = 0$, $\lambda_C = 0.01$, $R_C = 0.4$, $\lambda_B = 0.05$, $R_B = 0.4$, $C_B = 0.001$, $C_C = 0.001$, $dt = 1/261$, $\Delta\tau = 1/261$, $K = 8$.

widens as the option gets deeper in-the-money. This is the pattern that is observed in Figure (3.8b).

Recovery Rate and Hazard Rate

In Figure (3.9) the sensitivity of the CVA due to changes in the recovery rate of the counterparty bond is presented. Given that the recovery rate determines the amount of instrument that can be recovered in case of a default, the term $(1 - R_C)$ estimates the loss that would arise in case of default. It is expected that higher recovery rates imply lesser losses and then lesser CVA. Figure (3.9a) shows this monotonic relationship which is also in line with the behavior seen in Equation (3.2.27).

The hazard rate of the counterparty bond measures the likelihood that the bond will default at a certain point of time. Hence, if the hazard rate increases, the probability of default of the bond also increases. Then, it is expected to see a higher CVA value when deriving the option price. Figure (3.10a) presents the CVA value for different hazard rates where the expected behavior is noticed.

3.5 Conclusion

Chapter 3 was dedicated to the development of a pricing model that considered at the same type the presence of transaction costs in the replication strategy and the probability of default of both the issuer and counterparty of the financial contract.

In Section 3.2 we presented all the steps required to develop the correspondent market model. We followed the seminal work of Ref [13] in which a replicant portfolio is created. We adapted the framework to consider the existence of constant transactions costs when buying or

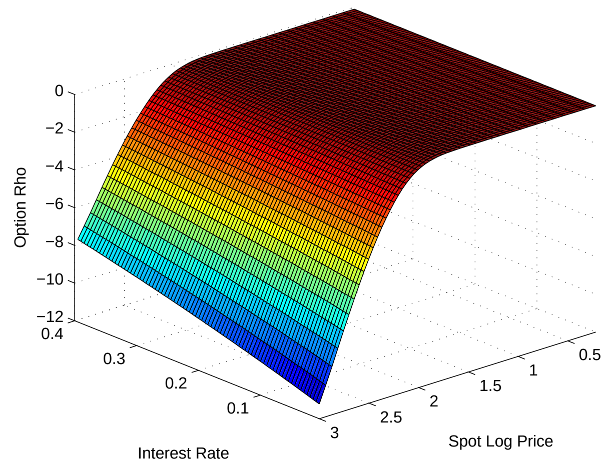
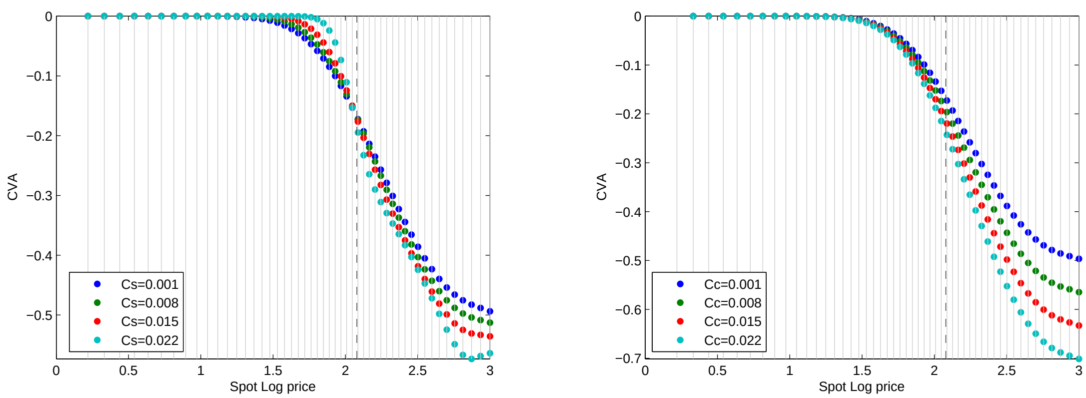


Figure 3.7: Option Rho

selling the underlying asset and both issuer and counterparty bonds. We obtained a parabolic PDE with nonlinear terms that arose from the presence of these costs.

In Section 3.3 we transformed the original equation and provide three sufficient conditions that the parameters must fulfilled in order to assure the existence of a convex solution. From a financial perspective, these conditions implied that the volatility parameter couldn't be either too little or too high and that the stock-rate growth under the risk-neutral measure had to be bounded. It is important to note that these conditions are in-line with a standard state of an asset so that the model's proof can be used under many different scenarios. To prove the existence of the solution, we proposed a fixed-point approach. Using the Schauder Fixed-Point theorem, we created an operator T and a defined a subset K such that T is a compact continuous mapping of K into itself with $T(K) \subset K$. Finally, we showed that the solution was in fact convex by noting that on each time step, the equation could be reduced to a Black-Scholes equation with dividend yield.

In Section 3.4 we applied an Euler scheme to find a solution of the original problem. We used a non-uniform grid where the spacing was fine near the strike and coarse away from the strike. We adapted the Euler method to consider this grid and defined the correspondent iterative method to obtain the desired solution. We priced an European call under different scenarios and analyzed the most important Greeks and different sensitivities of the price when changing all the different parameters of the equation. The numerical results showed that Delta and Gamma behave similarly as in a plain vanilla option but Vega and Rho presented differences in terms of the usual behavior. We analyzed that the presence of transaction costs make an impact in the way volatility and risk-free interest rate affects the option price. We also realized that the spot financing cost q_S has to be linked with the risk-free interest rate to assure consistent results.



(a) BK_{TC} (b) BK_{TC}

Figure 3.8: $r = 0.05$, $q_S = 0.05$, $\gamma_S = 0.03$, $\sigma = 0.1$, $S_f = 0$, $\lambda_C = 0.01$, $R_C = 0.4$, $\lambda_B = 0.05$, $R_B = 0.4$, $C_B = 0.001$, $dt = 1/261$, $\Delta\tau = 1/261$, $K = 8$. $C_C = 0.001$ for (3.8a) and $C_S = 0.002$ for (3.8b).

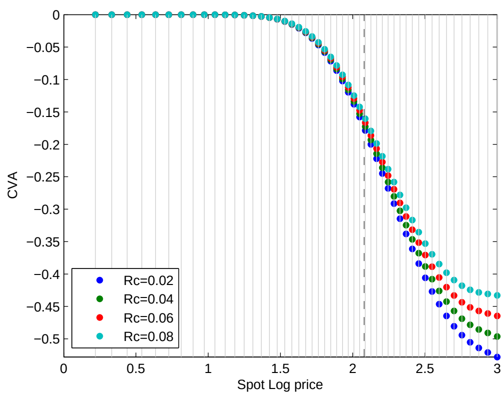
3.6 Resumen del capítulo

El Capítulo 3 presenta el desarrollo del modelo de valuación de opciones financieras que considera al mismo tiempo la existencia de costos de transacción en la estrategia de replicación como así tambien la probabilidad de default de tanto el emisor del contrato como de su contraparte.

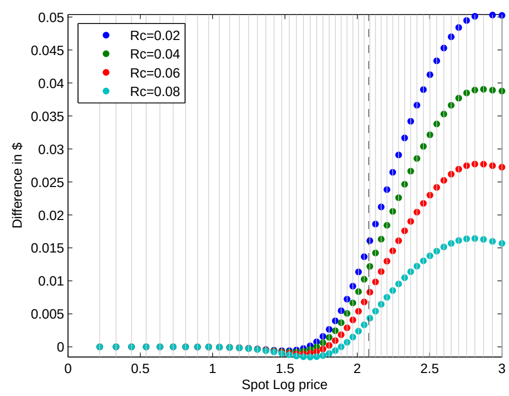
En la Sección 3.2 desarrollamos todos los pasos necesarios para generar el modelo de valuación siguiendo el reconocido trabajo de Ref [13]. Adaptamos su esquema de valuación para permitir la presencia de costos de transacción constantes al comprar o vender el activo subyacente, el bono del emisor o el de su contraparte en la estrategia de replicación. Así, obtuvimos una ecuación diferencial parabólica con terminos no lineales producto de la inclusión de los mencionados costos de transacción.

En la Sección 3.3 transformamos la ecuación original mediante cambios de variables y determinamos tres condiciones suficientes para los parámetros del modelo para asi poder asegurar la existencia de una solución convexa. Desde una perspectiva financiera, estas condiciones implican que el parámetro de volatilidad no puede ser ni muy chico ni muy grande y que la tasa de crecimiento bajo la medida libre de riesgo de el activo subyacente debe estar acotada. Es importante notar que las tres condiciones corresponden a estados naturales del activo subyacente para lo cual estas condiciones resultan ser válidas en distintos escenarios. Para probar la existencia de solución, propusimos un enfoque de punto fijo utilizando el teorema de punto fijo de Schauder. Finalmente, mostramos que la solución es efectivamente convexa notando que en cada paso temporal la ecuación puede ser reducida a una ecuación de Black-Scholes con tasa de dividendos.

En la Sección 3.4 desarrollamos un esquema numérico de tipo Euler para encontrar una solución aproximada del problema original. Usamos una malla no uniforme donde el espaciado

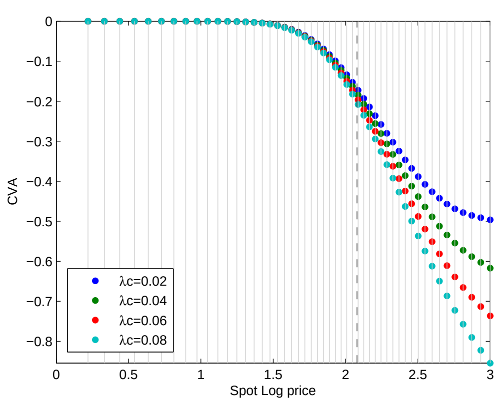


(a) BK_{TC}

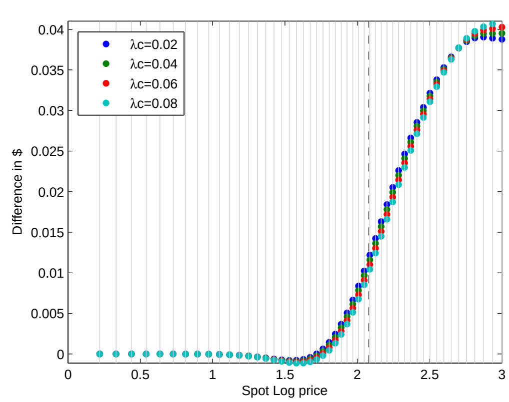


(b) Difference between BK and BK_{TC}

Figure 3.9: $C_S = 0.002$, $r = 0.05$, $q_S = 0.05$, $\gamma_S = 0.03$, $\sigma = 0.1$, $S_f = 0$, $\lambda_C = 0.01$, $\lambda_B = 0.05$, $R_B = 0.4$, $C_B = 0.001$, $C_C = 0.001$, $dt = 1/261$, $\Delta\tau = 1/261$, $K = 8$



(a) BK_{TC}



(b) Difference between BK and BK_{TC}

Figure 3.10: $C_S = 0.002$, $r = 0.05$, $q_S = 0.05$, $\gamma_S = 0.03$, $\sigma = 0.1$, $S_f = 0$, $\lambda_B = 0.05$, $R_B = 0.4$, $R_C = 0.4$, $C_B = 0.001$, $C_C = 0.001$, $dt = 1/261$, $\Delta\tau = 1/261$, $K = 8$.

es mas fino cerca del strike y mas grueso fuera del mismo. Adaptamos el método de Euler para considerar esta malla y definimos el correspondiente método iterativo para encontrar la solución deseada. Dado el esquema desarrollado, valuamos una opcion de tipo call europea bajo diversos escenarios. Los resultados numéricos mostraron que tanto Delta como Gamma se comportan similarmente a la opción financiera del modelo estandar. Sin embargo, la diferencia se observa al analizar Vega and Rho notando que la presencia de costos de transacción impacta en la forma en la que las volatilidades y la tasa de interés libre de riesgo afectan el precio de la opción. Además notamos que el parámetro de costo de financiación tiene que estar relacionado con la tasa de interés libre de riesgo para obtener resultados consistentes.

Bibliography

[1] Robert A Adams and John JF Fournier, *Sobolev Spaces*, vol. 140, Academic Press, 2003.

[2] Shahram Alavian, Jie Ding, Peter Whitehead, and Leonardo Laudicina, *Counterparty Valuation Adjustment (CVA)*, Available at SSRN: https://ssrn.com/abstract=1310226 or http://dx.doi.org/10.2139/ssrn.1310226 (2008).

[3] William F Ames, *Numerical methods for partial differential equations*, Academic Press, 2014.

[4] P Amster, CG Averbuj, MC Mariani, and D Rial, *A Black–Scholes option pricing model with transaction costs*, Journal of Mathematical Analysis and Applications **303** (2005), no. 2, 688–695.

[5] Pablo Amster, *Topological methods in the study of boundary value problems*, Springer, 2014.

[6] Pablo Amster and Andres P Mogni, *Adapting the CVA model to Leland's framework*, arXiv preprint arXiv:1802.04837 (2018).

[7] ______, *On a pricing problem for a multi-asset option with general transaction costs*, arXiv preprint arXiv:1704.02036v2 (2018).

[8] Martin Baxter and Andrew Rennie, *Financial calculus: an introduction to derivative pricing*, Cambridge University Press, 1996.

[9] Jérôme Bodeau, Gaël Riboulet, and Thierry Roncalli, *Non-uniform grids for PDE in finance*, (2000).

[10] Damiano Brigo and Agostino Capponi, *Bilateral counterparty risk valuation with stochastic dynamical models and application to credit default swaps*, arXiv preprint arXiv:0812.3705 (2008).

[11] Damiano Brigo and Kyriakos Chourdakis, *Counterparty risk for credit default swaps: Impact of spread volatility and default correlation*, International Journal of Theoretical and Applied Finance **12** (2009), no. 07, 1007–1026.

[12] Damiano Brigo, Andrea Pallavicini, and Vasileios Papatheodorou, *Arbitrage-free valuation of bilateral counterparty risk for interest-rate products: impact of volatilities and correlations*, International Journal of Theoretical and Applied Finance **14** (2011), no. 06, 773–802.

[13] Christoph Burgard and Mats Kjaer, *Partial differential equation representations of derivatives with bilateral counterparty risk and funding costs*, The Journal of Credit Risk **7** (2011), no. 3, 75.

[14] Stéphane Crépey, *A BSDE approach to counterparty risk under funding constraints*, Working paper (2011).

[15] ———, *Bilateral counterparty risk under funding constraintsPart 1: Pricing*, Mathematical Finance **25** (2015), no. 1, 1–22.

[16] ———, *Bilateral counterparty risk under funding constraintsPart 2: CVA*, Mathematical Finance **25** (2015), no. 1, 23–50.

[17] Jim Douglas and Henry H Rachford, *On the numerical solution of heat conduction problems in two and three space variables*, Transactions of the American Mathematical Society **82** (1956), no. 2, 421–439.

[18] L.C. Evans, *Partial Differential Equations*, Graduate Studies in Mathematics, American Mathematical Society, 1998.

[19] Gabriele Fiorentini, Angel Leon, and Gonzalo Rubio, *Estimation and empirical performance of heston's stochastic volatility model: the case of a thinly traded market*, Journal of Empirical Finance **9** (2002), no. 2, 225–255.

[20] Ionut Florescu, Maria C Mariani, and Indranil Sengupta, *Option pricing with transaction costs and stochastic volatility*, Electronic Journal of Differential Equations **2014** (2014), no. 165, 1–19.

[21] S Foulon et al., *ADI finite difference schemes for option pricing in the Heston model with correlation*, International Journal of Numerical Analysis & Modeling **7** (2010), no. 2, 303–320.

[22] Andrew Green, *XVA: Credit, Funding and Capital Valuation Adjustments*, John Wiley & Sons, 2015.

[23] Jon Gregory, *Being two-faced over counterparty credit risk*, Risk **22** (2009), no. 2, 86.

[24] ———, *The XVA challenge: Counterparty credit risk, funding, collateral, and capital*, John Wiley & Sons, 2015.

[25] Espen Gaarder Haug, *The complete guide to option pricing formulas*, McGraw-Hill Companies, 2007.

[26] Steven L Heston, *A closed-form solution for options with stochastic volatility with applications to bond and currency options*, The Review of Financial Studies **6** (1993), no. 2, 327–343.

[27] J.C. Hull, *Options, futures, and other derivatives*, Pearson Education, 2017.

[28] Cyril Imbert and Luis Silvestre, *An introduction to fully nonlinear parabolic equations*, An introduction to the Kähler-Ricci flow, Springer, 2013, pp. 7–88.

[29] KJ In't Hout and S Foulon, *ADI finite difference schemes for option pricing in the Heston model with correlation*, International Journal of Numerical Analysis and Modeling **7** (2010), no. 2, 303–320.

[30] KJ In't Hout and BD Welfert, *Stability of ADI schemes applied to convection–diffusion equations with mixed derivative terms*, Applied Numerical Mathematics **57** (2007), no. 1, 19–35.

[31] Darae Jeong and Junseok Kim, *A comparison study of ADI and operator splitting methods on option pricing models*, Journal of Computational and Applied Mathematics **247** (2013), 162–171.

[32] OA Ladyzhenskaya, VA Solonnikov, and NN Uralceva, *Linear and quasilinear equations of parabolic type*, Providence, RI: American Mathematical Society, 1968.

[33] Hayne E Leland, *Option pricing and replication with transactions costs*, The Journal of Finance **40** (1985), no. 5, 1283–1301.

[34] Gary M Lieberman, *Second order parabolic differential equations*, World scientific, 1996.

[35] Keith W Morton and David Francis Mayers, *Numerical solution of partial differential equations: an introduction*, Cambridge University Press, 2005.

[36] Donald W Peaceman and Henry H Rachford, Jr, *The numerical solution of parabolic and elliptic differential equations*, Journal of the Society for industrial and Applied Mathematics **3** (1955), no. 1, 28–41.

[37] William H Press, *Numerical recipes 3rd edition: The art of scientific computing*, Cambridge University Press, 2007.

[38] Indranil SenGupta, *Option pricing with transaction costs and stochastic interest rate*, Applied Mathematical Finance **21** (2014), no. 5, 399–416.

[39] Daniel Ševčovič and Magdaléna Žitňanská, *Analysis of the nonlinear option pricing model under variable transaction costs*, Asia-Pacific Financial Markets (2016), 1–22.

[40] Steven E Shreve, *Stochastic calculus for finance ii: Continuous-time models*, vol. 11, Springer Science & Business Media, 2004.

[41] Gordon D Smith, *Numerical solution of partial differential equations: finite difference methods*, Oxford University Press, 1985.

[42] John C Strikwerda, *Finite difference schemes and partial differential equations*, vol. 88, Siam, 2004.

[43] Domingo Tavella and Curt Randall, *Pricing financial instruments: The finite difference method*, vol. 13, John Wiley & Sons, 2000.

[44] Paul Wilmott, Jeff Dewynne, and Sam Howison, *Option pricing: mathematical models and computation*, Oxford Financial Press, 1993.

[45] Valeri Zakamouline, *Hedging of option portfolios and options on several assets with transaction costs and nonlinear partial differential equations*, International Journal of Contemporary Mathematical Sciences **3** (2008), no. 4, 159–180.

[46] ———, *Option pricing and hedging in the presence of transaction costs and nonlinear partial differential equations*, Available at SSRN: https://ssrn.com/abstract=938933 or http://dx.doi.org/10.2139/ssrn.938933 (2008).