COMENIUS UNIVERSITY IN BRATISLAVA
FACULTY OF MATHEMATICS, PHYSICS, AND INFORMATICS
DEPARTMENT OF APPLIED MATHEMATICS AND STATISTICS

Qualitative and Quantitative Analysis of Nonlinear Parabolic Equations with Application in Finance

PhD Thesis

Study program:	Applied mathematics
Field of study:	22- Mathematics
Department:	Department of Applied Mathematics and Statistics
Supervisor:	prof. RNDr. Daniel Ševčovič, DrSc.

BRATISLAVA 2023 **M.Sc., Cyril Izuchukwu Udeani**

Univerzita Komenského v Bratislave
Fakulta matematiky, fyziky a informatiky

ZADANIE ZÁVEREČNEJ PRÁCE

Meno a priezvisko študenta:	Cyril Izuchukwu Udeani
Študijný program:	aplikovaná matematika (Jednoodborové štúdium, doktorandské III. st., denná forma)
Študijný odbor:	matematika
Typ záverečnej práce:	dizertačná
Jazyk záverečnej práce:	anglický
Sekundárny jazyk:	slovenský

Názov: Qualitative and Quantitative Analysis of Nonlinear Parabolic Equations with Applications in Finance
Kvalitatívna a kvantitatívna analýza nelineárnych parabolických rovníc s aplikáciami vo financiách

Anotácia: Kvalitatívna analýza a numerická analýza nelineárnych a nelokálnych parciálnych diferenciálnych rovníc vyskytujúcich sa v matematickej teórií oceňovania finančných derivátov a optimálneho manažentu portfólia. Analýza nelineárnych Hamilton-Jacobi-Bellmanovych rovníc .

Literatúra: S. Kilianová and D. Ševčovič: A Transformation Method for Solving the Hamilton-Jacobi-Bellman Equation for a Constrained Dynamic Stochastic Optimal Allocation Problem, ANZIAM Journal (55) 2013, 14-38.

S. Kilianová and D. Ševčovič: Dynamic intertemporal utility optimization by means of Riccati transformation of Hamilton-Jacobi Bellman equation, Japan Journal of Industrial and Applied Mathematics, 36(2), 2019, 497-517.

J. Cruz and D. Ševčovič: On solutions of a partial integro-differential Black-Scholes equation in Bessel potential spaces, Japan Journal of Industrial and Applied Mathematics 37 (2020), 691-721.

Školiteľ:	prof. RNDr. Daniel Ševčovič, DrSc.
Katedra:	FMFI.KAMŠ - Katedra aplikovanej matematiky a štatistiky
Vedúci katedry:	doc. Mgr. Igor Melicherčík, PhD.
Dátum zadania:	24.02.2012

Dátum schválenia: 16.01.2019

prof. RNDr. Daniel Ševčovič, DrSc.
garant študijného programu

študent

školiteľ

Abstract

In this thesis, we investigate the qualitative and quantitative analysis of nonlinear parabolic equations arising from finance. First, we study the existence and uniqueness of solutions of nonlinear partial integro-differential equations (PIDEs) arising from the financial market. We consider Black–Scholes models for pricing options on underlying assets following a Lévy stochastic process with jumps. The existence and uniqueness results of the PIDE are presented in the scale of Bessel potential spaces using the theory of abstract semilinear parabolic equations in high-dimensional spaces. As an application in the one-dimensional space, we consider a general shift function arising from nonlinear option pricing models taking into account a large trader stock-trading strategy. We consider a PIDE, where the shift function may depend on a prescribed large investor stock-trading strategy function. Second, we analyze problems arising from stochastic dynamic optimization, which leads to a solution of a fully nonlinear evolutionary Hamilton-Jacobi-Bellman (HJB) equation. We consider the HJB equation arising from portfolio optimization selection, where the goal is to maximize the conditional expected value of the terminal utility of the portfolio. After a suitable transformation, the fully nonlinear HJB equation is transformed into a quasilinear parabolic equation whose diffusion function is obtained as the value function of a specific conic programming problem. We employ the monotone operator technique, Banach's fixed point theorem, and Fourier transform to obtain the existence and uniqueness of a solution to the general form of the transformed parabolic equation in a suitable Sobolev space in an abstract setting. We also presented some financial applications of the proposed result in one-dimensional space. Furthermore, the behavior of the solution corresponding to the nonlinear HJB equation is studied. We analyze the behavior of the solution with respect to two decision sets. Finally, we present numerical analyses of the parabolic equations using deep learning. Specifically, we employ the physics-informed deep operator network (PI-DeepONet) to approximate the solution operator

of the parabolic equation associated with the HJB equation. Our qualitative analysis shows that PI-DeepONet can effectively learn the solution operator of the associated HJB equation.
Keywords: **Hamilton-Jacobi-Bellman equation; Maximal monotone operator; Dynamic stochastic portfolio optimization; Lévy measure; Option pricing; Bessel potential spaces; Deep learning, PI-DeepONet**

Abstrakt

V tejto práci sa zaoberáme kvalitatívnou a kvantitatívnou analýzou nelineárnych parabolických rovníc pochádzajúcich z matematickej teórie financií. Najprv študujeme existenciu a jednoznačnosť riešení nelineárnych parciálnych integro-diferenciálnych rovníc (PIDE) vznikajúcich vo finančnom modelovaní trhu. Uvažujeme Black–Scholes modely pre oceňovanie opcií na podkladové aktíva, ktoré sledujú Lévyho stochastické procesy so skokmi. Výsledky existencie a jednoznačnosti PIDE sú dokazované v škále Besselových potenciálnych priestorov s pomocou teórie abstraktných semilineárnych parabolických rovníc vo viacrozmerných priestoroch. Ako aplikáciu v jednorozmernom priestore uvažujeme všeobecnú funkciu posunu vyplývajúcu z nelineárnych modelov oceňovania opcií zohľadňujúcich stratégiu obchodovania s akciami veľkého obchodníka. Uvažujeme o PIDE, kde funkcia posunu môže závisieť od predpísanej funkcie stratégie obchodovania s akciami veľkého investora. Po druhé, analyzujeme problémy vyplývajúce zo stochastickej dynamickej optimalizácie, ktorá vedie k riešeniu plne nelineárnej evolučnej Hamilton-Jacobi-Bellmanovej (HJB) rovnice. Uvažujeme HJB rovnicu, ktoré vyplýva z dynamickej ptimalizácie portfólia, kde cieľom je maximalizovať podmienenú očakávanú hodnotu konečnej užitočnosti portfólia. Po vhodnej transformácii je plne nelineárna HJB rovnica transformovaná na kvázilineárnu parabolickú rovnicu, ktorej difúzna funkcia je získaná ako hodnotová funkcia konkrétneho kónického programovacieho problému. Používame techniku monotónneho operátora, Banachovu vetu o pevnom bode a Fourierovu transformáciu, aby sme získali existenciu a jedinečnosť riešenia všeobecného tvaru transformovanej parabolickej rovnice vo vhodnom Sobolevovom priestore v abstraktnom prostredí. Uviedli sme aj niektoré finančné aplikácie navrhovaného výsledku v jednorozmernom priestore. Ďalej sa študuje správanie riešenia zodpovedajúceho nelineárnej rovnici HJB. Analyzujeme správanie sa riešenia vzhľadom na dve rozhodovacie množiny. Nakoniec uvádzame numerické analýzy parabolických rovníc pomocou hlbokého učenia. Konkrétne využívame fyzikálne informovanú sieť hĺbkových operátorov (PI-DeepONet) na aproximáciu operátora riešenia parabolickej rovnice spojenej s rovnicou HJB. Naša kvalitatívna analýza ukazuje, že PI-DeepONet sa dokáže efektívne naučiť operátora riešenia súvisiacej rovnice HJB.

Kľúčové slová: Hamilton-Jacobi-Bellmanova rovnica; Maximálny monotónny operátor; Dynamická stochastická optimalizácia portfólia; Lévyho miera; Cena opcie; Besselove potenciálne priestory; Hlboké učenie, PI-DeepONet

Acknowledgements

Completing this Ph.D. thesis has been an arduous yet fulfilling journey, and I owe a debt of gratitude to those who have played a pivotal role in its accomplishment.

I am profoundly grateful to my supervisor, prof. RNDr. Daniel Ševčovič, DrSc., whose expertise, guidance, and unwavering support have been instrumental throughout this research process. Your mentorship has not only shaped the academic rigor of this thesis but has also been a source of inspiration for my own growth as a researcher. I would also want to appreciate his insights, encouragement, and overwhelming support, not just in this project but generally in life. I am privileged to have him as my supervisor.

I extend heartfelt thanks to my esteemed reviewers and editors for their insightful feedback, scholarly input, and constructive criticism. Your expertise has enriched the quality of this work and I appreciate the time and effort you dedicated to reviewing and advising on my research.

The financial support provided by the Slovak Research and Development Agency under the project APVV-20-0311 and VEGA 1/0062/18 has been crucial in facilitating my research. I am thankful for the resources and opportunities this funding has provided, allowing me to delve deeper into my chosen field.

To my colleagues and fellow researchers, your camaraderie, intellectual exchanges, and shared experiences have created an enriching academic environment. The collaborative spirit within our academic community has contributed significantly to the development of my ideas and perspectives. Specifically, I would like to thank Rev. Sister, Professor Maria Amarakristi Onyido for her unwavering support and encouragement, I am deeply grateful. Your belief in me has been a constant source of motivation, and I appreciate the understanding and patience you have shown during this demanding period.

I am indebted to my friends and family for their unwavering support and encouragement. I would like to thank my mother, especially for her support and constant encouragement. Your belief in my abilities and understanding during the challenging phases of this journey have been my pillars of strength.

A special acknowledgment goes to my lovely wife Cynthia Tochukwu Nnolum for her love, patience, understanding, and constant encouragement. Your presence has been my source of solace and inspiration, and I am grateful for the sacrifices made to accommodate the demands of my academic pursuits. Your support has been my rock, and I am grateful for the sacrifices you made to help me achieve this milestone.

This thesis is not just the culmination of my individual efforts, but a collective achievement made possible by the support of these incredible individuals. Their contributions have left an indelible mark on my academic journey.

Dedication

To my family and lovely wife, Cynthia.

Contents

List of Symbols **1**

1 Introduction **2**

1.1 Objective . . . 2

1.2 Background . . . 3

1.3 Thesis Overview . . . 10

1.3.1 Detailed Structure . . . 10

1.4 List of Publications . . . 11

2 Monotone Operators and their Properties **12**

2.1 Preliminaries and basic definitions . . . 12

2.2 Analytic Semigroups . . . 15

2.3 Elementary properties of monotone operators . . . 17

2.4 Existence results for parabolic equations . . . 18

3 PIDEs and their Applications to Option Pricing **21**

3.1 Background and Motivation . . . 22

3.2 Definitions and basic properties . . . 23

3.2.1 Examples of Lévy processes in finance . . . 23

3.2.2 Admissible activity Lévy measures . . . 25

3.3 Main results . . . 26

3.3.1 Option Pricing Under Stock-Trading Strategy . . . 27

3.4 Discussion to Chapter 3 . . . 30

4 Application of Maximal monotone operator to Portfolio Management **32**

4.1 Background and motivation . . . 33

4.2 Riccati transformation . . . 34

4.3 Preliminaries, definitions, and main results 35
4.3.1 Main results 36
4.4 Behavior of the solution with respect to the decision set 38
4.4.1 Properties of the value function 38
4.5 Numerical examples 39
4.5.1 Qualitative behavior of the solution 43
4.5.2 Examples 45
4.6 Discussion to Chapter 4 47

5 Conventional numerical method **49**
5.1 Finite Difference Method 49
5.1.1 Fundamental Concept of FDM 50
5.2 Application to Option Pricing 51
5.3 Discussion to Chapter 5 53

6 Application of Deep Learning for Solving nonlinear parabolic PDEs **56**
6.1 Introduction 56
6.2 Deep Neural Networks 57
6.2.1 Artificial Neural Networks 57
6.2.2 Physics-informed Neural Networks (PINNs) 59
6.2.3 Automatic Differentiation 61
6.3 Comparison between Traditional Neural Network, PINNs, and and Finite Difference Approximation 62
6.4 Application of PI-DeepONet to portfolio managment 64
6.4.1 Methodology of PI-DeepONet 64
6.4.2 Problem formulation 66
6.5 Results and Discussion to Chapter 6 68

7 Concluding Remarks and Prospects **70**

8 Appendices **72**
8.1 Appendix A 72
8.1.1 Proof of Propositions 1 72
8.1.2 Proof of Proposition 2 74
8.1.3 Proof of Theorem 5 75
8.2 Appendix B 76
8.2.1 Proof of Theorem 6 76
8.2.2 Proof of Theorem 7 80
8.2.3 Proof of Theorem 8 81
8.2.4 Proof of Proposition 3 82

List of Figures **83**

9 References **85**

List of Symbols

$\mathbb{R}^n$	Set of all n-dimensional real vectors
$\langle ., . \rangle$	Duality pairing between a vector space and its dual
$\hat{f}$	Fourier transform of a function f
$\mathcal{E}$	Bounded convex set of mean returns
$\mathbf{\Sigma}$	$n \times n$ positive definite covariance matrix
$\hookrightarrow$	Continuous embedding
$\\|\cdot\\|$	Norm in an infinite dimensional space
$\lvert \cdot \rvert$	Euclidean norm in $\mathbb{R}^n$
$a \cdot b$	Euclidean product in $\mathbb{R}^n$ with the norm $\lvert z \rvert = \sqrt{z \cdot z}$
$C^k(\Omega)$	kth continuously differentiable functions on Ω, $0 \leq k \leq \infty$
$C_0^\infty(\Omega)$	C^∞ functions on Ω with compact support
∇	Gradient of a scalar-valued function
Δ	Laplacian operator
$L^p(\Omega)$ $(1 \leq p < \infty)$	Function spaces $f : \Omega \to \mathbb{R}$ with $\int_\Omega \lvert f(x) \rvert^p < \infty$
$W^{m,p}(\Omega)$	Function spaces with $f \in L^p(\Omega)$ and $\partial^\alpha f \in L^p(\Omega)$, $\alpha = (\alpha_1, \cdots, \alpha_n)$
$\mathscr{L}_\alpha^p(\Omega)$	Bessel potential space containing functions φ with $\varphi = G_\alpha * f$, $f \in L^p(\Omega)$
G_α	Bessel kernel of order α
$Rg(A)$	Range of a map A

CHAPTER 1

Introduction

1.1 Objective

The objective of this thesis is to investigate the qualitative and quantitative analysis of fully nonlinear partial integro-differential equations (PIDEs) of parabolic type arising from finance. We consider a fully nonlinear Hamilton-Jaccobi-Belman (HJB) equation describing portfolio optimization problems, as well as PIDE related to the nonlinear Black–Scholes equation for pricing vanilla options. The nonlinear generalization of the Black–Scholes and HJB equations can be transformed into the quasilinear parabolic equation for the unknown function $\varphi = \varphi(x, \tau)$ representing either the Gamma function of the portfolio $\varphi = S\partial_S^2 V$ or relative risk aversion function $\varphi = -\frac{\partial_x^2 V}{\partial_x V}$, respectively (see, Ševčovič [59]). The following equation describes the resulting quasilinear parabolic equation:

$$\partial_t \varphi = \mathcal{L}\alpha(\cdot, \varphi) + F(x, \varphi, \nabla\varphi), \quad x \in \mathbb{R}^n, t \in [0, T), \tag{1.1}$$

where $\mathcal{L}$ represent integral and differential operator (e.g., Laplacian); α is some suitable value function. Our goal is to study the qualitative and quantitative properties of solutions to such parabolic equations. First, we consider HJB equations with application in optimal portfolio management under suitable assumptions on the utility function and establish its solution in suitable Sobolev spaces. Then, for the parabolic equation for option pricing, we consider nonlocal nonlinear equation corresponding to the nonlinear Black–Scholes equation and establish its solution in higher-dimensional space in Bessel potential space. Our approaches to proving the existence and uniqueness of solutions are two-fold. First, we employ the monotone operator technique to establish the solution of the HJB equation in Sobolev spaces. For the PIDE corresponding to the nonlinear Black–Scholes equation, we employ the theory of abstract semilinear parabolic equations to prove the existence and uniqueness of solutions in the scale of Bessel potential spaces.

1.2 Background

In the past decades, the theory of differential equations has been extensively studied, and the methods for their solutions strongly depend on the specific equation. A comprehensive study on these methods are presented in [7, 21, 54]. Solving parabolic equations using the theory of monotone operators is a powerful mathematical approach that has gained significant attention in the field of partial differential equations [66]. This method, often referred to as the monotone operator technique, offers a systematic framework for tackling a wide range of parabolic equations with various boundary and initial conditions. A monotone operator is a mathematical concept used to describe operators that preserve the order of elements in a given space. It plays a crucial role in the study of parabolic equations because it ensures that the equation is well posed and has a stable solution. The monotone operator technique involves transforming the original parabolic equation into an equivalent fixed-point problem involving a monotone operator. The idea is to construct a sequence of approximations that converges to the solution of the original equation. The main advantages of monotone operator techniques are as follows. The monotone operator technique guarantees the well-posedness of the solution, ensuring that the solution exists, is unique, and varies continuously with respect to the data. Most times, this technique leads to stable numerical algorithms for solving parabolic equations. This stability is essential for accurate and robust numerical solutions. Furthermore, the sequence of approximations generated by the monotone scheme converges to the true solution of the parabolic equation under certain conditions. Although this approach is valuable as it combines mathematical rigor with practical applicability, like any mathematical method, its success depends on careful application, analysis, and adaptation to specific problems. Besides the monotone operator technique, the semigroup theory provides a powerful framework for studying the existence, uniqueness, and regularity of solutions to various PDEs, including the parabolic type. Some notable works and references that contribute to this field can be found in [7, 21, 27]. A semigroup is a mathematical structure that consists of a set (often functions) and an associative binary operation that combines elements of the set. Semigroup of linear operator theory provides a systematic approach to solving parabolic PDEs by viewing them as evolution equations [52]. In this context, the semigroup represents the evolution operator that maps the initial condition at the initial time to the solution at any later time. In general, the semigroup associated with a parabolic PDE should satisfy some key properties. The semigroup should be positive preserving, i.e., it should map nonnegative initial conditions to nonnegative solutions. It should be Hölder continuous in the underlying Banach space, i.e., it should be continuous with respect to the Hölder norm in the underlying Banach space, ensuring that small changes in the initial condition lead to small changes in the solution. Inspired by the above properties, in the qualitative part of this thesis, we employ the monotone operator technique and semigroup theory to establish the existence and uniqueness of solutions to the parabolic equation (1.1).

As mentioned above, the parabolic equation (1.1) describes the nonlinear Black–Scholes equation for option pricing as well as the nonlinear HJB equation for selecting optimal portfolios [59]. Because of its simplicity and the existence of an analytical formula to price derivative securities, the classical Black–Scholes model has been widely used in finance, especially for pricing vanilla options. This model usually relies on restrictive assumptions like market completeness and the assumption that the underlying asset price follows a geometric Brownian motion. However, the assumption that an investor can trade a large amount of assets without affecting the underlying asset price is usually not satisfied, especially in illiquid markets. Additionally, the linear Black–Scholes equation usually leads to an undesirable property since it provides a solution corresponding to a perfectly replicated portfolio. For this reason, several generalizations have been made to relax some of these assumptions. For example, these assumptions were relaxed by (i) considering the presence of transaction costs [6, 42], (ii) feedback and illiquid market effects due to large traders choosing given stock-trading strategies [26, 25, 65], and (iii) the risk from the unprotected portfolio [30]. In these generalizations, the constant volatility was replaced by a nonlinear function depending on the second derivative of the option price. Frey and Stremme derived a nonlinear Black–Scholes model that plays an essential role in the class of the generalized Black–Scholes equation with such a nonlinear diffusion function [24, 26, 30]. In this model, the asset dynamics considers the presence of feedback effects due to a large trader choosing his stock-trading strategy [65]. Another important direction in generalizing the original Black–Scholes equation arises from the fact that the sample paths of a Brownian motion are continuous. However, the real stock price of a typical company exhibits random jumps on the intraday scale, making price trajectories discontinuous. In the classical Black–Scholes model, the logarithm of the price process follows a normal distribution. However, the empirical distribution of stock returns shows fat tails. When calibrating the theoretical prices to the market prices, the implied volatility is not constant as a function of strike price nor a function of time to maturity, contradicting the prediction of the Black–Scholes model. On the other hand, the models with jumps and diffusion can solve the problems inherent to the Black–Scholes model. Jump models also play an essential role in the option market. In the Black–Scholes model, the market is complete, implying that every payoff can exactly be replicated;

meanwhile, there is no perfect hedge in jump models, making the way of options not redundant.

Recently, the relationships between more general nonlocal operators and jump processes have been widely investigated. For instance, there is an actual connection between the solution to PIDEs and properties of the corresponding Markov jump process (c.f., Abels and Kassmann [2]; Florescu and Mariani [23]). Furthermore, the role of PIDEs has also been recently investigated in various fields, such as pure mathematics, biological sciences, and economics [3, 4, 73]. PIDE problems arising from financial mathematics, especially from option pricing models, have been of great interest to many researchers. In most cases, standard methods for solving these problems lead to the study of parabolic equations. Mikulevičius and Pragaraustas [49] investigated solutions of the Cauchy problem to the parabolic PIDE with variable coefficients in Sobolev spaces. They employed their results to obtain solutions of the corresponding martingale problem. Crandal *et al.* [28] employed the notion of a viscosity solution to investigate the qualitative results. Soner *et al.* [14] and Barles *et al.* [8] extended and generalized their results for the first and second-order operators, respectively. Florescu and Mariani [23] employed the Schaefer fixed point argument to establish the existence of a weak solution of the generalized PIDE. Amster *et al.* [58] used the notion of upper and lower solutions to obtain the solution of such PIDEs. They proved the existence of solutions in a general domain for multiple assets and the regime-switching jump-diffusion model. Cont *et al.* [17] investigated the actual connection between option pricing in exponential Lévy models and the corresponding PIDEs for European options and those with single or double barriers. They discussed and established the conditions for which prices of options are classical solutions of the corresponding PIDE.

In this thesis, we establish a certain PIDE for option pricing in illiquid market by assuming that the stock price follows a certain dynamics. We also present the existence of a solution and localization results of the associated PIDE in some suitable spaces. The qualitative properties of solutions to the nonlocal linear and nonlinear PIDE are investigated and established in the scale of Bessel potential spaces using the theory of abstract semilinear parabolic equation. Regarding the PIDE for option pricing, we relax the liquid and complete market assumptions and extend the models that study market illiquidity to the case where the underlying asset price follows a Lévy stochastic process with jumps. As a result, we establish the corresponding PIDE for option pricing under suitable assumptions. Then, the qualitative properties of solutions to nonlocal linear and nonlinear PIDEs are presented using the theory of abstract semilinear parabolic equation in the scale of Bessel potential spaces. The existence and uniqueness of solutions to the PIDE for a general class of the so-called admissible Lévy measures satisfying suitable growth conditions at infinity and origin are also established in multidimensional space. Additionally, the qualitative properties of solutions to the generalized PIDE are investigated by considering a general shift function arising from nonlinear option pricing models, which takes into account a large trader stock-trading strategy with the underlying asset price following the Lévy process. Various numerical experiments are presented to illustrate the influence of a large trader and the intensity of jumps in the option price.

Furthermore, it is known that the fully nonlinear HJB equation plays an essential role in finance. For example, it gives the necessary and sufficient conditions for optimal control with respect to the value function [60]. Recent studies on such parabolic equations employed the method of the upper and lower solution [54, 70]. Specifically, Macová and Ševčovič [45] analyzed the solutions to a fully nonlinear parabolic equation representing the problem of optimal portfolio construction. They showed the formulation of the problem of optimal stock-to-bond proportion in the management of a pension fund portfolio in terms of the solutions to the HJB equation. Federicol et al. [22] investigated the utility maximization problem for an investment-consumption portfolio when the current utility depends on the wealth process - regularity of solutions to the HJB equation. They defined a dual problem and treated it by means of dynamic programming and showed that the viscosity solutions of the associated HJB equation belong to a class of smooth functions. Ishimura and Ševčovič [29] recently analyzed solutions to the HJB equation (4.7) with range bounds on the optimal response variable. They constructed monotone traveling wave solutions and identified parametric regions for which the traveling wave solutions have positive or negative wave speeds. More recently, Abe and Ishimura [1] introduced the Riccati transformation method for solving the full nonlinear HJB equations, which was later generalized by Kilianová and Ševčovič [35]. In their later study, Kilianová and Ševčovič [35] investigated solutions of a fully nonlinear HJB equation for a constrained dynamic stochastic optimal allocation problem. However, no attempt has been made to solve the fully nonlinear HJB equation arising in portfolio optimization in a suitable Sobolev space using the monotone operator technique. In this thesis, we also study the qualitative and quantitative properties of the HJB equation with application in finance in high dimensions using the monotone operator argument. We employ the monotone operator technique because it plays a crucial role in establishing constructive proofs for existence theorems and leads to various comparison results, which are effective tools for studying qualitative properties of solutions. We present the existence and uniqueness of the solution to the nonlinear HJB equation arising from stochastic dynamic programming using a combination of the monotone operator technique, Fourier transform, and Banach fixed point argument. We consider the fully nonlinear HJB equation arising from the portfolio selection problem, where the goal of an investor is to optimize the conditional expected value of the terminal utility of the portfolio. Such a nonlinear parabolic equation is presented in an abstract setting, which can also be viewed as a nonlinear PIDE. The existence result of such a nonlinear parabolic equation presented in an abstract setting is established in Sobleve spaces with some shift/perturbation in the main operator of the underlying equation. This Cauchy problem for the nonlinear parabolic equation corresponds to equation (1.1) using a suitable Ricatti transformation.

Numerous studies have shown that most differential equations arise from many scientific and engineering fields for modeling physical phenomena. However, most of these differential equations are analytically intractable, especially in high-dimensional space. In the numerical approach of solving most differential equations, several traditional numerical methods, such as the finite volume method, finite difference method, finite element method, and spectral methods (e.g., Fourier-spectral method), have been widely used to solve complicated parabolic equations, including those with complex geometries, nonlinearities, and variable coefficients. The choice of a numerical method usually depends on various factors, including the complexity of the problem, the geometry, the stability requirements, and computational resources. Complex parabolic equations often require careful consideration of the numerical scheme, as well as techniques to handle nonlinearities, boundary conditions, and stability issues. Although these traditional methods, such as the finite difference method, finite elements, finite volume method, method of lines, spectra methods, and adaptive mesh refinement methods, have provided accurate and reliable results, they require high computational resources, and a slight change in the parameter of the equation could lead to independent simulations. Despite the fact that these classical methods have been extensively studied, their convergence properties have not been properly investigated. Additionally, in the numerical solution of partial differential equation (PDE) problems through the discretization process using finite difference approximations, the algebraic systems generated are finalized using an iterative method. In order to overcome these challenges, many researchers have replaced traditional numerical discretization methods with artificial neural networks (ANNs) to approximate the PDE solution. Recently, deep neural networks (DNNs) have been widely used to solve classical applied mathematical problems, including PDEs, utilizing machine learning and artificial intelligence approaches [33]. The neural networks and deep learning methods have gained significant attention for solving various PDEs, especially of parabolic type, because of their capability to approximate and solve problems with complex geometries. Due to significant nonlinearities, convection dominance, or shocks, some PDEs are difficult to solve using standard numerical approaches. Recent studies have shown that deep learning is a promising method for building meta-models for fast predictions of dynamic systems. In particular, NNs have proven to represent the underlying nonlinear input-output relationship in complex systems. To this end, deep learning has recently emerged as a new paradigm of scientific computing thanks to the universal approximation theorem and great expressivity of neural networks [15]. Solving parabolic PDEs that describe phenomena evolving over time with spatial dependencies, such as heat diffusion or option pricing in finance, can be challenging, but ANNs and deep learning methods offer a promising approach. The usual approach of solving any parabolic PDE requires finding a function that satisfies the equation and its boundary/initial conditions. ANNs can be used to represent this unknown function. The input to the network can be the spatial coordinates and time, and the output can be the value of the solution at that point in space and time. In most cases of training the network, labeled data are usually required, which consists of input points (spatial and temporal coordinates) and corresponding known solution values. These training data are usually generated by discretizing the PDE domain and solving the PDE numerically using traditional methods to obtain ground truth solutions. The advantages of using ANNs and deep learning for solving parabolic PDEs include their ability to handle complex geometries and adapt to a wide range of boundary and initial conditions. They can also be faster than traditional numerical methods in most cases. Although ANNs offer promising solutions for solving parabolic equations, they come with several limitations and challenges. For instance, ANNs are often considered as "black-box" models, i.e., one may not really know how all the individual neurons work together to obtain the final output. This makes it challenging to interpret the learned solution and gain insights into the underlying physical processes. This lack of interpretability can be a significant drawback in scientific and engineering applications where understanding physics is crucial.

In an attempt to approximate the solution of PDEs, several researchers employed the deep Galerkin method [33], which employs DNNs to solve high-dimensional PDEs. Another approach called the deep Ritz method [72], which defined the loss as the energy of the solution of the problem, was also introduced. More recently, a more suitable technique called physics-informed neural networks (PINNs) was introduced to approximate the solution of PDEs [57]. PINN is a special class of ANNs that combine deep learning techniques with domain-specific knowledge. The primary aim of PINNs is to harness the power of neural networks while ensuring that the solution adheres to the underlying physics of the problem. PINNs are designed to embed the governing physics of a problem directly into the neural network architecture and loss function. This is usually achieved by adding terms to the loss function that enforce PDE constraints and boundary/initial conditions. One of the key advantages of PIN is that it incorporates prior knowledge about the underlying physics, making it well-suited for problems where the governing equations are known but difficult to solve analytically. Although PINN offers several advantages, it comes with many limitations. For example, PINNs are more effective at interpolation within the range of training data, but extrapolating to regions far from the training data can be challenging. In other to overcome these limitations, several researchers have recently introduced and analyzed different variants of PINNs [31, 34, 74]. For instance, Kharazmi et al. [34] proposed a variant of PINN called a variational PINN (hp-VPINN) in which a Galerkin approach is used on collocation points. Many other collocation-based PINN approaches, such as conservative PINN [31], were also introduced. Another variant of PINN called physics-constrained neural networks (PCNNs) [74] was also introduced. Unlike PINN, which incorporates both the PDE and boundary/initial conditions (soft BC) into the training loss function, PCNNs are data-free neural networks that enforce the initial and boundary conditions (hard BC) through a custom neural network architecture while embedding the PDE in the training loss. Furthermore, PINN has achieved great success than its variants, which can be seen from the significant increase in the PINN citation [57]. Although PINNs are faster than traditional numerical methods, they also have some limitations in terms of changes in the underlying parameters governing the PDE. In other words, a slight change in the underlying parameters could result in a retraining of the model. This is also applicable to traditional numerical methods in which a slight change in the input parameter will lead to a new independent simulation.

To overcome the shortcomings of PINNs, the concept of deep operator network (DeepONet) was further introduced [44]. DeepONet is a neural network-based model that can learn linear and nonlinear PDE solution operators with a small generalization error via universal approximation theorem for operators. DeepONet consists of two parts: a deep neural network that learns the solution of the PDE and an operator network that enforces the PDE at each iteration. The operator network acts as a constraint to ensure that the neural network outputs satisfy the underlying PDE. DeepONet maps input functions with infinite dimensions to output functions with infinite dimensions. It can efficiently and accurately solve PDE with any initial and boundary conditions without retraining the network. Moreover, several authors have introduced various forms of DeepONet, including DeepONet with proper orthogonal decomposition (POD-DeepONet), Bayesian DeepONet, neural operator with coupled attention, multiscale DeepONet, and physics-informed DeepONet (PI-DeepONet) (see [31, 34, 44, 74] and the references therein). PI-DeepONet approximates the PDE solution operator using two networks: one network that encodes the discrete input function space (i.e., branch net) and one that encodes the domain of the output functions (i.e., trunk net) (cf. [44]). PI-DeepONet is a variant of DeepONet that incorporates known physics (or governing equations) into the neural network architecture. It can effectively approximate the solution of different PDEs without requiring a large amount of training data by introducing a regularization mechanism that biases the outputs of DeepONet models to ensure physical consistency. PI-DeepONet can efficiently solve parametric linear and nonlinear PDEs compared to other variants of PINN since it can take source term parameters (including other parameters) as the input variables. This approach can improve the accuracy of the solution and reduce the amount of data needed for training. Moreover, PI-DeepONet can break the curse of dimensionality in the input space, making it more suitable than other traditional approaches.

Inspired by the above development and studies, in the numerical part of this thesis, we employ deep learning techniques to solve the parabolic equation corresponding to the Cauchy problem of the underlying HJB equation. Specifically, we employ the PI-DeepONet to approximate the solution operator of the HJB equation. We consider a fully nonlinear HJB equation arising from the stochastic optimization problem, where the goal of an investor is to maximize the conditional expected value of the terminal utility of a portfolio. PI-DeepONet is used to learn the optimal portfolio strategy for an investor by approximating the solution operator of the transformed quasilinear equation. The HJB equation is first transformed using the Ricatti transform introduced by Ishimura and Ševčovič [29] and Ševčovič and Kilianová [35] into a quasilinear parabolic equation, which is then approximated using PI-DeepONet. The neural network architecture is trained using a combination of supervised learning and physics-informed learning. The supervised learning part involves minimizing the mean squared error between the neural network predictions and a set of training data points. The physics-informed learning part involves enforcing the PDE constraint using the operator network at each iteration. This results in a more accurate solution that satisfies the underlying PDE, even when the input parameters governing the PDE change.

1.3 Thesis Overview

1.3.1 Detailed Structure

The remainder of this thesis is organized as follows. Chapter 2 discusses the concept of monotone operators, its characterizations in function spaces, and existence results for parabolic equations. The concept of semigroup theory for solving parabolic equations is also presented. Chapter 3 presents the qualitative analysis of nonlinear PIDEs (1.1) associated with the Black–Scholes equation for pricing options in multidimensional cases. The existence and uniqueness of the solution to this equation are established in the scale of Bessel potential spaces using the theory of abstract semilinear parabolic equation. Chapter 4 presents existence results for the nonlinear parabolic equation (1.1) associated with the HJB equation for selecting optimal portfolios. The behaviors of solutions to such parabolic equations with respect to different decision sets are also discussed. In Chapter 5, we present a detailed discussion of traditional numerical methods for solving PDEs. Chapter 6 presents the numerical analysis of the parabolic equation associated with the HJB equation using PI-DeepONet. Finally, Chapter 7 presents the conclusion and future studies.

1.4 List of Publications

Below is the list of publications arising from direct results of the thesis.

1. Udeani Cyril Izuchukwu, and Ševčovič Daniel. "Application of maximal monotone operator method for solving Hamilton–Jacobi–Bellman equation arising from optimal portfolio selection problem." Japan Journal of Industrial and Applied Mathematics 38, no. 3 (2021): 693-713.

2. Ševčovič Daniel, and Cyril Izuchukwu Udeani. "Multidimensional linear and nonlinear partial integro-differential equation in Bessel potential spaces with applications in option pricing." Mathematics 9, no. 13 (2021): 1463.

3. Bello, AU, Omojola, MT, Onyido, MA, Uba, MO, and Udeani, CI. "New Method for Computing Zeros of Monotone Maps in Lebesgue Spaces With Applications to Integral Equations, fixed Points, Optimization, and Variational Inequality Problems." Acta Mathematica Universitatis Comenianae 91, no.2 (2022).

4. Cruz Jose, Grossinho Maria, Ševčovič Daniel, and Udeani Cyril Izuchukwu. "Linear and Nonlinear Partial Integro-Differential Equations arising from Finance." In: Understanding Integro-Differential Equations, Understanding Integro-Differential Equations, Eds: J. Vasundhara Devi, Z. Drici, F. A. McRae, Nova Science Publishers, Inc., Hauppauge, 2023, pp. 191- 256, ISBN: 979-8-89113-040-1.

5. Ševčovič Daniel and Udeani Cyril Izuchukwu. "Learning the solution operator of a nonlinear parabolic equation using physics-informed deep operator network." ECMI2023 Proceedings, Springer (accepted), (2023).

6. Ševčovič Daniel, and Udeani Cyril Izuchukwu. "Hamilton-Jacobi-Bellman Equation Arising from Optimal Portfolio Selection Problem." (Submitted to ECMI2023 Proceedings, Springer), (2023).

7. Onyido Maria A, Uba Markjoe O, Udeani Cyril I, and Nwokoro Peter U. "A hybrid scheme for fixed points of a countable family of generalized nonexpansive-type maps and finite families of variational inequality and equilibrium problems with applications." Carpathian Journal of Mathematics 39, no. 1 (2023): 281–292.

CHAPTER 2

Monotone Operators and their Properties

This chapter is devoted to the study of monotone operators, their characterizations in function spaces, and existence results for nonlinear parabolic partial differential equations. Specifically, we discuss the theory of Banach spaces, Sobolev spaces, Hölder space, and the scale of Bessel potential spaces.

2.1 Preliminaries and basic definitions

In this section, we present the basic notions, definitions, and properties of function spaces used in this thesis. Specifically, we discuss the theory of Banach spaces, Sobolev spaces, Hölder spaces, and the scale of Bessel potential spaces.

Let V be a complete normed vector space (i.e., Banach space) with the norm $\|\cdot\|$. We denote the dual space of V by V', as the space of all continuous linear functional on V, and the norm of V' is defined by

$$\|f\|_{V'} = \sup_{\substack{\|x\|\leq 1 \\ x\in V}} |f(x)| = \sup_{\substack{\|x\|\leq 1 \\ x\in V}} f(x)$$

for any $f \in V'$. In the sequel, we shall denote the duality pairing between the spaces V and V' by $\langle .,. \rangle$, i.e., the value of a functional $F \in V'$ at $u \in V$ is denoted by $\langle F, u\rangle$. Note that the bidual or the second dual V'' can also be constructed in a similar manner. Moreover, let $f \in V'$ and $x \in V$ be given, we can define $\phi(x) \in V''$ by $\phi(x)(f) \equiv f(x)$, which fulfills $\|\phi(x)\| \leq \|x\|$. For each $x \in V$, we can identify $f \in V'$ with $f(x) = \|x\|$ and $\|f\| = 1$, which implies that $\|\phi(x)\| = \|x\|$. Clearly, the function ϕ is linear, which implies that the map $\phi : V \to V''$ is a linear isometry of V onto a closed subspace of

V'', denoted by $V \hookrightarrow V''$. Now, if ϕ is onto the whole space V'', then V is reflexive. In a more general setting, a reflexive Banach space is defined as follows.

Definition 1. *[13] Let V be a Banach space and let $j : V \to V''$ be a cannonical injection from V into V''. Then, V is said to be reflexive if j is surjective, i.e., $j(V) = V''$.*

The following remarks are consequences of the above definition.

Remark 1. *(i) If V is reflexive, then we can identify V'' with V.*

(ii) A closed subspace and dual space of a reflexive space is reflexive.

(iii) Every finite-dimensional space is reflexive since $dimV = dimV' = dimV''$.

The following theorem presents the necessary and sufficient condition for a Banach space to be reflexive.

Theorem 1. *[7, 13, 66] A Banach space is reflexive if and only if it is sequentially weakly compact, i.e., every bounded sequence contains a weakly convergent subsequence.*

For more details on the proof and application of Theorem 1, see [7, 11, 13, 66].

Definition 2. *[13] A metric space X is separable if it contains a countable dense subset.*

It is worth noting that every finite-dimensional space is separable, and several important spaces used in the analysis of differential equations are separable and reflexive.

Next, we define and characterize some important spaces used in this thesis. Let Ω be an open Lebesgue measurable subset of $\mathbb{R}^n$ with smooth boundary $\partial\Omega$ and let $1 \leq p \leq \infty$. Then, $L^p(\Omega)$ $(1 \leq p < \infty)$ is the set of equivalence classes of measurable functions $f : \Omega \to \mathbb{R}$ such that $\int_\Omega |f(x)|^p dx < \infty$. Moreover, $L^\infty(\Omega)$ is the space consisting of all essentially bounded function $f : \Omega \to \mathbb{R}$ such that $|f(x)| \leq C$ a.e on Ω for some constant $C < \infty$.

Remark 2. *[13]*

(i) $L^p(\Omega)$ spaces are reflexive (for $1 < p < \infty$) and separable (for $1 \leq p < \infty$).

(ii) $L^\infty(\Omega)$ is neither reflexive nor separable.

Now, let us introduce some function spaces whose derivative belongs to L^p. First, we denote the space of test functions on Ω by $C_0^\infty(\Omega)$ (i.e., the class of all C^∞ functions on $\mathbb{R}^n$ with compact support), which is dense in every $L^p(\Omega)$ for $1 \leq p < \infty$. Here, the support of a continuous function $f : \mathbb{R}^n \to \mathbb{R}$ is defined as $suppf = \overline{\{x \in \mathbb{R}^n : f(x) \neq 0\}}$. Moreover, we introduce the Sobolev space of positive integer orders, consisting of real-valued functions defined on Ω that satisfy some integrability properties with their distributional derivatives.

Definition 3. *[66] Let Ω be an open subset of $\mathbb{R}^n, 1 \le p \le \infty$, and $0 \le m$(integers). Then, $W^{m,p}(\Omega)$ is a linear space of all functions $f \in L^p(\Omega)$ for which $\partial^\alpha f \in L^p(\Omega)$ for all multi-indices $\alpha = (\alpha_1, \cdots, \alpha_n)$ of nonegative integers with $|\alpha| = \alpha_1 + \cdots + \alpha_n \le m$, where $\partial^\alpha = \partial_1^{\alpha_1}\partial_2^{\alpha_2}\cdots\partial_n^{\alpha_n}$. Such a space is called a Sobolev space of order m and integrability p with the norms*

$$\|f\|_{m,p} = (\sum_{|\alpha|\le m} \|\partial^\alpha f\|^p_{L^p(\Omega)})^{1/p},\ 1 \le p < \infty,$$

$$\|f\|_{m,\infty} = \max_{|\alpha|\le m} \|\partial^\alpha f\|_{L^\infty(\Omega)}.$$

From the above definition, it is clear that $W^{0,p}(\Omega) = L^p(\Omega)$. Note that if $1 \le p < \infty$, then $C^\infty(\Omega) \cap W^{m,p}(\Omega)$ is dense in $W^{m,p}(\Omega)$. In other words, smooth functions are dense in $W^{m,p}(\Omega)$. For more details, see [66, Theorem 4.1]. The following remark demonstrates the elementary properties of Sobolev spaces.

Remark 3. *[13, 66]*

(i) $W^{m,p}(\Omega)$ *is a Banach space.*

(ii) if $1 \le p < \infty$, *then* $W^{m,p}(\Omega)$ *is separable.*

(iii) if $1 < p < \infty$, *then* $W^{m,p}(\Omega)$ *is reflexive.*

(iv) $W^{m,2}(\Omega)$, *usually denoted as* $H^m(\Omega)$, *is a Hilbert space with the scalar product* $(u, v) := \sum_{|\alpha|\le m} \int_\Omega D^\alpha f \overline{D^\alpha g} dx, \forall f, g \in W^{m,2}(\Omega)$, *where D denotes the derivative in the sense of distribution.*

Moreover, the following continuous embeddings hold:

$$W^{m,p}(\Omega) \hookrightarrow L^q(\Omega) \quad \text{if } \frac{1}{p} \ge \frac{1}{q} \ge \frac{1}{p} - \frac{m}{n} > 0,$$

$$W^{m,p}(\Omega) \hookrightarrow C^k(\Omega) \quad \text{if } m - \frac{n}{p} > k,$$

where the last embedding indicates that every function in $W^{m,p}(\Omega)$ has a continuous representation for $m - \frac{n}{p} > 0$.

Next, we introduce a class of Sobolev-type spaces of noninteger order. Let us recall the convolution operator $(G * f)(x) = \int_{\mathbb{R}^n} G(x-y) f(y) \mathrm{d}y$.

Definition 4. *Let $1 < p < \infty, \alpha \ge 0$. Then, the Bessel potential space, denoted as $\mathscr{L}^p_\alpha(\Omega)$, consists of all functions φ such that $\varphi = G_\alpha * f$ for some $f \in L^p(\Omega)$, where G_α is the Bessel kernel of order α given by*

$$G_\alpha(x) = \frac{1}{(4\pi)^{n/2}\Gamma(\alpha/2)} \int_0^\infty y^{-1+(\alpha-n)/2} e^{-(y+|x|^2/(4y))} \mathrm{d}y$$

for $x \in \mathbb{R}^n/\{0\}$.

The Bessel potential space $\mathscr{L}^p_\alpha(\Omega)$ is a natural extension of the Sobolev space $W^{\alpha,p}(\Omega)$. They are Banach spaces for $1 \leq p < \infty$. The Bessel potential space $\mathscr{L}^p_\alpha(\Omega)$ can be constructed by interpolating the Sobolev spaces of integral order or using the Fourier transformation. If it is constructed using the Fourier transformation, then its norm is given as follows:

$$\|u\|_{\mathscr{L}^p_\alpha} = \|\mathcal{F}^{-1}[(1+|\xi|^2)^{\alpha/2}\mathcal{F}(u)(\xi)]\|_{L^p}, \forall u \in \mathscr{L}^p_\alpha,$$

where $\mathcal{F}^{-1}$ is the inverse transformation and $|\xi| = (\xi_1^2 + \cdots + \xi_n^2)^{1/2}$. The following remark shows the basic properties of the Bessel kernel $G_\alpha, \alpha > 0$.

Remark 4. *Let* $f \in L^p(\mathbb{R}^n)$ *and the Fourier transform of* f *be given as*

$$\hat{f}(\xi) = \int_{\mathbb{R}^n} f(y)e^{-2\pi i x y}dy.$$

Then,

(i) G_α *is monotone decreasing, nonnegative, and* $\int_{\mathbb{R}^n} G_\alpha(y)dy = 1$.

(ii) $\hat{G}_\alpha(\xi) = (1+|\xi|^2)^{-\alpha/2}$, $\xi \in \mathbb{R}^n$.

We remark here that the Bessel potential spaces are separable for $1 \leq p < \infty$, reflexive for $1 < p < \infty$, and the Hilbert spaces for $p = 2$.

2.2 Analytic Semigroups

This section presents the basic definitions and characterization of analytic semigroup used in this thesis.

Definition 5. *[27, Definition 1] An analytic semigroup is a family of bounded linear operators* $\{S(t), t \geq 0\}$ *in a Banach space* X *satisfying the following conditions:*

(i) $S(0) = I, S(t)S(s) = S(s)S(t) = S(t+s)$, *for all* $t, s \geq 0$;

(ii) $S(t)u \to u$ *when* $t \to 0^+$ *for all* $u \in X$;

(iii) $t \to S(t)u$ *is a real analytic function on* $0 < t < \infty$ *for each* $u \in X$.

The associated infinitesimal generator A is defined as follows: $Au = \lim_{t\to 0^+} \frac{1}{t}(S(t)u - u)$, and its domain $D(A) \subseteq X$ consists of elements $u \in X$ for which the limit exists in the space X.

Definition 6. *[27] Let $S_{a,\varphi} = \{\lambda \in \mathbb{C} : \varphi \leq \arg(\lambda - a) \leq 2\pi - \varphi\}$ be a sector of complex numbers. A closed densely defined linear operator $A : D(A) \subset X \to X$ is called a sectorial operator if there exists a constant $M \geq 0$ such that*

$$\|(A - \lambda)^{-1}\| \leq M/|\lambda - a|$$

for all $\lambda \in S_{a,\varphi} \subset \mathbb{C} \setminus \sigma(A)$, where $\sigma(A)$ denotes the spectrum of the operator A.

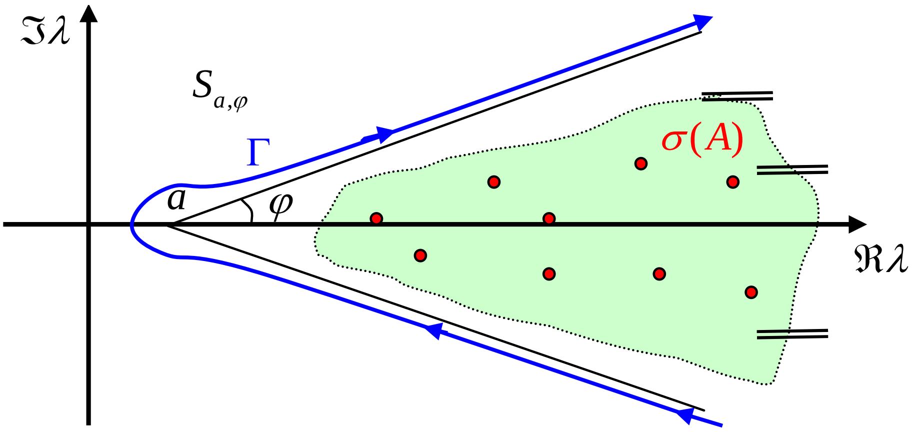


Figure 2.1: Sector $S_{a,\varphi}$ in the complex plane [61]

Note that if A is a bounded linear operator on a Banach space, then A is sectorial. Moreover, if A is a self-adjoint densely defined operator in a Hilbert space and bounded below, then A is sectorial.

Suppose an operator A is sectorial in a Banach space X, then $-A$ is a generator of an analytic semigroup $\{e^{-At}, t \geq 0\}$ acting on X (c.f., [27, Chapter I]). For any $\gamma > 0$, we can define the operator $A^{-\gamma} : X \to X$ as

$$A^{-\gamma} = \frac{1}{\Gamma(\gamma)} \int_0^\infty \xi^{\gamma-1} e^{-A\xi} \mathrm{d}\xi.$$

Then, the fractional power space $X^\gamma = D(A^\gamma)$ is the domain of the operator $A^\gamma = (A^{-\gamma})^{-1}$, i.e.,

$$X^\gamma = \left\{u \in X : \exists f \in X, u = A^{-\gamma} f\right\}, \tag{2.1}$$

with the norm $\|u\|_{X^\gamma} = \|A^\gamma u\|_X = \|f\|_X$. Moreover, we have continuous embedding: $D(A) \equiv X^1 \hookrightarrow X^{\gamma_1} \hookrightarrow X^{\gamma_2} \hookrightarrow X^0 \equiv X$, for $0 \leq \gamma_2 \leq \gamma_1 \leq 1$.

According to [27, Section 1.6], $A = -(\sigma^2/2)\Delta$ is a sectorial operator in the Lebesgue space $X = L^p(\mathbb{R}^n)$ for any $p \geq 1, n \geq 1$, and $D(A) \subset W^{2,p}(\mathbb{R}^n)$, where in our application σ^2 denotes historical volatility, and Δ is the Laplacian. Thus, it follows from [67, Chapter 5] that the space $X^\gamma, \gamma > 0$, can be identified with the Bessel potential space $\mathscr{L}^p_{2\gamma}(\mathbb{R}^n)$, where

$$\mathscr{L}^p_{2\gamma}(\mathbb{R}^n) := \{u \in X : \exists f \in X, u = G_{2\gamma} * f\}.$$

and $G_{2\gamma}$ is the Bessel potential function. The norm of $u = G_{2\gamma} * f$ is given by $\|u\|_{X^\gamma} = \|f\|_{L^p}$. The space X^γ is continuously embedded in the fractional Sobolev-Slobodeckii space $W^{2\gamma,p}(\mathbb{R}^n)$ (c.f., [27, Section 1.6]).

2.3 Elementary properties of monotone operators

This section presents and discusses some elementary properties of monotone operators used in this thesis.

Definition 7. *Let E and F be two Banach spaces. An unbounded linear operator from E into F is a linear map $A : D(A) \subset E \to F$ defined on a linear subspace $D(A) \subset E$ with values in F, where $D(A)$ is the domain of A. Moreover, the operator A is bounded (or continuous) if $D(A) = E$ and there exists a constant $C \geq 0$ such that*

$$\|Au\| \leq C\|u\|, \ \forall u \in E.$$

Now, suppose H is a Hilbert space and $A : D(A) \subset H \to H$ is an unbounded linear operator with $\overline{D(A)} = H$, in which H' is identified with H. Then, A^* can be identified as an unbounded linear operator in H. Moreover,

(i) A is symmetric if $(Au, v) = (u, Av), \forall u, v \in D(A)$.

(ii) A is self-adjoint if $D(A^*) = D(A)$ and $A^* = A$.

It is worth noting that the notion of symmetric and self-adjoint operators coincides if A is a bounded linear operator. Meanwhile, if A is unbounded, there is a slight difference between symmetric and self-adjoint operators. Generally, any self-adjoint operator is symmetric, but the converse is false. An operator A is symmetric if and only if $A \subset A^*$, i.e., $D(A) \subset D(A^*)$. Moreover, if A is maximal monotone, then A is symmetric if and only if A is self-adjoint (see [13, Proposition 7.6 1]). Such an unbounded operator A is said to be *accretive* if $(Au, u) \geq 0, \forall u \in D(A)$, and it is $m - accretive$ if, in addition, $Rg(A + I) = H$, where I denotes the identity map.

Next, we introduce and discuss the notion of monotone operators and their fundamental properties in a reflexive Banach space.

Definition 8. *Let V be a reflexive Banach space and $\mathcal{A} : V \to V'$ be an operator, and $\langle ., \ . \rangle_{V',V}$ be the scalar duality in V. Then, $\mathcal{A}$ is*

(i) monotone if $\langle \mathcal{A}(u) - \mathcal{A}(v), u - v \rangle \geq 0, \forall u, v \in V$;

(ii) hemicontinuous if for each $u, v \in V$, the real-valued function $t \mapsto \mathcal{A}(u + tv)(v)$ is continuous;

(iii) *type* M *if* $u_n \rightharpoonup u, \mathcal{A}(u_n) \rightharpoonup f$, *and* $\limsup \mathcal{A}u_n(u_n) \leq f(u)$ *implies* $\mathcal{A}(u) = f$*;*

(iv) *strictly monotone if* $\langle \mathcal{A}(u) - \mathcal{A}(v), u - v\rangle > 0$*, for* $u \neq v$ *in* V*;*

(v) *strongly monotone if there exists a constant* $k > 0$ *such that* $\langle \mathcal{A}(u)-\mathcal{A}(v), u-v\rangle \geq k\|u - v\|_V^2, \forall u, v \in V$*;*

(vi) *coercive if* $\frac{\langle \mathcal{A}(u),u\rangle}{\|u\|_V} \to \infty$ *as* $\|u\|_V \to \infty$*;*

(vii) *maximal monotone if it is monotone and* $\mathcal{A}(u) = f$ *if* $\langle \mathcal{A}(v)-f, u-v\rangle \geq 0, \forall v \in V$*.*

From the above definitions, we deduce that the operator $\mathcal{A} : V \to V'$ is demicontinuous if it is type M and bounded and type M if it is hemicontinuous and monotone. According to Browder and Rockafellar [66], monotonicity implies local boundedness, which means that for monotone operators, demicontinuous and hemicontinuity are equivalent. Furthermore, strong monotonicity implies coerciveness. The definition of maximal monotone operator simply means no proper monotone extension. It is worth noting that monotonicity and hemicontinuity imply maximal monotone, which in turn implies type M. The monotonicity and maximal monotonicity of subdifferential of a function can be characterized as follows. If a function is convex and proper, then its subdifferential is monotone [51]. Moreover, if a function is convex, closed, and proper, then its subdifferential is maximal monotone.

2.4 Existence results for parabolic equations

This section is devoted to the characterization of solutions of parabolic differential equations for monotone operators.

Let V be a reflexive Banach space with dual V', and let $\mathcal{V} = L^p((0,T);V)$ for $1 < p < \infty$, with a dual $\mathcal{V}' = L^p((0,T);V')$. Suppose H is a Hilbert space in which its dual is identified by the Riesze representation. Assume that V is dense and continuously embedded in H. In other words, V, H, and V' satisfy the Gelfand triple, i.e., $V \hookrightarrow H \hookrightarrow V'$. Define the mapping $\mathcal{A} : V \to V'$ and let $u_0 \in H, f \in \mathcal{V}'$. Now, consider the Cauchy problem of finding $u \in \mathcal{V}$ such that

$$\partial_\tau u + \mathcal{A}(u) = f \text{ in } \mathcal{V}', \qquad u(0) = u_0. \tag{2.2}$$

Since $f \in \mathcal{V}'$ and $\mathcal{A}(u) \in \mathcal{V}'$, then $u' \in \mathcal{V}'$. This indicates that u is continuous into H, and the condition $u(0) = u_0$ makes sense. Then, the following results hold.

Theorem 2. *[66] Assume that* V, H*, and* $\mathcal{V}$ *satisfy the above settings. Suppose the operator* $\mathcal{A} : V \to V'$ *is given such that its realization in* $\mathcal{V}$ *is of type* M*, bounded, and coercive with*

$$\langle \mathcal{A}(u), u\rangle \geq \|u\|_{\mathcal{V}}^p, \quad u \in \mathcal{V}.$$

Then, for each $f \in \mathcal{V}$, and $u_0 \in H$, there exists a unique solution to the Cauchy problem (2.2).

Proof: The proof of Theorem 2 is based on the Galerkin method. For more details, see [7, 66]. The next lemma considers the assumption on the operator $\mathcal{A} : V \to V'$, which leads to the corresponding assumption on its realization in $\mathcal{V}$. Without loss of generality, also denote the realization in $\mathcal{V}$ as $\mathcal{A}$.

Lemma 1. *[13, 66]*

(i) Suppose $\mathcal{A} : V \to V'$ is demicontinuous, then for each measurable function $\omega : [0,T] \to V$, $\mathcal{A}(\omega(\cdot)) : [0,T] \to \mathcal{V}'$ is also measurable.

(ii) Suppose also that $\mathcal{A}$ is bounded with

$$\|\mathcal{A}(u)\| \leq C\|u\|^{p-1}, \quad u \in \mathcal{V}.$$

Then, its realization $\mathcal{A} : \mathcal{V} \to \mathcal{V}'$ is demicontinuous.

(iii) Moreover, $\mathcal{A}$ is $V-$monotone if and only if its realization is $\mathcal{V}$- monotone.

Remark 5. *We note here that Theorem 2 also holds and the solution of (2.2) is unique if $\mathcal{A} : V \to V'$ is monotone, hemicontinuous, bounded, and coercive with an estimate on V. This notion can then be extended to cover a family of such operators and related Cauchy problems. This scenario is described with the relaxation of the coercive assumption in Theorem 3.*

Theorem 3. *[7, 66] Let V be a separable reflexive Banach space, dense and continuous in a Hilbert space H which is identified with its dual, so $V \hookrightarrow H \hookrightarrow V'$. Let $p \geq 2$ and set $\mathcal{V} = L^p((0,T);V)$. Assume a family of operators $\mathcal{A}(\tau,.) : V \to V', 0 \leq \tau < T$, is given such that*

(i) for each $u \in V$, the function $\mathcal{A}(.,u) : [0,T] \to V'$ is measurable,

(ii) for a.e $\tau \in [0,T]$, the operator $\mathcal{A}(\tau,.) : V \to V'$ is monotone, hemicontinuous and bounded by $\|\mathcal{A}(\tau,u)\| \leq C(\|u\|^{p-1} + k(\tau)), u \in V, 0 \leq \tau < T$, where $k \in L^{p'}(0,T)$,

(iii) and there exists $\lambda > 0$ such that $\langle \mathcal{A}(\tau,u), u\rangle \geq \lambda\|u\|^p - k(\tau), u \in V, 0 \leq \tau < T$.

Then for each $\tilde{f} \in \mathcal{V}'$ and $u_0 \in H$, there exists a unique solution $u \in \mathcal{V}$ of the Cauchy problem

$$\partial_\tau u(\tau) + \mathcal{A}(\tau, u(\tau)) = \tilde{f}(\tau) \text{ in } \mathcal{V}', \quad u(0) = u_0. \tag{2.3}$$

The idea presented in Theorem 3 can also be extended to the case when $\mathcal{A}$ is a derivative to obtain a strong solution of (2.3). This scenario can be seen as a nonlinear analog of Theorem 3, which also corresponds to the solution obtained using the theory of semigroup on H.

The following corollary shows that the coercive assumption can be relaxed elementary when the operator satisfies linear growth.

Corollary 1. *Suppose that the assumptions of Theorem 3 are satisfied with $p = 2$ and there exist $\lambda, \alpha > 0$ such that*

$$\mathcal{A}(\tau, v)(v) + \lambda \|V\|_H^2 \geq \alpha \|v\|^2, \quad a.e\ \tau \in [0, T], v \in V.$$

Then the Cauchy problem (2.3) has a unique solution for each $f \in \mathcal{V}$ and $u_0 \in H$.

Theorem 3 forms the basis for establishing our existence and uniqueness results for the fully nonlinear HJB equation (4.1), which corresponds to the nonlinear parabolic equation (1.1).

CHAPTER 3

PIDEs and their Applications to Option Pricing

This chapter is devoted to analyzing solutions of nonlinear partial integro-differential equations (PIDEs) arising from financial modeling in multidimensional spaces. We employ the theory of abstract semilinear parabolic equations in order to prove the existence and uniqueness of solutions in the scale of Bessel potential spaces. We consider a wide class of Lévy measures satisfying suitable growth conditions near the origin and infinity. The novelty of the chapter is the generalization of existing results in one dimension to the multidimensional case. We consider Black–Scholes models for option pricing on underlying assets following a Lévy stochastic process with jumps. As an application to option pricing in the one-dimensional space, we consider a general shift function arising from a nonlinear option pricing model taking into account a large trader stock-trading strategy. We prove the existence and uniqueness of a solution to the nonlinear PIDE in which the shift function may depend on a prescribed large investor stock-trading strategy function (see Ševčovič and Udeani [62]). The main results of this chapter are contained in our paper[62]

[62] D. Ševčovič and C. I. Udeani. Multidimensional Linear and Nonlinear Partial Integro-Differential Equation in Bessel Potential Spaces with Applications in Option Pricing. *Mathematics,* **9** (13) (2021) 1463.

3.1 Background and Motivation

In this chapter, we investigate the existence and uniqueness of a solution to a nonlocal equation of the form:

$$\begin{aligned}\frac{\partial u}{\partial \tau} &= \Delta u + \int_{\mathbb{R}^n} [u(\tau, x+z) - u(\tau, x) - z \cdot \nabla u(\tau, x)] \, \nu(\mathrm{d}z) + g(\tau, x, u, \nabla u), \quad (3.1)\\ & u(0, x) = u_0(x), \quad x \in \mathbb{R}^n, \tau \in (0, T).\end{aligned}$$

Here, g is a given sufficiently smooth function; ν denotes a positive measure on $\mathbb{R}^n$ such that its Radon–Nikodym derivative is a nonnegative Lebesgue measurable function h in $\mathbb{R}^n$, i.e., $\nu(\mathrm{d}z) = h(z)\mathrm{d}z$.

The nonlocal equation (3.1) generalizes known results by Cruz and Ševčovič [18] to the multidimensional space in the scale of Bessel potential Sobolev spaces. Moreover, we analyze the following generalization of the nonlocal equation (3.1):

$$\frac{\partial u}{\partial \tau} = \frac{\sigma^2}{2} \Delta u + \int_{\mathbb{R}^n} [u(\tau, x+\xi) - u(\tau, x) - \xi \cdot \nabla u(\tau, x)] \, \nu(\mathrm{d}z) + g(\tau, x, u, \nabla u), \quad (3.2)$$

where $\xi = \xi(\tau, x, z)$ is a shift function, which may depend on the variables $\tau > 0, x, z \in \mathbb{R}$. An application for such a general shift function ξ can be found in nonlinear option pricing models considering a large trader stock-trading strategy with the underlying asset price dynamic following the Lévy process (c.f., Cruz and Ševčovič [19]). If $\xi(x, z) \equiv z$, then (3.2) reduces to equation (3.1). The nonlinearity g often arises from applications occurring in pricing, e.g., XVA derivatives (c.f., Arregui *et al.* [4, 5]) or applications of the penalty method for American option pricing under a PIDE model (c.f., Cruz and Ševčovič [18]).

The primary aim of this chapter is to investigate the solution of PIDE (3.1) in the framework of Bessel potential spaces for a multidimensional case, $n \geq 1$. These spaces form a nested scale $\{X^\gamma\}_{\gamma \geq 0}$ of Banach spaces satisfying $X^{\gamma_1} \hookrightarrow X^{\gamma_2}$ for any $1 \geq \gamma_1 \geq \gamma_2 \geq 0$, and $X^1 \equiv D(A), X^0 \equiv X$. The operator A is sectorial in the space X with a dense domain $D(A) \subset X$ (c.f., Henry [27]). An example of such an operator is the Laplace operator, i.e., $A = -\Delta$ in $\mathbb{R}^n$ with the domain $D(A) \equiv W^{2,p}(\mathbb{R}^n) \subset X \equiv L^p(\mathbb{R}^n)$. It is worth noting that if $A = -\Delta$, then X^γ is embedded in the Sobolev–Slobodecki space $W^{2\gamma,p}(\mathbb{R}^n)$, which is a space consisting of all functions such that 2γ-fractional derivative belongs to the Lebesgue space $L^p(\mathbb{R}^n)$ of p-integrable functions (c.f., [27]). We investigate solutions to the PIDE (3.2) for a wide class of Lévy measures ν satisfying suitable growth conditions near $\pm\infty$ and origin in a higher dimensional space.

3.2 Definitions and basic properties

This section presents some basic definitions and properties of Lévy measures, as well as the notion of admissible activity Lévy measures used in this thesis. Recall that we denote the Euclidean norm and the norm in infinite dimensional spaces (e.g., $L^p(\mathbb{R}^n)$, X^γ) as $|\cdot|$ and $\|\cdot\|$, respectively; $a \cdot b$ denotes the usual Euclidean product in $\mathbb{R}^n$ with the norm $|z| = \sqrt{z \cdot z}$.

Definition 9. *[62] A Lévy process on $\mathbb{R}^n$ is a stochastic (right continuous) process $X = \{X_t, t \geq 0\}$ having the left limit with independent stationary increments. It is uniquely characterized by its Lévy exponent ϕ:*

$$\mathbb{E}_x(e^{iy \cdot X_t}) = e^{-t\phi(y)}, \ y \in \mathbb{R}^n.$$

The subscript x in the expectation operator $\mathbb{E}_x$ indicates that the process X_t starts from a given value x at the origin $t = 0$. The Lévy exponent ϕ has the following unique decomposition:

$$\phi(y) = ib \cdot y + \sum_{i,j=1}^{n} a_{ij} y_i y_j + \int_{\mathbb{R}^n} \left(1 - e^{iy \cdot z} + iy \cdot z 1_{|z| \leq 1}\right) \nu(\mathrm{d}z),$$

where $b \in \mathbb{R}^n$ is a constant vector; (a_{ij}) is a constant matrix, which is positive semidefinite; $\nu(\mathrm{d}z)$ is a nonnegative measure on $\mathbb{R}^n \setminus \{0\}$ such that $\int_{\mathbb{R}^n} \min(1, |z|^2)\nu(\mathrm{d}z) < \infty$ (c.f., [53]).

Next, we introduce and discuss the concept of exponential Lévy models. Exponential Lévy models form a class of stochastic processes used in mathematical finance to model the dynamics of underlying asset prices. The underlying process is a stochastic process whose logarithm is a Lévy process, i.e., a process that has stationary and independent increments with jumps of random size. The exponential of this process then becomes a stochastic process that has the property of multiplicative decomposition, i.e., it can be written as the product of two independent processes: one that has deterministic growth (e.g., a deterministic interest rate) and one that has stochastic volatility (e.g., a Lévy process).

3.2.1 Examples of Lévy processes in finance

There are two major types of exponential Lévy models considered in the literature: jump-diffusion and infinite activity pure jump models. In jump-diffusion models, the log-price is represented as a Lévy process with a nonzero diffusion part ($\sigma > 0$) and a jump process with finite activity (i.e., $\nu(\mathbb{R}) < \infty$). In contrast, there is no diffusion part in infinite activity pure jump models and only a jump process with infinite activity (i.e.,

$\nu(\mathbb{R}) = \infty$). Next, we discuss different types of exponential Lévy models that differ in the choice of the Lévy measure.

Jump-Diffusion models

A Lévy process with jump-diffusion has the following general form:

$$X_t = \gamma t + \sigma W_t + \sum_{i=1}^{N_t} Y_i,$$

where $\sigma > 0$, and N_t is a Poisson process with intensity λ that counts the jumps of X_t, and $Y_i, i = 1, \cdots, N_t$, are independent and identically distributed random variables with distribution μ. The Lévy measure ν is $\lambda\mu$, and the drift γ is given by

$$\gamma = -\frac{\sigma^2}{2} - \int_{\mathbb{R}} \left(e^z - 1 - z 1_{|z|\leq 1}\right) \nu\left(\mathrm{d}z\right).$$

Merton's model

This is the first jump-diffusion model proposed by Merton [47] in the context of financial applications. The random variables $Y_i, i = 1, 2, 3...$, are normally distributed with mean m and variance δ^2. It has the following Lévy density:

$$\nu(\mathrm{d}z) = \lambda \frac{1}{(2\pi\delta^2)^{n/2}} e^{-\frac{|z-m|^2}{2\delta^2}} \mathrm{d}z\,, \tag{3.3}$$

where the parameters $m \in \mathbb{R}^n, \lambda, \delta > 0$, are given. It is worth noting that Merton's measure is a finite activity Lévy measure, i.e., $\nu(\mathbb{R}^n) = \int_{\mathbb{R}^n} \nu(\mathrm{d}z) < \infty$, with finite variation $\int_{|z|\leq 1} |z|\nu(\mathrm{d}z) < \infty$.

Infinite activity

The variance Gamma and normal inverse Gaussian (NIG) processes are obtained by a subordination of a Brownian motion and a tempered α-stable process; variance Gamma and NIG processes correspond to $\alpha = 0$ and $\alpha = 1/2$, respectively. These models are widely used in finance because of the existence of the probability density of the subordinator in a closed form for these values of α (for more details on the probability density, see [16]).

Variance Gamma process

This is a process of infinite activity and finite variation ($\int_{|z|\leq 1} |z|\nu(\mathrm{d}z) < \infty$) that is widely used in financial modeling. It has the following Lévy density:

$$\nu(dz) = \frac{1}{\kappa|z|}e^{Az-B|z|}$$

where

$$A = \frac{\theta}{\sigma^2}, B = \frac{\sqrt{\theta^2 + 2\frac{\sigma^2}{\kappa}}}{\sigma^2}.$$

Here, σ and θ are parameters related to the volatility and drift of the Brownian motion, respectively; κ is the parameter related to the variance of the subordinator, which is a Gamma process (see [16]).

Normal inverse Gaussian model

The NIG model is a process of infinite activity and infinite variation. It has the following Lévy density [16]

$$\nu(dz) = \frac{C}{|z|}e^{Az}K_1(B|z|)$$

where

$$C = \frac{\sqrt{\theta^2 + \frac{\sigma^2}{\kappa}}}{2\pi\sigma\sqrt{\kappa}}, A = \frac{\theta}{\sigma^2}, B = \frac{\sqrt{\theta^2 + \frac{\sigma^2}{\kappa}}}{\sigma^2},$$

where θ, σ, and κ have the same meaning as in the variance Gamma process.

3.2.2 Admissible activity Lévy measures

This subsection presents the notion of an admissible activity Lévy measure introduced by Cruz and Ševčovič [18, 19] for the one-dimensional case $n = 1$, which was later extended by Ševčovič and Udeani [62] for the multidimensional case $n \geq 1$.

Definition 10. *[62, Definition 1] A measure ν in $\mathbb{R}^n$ is called an admissible activity Lévy measure if there exists a nonnegative Lebesgue measurable function $h : \mathbb{R}^n \to \mathbb{R}$ such that $\nu(\mathrm{d}z) = h(z)\mathrm{d}z$ with*

$$0 \leq h(z) \leq C_0|z|^{-\alpha}e^{-D|z|-\mu|z|^2}, \tag{3.4}$$

for all $z \in \mathbb{R}^n$ and the shape parameters $\alpha, \mu \geq 0, D \in \mathbb{R}$ ($D > 0$ if $\mu = 0$), where $C_0 > 0$ is a positive constant.

The additional conditions $\int_{\mathbb{R}^n} \min(|z|^2, 1)\nu(\mathrm{d}z) < \infty$ and $\int_{|z|>1} e^z \nu(\mathrm{d}z) < \infty$ are satisfied provided that ν is an admissible Lévy measure with shape parameters $\alpha < n+2$, and either $\mu > 0, D^{\pm} \in \mathbb{R}$, or $\mu = 0$ and $D^- + 1 < 0 < D^+$. For the Merton model, we have $\alpha = 0, D^{\pm} = 0$ and $\mu = 1/(2\delta^2) > 0$. Meanwhile, for the Kou model, we have $\alpha = \mu = 0, D^+ = \lambda^-, D^- = -\lambda^+$. For the variance Gamma process, we have $\alpha = 1, \mu = 0, D^{\pm} = A \pm B$.

3.3 Main results

This section is devoted to our main results on the nonlocal equation in the scale of Bessel potential space in higher-dimensional space.

Now, consider the following mapping

$$Q(u,\xi) = u(x+\xi(x)) - \xi(x)\frac{\partial u}{\partial x}, \quad x \in \mathbb{R}.$$

where ξ is a shift function depending on the variable $x \in \mathbb{R}$. Then, by [18, Lemma 3.4], there is a constant $\hat{C}$ such that for any $\xi_1, \xi_2 \in L^\infty(\mathbb{R})$ and u with $\partial_x u \in X^{\gamma-1/2}$, $\frac{1}{2} < \gamma < 1$, we have

$$\|Q(u,\xi_1) - Q(u,\xi_2)\|_{L^p(\mathbb{R})} \le \hat{C}\|\xi_1 - \xi_2\|_\infty^{2\gamma-1}(\|\xi_1\|_\infty + \|\xi_2\|_\infty)\|\partial_x u\|_{X^{\gamma-1/2}}.$$

Next, we present a more general result of the above statement. The statement of the proposition and its proof is contained in our recent publication (see, Ševčovič and Udeani [62]).

Proposition 1. *[62, Ševčovič and Udeani] Let $\xi \in L^\infty(\mathbb{R}^n)$ and define the mapping $Q(u,\xi)$ by*

$$Q(u,\xi) = u(x+\xi(x)) - \xi(x)\cdot\nabla_x u(x), \quad x \in \mathbb{R}^n.$$

Then, there exists a constant $\hat{C} > 0$ such that for any vector valued functions $\xi_1, \xi_2 \in L^\infty(\mathbb{R}^n)$, and u such that $\nabla_x u \in (X^{\gamma-1/2})^n$, $1/2 \le \gamma < 1$, the following estimate holds:

$$\|Q(u,\xi_1) - Q(u,\xi_2)\|_{L^p(\mathbb{R}^n)} \le \hat{C}\|\xi_1 - \xi_2\|_\infty^{2\gamma-1}(\|\xi_1\|_\infty + \|\xi_2\|_\infty)\|\nabla_x u\|_{X^{\gamma-1/2}}.$$

Proposition 1 shows the estimate when the nonlocal term depends only on x, i.e., for the case $\xi = \xi(x)$. It generalizes the known results by Cruz and Ševčovič [19, 18] for the one-dimensional case $n = 1$. For the proof of Proposition 1, see Appendix 8.1

Corollary 2. *Let u be such that $\nabla_x u \in (X^{\gamma-1/2})^n$ where $1 > \gamma \ge 1/2$. Then, for any $\xi \in \mathbb{R}^n$, the following pointwise estimate holds:*

$$\|Q(u,\xi)\|_{L^p(\mathbb{R}^n)} \le C_0|\xi|^{2\gamma}\|\nabla_x u\|_{X^{\gamma-1/2}}.$$

Proposition 2. *[62, Ševčovič and Udeani] Suppose that the shift mapping $\xi = \xi(x,z)$ satisfies $\sup_{x\in\mathbb{R}}|\xi(x,z)| \le C_0|z|^\omega(1+e^{D_0|z|})$ for some constants $C_0 > 0, D_0 \ge 0, \omega > 0$ and any $z \in \mathbb{R}$. Assume ν is a Lévy measure with the shape parameters α, D, and either $\mu > 0, D \in \mathbb{R}$, or $\mu = 0$ and $D > D_0 \ge 0$. Assume $1/2 \le \gamma < 1$, and $\gamma > (\alpha - n)/(2\omega)$. Then, there exists a constant $C_0 > 0$ such that*

$$\|f(u)\|_{L^p} \le C_0\|\nabla_x u\|_{X^{\gamma-1/2}},$$

provided that $\nabla_x u \in (X^{\gamma-1/2})^n$. If $u \in X^\gamma$, then $\|f(u)\|_{L^p} \le C\|u\|_{X^\gamma}$, i.e., $f : X^\gamma \to X$ is a bounded linear operator.

Proposition 2 shows the case when the nonlocal integral term depends on x and z variables. It generalizes the result in [19, Lemma 3.4] due to Cruz and Ševčovič proven for the case where $\xi(x,z) \equiv z$. The proof of Proposition 2 is contained in our recent paper [62, Proposition 2]. For more details on the proof of Proposition 1, see Appendix 2.

Theorem 4. *Suppose that the shift mapping $\xi = \xi(x,z)$ satisfies $\sup_{x\in\mathbb{R}} |\xi(x,z)| \le C_0|z|^\omega(1+e^{D_0|z|}), z \in \mathbb{R}^n$, for some constants $C_0 > 0, D_0 \ge 0, \omega > 0$. Assume ν is an admissible activity Lévy measure with the shape parameters α, D, and, either $\mu > 0, D \in \mathbb{R}$, or $\mu = 0, D > D_0 \ge 0$. Assume $1/2 \le \gamma < 1$ and $\gamma > (\alpha - n)/(2\omega)$, $n \ge 1$. Suppose that $g(\tau, x, u, \nabla_x u)$ is Hölder continuous in the τ variable and Lipschitz continuous in the remaining variables, respectively. Assume $u_0 \in X^\gamma$, and $T > 0$. Then, there exists a unique mild solution u to PIDE (3.2) that satisfies $u \in C([0,T], X^\gamma) \cap C^1([0,T], X)$.*

Theorem 4 is a direct consequence of Propositions 2. It is a nontrivial generalization of the result shown by Ševčovič and Cruz [18] in the one space dimension $n = 1$. The proof of Theorem 4 is based on the result due to Henry [27, Proposition 3.5].

3.3.1 Option Pricing Under Stock-Trading Strategy

This section is devoted to the application of the results to the pricing of options under a stock trading strategy.

The proof of Theorem 4 is based on Propositions 2 and [27, Proposition 3.5]. It is a nontrivial generalization of the result shown by Ševčovič and Cruz [18] in one-dimensional space, i.e., $n = 1$.

In finance, it is known that the price $V = V(t,S)$ of an option on an underlying asset price $S > 0$ at time $t \in [0,T]$ can be obtained from the solution to the linear Black–Scholes parabolic equation of the form:

$$\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + rS\frac{\partial V}{\partial S} - rV = 0, \quad V(T,S) = \Phi(S), \tag{3.5}$$

where $r > 0$ is a parameter representing the risk-free interest rate of a bond; $\sigma > 0$ is the historical volatility of the underlying asset price S time series. We assume that the underlying asset price follows the geometric Brownian motion $dS/S = \mu dt + \sigma dW$, where $\{W_t, t \ge 0\}$ denotes the standard Wiener process. The terminal condition $\Phi(S)$ represents the pay-off diagram at maturity $t = T$, $\Phi(S) = (S-K)^+$ (call option case) or $\Phi(S) = (K-S)^+$ (put option case). Equation (3.5) can be transformed using $x = \ln\frac{S}{K}$, $\tau = T - t$, and $V(t,S) = e^{-r\tau}u(\tau,x)$ into parabolic equation

$$\frac{\partial u}{\partial \tau} = \frac{\sigma^2}{2}\frac{\partial^2 u}{\partial^2 x} + \left(r - \frac{\sigma^2}{2}\right)\frac{\partial u}{\partial x} - ru. \tag{3.6}$$

The multidimensional generalization of (3.5), where the option price $V(t, S_1, \cdots, S_n)$ depends on the vector of n underlying stochastic assets $S = (S_1, \cdots, S_n)$ with several volatilities σ_i and mutual correlations $\varrho_{ij}, i, j = 1, \cdots, n$, is given by

$$\frac{\partial V}{\partial t} + \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n}\rho_{ij}\sigma_i\sigma_j S_i S_j \frac{\partial^2 V}{\partial S_i \partial S_j} + r\sum_{i=1}^{n} S_i \frac{\partial V}{\partial S_i} - rV = 0, \quad V(T, S) = \Phi(S). \tag{3.7}$$

It is worth noting that equations (3.5) and (3.7) can be transformed into equation (3.1) defined on the whole space $\mathbb{R}^n$ (c.f., Ševčovič, Stehlíková, Mikula [63, Chapter 4, Section 5]). According to stock markets analysis, the models (3.5) and (3.7) were derived under several restrictive assumptions like perfect replication of a portfolio and its liquidity, completeness and frictionless of the financial market, and absence of transaction costs. However, these assumptions are often violated in financial markets. Because of this, several authors have investigated these models in order to relax these assumptions (see for [6, 30, 42, 43, 64]. According to Cruz and Ševčovič [19], the Black–Scholes model can incorporate the effect of a large trader and the assumption on the liquidity of the market can be relaxed by assuming that the underlying asset price follows a Lévy stochastic process with jumps. This results in the following nonlinear PIDE:

$$\frac{\partial V}{\partial t} + \frac{1}{2}\frac{\sigma^2 S^2}{(1 - \varrho S \partial_S \phi)^2}\frac{\partial^2 V}{\partial S^2} + rS\frac{\partial V}{\partial S} - rV + \int_{\mathbb{R}} \left(V(t, S + H) - V(t, S) - H\frac{\partial V}{\partial S}\right)\nu(\mathrm{d}z) = 0, \tag{3.8}$$

where the shift function $H = H(\phi, S, z)$ depends on the large investor stock-trading strategy function $\phi = \phi(t, S)$, which is a solution to the following implicit algebraic equation:

$$H = \rho S(\phi(t, S + H) - \phi(t, S)) + S(e^z - 1). \tag{3.9}$$

The large trader strategy function ϕ may depend on the derivative $\partial_S V$ of the option price V, e.g., $\phi(t, S) = \partial_S V(t, S)$. Moreover, in our application, we assume the trading strategy function $\phi(t, S)$ is prescribed and it is globally Hölder continuous. For more details on how (3.8) can be transformed into a linear or nonlinear parabolic PIDE with respect to the value of ρ, see our recent paper [62, Ševčovič and Udeani].

Nevertheless, in our application, we assume the trading strategy function $\phi(t, S)$ is prescribed and it is globally Hölder continuous.

If $\rho = 0$, then $H = S(e^z - 1)$. Using the standard transformation $\tau = T - t, x = \ln(\frac{S}{K})$ and setting $V(t, S) = e^{-r\tau}u(\tau, x)$, then equation (3.8) can be reduced to a linear PIDE of the form (3.1) in the one-dimensional space ($n = 1$).

If $\rho > 0$, then (3.8) can be transformed into a nonlinear parabolic PIDE. Indeed, suppose that the transformed large trader stock-trading strategy $\psi(\tau, x) = \phi(t, S)$. Then, $V(t, S)$ solves equation (3.8) if and only if the transformed function $u(\tau, x)$ is a solution to the nonlinear parabolic equation:

$$
\begin{aligned}
\frac{\partial u}{\partial \tau} &= \frac{\sigma^2}{2} \frac{1}{(1-\rho\partial_x\psi)^2} \frac{\partial^2 u}{\partial^2 x} + \left(r - \frac{\sigma^2}{2} \frac{1}{(1-\rho\partial_x\psi)^2} - \delta(\tau, x) \right) \frac{\partial u}{\partial x} \\
&+ \int_{\mathbb{R}} \left(u(\tau, x+\xi) - u(\tau, x) - \xi \frac{\partial u}{\partial x}(\tau, x) \right) \nu(\mathrm{d}z), \quad u(0, x) = \Phi(Ke^x), \qquad (3.10)
\end{aligned}
$$

$\tau \in [0, T], x \in \mathbb{R}$. The shift function $\xi(\tau, x, z)$ is a solution to the algebraic equation:

$$
e^{\xi} = e^z + \rho(\psi(\tau, x+\xi) - \psi(\tau, x)), \qquad (3.11)
$$

and $\delta = \int_{\mathbb{R}} (e^{\xi} - 1 - \xi) \nu(\mathrm{d}z) = \int_{\mathbb{R}} (e^z - 1 - \xi + \rho(\psi(\tau, x+\xi) - \psi(\tau, x))) \nu(\mathrm{d}z)$. For small values of $0 < \rho \ll 1$, we can construct the first order asymptotic expansion $\xi(\tau, x, z) = \xi_0(\tau, x, z) + \rho\xi_1(\tau, x, z)$. For $\rho = 0$, we obtain $\xi_0(\tau, x, z) = z$. Hence

$$
e^{z+\rho\xi_1} = e^z + \rho(\psi(\tau, x+z+\rho\xi_1) - \psi(\tau, x)).
$$

Taking the first derivative of the above implicit equation with respect to ρ and evaluating it at the origin $\rho = 0$, we obtain $\xi_1 = e^{-z}(\psi(\tau, x+z) - \psi(\tau, x))$, i.e.,

$$
\xi(\tau, x, z) = z + \rho e^{-z}(\psi(\tau, x+z) - \psi(\tau, x)). \qquad (3.12)
$$

As a consequence, we obtain the following lemma. For more details on its proof and applications, see [62, Ševčovič and Udeani].

Lemma 2. *[62] Assume that the stock-trading strategy $\phi = \phi(t, S)$ is a globally ω-Hölder continuous function, $0 < \omega \le 1$. Then, the transformed function $\psi(\tau, x) = \phi(t, S)$, with $\tau = T - t$ and $x = \ln S$, is ω-Hölder continuous, and the first order asymptotic expansion $\xi(\tau, x, z)$ of the nonlinear algebraic equation (3.11) is ω-Hölder continuous in all variables. Furthermore, there exists a constant $C_0 > 0$ such that $\sup_{\tau,x} |\xi(\tau, x, z)| \le C_0 |z|^{\omega}(1 + e^{|z|})$ for any $z \in \mathbb{R}$.*

Next, consider a simplified linear approximation of (3.8), where we set $\rho = 0$ in the diffusion function with the assumption that the shift function H depends on ρ. Then, the transformed Cauchy problem for the solution u with the first-order approximation of the shift function ξ is given by

$$
\begin{aligned}
\frac{\partial u}{\partial \tau} &= \frac{\sigma^2}{2} \frac{\partial^2 u}{\partial^2 x} + \left(r - \frac{\sigma^2}{2} + \delta(\tau, x) \right) \frac{\partial u}{\partial x} \\
&+ \int_{\mathbb{R}} \left(u(\tau, x+\xi) - u(\tau, x) - \xi \frac{\partial u}{\partial x}(\tau, x) \right) \nu(\mathrm{d}z), \qquad (3.13)
\end{aligned}
$$

$\tau \in [0,T], x \in \mathbb{R}$, where $\xi(\tau,x,z) = z + \rho(\psi(\tau,x+z) - \psi(\tau,x))$. Consequently, we obtain the following theorem.

Theorem 5. *[62] Assume the transformed stock-trading strategy function $\psi(\tau,x)$ is globally ω-Hölder continuous in both variables. Suppose that ν is an admissible Lévy measure with the shape parameters $\alpha < 3, D \in \mathbb{R}$, where either $\mu > 0$, or $\mu = 0$ and $D > 1$. Let X^γ be the space of Bessel potentials space $\mathscr{L}^p_{2\gamma}(\mathbb{R})$, where $\frac{\alpha-1}{2\omega} < \gamma < \frac{p+1}{2p}$ and $\frac{1}{2} \leq \gamma < 1$. Let $T > 0$. Then, the linear PIDE (3.13) has a unique mild solution u with the property that the difference $U = u - u^{BS}$ belongs to the space $C([0,T], X^\gamma) \cap C^1([0,T], X)$.*

In the proof of Theorem 5, we note that the initial condition $u(0,\cdot)$ may not belong to X^γ since it is not smooth for $x = 0$ and grows exponentially for $x \to \infty$ (call option) or $x \to -\infty$ (put option). However, the shifted function $U = u - u^{BS}$ satisfies $U(0,\cdot) \equiv 0$, and so the initial condition $U(0,\cdot)$ belongs to X^γ. Furthermore, the shift function u^{BS} enters the governing PIDE as it includes the term $f(u^{BS}(\tau,\cdot))$ on the right-hand side. Moreover, the shift term $f(u^{BS}(\tau,\cdot))$ is singular for $\tau \to 0^+$ since $u^{BS}(0,x)$ is not sufficiently smooth for $x = 0$. In this thesis [62], we show the Hölder estimates, which are sufficient for proving the main result of Theorem 5. The exponential growth of the function u^{BS} can be overcome since $\tilde{f}(e^x) = 0$, where $\tilde{f}(u) = f(u) - \delta(\tau,\cdot)\partial_x u$, i.e.,

$$\tilde{f}(u)(x) = \int_{\mathbb{R}} \left(u(x+\xi) - u(x) - (e^\xi - 1)\partial_x u(x) \right) \nu(\mathrm{d}z).$$

Remark 6. *It is worth noting that the call/put option pay-off functions $\Phi(S) = \Phi(Ke^x) = (S-K)^+ = K(e^x-1)^+$ / $\Phi(S) = \Phi(Ke^x) = (K-S)^+ = K(1-e^x)^+$ need not belong to the Banach space X^γ. Therefore, to overcome this problem and formulate the existence and uniqueness of a solution to the PIDE (3.13), one can employ the idea of [18] by shifting the solution u by u^{BS}. Here, $u^{BS}(\tau,x) = e^{r\tau}V^{BS}(T-\tau, Ke^x)$ is an explicit solution to the linear Black-Scholes equation without the nonlocal part. In other words, u^{BS} solves the linear parabolic equation:*

$$\frac{\partial u^{BS}}{\partial \tau} - \frac{\sigma^2}{2}\frac{\partial^2 u^{BS}}{\partial x^2} - \left(r - \frac{\sigma^2}{2}\right)\frac{\partial u^{BS}}{\partial x} = 0, \quad u^{BS}(0,x) = \Phi(Ke^x),\ \tau \in (0,T), x \in \mathbb{R}. \tag{3.14}$$

Recall that $u^{BS}(\tau,x) = Ke^{x+r\tau}N(d_1) - KN(d_2)$ (call option case), where $d_{1,2} = (x + (r \pm \sigma^2/2)\tau)/(\sigma\sqrt{\tau})$ (c.f., [42, 63]). Here, $N(d) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{d} e^{-\xi^2/2}\mathrm{d}\xi$ is the cumulative density function of the normal distribution.

3.4 Discussion to Chapter 3

In this chapter, we studied the existence and uniqueness of solutions of nonlinear partial integro-differential equations (PIDEs) arising from financial modeling in multidimensional spaces. Such a PIDE models the well-known Black-Scholes equation for pricing call/put option, which corresponds to the parabolic equation of the form (1.1). We relaxed and generalized some of the existing assumptions on the Black-Scholes model. We employed the theory of abstract semilinear parabolic equations in order to prove the existence and uniqueness of solutions in the scale of Bessel potential spaces. Specifically, the corresponding nonlocal PIDE generalizes known results by Cruz and Ševčovič [18]. The novelty of this chapter is the generalization of existing results in one dimension to the multidimensional case. We considered a wide class of admissible Lévy measures satisfying suitable growth conditions near the origin and infinity. We investigated solutions to the nonlocal equation in which the shift function may depend on a prescribed large investor stock-trading strategy function. We showed the Hölder estimates, which are sufficient for proving the main result of the theorem. As an application to option pricing in the one-dimensional space, we considered a general shift function arising from a nonlinear option pricing model taking into account a large trader stock-trading strategy.

CHAPTER 4

Application of Maximal monotone operator to Portfolio Management

It is known that nonlinear parabolic equations model several physical processes that are of diffusion type, e.g., heat propagation, filtration, dynamics of biological groups, and optimal control problems. Hamilton-Jacobi-Bellman (HJB) equation, as a nonlinear partial differential equation, plays a crucial role in optimal control theory since it provides a necessary and sufficient condition for optimality. Several studies have shown that such an HJB equation can be modeled as a nonlinear diffusion equation. In this chapter, we investigate the existence and uniqueness of a solution to fully nonlinear parabolic HJB arising from optimal portfolio management. Our approach in establishing the existence results is to first transform the nonlinear HJB into a nonlinear diffusion equation, which in turn represents a nonlocal PIDE in Sobolev spaces. We consider the HJB equation arising from portfolio selection problems, where the goal of an investor is to maximize the conditional expected value of the terminal utility of the portfolio (see Udeani and Ševčovič [68]). The main results of this chapter are contained in our paper[68]

[68] C. I. Udeani and D. Ševčovič. Application of maximal monotone operator method for solving Hamilton–Jacobi–Bellman equation arising from optimal portfolio selection problem. *Japan Journal of Industrial and Applied Mathematics, Springer.* **5** (2021), pp 1–21.

4.1 Background and motivation

The goal of this chapter is to study the existence and uniqueness of a solution $\varphi = \varphi(\tau, x)$ to the Cauchy problem for the nonlinear parabolic equation

$$\partial_\tau \varphi - \Delta\alpha(\tau, \varphi) = g_0(\tau, \varphi) + \nabla \cdot \boldsymbol{g}_1(\tau, \varphi), \tag{4.1}$$

$$\varphi(\cdot, 0) = \varphi_0, \tag{4.2}$$

where $\tau \in (0, T), x \in \mathbb{R}^d, d \geq 1$, and $g_0, g_{1j} : [0, T] \times H \to H, j = 1, \cdots, n$, are globally Lipschitz continuous functions (see, the paper ([68, Udeani and Ševčovič]). Here, $H = L^2(\mathbb{R}^d)$. The diffusion function $\alpha = \alpha(x, \tau, \varphi)$ is assumed to be globally Lipschitz continuous and strictly increasing in the φ-variable. An example of such a Lipschitz continuous function $\alpha(x, \tau, \varphi)$ is the value function of the following parametric optimization problem:

$$\alpha(x, \tau, \varphi) = \min_{\boldsymbol{\theta} \in \triangle} \left(-\mu(x, t, \boldsymbol{\theta}) + \frac{\varphi}{2} \sigma(x, t, \boldsymbol{\theta})^2 \right), \quad \tau \in (0, T), x \in \mathbb{R}^d, \varphi > \varphi_{min}\,, \tag{4.3}$$

where μ, σ^2 are given C^1 functions, representing the drift and volatility, respectively, and $\triangle \subset \mathbb{R}^n$ is a compact decision set. Depending on the structure of the decision set $\triangle$, the function α is $C^{1,1}$ smooth if $\triangle$ is a convex set. But it can only be $C^{0,1}$ smooth if $\triangle$ is not connected.

Our motivation for studying the nonlinear parabolic equation of the form (4.1) for $d = 1$ arises from the dynamic stochastic programming, where the goal is to maximize the conditional expected value of the terminal utility of the portfolio:

$$\max_{\boldsymbol{\theta}|_{[0,T)}} \mathbb{E} \left[u(x_T^{\boldsymbol{\theta}}) \,\middle|\, x_0^{\boldsymbol{\theta}} = x_0 \right], \tag{4.4}$$

on a finite time horizon $[0, T]$. Here, $u : \mathbb{R} \to \mathbb{R}$ is a given increasing terminal utility function, and x_0 is a given initial state condition of the process $\{x_t^{\boldsymbol{\theta}}\}$ at $t = 0$. The underlying stochastic process $\{x_t^{\boldsymbol{\theta}}\}$ with a drift $\mu(x, t, \boldsymbol{\theta})$ and volatility $\sigma(x, t, \boldsymbol{\theta})$ is assumed to satisfy the following Itô's stochastic differential equation:

$$dx_t^{\boldsymbol{\theta}} = \mu(x_t^{\boldsymbol{\theta}}, t, \boldsymbol{\theta}_t) dt + \sigma(x_t^{\boldsymbol{\theta}}, t, \boldsymbol{\theta}_t) dW_t\,, \tag{4.5}$$

where the control process $\{\boldsymbol{\theta}_t\}$ is adapted to the process $\{x_t, t \geq 0\}$. Here, $\{W_t, t \geq 0\}$ is the standard one-dimensional Wiener process. The control parameter $\boldsymbol{\theta}$ is assumed to belong to a given compact subset $\triangle$ in $\mathbb{R}^n$. An example of such a decision set is a compact convex simplex $\triangle \equiv \mathcal{S}^n = \{\boldsymbol{\theta} \in \mathbb{R}^n \mid \boldsymbol{\theta} \geq \mathbf{0}, \mathbf{1}^T\boldsymbol{\theta} = 1\} \subset \mathbb{R}^n$, where

$\mathbf{1} = (1, \cdots, 1)^T \in \mathbb{R}^n$ or $\triangle \subset \mathcal{S}^n$ can be finite discrete set. Consider the value function

$$V(x,t) := \sup_{\boldsymbol{\theta}|_{[t,T)}} \mathbb{E}\left[u(x_T^{\boldsymbol{\theta}})|x_t^{\boldsymbol{\theta}} = x\right]. \tag{4.6}$$

subject to the terminal condition $V(x,T) = u(x)$. According to the theory of stochastic dynamic programming [12], such a value function (4.6) solves the fully nonlinear HJB equation describing the optimal portfolio selection strategy, which is given by

$$\partial_t V + \max_{\boldsymbol{\theta}\in\triangle}\left(\mu(x,t,\boldsymbol{\theta})\,\partial_x V + \frac{1}{2}\sigma(x,t,\boldsymbol{\theta})^2\,\partial_x^2 V\right) = 0\,, \tag{4.7}$$

$$V(x,T) = u(x), \tag{4.8}$$

where $x \in \mathbb{R}, t \in [0,T)$.

Next, we illustrate the transformation of the nonlinear HJB equation (4.7) into a quasilinear parabolic equation using the Ricatti transformation function.

4.2 Riccati transformation

This section presents how the HJB equation (4.7) can be transformed into a quasilinear PDE using the so-called Ricatti transformation techniques. Such a transformed parabolic equation corresponds to the Cauchy problem (4.1), which is obtained after some perturbation in the main operator. The Riccati transformation φ of the value function V can be defined based on the approach introduced by Abe and Ishimura [1], Ishimura and Ševčovič [29], Ševčovič and Macová [45], and Kilianová and Ševčovič [35] as follows:

$$\varphi(x,\tau) = -\frac{\partial_x^2 V(x,t)}{\partial_x V(x,t)}, \quad \text{where} \quad \tau = T - t. \tag{4.9}$$

Suppose the value function $V(x,t)$ is increasing in the x-variable. In other words, assume that the terminal utility function $u(x)$ is increasing. Then, the HJB equation (4.7) can be rewritten as follows:

$$\partial_t V - \alpha(\cdot,\varphi)\partial_x V = 0, \qquad V(\cdot,T) = u(\cdot), \tag{4.10}$$

where $\alpha(x,\tau,\varphi)$ is the value function of the following parametric optimization problem:

$$\alpha(x,\tau,\varphi) = \min_{\boldsymbol{\theta}\in\triangle}\left(-\mu(x,t,\boldsymbol{\theta}) + \frac{\varphi}{2}\sigma(x,t,\boldsymbol{\theta})^2\right), \quad \tau = T - t\,. \tag{4.11}$$

Remark 7. *It is worth noting that the optimization problem (4.11) is related to the classical Markowitz model on optimal portfolio selection problem formulated as max-*

imization of the mean return $\mu(\boldsymbol{\theta}) \equiv \boldsymbol{\mu}^T\boldsymbol{\theta}$ *under the volatility constraint* $\frac{1}{2}\sigma(\boldsymbol{\theta})^2 \equiv \frac{1}{2}\boldsymbol{\theta}^T\boldsymbol{\Sigma}\boldsymbol{\theta} \leq \frac{1}{2}\sigma_0^2$, *i.e.,*

$$\max_{\boldsymbol{\theta}\in\triangle} \boldsymbol{\mu}^T\boldsymbol{\theta}, \quad s.t.\ \ \frac{1}{2}\boldsymbol{\theta}^T\boldsymbol{\Sigma}\boldsymbol{\theta} \leq \frac{1}{2}\sigma_0^2,$$

where the decision set is the simplex $\triangle = \{\boldsymbol{\theta} \in \mathbb{R}^n \mid \boldsymbol{\theta} \geq \mathbf{0}, \mathbf{1}^T\boldsymbol{\theta} = 1\}$. *The Lagrange multiplier for the volatility constraint can be viewed as the parameter* φ *entering the parametric optimization problem (4.11).*

Next, let $\partial_x\alpha$ be the total differential of the function $\alpha(x,\tau,\varphi)$, where $\varphi = \varphi(x,\tau)$, i.e.,

$$\partial_x\alpha(x,\tau,\varphi) = \alpha'_x(x,\tau,\varphi) + \alpha'_\varphi(x,\tau,\varphi)\,\partial_x\varphi.$$

Here, α'_x and α'_φ are partial derivatives of α with respect to x- and φ- variables, respectively. The following remark is a consequence of the result presented by Kilianová and Ševčovič [36].

Remark 8. *The relationship between the transformed function* φ *and the value function* V *is given by the result due to Kilianová and Ševčovič [36]. With regard to [36, Theorem 4.2], an increasing value function* $V(x,t)$ *in the* x*-variable is a solution to the HJB equation (4.7) if and only if the transformed function* $\varphi(x,\tau) = -\partial_x^2 V(x,t)/\partial_x V(x,t)$, $t = T-\tau$, *is a solution to the Cauchy problem for the quasilinear parabolic PDE:*

$$\partial_\tau\varphi - \partial_x^2\alpha(\cdot,\varphi) = -\partial_x\left(\alpha(\cdot,\varphi)\varphi\right), \tag{4.12}$$
$$\varphi(x,0) = \varphi_0(x) \equiv -u''(x)/u'(x), \quad (x,\tau) \in \mathbb{R}\times(0,T). \tag{4.13}$$

It is worth noting that the Cauchy problem for the quasilinear parabolic PDE (4.12) is equivalent to the nonlinear parabolic equation (4.1) in one-dimensional space. This is obtainable after some shift/perturbation in the main operator of the transformed equation (4.12).

4.3 Preliminaries, definitions, and main results

First, we present some basic definitions and the underlying settings used in this chapter. Then, the existence and uniqueness of the solution to the parabolic equation (4.1) is established in high-dimensional spaces.

Next, we recall the following settings, which form the basis of the Sobleve spaces used in this thesis.

Definition 11. *Let* $H = L^2(\mathbb{R}^d) = \{f : \mathbb{R}^d \to \mathbb{R}, \|f\|_{L^2}^2 = \int_{\mathbb{R}^d}|f(x)|^2dx < \infty\}$ *be a Hilbert space endowed with the inner product* $(f,g) = \int_{\mathbb{R}^d} f(x)g(x)dx$. *Then, the*

Banach spaces V and V' are defined as follows:

$$V = H^1(\mathbb{R}^d), \quad V' = H^{-1}(\mathbb{R}^d).$$

The Sobolev spaces $H^s(\mathbb{R}^d)$ are defined by means of the Fourier transform

$$\hat{f}(\xi) = \frac{1}{(2\pi)^{d/2}} \int_{\mathbb{R}^d} e^{-ix\cdot\xi} f(x)dx, \quad \xi = (\xi_1, \xi_2, ..., \xi_d)^T \in \mathbb{R}^d,$$

$$H^s(\mathbb{R}^d) = \{f : \mathbb{R}^d \to \mathbb{R}, (1 + |\xi|^2)^{s/2}\hat{f}(\xi) \in L^2(\mathbb{R}^d)\}, \ s \in \mathbb{R},$$

endowed with the norm $\|f\|^2_{H^s} = \int_{\mathbb{R}^d}(1 + |\xi|^2)^s|\hat{f}(\xi)|^2 d\xi$, where $|\xi| = (\xi_1^2 + \cdots + \xi_d^2)^{1/2}$. Moreover, $V, H,$ and V' forms the Gelfand triple (V, H, V'), i.e., $V \hookrightarrow H \hookrightarrow V'$, where V' is the dual space of V.

Define the linear operator $A : V \to V'$ as follows

$$A\psi = \psi - \Delta\psi.$$

Then, A is a self-adjoint operator in the Hilbert space $H = L^2(\mathbb{R}^d)$ with the following Fourier transform representation:

$$\widehat{A\psi}(\xi) = (1 + |\xi|^2)\hat{\psi}(\xi).$$

where $\xi \in \mathbb{R}^d$ and $|\xi| = (\xi_1^2 + \cdots + \xi_d^2)^{1/2}$. Furthermore, the fractional powers of A are defined by $\widehat{A^s\psi}(\xi) = (1 + |\xi|^2)^s\hat{\psi}(\xi), \ s \in \mathbb{R}$. In particular,

$$\widehat{A^{\pm 1/2}\psi}(\xi) = (1 + |\xi|^2)^{\pm 1/2}\hat{\psi}(\xi),$$

and $A^{-1/2}$ is a self-adjoint operator in the Hilbert space $H = L^2(\mathbb{R}^d)$. Moreover, $A^{-1} = A^{-1/2}A^{-1/2}$.

In this thesis, we denote the duality pairing between the spaces V and V' by $\langle ., . \rangle \equiv \langle \cdot, \cdot \rangle_{V',V}$, i.e., the value $F(u)$ of a functional $F \in V'$ at $u \in V$ is denoted by $\langle F, u \rangle$.

4.3.1 Main results

This section is devoted to the existence results for the nonlinear parabolic equation (4.1).

Theorem 6. *([68, Ševčovič and Udeani]) Assume that the above settings on H and V hold. Assume that $g_0, g_{1j} : [0, T] \times H \to H, j = 1, \cdots, d$, be globally Lipschitz continuous functions. Suppose $\alpha \in C^{0,1}(\mathcal{D})$ is such that there exist constants $\omega, L, L_0 > 0$ such* *that* $0 < \omega \le \alpha'_\varphi(x,\tau,\varphi) \le L$, $|\nabla_x \alpha(x,\tau,\varphi)| \le p(x,\tau) + L_0|\varphi|$, $\alpha(x,\tau,0) = h(x,\tau)$ *for a.e.* $(x,\tau,\varphi) \in \mathcal{D}$, *where* $\mathcal{D} = \mathbb{R}^d \times (0,T) \times (\varphi_{min}, \infty)$, *and* $p, h \in L^\infty((0,T);H)$. *Then for any* $T > 0$ *and* $\varphi_0 \in H$, *there exists a unique solution* $\varphi \in \mathcal{V}$ *of the Cauchy problem*

$$\partial_\tau \varphi + A\alpha(\cdot,\tau,\varphi) = g_0(\tau,\varphi) + \nabla \cdot \boldsymbol{g}_1(\tau,\varphi), \qquad \varphi(\cdot,0) = \varphi_0(\cdot). \tag{4.14}$$

Theorem 6 shows the existence and uniqueness of the solution to the parabolic equation in an abstract setting, which corresponds to fully nonlinear evolutionary HJB (4.7). The proof of Theorem 6 is based on the concept of the monotone operator technique in higher-dimensional space. We employed Banach's fixed point theorem to obtain the uniqueness of the solution to the general form of the transformed parabolic equation in a suitable Sobolev space. The complete proof of Theorem 6 is contained in our recent paper (see, [68, Ševčovič and Udeani]), which is attached in the Appendix 8.2.

Under assumption of the Theorem 6, we have $\alpha(\cdot,0), g_0(\cdot,0), g_{1j}(\cdot,0) \in \mathcal{H}$. Here, the space $\mathcal{X} = L^\infty((0,T);V')$ is endowed with the norm

$$\|\varphi\|^2_{\mathcal{X}} = \sup_{\tau\in[0,T]} \|\varphi(\tau)\|^2_{V'}, \ \forall \varphi \in \mathcal{X}.$$

Consequently, we have the following theorem.

Theorem 7. *([68, Ševčovič and Udeani]) Suppose that the functions* $\alpha, g_0, g_{1j}, j = 1, \cdots, d$, *fulfills the assumptions of Theorem 6. Then, the unique solution* $\varphi \in \mathcal{V}$ *to the Cauchy problem is absolutely continuous, i.e.,* $\varphi \in C([0,T];H)$. *Moreover, there exist a constant* $\tilde{C} > 0$, *such that the unique solution satisfies the following inequality:*

$$\|\varphi\|^2_{\mathcal{X}} + \|\varphi\|^2_{\mathcal{H}} \le \tilde{C}\big(\|\varphi_0\|^2_{V'} + \|\alpha(\cdot,0)\|^2_{\mathcal{H}} + \|g_0(\cdot,0)\|^2_{\mathcal{H}} + \sum_{j=1}^{d} \|g_{1j}(\cdot,0)\|^2_{\mathcal{H}}\big), \tag{4.15}$$

where $\alpha(x,\tau,0) = h(x,\tau)$ for a.e. $(x,\tau,\varphi) \in \mathcal{D}$ and $h \in L^\infty((0,T);H)$, $\mathcal{D} = \mathbb{R}^d \times (0,T) \times (\varphi_{min},\infty)$. Here, $\mathcal{H}$ is a Hilbert space endowed with the norm

$$\|\varphi\|^2_{\mathcal{H}} = \int_0^T \|\varphi(\tau)\|^2_H d\tau, \ \forall \varphi \in \mathcal{H}.$$

Theorem 7 shows that the unique solution to the Cauchy problem (4.14) is absolutely continuous under the assumption of Theorem 7. It also shows that the unique solution $\varphi \in \mathcal{V}$ satisfies the energy estimate (4.15). The proof of Theorem 7 is contained in our paper (see, [68, Ševčovič and Udeani]), which is attached in the Appendix 8.2.

Theorem 8. *[68, Ševčovič and Udeani] Let the decision set* $\triangle \subset \mathbb{R}^n$ *be compact and the* *function $u : \mathbb{R} \to \mathbb{R}$ be an increasing utility function such that $\varphi_0(x) = -u''(x)/u'(x)$ belongs to the space $L^2(\mathbb{R}) \cap L^\infty(\mathbb{R})$. Suppose that the drift $\mu(x, \boldsymbol{\theta})$ and volatility function $\sigma^2(\boldsymbol{\theta}) > 0$ are C^1 continuous in the x and $\boldsymbol{\theta}$ variables, and the value function $\alpha(x, \varphi)$ given in (4.3) satisfies $p \in L^2(\mathbb{R}) \cap L^\infty(\mathbb{R}), h \in L^\infty(\mathbb{R})$, and $\partial_x^2 h \in L^2(\mathbb{R})$, where*

$$p(x) = \max_{\boldsymbol{\theta} \in \triangle} |\partial_x \mu(x, \boldsymbol{\theta})|, \quad h(x) = -\max_{\boldsymbol{\theta} \in \triangle} \mu(x, \boldsymbol{\theta}).$$

Then, for any $T > 0$, there exists a unique solution $\varphi \in C([0,T]; H) \cap \mathcal{H}$ of the Cauchy problem

$$\partial_\tau \varphi - \partial_x^2 \alpha(\cdot, \varphi) = -\partial_x \left(\alpha(\cdot, \varphi)\varphi\right), \quad \varphi(x, 0) = \varphi_0(x), \quad (x, \tau) \in \mathbb{R} \times (0, T), \tag{4.16}$$

satisfying $\varphi \in C([0,T]; H) \cap L^2((0,T); V) \cap L^\infty((0,T) \times \mathbb{R})$.

Theorem 8 shows the necessary and sufficient conditions (under certain assumptions) on the utility function, drift function, and volatility function for the existence and uniqueness of solution of the Cauchy problem. The complete proof of Theorem 8 is contained in our recent paper (see, [68, Ševčovič and Udeani]), which is attached in the Appendix 8.2.

4.4 Behavior of the solution with respect to the decision set

This section is devoted to the qualitative behavior of the solution of the HJB and the corresponding quasilinear parabolic equation with respect to the decision set. It first presents the properties of the value function of the parametric programming, which gives sufficient conditions for establishing our existence results. Then, the relationships between the decision set and the corresponding solution to the Cauchy problem are discussed. Specifically, we investigate and analyze the behavior of the solution $\varphi = \varphi(x, \tau)$ to the Cauchy problem (4.12), which corresponds to the HJB equation (4.7).

4.4.1 Properties of the value function

Proposition 3. *([68, Ševčovič and Udeani]) Let $\triangle \subset \mathbb{R}^n$ be a given compact decision set. Assume that the functions $\mu(x, t, \boldsymbol{\theta})$ and $\sigma(x, t, \boldsymbol{\theta})^2$ are globally Lipschitz continuous in $x \in \mathbb{R}^d, t \in [0, T]$ and $\boldsymbol{\theta} \in \triangle$ variables, and there exist positive constants $\omega, L > 0$ such that $\omega \le \frac{1}{2}\sigma(x, t, \boldsymbol{\theta})^2 \le L$ for any $x \in \mathbb{R}^d, t \in [0, T]$, and $\boldsymbol{\theta} \in \triangle$.*

Then $\alpha \in C^{0,1}(\mathcal{D})$. Moreover, the function α is strictly increasing, and

$$0 < \omega \le \frac{\alpha(x,\tau,\varphi_2) - \alpha(x,\tau,\varphi_1)}{\varphi_2 - \varphi_1} \le L, \quad \text{for any } (x,\tau,\varphi_i) \in \mathcal{D}, \tag{4.17}$$

i.e., $\omega \le \alpha'_\varphi(x,\tau,\varphi) \le L$, and

$$|\nabla_x \alpha(x,\tau,\varphi)| \le p(x,\tau) + L_0|\varphi|, \tag{4.18}$$

for a.e. $(x,\tau,\varphi) \in \mathcal{D}$, where $p(x,\tau) := \max_{\boldsymbol{\theta}\in\triangle} |\nabla_x \mu(x,t,\boldsymbol{\theta})|$ and $L_0 := \max_{\boldsymbol{\theta}\in\triangle, t\in[0,T], x\in\mathbb{R}^d} |\nabla_x \sigma^2(x,t,\theta)|$ where $t = T - \tau$.

Proposition 3 shows the qualitative properties of the value function and sufficient conditions imposed on the decision set $\triangle$ and functions μ and σ that guarantee higher smoothness of the value function α. It shows the conditions for which the value function α belongs to $C^{0,1}(\mathcal{D})$, where $\mathcal{D} = \mathbb{R}^d \times (0,T) \times (\varphi_{min}, \infty)$, which is crucial in the assumption of Theorem 6. Moreover, Proposition 3 relaxes the assumption of Lipschitz continuity of the drift function $\mu(x,t,\boldsymbol{\theta})$ and volatility function $\sigma(x,t,\boldsymbol{\theta})^2$ in the θ-variable. In other words, it shows that we only require the boundedness of the functions $\mu(x,t,\boldsymbol{\theta})$ and $\sigma(x,t,\boldsymbol{\theta})^2$ in the θ-variable for the value function α to belong to $C^{0,1}(\mathcal{D})$, where $\mathcal{D} = \mathbb{R}^d \times (0,T) \times (\varphi_{min}, \infty)$. The complete proof of Proposition 3 is contained in our recent paper (see, [68, Ševčovič and Udeani]), which is attached in the appendix.

Theorem 9. *[37, Theorem 1] Suppose that $\triangle \subset \mathbb{R}^n$ is a convex compact set, and the functions $\mu(x,t,\boldsymbol{\theta})$ and $\sigma(x,t,\boldsymbol{\theta})^2$ are $C^{1,1}$ smooth such that the objective function $f(x,t,\varphi,\boldsymbol{\theta}) := -\mu(x,t,\boldsymbol{\theta}) + \frac{\varphi}{2}\sigma(x,t,\boldsymbol{\theta})^2$ is strictly convex in the variable $\boldsymbol{\theta} \in \triangle$ for any $\varphi \in (\varphi_{min}, \infty)$, then the function α belongs to the space $C^{1,1}(\mathcal{D})$.*

Theorem 9 was proved in [37]. It gives sufficient conditions imposed on the decision set $\triangle$ and functions μ and σ guaranteeing higher smoothness of the value function α. Its proof is based on the classical envelope theorem due to Milgrom and Segal [50] and the result on Lipschitz continuity of the minimizer $\hat{\boldsymbol{\theta}} = \hat{\boldsymbol{\theta}}(x,\tau,\varphi)$ belonging to a convex compact set $\triangle$ due to Klatte [40].

4.5 Numerical examples

This subsection presents examples of the value function $\alpha = \alpha(\varphi)$ to the parametric optimization problem with different decision sets. First, we consider a simple decision set $\triangle = \{\boldsymbol{\theta} \in \mathbb{R}^2, \boldsymbol{\theta} \ge 0, \mathbf{1}^T\boldsymbol{\theta} = 1\}, n = 2, \mu(\boldsymbol{\theta}) = \boldsymbol{\mu}^T\boldsymbol{\theta}, \sigma^2(\boldsymbol{\theta}) = \boldsymbol{\theta}^T\boldsymbol{\Sigma}\boldsymbol{\theta}$, where $\boldsymbol{\Sigma}$ is a positive definite covariance matrix, and $\boldsymbol{\mu}$ is a positive vector of mean return. The value function $\alpha = \alpha(\varphi)$ can be explicitly expressed as follows:

$$\alpha(\varphi) = \begin{cases} E^-\varphi + D^-, & \text{if } 0 < \varphi \leq \varphi_*^-, \\ A - \frac{B}{\varphi} + C\varphi, & \text{if } \varphi_*^- < \varphi < \varphi_*^+, \\ E^+\varphi + D^+, & \text{if } \varphi_*^+ \leq \varphi. \end{cases}$$

Here, $(\varphi_*^-, \varphi_*^+)$ is the maximal interval where the optimal value $\hat{\boldsymbol{\theta}}(\varphi) \in \triangle$ of the function $\boldsymbol{\theta} \mapsto -\boldsymbol{\mu}^T\boldsymbol{\theta} + \frac{\varphi}{2}\boldsymbol{\theta}^T\boldsymbol{\Sigma}\boldsymbol{\theta}$ is strictly positive $(\hat{\boldsymbol{\theta}}(\varphi) > 0)$ for $\varphi \in (\varphi_*^-, \varphi_*^+)$, and $C, E^\pm > 0$, $B \geq 0$, $A, D^\pm$ are constants explicitly depending on the covariance matrix $\boldsymbol{\Sigma}$ and the vector of mean return $\boldsymbol{\mu}$ such that the function α is C^1 continuous at $\varphi_*^\pm$, i.e., $E^\pm = B/(\varphi_*^\pm)^2 + C$ and $D^\pm = A - B/\varphi_*^\pm + C\varphi^\pm - E^\pm\varphi_*^\pm$. It is clear that α is only $C^{1,1}$ continuous function with two points $\varphi_*^\pm$ of discontinuity of the second derivative α''.

Furthermore, consider a decision set consisting of finite number of points. Then, the value function $\alpha(\varphi)$ corresponding to such decision set is only piece-wise linear. In other words, if $\hat{\triangle} = \{\boldsymbol{\theta}^1, \cdots, \boldsymbol{\theta}^k\} \subset \{\boldsymbol{\theta} \in \mathbb{R}^2, \boldsymbol{\theta} \geq 0, \mathbf{1}^T\boldsymbol{\theta} = 1\}$, then $\alpha(\varphi) = \min_{i=1,\cdots,k} \alpha^i(\varphi)$, where $\alpha^i(\varphi) = E^i\varphi + D^i$ is a linear function with the slope $E^i = (1/2)(\boldsymbol{\theta}^i)^T\boldsymbol{\Sigma}\boldsymbol{\theta}^i > 0$ and intercept $D^i = -\boldsymbol{\mu}^T\boldsymbol{\theta}^i$.

Figure 4.2 (a) shows the graph of the value function α corresponding to the Slovak pension fund system with the two types of decision sets. According to the data-set obtained from [38], the portfolio consists of the stock index (with a high mean return $\mu_s = 0.10$ and high volatility $\sigma_s = 0.3$) and bonds (with mean return $\mu_b = 0.03$ and very low volatility $\sigma_s = 0.01$). The returns on stocks index and bonds are negatively correlated $\varrho = -0.15$. $\boldsymbol{\mu} = (\mu_s, \mu_b)^T = (0.1, 0.05)^T$. Then $\boldsymbol{\Sigma}_{11} = \sigma_2^2, \boldsymbol{\Sigma}_{22} = \sigma_b^2, \boldsymbol{\Sigma}_{12} = \boldsymbol{\Sigma}_{21} = \varrho\sigma_s\sigma_b$. As shown in Figure 4.2 (a), the solid blue line corresponds to the convex compact decision set $\triangle = \{\boldsymbol{\theta} \in \mathbb{R}^2, \boldsymbol{\theta} \geq 0, \mathbf{1}^T\boldsymbol{\theta} = 1\}$. The piece-wise linear value function α (dotted red line) corresponds to the discrete decision set $\hat{\triangle} = \{\boldsymbol{\theta}^1, \boldsymbol{\theta}^2, \boldsymbol{\theta}^3\} \subset \triangle$. It represents the Slovak pension fund system consisting of three funds: growth funds with $\boldsymbol{\theta}^1 = (0.8, 0.2)^T$ (80% of stocks and 20% of bonds), balanced funds with $\boldsymbol{\theta}^2 = (0.5, 0.5)^T$ (equal proportion of stocks and bonds), and conservative funds with $\boldsymbol{\theta}^3 = (0, 1)^T$ (only bonds) (c.f. [38]). Figure 4.2 (b) shows the graph of the second derivative $\alpha''_\varphi(\varphi)$ of the value function $\alpha(\varphi)$ corresponding to the convex compact decision set $\triangle$. It has the first point of discontinuity φ_*^- close to the value 2. For $n > 2$, the number of discontinuities of α''_φ increases (c.f. [35]). Figure 4.3 shows another example of the value function and its second derivative for the portfolio consisting of five stocks (BASF, Bayer, Degussa-Huls, FMC, and Schering) entering DAX30 German stocks index. The covariance matrix $\boldsymbol{\Sigma}$ is taken from [20]. We set the vector of yields $\boldsymbol{\mu} = (0.03, 0.02, 0.04, 0.01, 0.01)^T$.

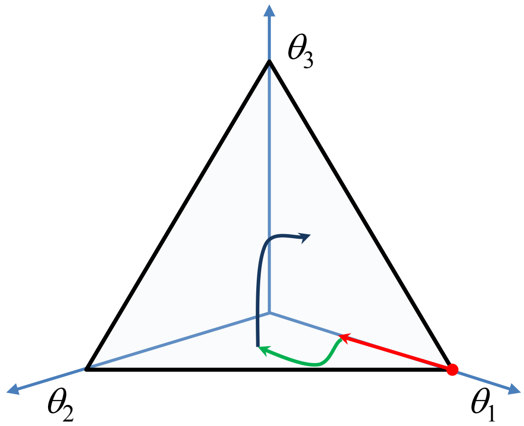


Figure 4.1: The path $\hat{\boldsymbol{\theta}}(\varphi)$ as a function of φ [61].

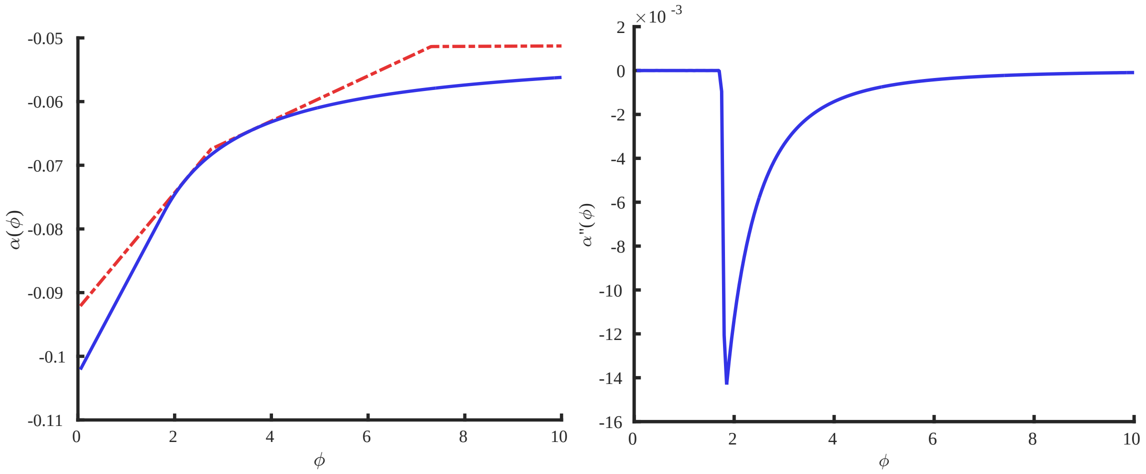
The minimizer $\hat{\boldsymbol{\theta}}(\varphi)$ of the convex optimization problem

$$\alpha(\varphi) = \min_{\boldsymbol{\theta}\in\triangle}\{-\boldsymbol{\mu}^T\boldsymbol{\theta} + \frac{\varphi}{2}\boldsymbol{\theta}^T\Sigma\boldsymbol{\theta}\}$$

when considered as a function of the risk aversion parameter φ is only Lipschitz continuous in φ. According to Millgrom–Segal envelope theorem, the derivative $\alpha'(\varphi)$ is given by $\alpha'(\varphi) = \frac{1}{2}\hat{\boldsymbol{\theta}}(\varphi)^T\Sigma\hat{\boldsymbol{\theta}}(\varphi)$. Figure 4.1 shows the path $\hat{\boldsymbol{\theta}}(\varphi)$ as a function of φ when increasing φ from $\varphi = 0$ to $\varphi \to \infty$. For small values of φ, only one asset with maximal mean return is active, i.e., $\theta_1 > 0, \theta_2 = \theta_3 = 0$. For intermediate values of φ, two assets are active $\theta_1 > 0, \theta_2 > 0, \theta_3 = 0$. Moreover, for larger values of φ, all three assets are active, i.e., $\theta_1 > 0, \theta_2 > 0, \theta_3 > 0$. The path $\varphi \mapsto \hat{\boldsymbol{\theta}}(\varphi)$ has a discontinuity in the first derivative when it leaves lower dimensional object (vertex, edge) and enter a higher-dimensional object volume.

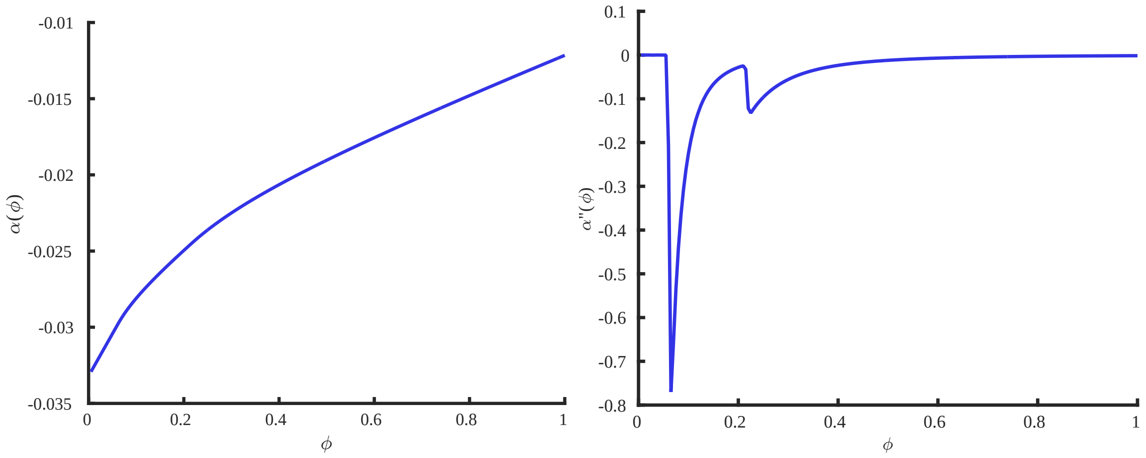
The advantage of the Riccati transformation of the original HJB is twofold. First, the diffusion function α can be computed in advance as a result of quadratic optimization problem when the vector $\boldsymbol{\mu}$ and the covariance matrix $\boldsymbol{\Sigma}$ are given or semidefinite programming problem when they belong to a uncertainity set of returns and covariance matrices (c.f. [39]). Figure 4.4 shows the vector of optimal weights $\boldsymbol{\theta}$, as a function of the parameter φ, obtained as the optimal solution to the quadratic optimization problem with the covariance matrix from [20] corresponding to the five assets (BASF, Bayer, Degussa–Huls, FMC, Scheringfrom) entering DAX30 index from 2008. There are more nontrivial weights θ_i when the parameter φ increases.

In contrast to the fully nonlinear character of the original HJB equation (4.7), the transformed equation (4.12) represents a quasilinear parabolic equation in the divergence form. Thus, efficient numerical schemes can be constructed for this class of equation. In our computational experiments, we employ the finite volume discretization scheme proposed and investigated by Kilianová and Ševčovič [35, 36, 37]). Figure 4.5



a) b)

Figure 4.2: a) A graph of the value function α, b) its second derivative $\alpha''(\varphi)$ for the portfolio consisting of the stocks index and bonds (c.f. [38]) for the convex compact decision set $\triangle$. The dotted line in a) corresponds to the discrete decision set $\hat{\triangle} = \{\boldsymbol{\theta}^1, \boldsymbol{\theta}^2, \boldsymbol{\theta}^3\} \subset \triangle$. Source: our computation is based on the method from [68].



a) b)

Figure 4.3: a) A graph of the value function $\alpha(\varphi)$, and b) the second derivative $\alpha''_\varphi(\varphi)$ corresponding to five stocks (BASF, Bayer, Degussa–Huls, FMC Scheringfrom) entering DAX30 index. Source: our computation is based on the method from [32, 68].

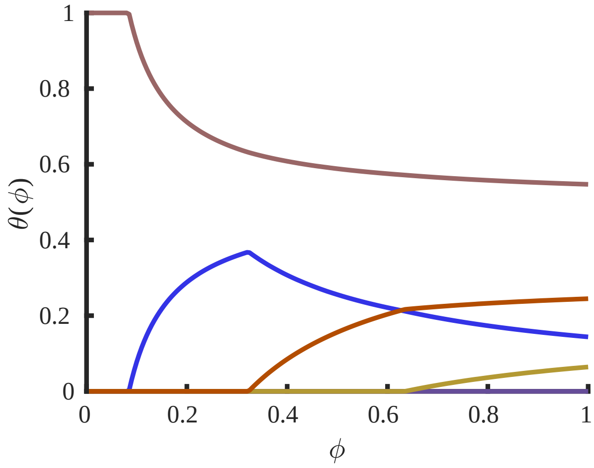


Figure 4.4: The optimal vector $\boldsymbol{\theta} = (\theta_1, \cdots, \theta_n)^T$ as a function of φ for the German DAX30 index. Source: our computation is based on the method from [35, 68]

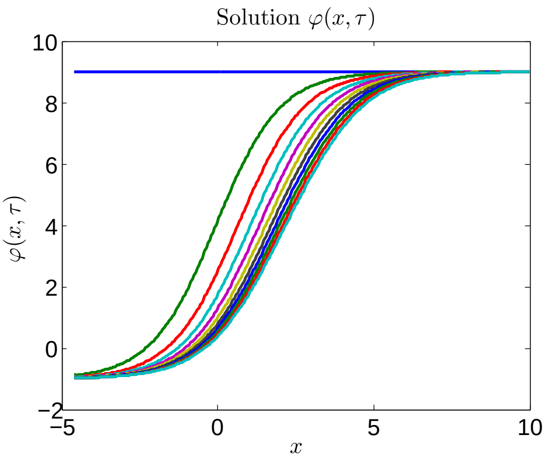


Figure 4.5: A solution $\varphi(x, \tau)$ for the DARA utility function with $a_0 = 9$, $a_1 = 8$, $x^* = 2$. Source: our computations based on the numerical method from [32, 68].

shows the results of time dependent sequence of profiles $\varphi(x, \tau)$ for a constant initial condition $\varphi_0 \equiv 9$. This figure also shows the solution profiles for the initial condition φ_0 attaining four decreasing values $\{9, 8, 7, 6\}$. It represents DARA utility function. The function $\varphi(x, \tau)$ is increasing in the x variable and decreasing in the $\tau = T - t$ variable. Therefore, the optimal vector $\boldsymbol{\theta}(x, \tau)$ contains more diversified portfolio of assets when x increases and the time $t \to T$ (see Figure 4.4). Furthermore, it is reasonable to invest in an asset with the highest expected return when the account value x is low, whereas an investor has to diversify the portfolio when x is large and time t is approaching terminal maturity T.

4.5.1 Qualitative behavior of the solution

Now, we study the behavior of the solution $\varphi = \varphi(x, \tau)$ to (4.12), where the value function $\alpha(x, \tau, \varphi)$ of the parametric optimization problem is subject to two decision sets. Suppose $\tilde{\triangle}$ and $\triangle$ are two decision sets such that $\tilde{\triangle} \subseteq \triangle$, corresponding to the value functions $\tilde{\alpha}(x, \tau, \tilde{\varphi})$ and $\alpha(x, \tau, \varphi)$, respectively. We study the behavior of the corresponding solutions $\tilde{\varphi}(x, \tau)$ and $\varphi(x, \tau)$ to the parabolic equation (4.12) with respect to the two decision sets. The goal is to examine if these solutions are comparable. First, we consider the case when the corresponding drift and volatility functions are constant. In this case, the solutions associated with the two decision sets are not generally comparable when one of the decision sets is a subset of the other. Meanwhile, we can only obtain a direct comparison between the two solutions if in addition one of the solutions is convex in the x-variable. Some counter examples are presented to demonstrate this behavior.

Let $\tilde{\triangle}$ and $\triangle$ be two decision sets such that $\tilde{\triangle} \subseteq \triangle$. Then, the corresponding HJB equations associated with decision sets $\tilde{\triangle}$ and $\triangle$ are given by

$$\partial_t \tilde{V} + \max_{\tilde{\boldsymbol{\theta}} \in \tilde{\triangle}} \left(\tilde{\mu}(x, t, \tilde{\boldsymbol{\theta}})\, \partial_x \tilde{V} + \frac{1}{2} \tilde{\sigma}(x, t, \tilde{\boldsymbol{\theta}})^2\, \partial_x^2 \tilde{V} \right) = 0, \tag{4.19}$$

$$\partial_t V + \max_{\boldsymbol{\theta} \in \triangle} \left(\mu(x, t, \boldsymbol{\theta})\, \partial_x V + \frac{1}{2} \sigma(x, t, \boldsymbol{\theta})^2\, \partial_x^2 V \right) = 0, \tag{4.20}$$

where $\tilde{V}(x, t)$ and $V(x, t)$ are the corresponding solutions for $t \in [0, T], x > 0$, satisfying the same terminal condition $\tilde{V}(x, T) = V(x, T) = u(x)$. Macova and Ševčovič [45] established that for $\tilde{\triangle} \subseteq \triangle$, $\tilde{\mu} = \mu$, and $\tilde{\sigma} = \sigma$, the value function V is a super-optimal solution to equation (4.19). In other words, they obtained that

$$\partial_t \tilde{V} + \max_{\tilde{\boldsymbol{\theta}} \in \tilde{\triangle}} \left(\tilde{\mu}(x, t, \tilde{\boldsymbol{\theta}})\, \partial_x \tilde{V} + \frac{1}{2} \sigma(x, t, \tilde{\boldsymbol{\theta}})^2\, \partial_x^2 \tilde{V} \right) \leq 0.$$

Consequently, they obtained that $\tilde{V}(x, t) \leq V(x, t)$, for any $t \in [0, T]$ and $x > 0$. In this study, we investigate the behavior of the corresponding parabolic equation, given as follows.

Let φ and $\tilde{\varphi}$ be solutions corresponding to the parabolic equation with $\triangle$ and $\tilde{\triangle}$, respectively, i.e.,

$$\partial_\tau \varphi - \partial_x^2 \alpha(\cdot, \varphi) = -\partial_x \left(\alpha(\cdot, \varphi) \varphi \right), \tag{4.21}$$

$$\varphi(x, 0) = \varphi_0(x), \quad (x, \tau) \in \mathbb{R} \times (0, T). \tag{4.22}$$

$$\partial_\tau \tilde{\varphi} - \partial_x^2 \alpha(\cdot, \tilde{\varphi}) = -\partial_x \left(\alpha(\cdot, \tilde{\varphi}) \tilde{\varphi} \right), \tag{4.23}$$

$$\tilde{\varphi}(x, 0) = \varphi_0(x), \quad (x, \tau) \in \mathbb{R} \times (0, T). \tag{4.24}$$

where $\alpha(x,\tau,\varphi)$ and $\tilde{\alpha}(x,\tau,\tilde{\varphi})$ are given, respectively, as follows:

$$\alpha(x,\tau,\varphi) = \min_{\boldsymbol{\theta}\in\triangle}\left(-\mu(x,t,\boldsymbol{\theta}) + \frac{\varphi}{2}\sigma(x,t,\boldsymbol{\theta})^2\right), \quad \tau = T-t, \tag{4.25}$$

$$\tilde{\alpha}(x,\tau,\tilde{\varphi}) = \min_{\boldsymbol{\theta}\in\tilde{\triangle}}\left(-\tilde{\mu}(x,t,\boldsymbol{\theta}) + \frac{\tilde{\varphi}}{2}\tilde{\sigma}(x,t,\boldsymbol{\theta})^2\right), \quad \tau = T-t. \tag{4.26}$$

The following lemma further illustrates the properties of the diffusion function α with respect to the two decision sets $\tilde{\triangle}$ and $\triangle$.

Remark 9. *Suppose $\tilde{\triangle}$ and $\triangle$ are two decision sets such that $\tilde{\triangle} \subseteq \triangle$, corresponding to the volatility functions $\tilde{\sigma}^2, \sigma^2$ and drift functions $\tilde{\mu}, \mu$, respectively. Then,*

(i) if $\sigma^2(x,t,\theta) \le \tilde{\sigma}^2(x,t,\theta) \implies \alpha(x,t,\varphi) \le \tilde{\alpha}(x,t,\varphi)$;

(i) $\mu(x,t,\theta) \le \tilde{\mu}(x,t,\theta) \implies \alpha(x,t,\varphi) \le \tilde{\alpha}(x,t,\varphi)$.

This can be obtained from the principle of a minimum of (4.25) on a subset of a set. Our goal is to investigate the relationship between the solutions of the two parabolic equations (4.25) and (4.26) with respect to the two decision sets. To this end, we present counter examples to demonstrate that these solutions need not be comparable based on the decision set.

4.5.2 Examples

In this subsection, we present some examples to demonstrate the behavior of the solutions to the parabolic equation corresponding to the fully nonlinear HJB equation (4.7). In the following two examples, we consider the case when the drift and volatility functions are constant. The first example shows that if one of the decision set is a subset of the other (say, $\tilde{\triangle} \subseteq \triangle$), the two corresponding solutions are comparable provided that one of the solutions is convex. However, the second example shows that the two solutions are not generally comparable. In what follows, we assume that $\tilde{\triangle} \subseteq \triangle$.

Example 1. *Consider a pair of the following parabolic equations:*

$$\begin{cases} \partial_\tau\varphi + \frac{\sigma^2}{2}\partial_x^2\varphi = 0, \\ \varphi(x,0) = x^2, \end{cases} \tag{4.27}$$

$$\begin{cases} \partial_\tau\tilde{\varphi} + \frac{\tilde{\sigma}^2}{2}\partial_x^2\tilde{\varphi} = 0, \\ \tilde{\varphi}(x,0) = x^2, \end{cases} \tag{4.28}$$

where $\alpha(\varphi) = \frac{\sigma^2}{2}$ and $\tilde{\alpha}(\tilde{\varphi}) = \frac{\tilde{\sigma}^2}{2}$ correspond to the decision sets $\triangle$ and $\tilde{\triangle}$, respectively. Here, σ and $\tilde{\sigma}$ are the volatilities corresponding to the decision sets $\triangle$ and $\tilde{\triangle}$, respectively. In both examples, we assume that the corresponding drifts are zero. Following a suitable approach for solving parabolic equations, we obtain solutions to parabolic equations (4.27) and (4.28), respectively, as follows:

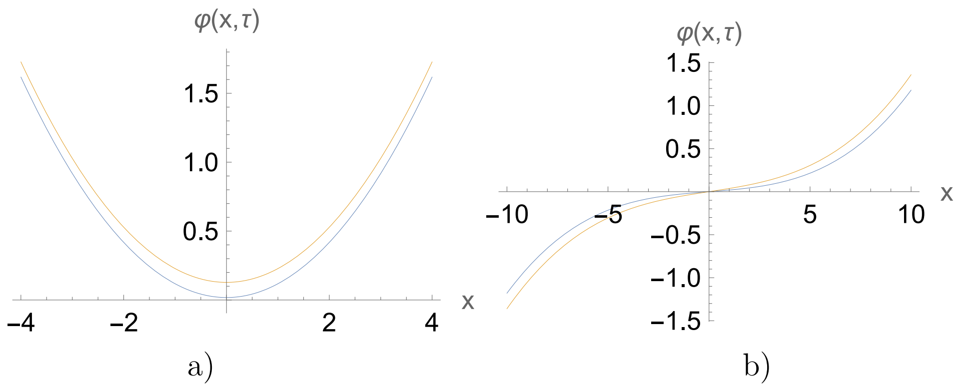


Figure 4.6: a) Graphs of the solution φ and $\tilde{\varphi}$ corresponding to the decision sets $\triangle$ and $\tilde{\triangle}$, respectively

$$\varphi(\tau, x) = x^2 + \sigma^2\tau, \qquad \tilde{\varphi}(\tau, x) = x^2 + \tilde{\sigma}^2\tau.$$

Since $\tilde{\triangle} \subseteq \triangle$, we have that $\alpha(x, \tau, \varphi) \leq \tilde{\alpha}(x, \tau, \tilde{\varphi})$ (see Remark 9). This implies that $\sigma^2 \leq \tilde{\sigma}^2$. Therefore, we obtain

$$\varphi(\tau, x) = x^2 + \sigma^2\tau \leq x^2 + \tilde{\sigma}^2\tau = \tilde{\varphi}(\tau, x), \text{ for all } x \in \mathbb{R}, \tau > 0.$$

Example 2. *Now, consider a pair of the following parabolic equations:*

$$\begin{cases} \partial_\tau \varphi + \frac{\sigma^2}{2}\partial_x^2 \varphi = 0, \\ \varphi(x, 0) = x^3, \end{cases} \tag{4.29}$$

$$\begin{cases} \partial_\tau \varphi + \frac{\sigma^2}{2}\partial_x^2 \varphi = 0, \\ \varphi(x, 0) = x^3. \end{cases} \tag{4.30}$$

It can easily be shown that the following functions φ and $\tilde{\varphi}$ solve the parabolic equations (4.29) and (4.30), respectively

$$\varphi(\tau, x) = x^3 + 3\sigma^2\tau x, \qquad \tilde{\varphi}(\tau, x) = x^3 + 3\tilde{\sigma}^2\tau x$$

for all $x \in \mathbb{R}$. Here, σ^2 and $\tilde{\sigma}^2$ are the volatilities corresponding to the decision sets $\triangle$ and $\tilde{\triangle}$, as stated in the previous example. However, the relationship $\varphi(\tau, x) \leq \tilde{\varphi}(\tau, x)$ holds only when $x > 0$, and it needs not be true for $x < 0$. Therefore, in general, the solutions φ and $\tilde{\varphi}$ are not comparable. In other words, we cannot really tell the behavior of the solution to such a parabolic equation corresponding to HJB equation when one of the decision sets is a subset of the other, suggesting that additional information is needed to obtain an appropriate comparison.

Figure 1(a) shows the behavior of the solutions $\varphi(x, \tau)$ and $\tilde{\varphi}(x, \tau)$ corresponding to the decision sets $\triangle$ and $\tilde{\triangle}$, as depicted in example (4.27). Since $\tilde{\triangle} \subseteq \triangle$, which implies that $\alpha(x, \varphi) \leq \tilde{\alpha}(x, \tilde{\varphi})$, we chose the volatilities, $\sigma^2 = 0.09$, $\tilde{\sigma}^2 = 0.64$, and $\tau = 2$. In Figure 4.6(a), the yellow curve denotes the solution $\tilde{\varphi}(x, \tau)$ corresponding to the decision set $\tilde{\triangle}$, whereas the blue curve denotes the solution $\varphi(x, \tau)$ correspond to the decision sets $\triangle$. For $x \in \mathbb{R}$, we found that the solutions corresponding to the two decision sets are comparable in this example. This could be because one of the solutions is convex in the x-variable. Figure 4.6(b) shows the behavior of the two solutions corresponding to the decision sets $\triangle$ and $\tilde{\triangle}$, as illustrated in example (4.29). Here, we set the volatilities $\sigma^2 = 0.04$, $\tilde{\sigma}^2 = 0.81$, and $\tau = 4$. The yellow line corresponds to the solution $\tilde{\varphi}(x,\tau)$ associated with the decision set $\tilde{\triangle}$, whereas the blue line corresponds to the decision sets $\triangle$. Thus, this figure demonstrates that the two solutions are not comparable.

In summary, given two decision sets $\tilde{\triangle}$ and $\triangle$ such that one of them is a subset of the other. We do not have a direct comparison for the corresponding solution to the associated parabolic equation of the form (4.12), when the corresponding volatility and drift functions are constant. In other words, if $\tilde{\triangle} \subseteq \triangle$, and the associated volatility and drift functions are constant, we do not have a direct comparison between the solution φ and $\tilde{\varphi}$ to the parabolic equations (4.21) and (4.23), respectively. Example 1 demonstrates that the convexity of one of the solutions is crucial to obtain a direct comparison. In contrast, Example 2 shows that there is no direct comparison if none of the solutions is convex with respect to the x-variable.

4.6 Discussion to Chapter 4

In this chapter, we investigated the existence and uniqueness of a solution to a fully nonlinear parabolic Hamilton-Jacobi-Bellman (HJB) equation arising from optimal portfolio management. Our approach in establishing the existence results is to first transform the HJB into a diffusion equation, which in turn represents a nonlinear partial differential equation (PIDE) in Sobolev spaces. The HJB equation represents a stochastic optimization problem, where the goal of an investor is to maximize the conditional expected value of the terminal utility of the portfolio. Then, the Cauchy problem is studied in high-dimensional spaces. We employ the monotone operator technique, Fourier transform approach, and Banach's fixed point theorem to obtain the existence and uniqueness of the solution to the general form of the transformed parabolic equation in high-dimensional spaces. The existence results are based on the properties of the value function of the parametric programming, which gives sufficient conditions (under certain assumptions) on the utility function, drift function, and volatility function for establishing our existence results. We also established the necessary and sufficient conditions on the utility function, drift function, and volatility function for the proof of our existence and uniqueness results. Furthermore, the relationships between the decision set and the corresponding solutions to the HJB equation with respect to the two decision sets are discussed. Specifically, we investigate and analyze the qualitative behavior of the solutions of the two equations with respect to two different decision sets. To study this relationship, we assume that one of the decision sets is properly contained in the other. Consequently, we found that, in general, there is no ordering between the corresponding solutions with respect to the decision set. Finally, some counter examples are presented to demonstrate this behavior.

CHAPTER 5

Conventional numerical method

This chapter presents a detailed discussion of conventional numerical methods for solving PDE, such as the finite difference and finite element methods. The finite difference approach forms the baseline of the deep learning method used in the next chapter.

5.1 Finite Difference Method

Although there are several numerical methods for solving differential equations [18, 58, 63], the finite difference method (FDM) is a versatile technique for solving parabolic equations and other types of PDEs, and it forms the foundation for many numerical solvers used in computational science and engineering. FDM has been widely used to approximate the solutions of differential equations arising from different fields (see [63]). The finite difference approach discretizes the continuous PDE into a set of algebraic equations defined on a grid. It uses direct discrete points system interpretation to define the equation and uses the combination of all the points to produce the system equation. The mathematical formulation of FDM involves approximating the derivatives in the differential equation using finite differences. The solution is then obtained by solving a system of linear equations that arise from the discretization process. The specific details of the method may vary depending on the type of differential equation and the boundary conditions (see [58, 63]). However, the following fundamental concepts are involved in solving parabolic equations using the finite difference approach.

5.1.1 Fundamental Concept of FDM

The specific formulation for solving parabolic equations (e.g., heat equation) using FDM is given as follows.

I. **Discretization of the Domain:** The first step in solving parabolic equations using FDM is to discretize the spatial and temporal domains. The spatial domain (e.g., a physical space or a rod) is divided into a grid of points and time domain is discretized into time steps. This creates a mesh or grid over which the approximate solutions are computed.

II. **Time Stepping:** Since parabolic equations involve time evolution, a time-stepping scheme, such as the implicit or explicit method, is chosen to update the solution at each time step. Implicit methods typically require solving a linear or nonlinear system of equations at each step, whereas explicit methods involve simple algebraic operations.

III. **Difference Equations:** The next step is to approximate the derivatives in the original PDE with finite difference approximations. For example, the spatial derivative can be approximated using central differences, forward differences, or backward differences. These approximations represent how the quantity changes from one grid point to another in space.

IV. **Initial and Boundary Conditions:** Parabolic equations usually require initial conditions (conditions at time $t = 0$) and boundary conditions (conditions along the spatial boundaries). These conditions are applied to the grid points at the initial time step and updated as the simulation progresses.

V. **Time-Marching Algorithm:** Then, a time-marching algorithm is used to advance the solution in time. At each time step, the values at grid points are computed based on the previous time step's values and the finite difference approximations. Implicit methods often involve solving linear systems (like tridiagonal systems), whereas explicit methods use simple update formulas.

VI. **Stability and Convergence:** This is the most crucial part of this technique. For an accurate approximation, the stability and convergence of the numerical scheme must be guaranteed. In other words, one must ensure that the numerical scheme is stable and convergent. Stability ensures that small perturbations do not lead to unbounded growth in the solution, and convergence guarantees that as the grid spacing and time step decrease, the numerical solution approaches the true solution of the PDE.

VII. **Error Analysis:** In this step, the accuracy of the numerical solution is evaluated by comparing it with the exact solution (if available) or by studying the behavior as the grid and time step are refined. Error analysis helps in estimating the numerical errors and improving the accuracy of the solution.

VIII. **Implementation:** This step involves implementing the finite difference method in a programming environment, such as Python or MATLAB, to execute the simulation. This is achieved by iterating through time steps until the desired simulation time is reached, thereby monitoring the solution's evolution.

5.2 Application to Option Pricing

In this section, we employ FDM to illustrate the numerical solution of the parabolic equation for option pricing. As a practical example, we consider the transformed equation (3.13), which corresponds to the classical Black-Scholes equation used in finance for pricing options. In the transformed Cauchy equation (3.13), we consider the case where $z = 0$, which means $\xi = 0$. We also assume that δ is a constant. For this illustration, we shall employ the implicit scheme for the time derivative, and a central discretization for the first order space derivative.

First, we restrict the theoretical infinite domain to the finite domain $[t_0, T] \times [L_1, L_2]$, with $L_1 < L_2$. Then, the region $[t_0, T] \times [L_1, L_2]$ is replaced by a discrete grid ($t_0 = 0$). For $n = 0, 1, \cdots, N_t \in \mathbb{N}$, define the discrete time step $\Delta t = \frac{T - t_0}{N_t}$ such that $t_n = t_0 + n\Delta t$. For $i = 0, 1, \cdots, N_x \in \mathbb{N}$, define the discrete space step $\Delta x = \frac{L_2 - L_1}{N_x}$ such that $x_i = L_1 + i\Delta x$. We divide the grid into equally spaced nodes of distance Δx and Δt in the x- and $t-$axes, respectively, with the mesh points of the form $(t_0 + n\Delta t, L_1 + i\Delta x)$. At this point, we are interested in the values of $u(t, x)$ in the mesh nodes. We let

$$u(t_0 + n\Delta t, L_1 + i\Delta x) = u_i^n,$$

where $i = 0, 1, \cdots, N_x$ and $n = 0, 1, \cdots, N_t$. In this simulation, our interest is to find the value of u at time t_0. The algorithm involves finding the values u^n given the knowledge of the values u^{n+1}. Then, the discretized form of the parabolic equation (3.13) becomes

$$\frac{u_i^{n+1} - u_i^n}{\Delta t} + (r - \frac{1}{2}\sigma^2 + \delta)\frac{u_{i+1}^n - u_{i-1}^n}{2\Delta x} + \frac{1}{2}\sigma^2 \frac{u_{i+1}^n + u_{i-1}^n - 2u_i^n}{\Delta x^2} = 0, \tag{5.1}$$

where δ is a constant, as defined in equation (3.13). Here, we set $\delta = 0.0025$. Then, by rearranging the terms in (5.1), we obtain

$$u_i^{n+1} = u_{i-1}^n \Big(\underbrace{(r - \frac{1}{2}\sigma^2 + \delta)\frac{\Delta t}{2\Delta x} - \frac{1}{2}\sigma^2 \frac{\Delta t}{\Delta x^2}}_{a} \Big) + u_i^n \Big(\underbrace{1 + \sigma^2 \frac{\Delta t}{\Delta x^2}}_{b} \Big) + u_{i+1}^n \Big(\underbrace{-(r - \frac{1}{2}\sigma^2 + \delta)\frac{\Delta t}{2\Delta x} - \frac{1}{2}\sigma^2 \frac{\Delta t}{\Delta x^2}}_{c} \Big).$$

Next, by renaming the coefficients, we have

$$u_i^{n+1} = a u_{i-1}^n + b u_i^n + c u_{i+1}^n,$$

which can be expressed in matrix form as follows:

$$\begin{pmatrix} u_1^{n+1} \\ u_2^{n+1} \\ \vdots \\ u_{N_x-2}^{n+1} \\ u_{N_x-1}^{n+1} \end{pmatrix} = \begin{pmatrix} b & c & 0 & \cdots & 0 \\ a & b & c & 0 & 0 \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & 0 & a & b & c \\ 0 & 0 & 0 & a & b \end{pmatrix} \cdot \begin{pmatrix} u_1^n \\ u_2^n \\ \vdots \\ u_{N_x-2}^n \\ u_{N_x-1}^n \end{pmatrix} + \begin{pmatrix} a u_0^n \\ 0 \\ \vdots \\ 0 \\ c u_{N_x}^n \end{pmatrix}$$

Then, the corresponding algebraic system is given by

$$u^{n+1} = \mathcal{D} u^n + B,$$

where $\mathcal{D}$ is a tridiagonal metrics, and B corresponds to the boundary terms. The algebraic system can then be solved for u^n by inverting the matrix $\mathcal{D}$. To solve this equation, we consider a call option with strike K at maturity T. It is worth noting that the stock price S_0 is not relevant for the algorithm and will be used to compute the option value at S_0. Next, we set $L_1 = \log(K/3)$ and $L_2 = \log(3K)$, which are obtained by choosing the computational region between $K/3$ and $3K$; $u_0^n = 0$ and $u_{N_x}^n = 3K - Ke^{-r(T-t)}$. For the PDE (3.13), we set $\delta = 0.0025$. The values of the parameter are summarized as follows. The interest rate r is set to 0.1; the volatility σ is set to 0.2, the strike price K is set to 100; the time to expiration T is set to 1. The initial value of the stock price S_0 is set to 100. These financial parameters play a crucial role in determining the behavior of the option price. For instance, a higher risk-free interest rate r tends to increase the present value of future cash flows, leading to a higher option price. Conversely, a lower interest rate r has the opposite effect. A higher volatility increases the potential price movements of the underlying asset, which generally leads to higher option prices. An increase in the initial stock price generally increases the value of a call option, as there is a higher probability that the option will end up in the money. The initial log stock price is used as a reference point in the calculations. It influences the starting point for the option price evolution but does not directly impact the sensitivity of the option price to changes in other parameters. A higher strike price for a call option tends to decrease its value as the option becomes less likely to end up in the money. The longer the time to expiration, the higher the option price, as there is more time for the underlying asset's price to move in a favorable direction.

The option price is calculated backward in time, starting from the expiration time and going back to the present. The final results are then plotted to visualize the option price surface and its evolution over time. The option price is obtained at the initial log stock price. Figure 5.1 shows the line plot demonstrating how the Black-Scholes model approximates the option price compared to its payoff at the initial time. As shown in Figure 5.1, the green dotted line corresponds to the option price, and the thick blue line represents the payoff. The 3D surface plot illustrates a comprehensive view of how the option price evolves over time and stock price (Figure 5.2). These visualizations provide insight into the behavior of option prices under the Black-Scholes model. The finite difference method allows for the estimation of option prices at different points in time, leading up to the present.

5.3 Discussion to Chapter 5

The finite difference method (FDM) is a versatile technique for solving parabolic equations (PDEs) and other types of PDEs, and it forms the foundation for many numerical solvers used in computational science and engineering. For several machine learning techniques for solving differential equations, once the FDM is established, it can be considered as a baseline for a machine learning approach. The idea is to replace the traditional finite difference scheme with a neural network that can learn the spatiotemporal patterns from data. This involves training a neural network to approximate the solution to the PDE given initial and boundary conditions. The neural network would take spatial and temporal coordinates as input and output of the corresponding solution. Training data would consist of pairs of input-output examples generated from the known solutions to the PDE. This approach is part of a broader field known as scientific machine learning, where machine learning techniques are applied to scientific problems. It is worth noting that the choice between traditional numerical methods like FDM and machine learning approaches depends on the problem at hand, the amount of available data, and the desired level of accuracy. Traditional methods often provide insights into the underlying physics, while machine learning methods can handle complex, high-dimensional data and may offer more flexibility in capturing intricate patterns. In this

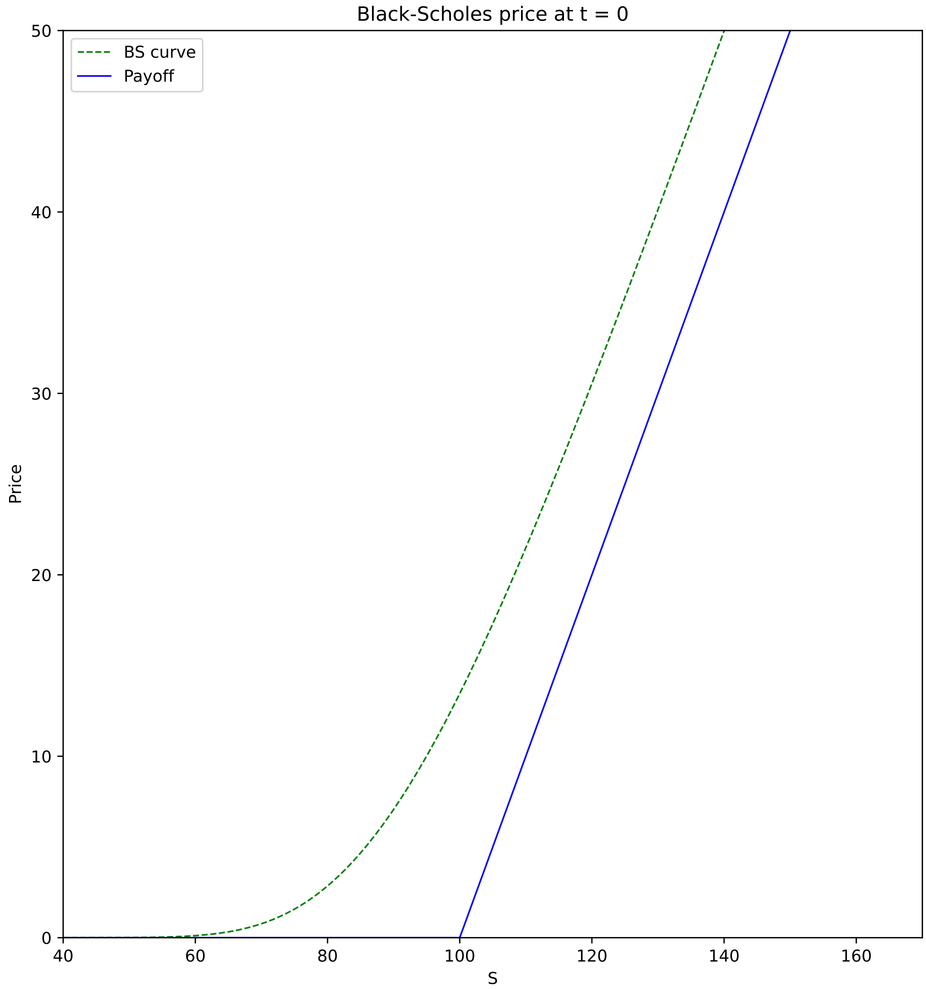


Figure 5.1: The Black-Scholes price at $t = 0$

thesis, the FDM approach forms the baseline of the deep learning approach used in Chapter 6 for solving the transform parabolic equation. In this chapter, we presented a detailed discussion of traditional numerical methods for solving PDE, such as the finite differences and finite element methods. As a practical example, we consider the transformed equation (3.13), which corresponds to the classical Black-Scholes equation used in finance for pricing options. We employ the implicit scheme for the time derivative and a central discretization for the first-order space derivative. The FDM technique allows for the estimation of option prices at different points in time, leading up to the present. The final results are then plotted to visualize the option price surface and its evolution over time.

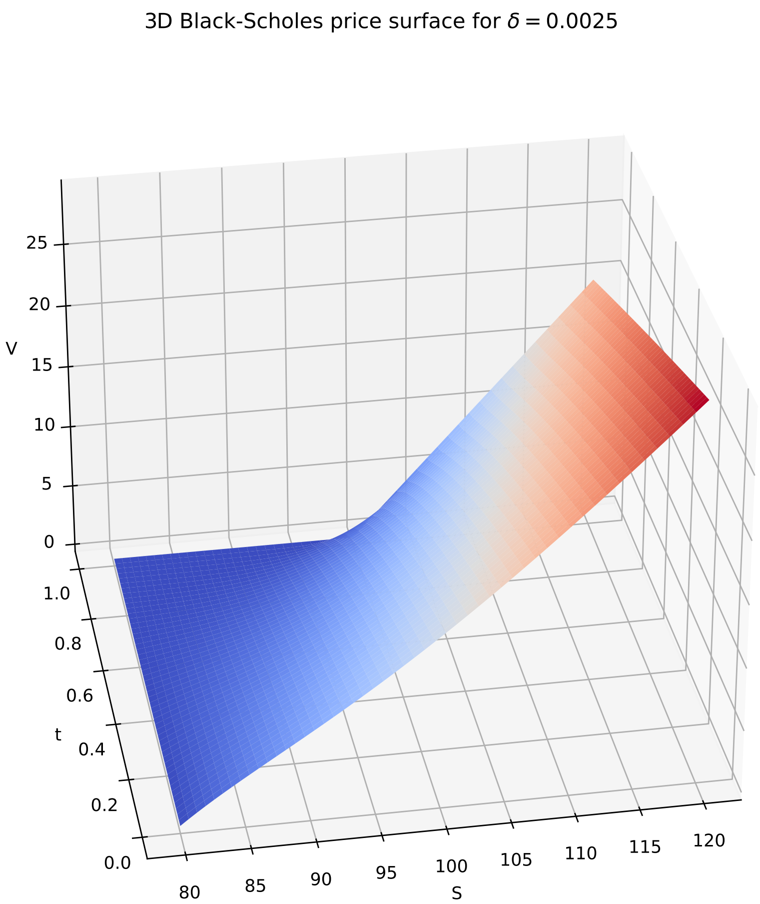


Figure 5.2: 3D Black-Scholes surface for $\delta = 0.0025$

CHAPTER 6

Application of Deep Learning for Solving nonlinear parabolic PDEs

This chapter is devoted to the application of deep learning techniques to solve the nonlinear parabolic equation (1.1). We consider the transformed part of this equation, which corresponds to Black-Scholes and HJB equations for option pricing and portfolio managment, respectively. Our goal is to employ physics-informed neural networks and physics-informed DeepONet to approximate the solution of the associated transformed parabolic equation. The main results of this chapter are contained in our paper [60]

6.1 Introduction

It is well-known that several differential equations arise from many scientific and engineering fields for modeling physical phenomena. However, most of these differential equations are analytically intractable, especially in high-dimensional space. The traditional methods for solving differential equations, including the finite volume method, finite difference method, finite element method, and spectral methods (e.g., Fourier-spectral method) often faced different challenges. For instance, they usually require high computational costs, and their convergence properties have not been properly investigated. Additionally, in the numerical solution of partial differential equation (PDE) problems through the discretization process using finite difference approximations, the algebraic systems generated are finalized using an iterative method. Although these methods are efficient and well-studied, they require much memory space

[60] D. Ševčovič and C. I. Udeani. Learning the solution operator of a nonlinear parabolic equation using physics-informed deep operator network. *ECMI2023 Proceedings, Springer.* (2023).

and time, leading to high computational errors. In order to overcome these challenges, many researchers have replaced traditional numerical discretization methods with artificial neural networks (ANNs) to approximate the PDE solution. Recently, deep neural networks (DNNs) have been widely used to solve classical applied mathematical problems, including PDEs, utilizing machine learning and artificial intelligence approaches [33]. Due to significant nonlinearities, convection dominance, or shocks, some PDEs are difficult to solve using standard numerical approaches. To this end, deep learning has recently emerged as a new paradigm of scientific computing thanks to the universal approximation theorem and great expressivity of neural networks [15]. Despite the significant breakthrough of machine learning in science, solving differential equations using deep learning naively usually leads to lack of interpretation, poor generalization, and lots of training data. Although some machine learning approaches, such as physics-informed neural networks, have proved to overcome the above challenges. However, a slight change in the underlying parameters governing the differential equation could result in the retraining of the model. Therefore, in this chapter, we employ the physics-informed DeepONet (PI-DeepONet) to approximate the solution operator of a fully nonlinear partial differential equation arising from finance. PI-DeepONet incorporates known physics into the neural network, which consists of a deep neural network that learns the solution of the PDE and an operator network that enforces the PDE at each iteration. As a model, we consider the HJB equation arising from the stochastic optimization problem, where the goal of an investor is to maximize the conditional expected value of the terminal utility of a portfolio. The fully nonlinear HJB equation is first transformed into a quasilinear parabolic equation using the Ricatti transform. Then, the solution of the transformed quasilinear equation is approximated using PI-DeepONet.

6.2 Deep Neural Networks

6.2.1 Artificial Neural Networks

Neural networks has been widely used as function approximators, which provides a new tool for machine learning and numerical analysis. ANNs, often called neural networks, are computing systems based on the collection of connected units or nodes called neurons. It consists of input, hidden, and output layers, which are connected with sets of weights and biases. The simplest ANN is a perceptron, and a network consisting of two or more hidden layers is called a multilayer perceptron. ANN is a machine learning used to solve large-scale machine learning problems. Multilayer perceptrons are the simplest ANNs, which are widely used to approximate functions. According to recent studies, ANN can efficiently be used to approximate the PDE

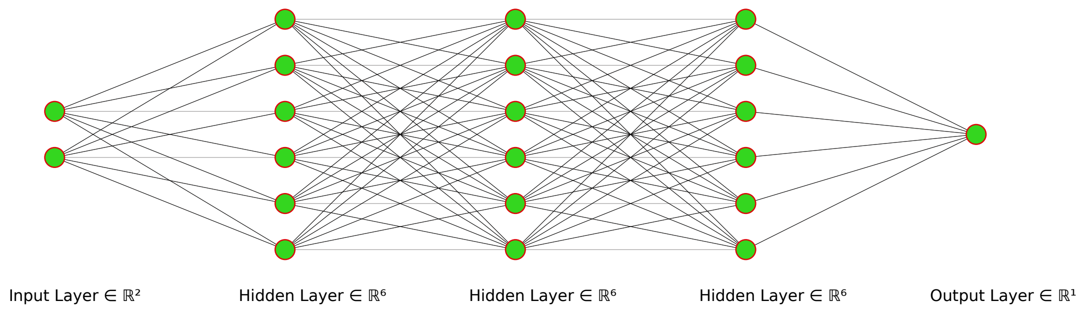


Figure 6.1: Neural network architecture with two input vectors, three hidden layers (each with six neurons), and one output vector.

solution because of its advantages, such as continuous and differentiable solutions, good interpolation properties, and less memory-intensive. It can also break the curse of dimensionality, which has been the shortfall of many traditional numerical methods. Figure 6.1 shows a simple neural network architecture consisting of one input layer (two input vectors), three hidden layers (each with six neurons), and one output layer, which are connected in a feedforward manner.

In the neural network, each node of is called a neuron, and each neuron is a function of all the neurons from the previous layers. Suppose the output from the previous layers is x, and let the function of this neuron be given by $h(x)$. Then,

$$h(x) = \sigma(\mathbf{w}^{\mathbf{T}} \cdot \mathbf{x} + \mathbf{b})$$

where σ is a nonlinear activation function, $\mathbf{w}$ and $\mathbf{b}$ are the weight and bias, which are the parameters of the neural networks. The basic concept in training every neural network model is to minimize these parameters. Neurons within the network collectively create a nonlinear mapping from the input to the output. This mapping is learned through a process known as backpropagation, which involves adjusting the weights of each neuron. It starts at the network's output and traverses the graph in the opposite direction, which continues until we reach the input layer. For training the neural network, a crucial component is a loss function that accepts the output vector from the neural networks and corresponding labels or correct values, measuring the disparity between them. Generally, the L_2 loss is commonly employed as a loss function in the context of continuous predictions. The process of updating each layer involves computing the gradients of the loss function concerning the weights, which usually requires change rule techniques.

6.2.2 Physics-informed Neural Networks (PINNs)

PINNs are scientific machine learning techniques that can approximate PDE solutions by training a neural network to minimize a loss function, which includes terms reflecting the initial and boundary conditions along the space-time domain's boundary and the PDE residual at the collocation points. The basic concept behind PINN training is that it can be considered an unsupervised learning that does not require labeled data, such as results from traditional simulations or experiments [57]. The PINN algorithm is a mesh-free technique that approximates PDE solutions by converting the problem of directly solving the governing equations into a loss function optimization problem. It works by integrating the mathematical model into the network and reinforcing the loss function with a residual term from the governing equation, which acts as a penalizing term to restrict the space of acceptable solutions. The neural network is designed to approximate the unknown function or solution to the differential equation. The input to the network typically includes the independent variables (e.g., time, space) involved in the differential equation. The output represents the approximation of the solution. The loss function consists of two main components: a data fidelity term and a physics-informed term. The data fidelity term ensures that the neural network solution matches the available data (if any). The physics-informed term enforces the differential equation itself, incorporating derivatives of the network output. During training, the neural network adjusts its parameters to minimize the combined loss function. The optimization process seeks to find a set of parameters that simultaneously satisfy the given data and the underlying physics described by the differential equation. The physics-informed term in the loss function involves taking partial derivatives of the neural network output with respect to the input variables and combining them with the differential equation. The process of computing the derivative with respect to the input values is carried out by automatic differentiation using exact expressions with floating point values rather than symbolic strings. This ensures that the neural network learns to satisfy the governing physics of the system.

Next, we discuss the approach and general implementation of PINN for solving differential equations. Consider the general differential equation of the form:

$$\frac{\partial u}{\partial t} + F(u; \gamma) = 0, \tag{6.1}$$

where $u(x, t)$ is the unknown solution of the equation, x and t are spatial and temporary variables, respectively, F denote a linear/nonlinear differential operator parameterized by γ. For a data-driven solution, the parameter γ is fixed, and the solution is learned using the network [57]. In contrast, in the data-driven discovery of PDEs, the network tries to find the best parameter γ that best describes the observed data. In this chapter of this thesis, we consider the data-driven solution technique. Since we are considering the data-driven solution approach, the parameter γ is fixed. Then, for a fix γ and using (6.1), we define the function f as

$$f := \frac{\partial u}{\partial t} + F(u; \gamma), \tag{6.2}$$

where f denotes the residual. It is worth noting that the function $f(x,t)$ and $u(t,x)$ have the same input parameters, i.e., x and t. This allows the incorporation of the PDE residual into a loss function to be minimized since PINNs require further differentiation to evaluate differential operator in the PDE. These derivatives can be easily computed through automatic differentiation with current state-of-the-art machine learning libraries [57]. The parameters of $u(t,x)$ and $f(t,x)$ are then learned by minimizing the corresponding loss function.

To solve such a parabolic equation using PINN, the neural network architecture takes x and t as inputs and outputs u, the solution of the equation, as output. The loss function consisting of the data fidelity and physics-informed parts is constructed as follows. The data fidelity term is constructed to ensure that the neural network solution matches the available data, usually the boundary and initial condition. Let $u_{obs}(x_i^u, t_i^u)$ denote the observed data at some specific points (x_i^u, t_i^u). Then, the data fidelity term of the loss function is given with respect to mean square errors loss as

$$MSE^u = \frac{1}{N^u} \sum_{i=1}^{N^u} \left| u_{pred}(x_i^u, t_i^u) - u_{obs}(x_i^u, t_i^u) \right|^2,$$

where $u_{pred}(x_i^u, t_i^u)$ denotes the predicted value evaluated at the initial and boundary conditions (x_i^u, t_i^u), and $u_{obs}(x_i^u, t_i^u)$ denotes the function values evaluated on the initial and boundary conditions.

The physics-informed term enforces the parabolic PDE, which involves taking partial derivatives of the neural network output and combining them with the PDE. The derivatives of the neural network can be obtained using automatic differentiation. The MSE loss of the physics is given by

$$MSE^f = \frac{1}{N^f} \sum_{i=1}^{N^f} \left| f(x_i^f, t_i^f) \right|^2,$$

where $\{x_i^f, t_i^f\}_{i=1}^{N^f}$ denotes the collocation points, where the PDE is been evaluated.

The total loss function is then obtained by combining the data fidelity and physics-informed terms. Thus, PINN is trained by minimizing the following MSE loss function using a suitable optimizer [57]

$$MSE = MSE^u + MSE^f.$$

The first term ensures that the initial and boundary conditions of the PDE are satisfied, whereas the second term encourages the PINN to learn the structural information expressed by the PDE during the training process. The neural network is trained on a dataset that contains both the observed data and the spatial-temporal domain of interest, and the network parameters are optimized to minimize the total loss. This approach leverages the flexibility of neural networks to approximate the solution to parabolic PDE while enforcing the governing physics.

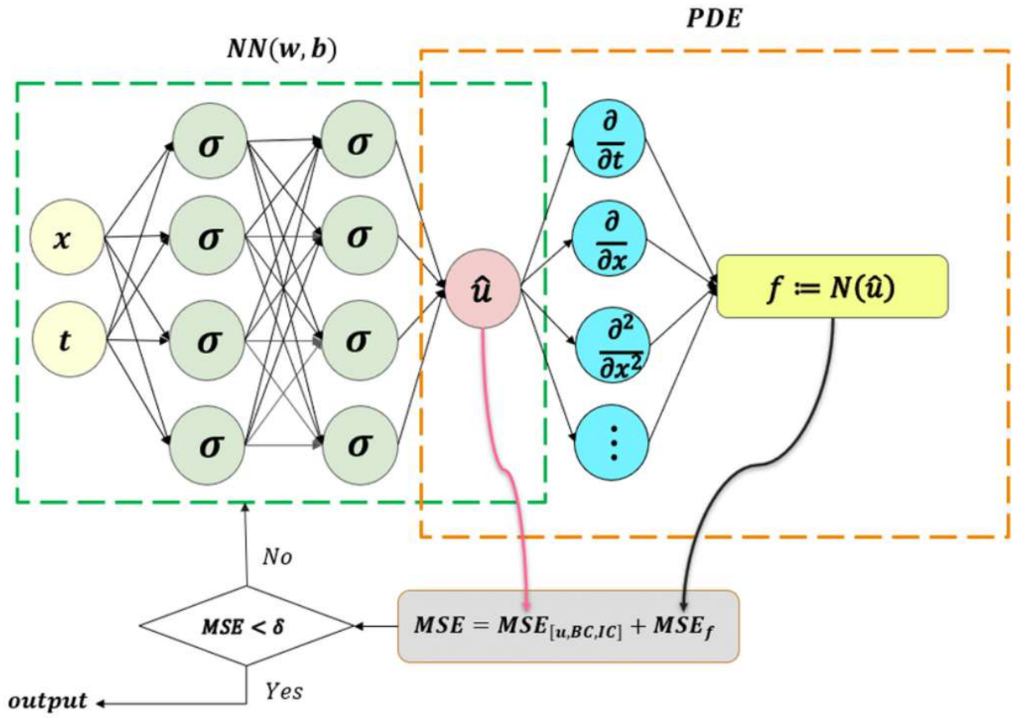


Figure 6.2: Simple architecture of physics informed neural network for solving partial differential equations [71]

6.2.3 Automatic Differentiation

To update the parameters of the deep neural networks during training, an automatic differentiation (AD) is used. AD is a technique used in deep learning to efficiently compute the gradients of the solutions with respect to the input parameters of the PDEs. It allows for efficient optimization of the parameters, which in turn leads to more accurate solutions for the PDEs. There are different approaches to using AD for solving PDEs in deep learning. One common method is to represent the solution of the PDE as a neural network, with the input being the parameters of the PDE and the output being the solution. The gradients of the output with respect to the input parameters can then be computed using AD, which can be used to optimize the parameters and improve the accuracy of the solution. Another approach is to use AD to compute the gradients of the loss function with respect to the input parameters of the PDE. This can be used in combination with numerical methods for solving PDEs, such as finite element methods or finite difference methods, to improve the accuracy of the solution. More specifically, AD works by computing the gradients of a function by recursively applying the chain rule of differentiation to the sub-expressions of the function. It is worth noting that this process can be done automatically by a computer algorithm without requiring the user to manually derive and code the gradients. In this thesis, we used adaptive moment estimation (Adam) and L-BFGS. Adam adapts the learning rate of each weight based on the historical gradients. It has been widely used for many deep learning applications because it employs the benefits of both momentum-based and RMSProp-based methods.

6.3 Comparison between Traditional Neural Network, PINNs, and and Finite Difference Approximation

In this section, we give a detailed comparison of traditional numerical methods (e.g., finite difference method), traditional neural networks, and PINNs. The choice between these methods depends on factors such as problem complexity, data availability, interpretability requirements, and computational resources. PINNs offer a middle ground by combining aspects of both neural networks and traditional numerical methods, making them particularly useful for problems where explicit physics-based constraints are crucial. Traditional numerical methods are reliable but may lack the flexibility and adaptability exhibited by PINNs in certain scenarios. Traditional neural networks, while powerful, may struggle with enforcing physics constraints without explicit guidance in the loss function. Table 6.1 shows a detailed comparison of the three methods.

Feature	**PINNs**	**Traditional NNs**	**FDM**
Physics Incorporation	Embeds physics equations in the loss function	May not explicitly incorporate physics	Directly solves PDEs based on discretization
Training Data	Requires limited training data (sparse points)	May require a large dataset	Does not require data; operates on a grid
Data Efficiency	Highly data-efficient due to physics-based constraints	May require a large dataset to generalize	Data-independent; depends on grid resolution
Mesh Independence	Exhibits mesh independence; suitable for irregular grids	May struggle with irregular meshes	Grid-dependent; regular grid often used
Boundary Conditions Handling	Enforces boundary conditions during training	Requires explicit enforcement or inclusion	Enforces boundary conditions directly
Adaptability to New Constraints	Easily adapts to additional physics constraints	May require significant architecture changes	Modifications needed for new constraints
Interpretability	Provides insights into physical quantities through the loss function	Interpretability may be challenging	Solution interpretation is straightforward
Computational Cost	Generally computationally efficient	Can be computationally expensive, especially for large networks	Computational cost depends on grid size
Applicability to Complex Geometries	Well-suited for problems with complex geometries	May struggle with complex geometries	May handle complex geometries with refinement
Versatility	Versatile across various physical systems	Limited to specific problems	Versatile but may require problem-specific discretization
Robustness to Noisy Data	Resilient to noise due to physics-based constraints	Sensitive to noisy training data	Sensitive to noise; filtering techniques used
Learning Curve	Moderate learning curve due to specific loss formulation	Standard learning curve for neural networks	Steeper learning curve due to numerical aspects
Problem Types	Well-suited for PDEs in physics and engineering	Applicable to a wide range of problems	Widely used for PDEs in various disciplines

Table 6.1: Comparison between PINNs, traditional neural networks, and traditional numerical methods based on PDE solution

6.4 Application of PI-DeepONet to portfolio managment

6.4.1 Methodology of PI-DeepONet

This section discusses the methodology of PI-DeepONet.

Consider the equation

$$\mathcal{F}(g, \varphi) = 0, \tag{6.3}$$

where $\mathcal{F}$ is a differential operator that denotes the governing PDE of some underlying physics laws, g denotes the source term of the PDE, and φ is the solution to the PDE. Let G be an operator between two infinite-dimensional function spaces defined by

$$G : g \to G(g),$$

where g and $G(g)$ are two functions. This mapping is called the solution operator of the PDE (6.3), which can be evaluated at a random location y. In learning an operator in a more general setting, the inputs usually consist of two independent parts: the input function $\{g(x_i)\}_{i=1}^{m}$ and the location variable (s) y. This learning can be done by directly using traditional neural networks like feedforward neural networks, recurrent neural networks, convolutional neural networks, or combining the two inputs as a single network input, ($i.e.$, $\{g(x_1), g(x_2), \cdots, g(x_m), y\}$). Meanwhile, it is not necessarily advisable to directly use recurrent neural networks or convolutional neural networks since the input does not have a definite structure. Therefore, it is recommended to use feedforward neural networks as the baseline model. The DeepONet consists of two subnetworks: branch and trunk nets. The branch net takes g, which represents a function evaluated at a collection of fixed sensors $\{x_i\}_{i=1}^{m}$, as the input and outputs a feature embedding of q dimensions. The trunk net takes the coordinates y as the inputs and also outputs a feature embedding of q dimensions. It is worth noting that the dimension of y needs not be equal to that of the input function g. This indicates that g and y need not be treated as a single network like in traditional neural networks. In general, the DeepONet network for learning an operator takes g and y as the inputs and outputs $G(g)(y)$. The final output of DeepONet is obtained by combining the outputs of the branch and trunk nets via dot product. Consequently, the physics-informed DeepONet is trained by minimizing the following loss function over all the input-output triplets $\{g_i, y_i, G(g_i)(y_i)\}_{i=1}^{N}$

$$\mathcal{L}(\theta) = \mathcal{L}_{Operator}(\theta) + \mathcal{L}_{Physics}(\theta), \tag{6.4}$$

where $\mathcal{L}_{Operator}(\theta) = \|G_\theta(g)(y) - G(g)(y)\|^2$, $\mathcal{L}_{Physics}(\theta) = \|\mathcal{F}[G_\theta(g)](y)\|^2$, and $\mathcal{F}$

is the differential operator for the governing PDE of the underlying physics laws, θ denote the set of weight matrix and bias vector in the networks. The first goal is to find such approximator $G_\theta(g)$, but thanks to the universal approximation theorem for the operator due to T. Chen and H. Chen [15], which guarantees the existence of such function, i.e., $G_\theta(g)(y) \approx G(g)(y) = \varphi(y) \in \mathbb{R}$. The final aim is to find the best parameters that minimize the loss function (6.4) using suitable optimization techniques. Figure 6.3 shows the schematic of DeepONet. The universal approximation theorem for operators is given as follows:

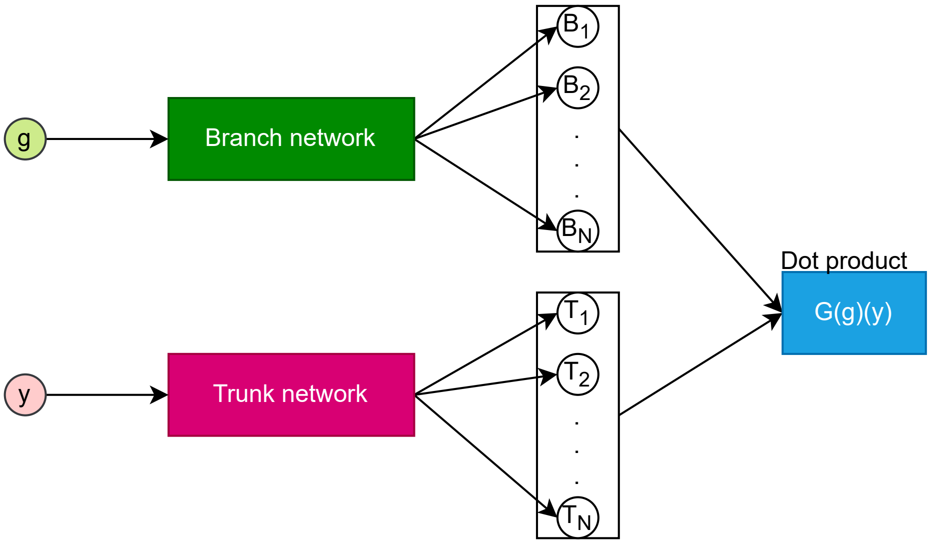


Figure 6.3: Schematic of DeepONet

Theorem 10. *[15](Universal Approximation Theorem for Operator)*
Let σ be a continuous non-polynomial function, and let X be a Banach space. Suppose $K_1 \subset X$ and $K_2 \subset \mathbb{R}^n$ are two compact sets, V is a compact set in $C(K_1)$, and G is nonlinear continuous operator that maps V into $C(K_2)$. Then, for any $\epsilon > 0$, there exist positive integers, $N, P,$ and m, constants $c_i^k, W_{bij}^k, b_{bi}^k, b_{tk} \in \mathbb{R}, W_{tk} \in \mathbb{R}^n, x_i \in K_1, i = 1, \cdots, N, k = 1, \cdots, P, j = 1, \cdots, m$ such that

$$|G(g)(y) - \underbrace{\sum_{k=1}^{P} \sum_{i=1}^{N} c_i^k \sigma \left(\sum_{j=1}^{m} W_{bij}^k g(x_i) + b_{bi}^k \right)}_{branch} \cdot \underbrace{\sigma(W_{tk} \cdot y + b_{tk})}_{trunk}| < \epsilon$$

for all $g \in V$ and $y \in K_1$, where $C(K_l)$ are the Banach space of all continuous functions on $K_l, l = 1, 2$ with the maxi-norm. Theorem 10 shows the stacked and unstacked DeepONet. It was proved in [15] with two-layer neural networks. It is worth noting that Theorem 10 holds when the Banach spaces $C(K_1)$ and $C(K_2)$ are replaced with $L^p(K_1)$ and $L^q(K_2), p, q \geq 1$. The stacked network has one trunk net and P

stacked branch nets, whereas the unstacked network has one trunk net and one branch net, which are independently fully connected. For more details see, T. Chen and H. Chen [15]. In this thesis, we use an unstacked PI-DeepONet to solve a parametric parabolic equation arising from portfolio selection problems. Figure 6.4.1 shows the architecture of PI-DeepONet used in this thesis.

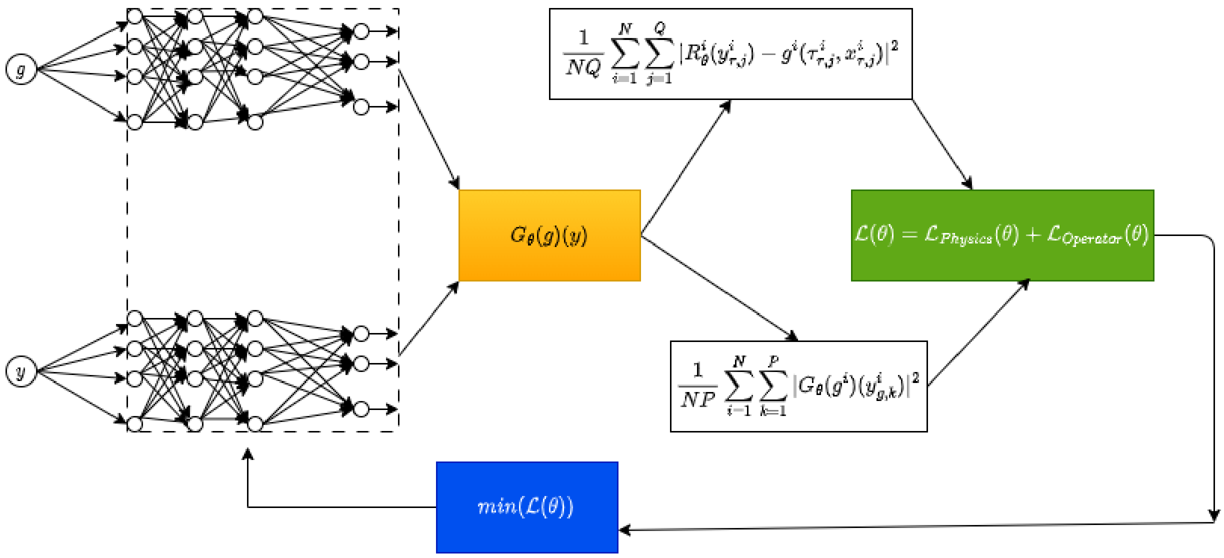


Figure 6.4: Simple architecture of physics informed DeepONet [60]

6.4.2 Problem formulation

Let $\Omega = [0, T] \times [0, 1]$. Our goal is to solve the following parabolic equation arising from mathematical finance using PI-DeepONet

$$\partial_\tau \varphi - \partial_x^2 \alpha(\varphi) = g(\tau, x), \ (\tau, x) \in \Omega, \tag{6.5}$$

with zero initial and boundary conditions, where $\varphi(\tau, x)$ is an unknown solution, and $g(\tau, x)$ is the source term function of two independent variables, τ and x, which is a parametric function that can take a wide range of values. It is worth noting that the domain of such a source term function could be infinite functional space. In this thesis, we consider a parabolic equation arising from an optimal control problem where the goal of an investor is to maximize the conditional expected value of terminal utility. It corresponds to the HJB parabolic equation (4.7), which can be obtained using a suitable Ricatti transformation (see Chapter 4) with some shift in the main operator (see Abe and Ishimura [1], Ishimura and Ševčovič [29], Ševčovič and Macová [45], and Kilianová and Ševčovič [35], Ševčovič and Udeani [68]). The solution φ can be viewed as the coefficient of absolute risk aversion of an investor. Without loss of generality, we denote $g = g(\tau, x)$. To employ the PI-DeepONet, we first define an operator that maps the input function to the PDE solution as

$$G(g) = \varphi.$$

The novelty of DeepONet is that it takes any arbitrary source term function as the input variables, making it more suitable than the PINN approach. Since φ is also a function, we can evaluate it at some point, say $y = (\tau, x)$, to obtain $G(g)(y) = \varphi(y)$. Here, $y = (\tau, x)$ denotes the points in the domain where the network predicts the solution of the PDE (6.5). In our approach, we will not use any input-output data, rather we only use the zero boundary and initial conditions. We will approximate the PDE solution operator $G(g)$ using two neural networks: branch and trunk nets. First, the input function of the branch net is discretized in a finite-dimensional space using a finite number of points called sensors. Then, the discretized input function is evaluated at m fixed sensors to obtain point-wise evaluations as the input function of the branch net. The trunk net takes the spatial and temporal coordinates and evaluates the solution operator to obtain the loss function. In general, the branch (with g as the input function) and trunk (with y as the input variable) networks are given by

$$b(\mathbf{g}(\tilde{\mathbf{x}})) = \mathbf{c} \cdot \sigma(\mathbf{W_b g}(\tilde{\mathbf{x}}) + \mathbf{b_b}), \quad (\mathbf{y}) = \sigma(\mathbf{W_t} \cdot \mathbf{y} + \mathbf{b_t})$$

respectively. Here, $\tilde{\mathbf{x}} = (\tau, \mathbf{x})$; c is some positive constant; σ is the activation function; $\mathbf{W_b}$ and $\mathbf{W_t}$ represent the weight matrices of branch and trunk networks, respectively; $\mathbf{b_b}$ and $\mathbf{b_t}$ represent the bias vector of the branch and trunk networks, respectively.

Now, let $g^i, i = 1, \cdots, N$, be any input function given that represents the source term in the PDE (6.5), then the given parabolic equation (6.5) becomes

$$g^i = \partial_\tau \varphi^i - \partial_x^2 \alpha(\varphi^i).$$

According to the approximation theorem for operators (see, Theorem 10), there exists $G_\theta(g^i)$ such that $G_\theta(g^i)(y) \approx G(g^i)(y) = \varphi^i(y)$. For a fixed i, the approximator in the DeepONet solution operator is the dot product of the outputs of the branch and trunk networks expressed as

$$G_\theta(g^i)(y) = \sum_{k=1}^{m*} \mathbf{b}_k(g^i(\tilde{\mathbf{x}})) \cdot \mathbf{t}_k(\mathbf{y}),$$

where $\mathbf{b}_k$ and $\mathbf{t}_k, k = 1, \cdots, m*$, denote the branch and trunk networks, respectively. Hence,

$$g^i \approx \partial_\tau G_\theta(g^i)(y) - \partial_x^2 \alpha(G_\theta(g^i)(y)).$$

Therefore, the physics loss, $\mathcal{L}_{Physics}$, which is calculated at the Q collocation points in the interior of the domain where the solution operator is evaluated, is obtained as

$$\mathcal{L}_{Physics}(\theta) = \frac{1}{NQ}\sum_{i=1}^{N}\sum_{j=1}^{Q}|R_\theta^i(y_{r,j}^i) - g^i(x_{r,j}^i)|^2.$$

Here, $R_\theta^i(y_{r,j}^i) = \partial_\tau G_\theta(g^i)(y_{r,j}^i) - \partial_x^2\alpha(G_\theta(g^i)(y_{r,j}^i))$ represents the residual satisfying the underlying PDE, and $y_{r,j}^i = (x_{r,j}^i, t_{r,j}^i)$ denotes the collocation points where the PDE is evaluated. The estimated PDE solution operator is differentiated with respect to the input variables using the so-called automatic differentiation.

Next, we use the zero boundary and initial conditions to obtain the second loss, $\mathcal{L}_{Operators}$, which is expressed as

$$\mathcal{L}_{Operator}(\theta) = \frac{1}{NP}\sum_{i=1}^{N}\sum_{k=1}^{P}|G_\theta(g^i)(y_{g,k}^i) - G_(g^i)(y_{g,k}^i)|^2 \tag{6.6}$$

where $y_{g,k}^i = (x_{g,k}^i, t_{g,k}^i)$ denotes points from the initial and boundary conditions. Since the boundary and initial conditions are zero, we have

$$\mathcal{L}_{Operator}(\theta) = \frac{1}{NP}\sum_{i=1}^{N}\sum_{k=1}^{P}|G_\theta(g^i)(y_{g,k}^i)|^2.$$

Hence, the total loss becomes

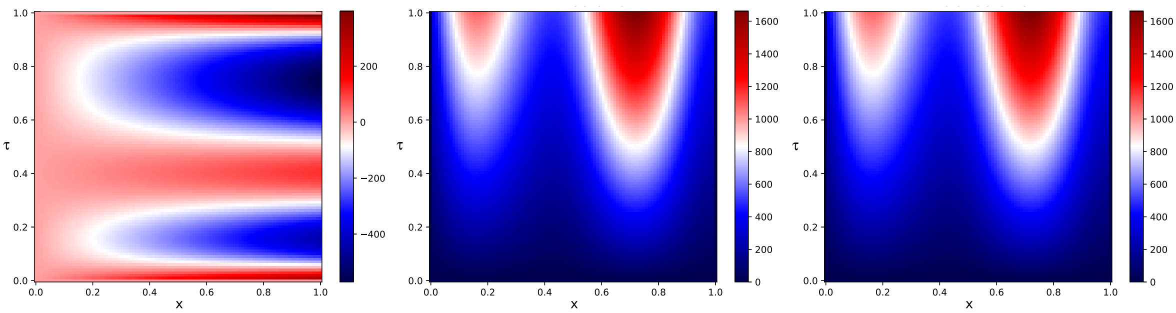
$$\mathcal{L}(\theta) = \mathcal{L}_{Physics}(\theta) + \mathcal{L}_{Operator}(\theta). \tag{6.7}$$

It follows that by minimizing the loss function (6.7) the network can effectively predict the solution of the HJB equation.

6.5 Results and Discussion to Chapter 6

The PI-DeepONet exhibits infinitesimal optimization and generalization errors, as it is easy to train and generalizes well to unseen data. In our approach, we did not use any input-output data, rather we only used the zero boundary and initial conditions. We approximate the PDE solution operator using branch and trunk nets. As a test example, we consider the diffusion function $\alpha(\varphi) = \varphi^2$. First, the input function of the branch net is discretized in a finite-dimensional space using a finite number of points called sensors. Then, the discretized input function is evaluated at fixed sensors to obtain point-wise evaluations. The trunk net takes the spatial and temporal coordinates and evaluates the solution operator to obtain the loss function. To generate our training data, we randomly sample $N = 500$ source term functions as input functions of the trunk net from a zero mean Gaussian process with an exponential quadratic kernel having a 0.2-length scale. The kernel function defines the covariance between two points in the process as a function of the distance between them. The parameter $l > 0$ determines how quickly the covariance between two points decays as the distance between them increases. In this study, we set $l = 0.2$. A smaller length scale results in a higher correlation between nearby points, whereas a larger length scale results in a lower correlation between nearby points. Then, the selected input functions are evaluated at $m = 100$ points as input sensors. The m outputs of the source term functions are sent to the branch network. Next, we select the $P = 100$ output sensors from the initial and boundary conditions, which are sent to the trunk nets. Our operator is then approximated by computing the dot product between the branch and trunk networks, and the corresponding operator loss is computed. After that, we select $Q = 100$ collocation points inside the domain, and the error related to the underlying physics is computed. Finally, the total loss is evaluated by combining the two losses, which are minimized using the Adam optimizer with a learning rate of 10^{-3}. Similarly, the test set is generated using the same approach. In Fig. 6.1, we compare a solution obtained by a physics-informed DeepONet method using the Relu activation function for 10000 iterations with a numerical solution constructed by means of the semi-implicit finite difference numerical method.

The PI-DeepONet exhibits infinitesimal optimization and generalization errors since it is easy to train and generalizes well to unseen data. In our simulation, we used the L_2 error, and its value is 6.14×10^{-4}. We used the Adam optimizer, which can easily track the global minimum, with a learning rate of 10^{-3}.



Input function g NN solution FDM numerical solution

Figure 6.5: Comparision of a PI-DeepONet solution and the numerical solution obtained by the finite difference method (FDM numerical solution). The right-hand side represents the input function g [60].

CHAPTER 7

Concluding Remarks and Prospects

This thesis aimed to investigate the qualitative and quantitative analysis of fully nonlinear nonlocal partial integro-differential equations (PIDEs) of parabolic type arising from financial mathematics. First, we studied the existence and uniqueness of solutions of nonlinear PIDEs arising from financial markets in multidimensional spaces. Such a PIDE models the well-known Black-Scholes equation for pricing call/put option. We relaxed and generalized some of the existing assumptions on the Black-Scholes model. The existence and uniqueness of solutions to the PIDEs was established in the scale of Bessel potential spaces using the theory of abstract semilinear parabolic equations. The novelty of this result is the generalization of existing results in one dimension to the multidimensional case. We considered a wide class of admissible Lévy measures satisfying suitable growth conditions near the origin and infinity. We investigate solutions to the nonlocal equation in which the shift function may depend on a prescribed large investor stock-trading strategy function. The Hölder estimate, which plays a crucial for proving the main result of the theorem, was also presented. Moreover, a general shift function arising from a nonlinear option pricing model taking into account a large trader stock-trading strategy was considered to demonstrate its application to option pricing in one-dimensional space.

Second, we investigated the existence and uniqueness of a solution to a fully nonlinear HJB equation arising from optimal portfolio management. The HJB equation represents a stochastic optimization problem, where the goal of an investor is to maximize the conditional expected value of the terminal utility of the portfolio. The nonlinear HJB equation was first transformed into a qausilinear parabolic equation. The existence and uniqueness of the solution to the transformed Cauchy problem was then proved in high-dimensional spaces. We employed the monotone operator technique, Fourier transform approach, and Banach's fixed point theorem to obtain the existence and uniqueness of the solution to the general form of the transformed parabolic equation in Sobolev spaces. The existence results are based on the properties of the value function of the parametric programming, which gives sufficient conditions (under certain assumptions) on the utility function, drift function, and volatility function for establishing our existence results. We also established the necessary and sufficient conditions on the utility function, drift function, and volatility function for the proof of our existence and uniqueness results. Furthermore, the relationships between the decision set and the corresponding solutions to the HJB equation with respect to the two decision sets are discussed. Specifically, we analyzed the qualitative behavior of the solutions of the two equations with respect to two different decision sets.

Furthermore, we presented a detailed discussion of traditional numerical methods for solving PDE, such as the finite differences, which form a baseline for our application of deep learning. As a practical illustration, we considered the transformed equation (3.13), which corresponds to the classical Black-Scholes equation used in finance for pricing options. We employed the implicit scheme for the time derivative and a central discretization for the first-order space derivative. The FDM technique allows one to estimate option prices at different points in time, leading up to the present. Finally we adopted a physics-informed DeepONet to approximate the solution operator of a parametric parabolic equation arising from portfolio selection problems. The input function of the branch net was discretized in a finite-dimensional space using a fixed number of sensors. The discretized input functions were evaluated at fixed sensors to obtain point-wise evaluations. The operator was approximated by computing the dot product between the branch and trunk networkss, and the corresponding operator loss was computed. We applied PI-DeepONet to solve the nonlinear parabolic equation obtained from the HJB equation.

In future studies, we will employ the PIDE approach to establish the existence and uniqueness solution to the fully nonlnear HJB parabolic equation. Here, we will consider an optimal control problem involving intertemporary utility. We will establish a suitable relationship between a fully nonlinear HJB and Black-Scholes equation for option pricing. Additionally, we will establish a suitable numerical scheme for solving the fully nonlinear PIDE and PDE of parabolic type in higher-dimensional space.

CHAPTER 8

Appendices

This chapter presents the proof of the main results arising from the thesis. We note here that these results and their proofs are contained in our recent publications (see Ševčovič and Udeani [68, 62], as well as J. Cruz, M. Grossinho, D. Ševčovič, and C. Udeani [32]. Appendix A contained the proof of results in Chapter 3, and Appendix B contains the proofs of our results in Chapter 4.

8.1 Appendix A

In this section, we present the proofs of our main results, which are contained in our paper (see Ševčovič and Udeani [62]).

8.1.1 Proof of Propositions 1

Proof. Let $u \in X$ be such that $\nabla_x u \in (X^{\gamma-1/2})^n$, i.e., $\partial_{x_i} u \in X^{\gamma-1/2}$ for each $i = 1, \cdots, n$. Then $\nabla_x u = A^{-(2\gamma-1)/2}\varphi = G_{2\gamma-1} * \varphi$ for some $\varphi \in (L^p(\mathbb{R}^n))^n$, and $\|\nabla_x u\|_{X^{\gamma-1/2}} = \|A^{(2\gamma-1)/2}\nabla_x u\|_X = \|\varphi\|_{L^p}$. Here, $\varphi = (\varphi_1, \cdots, \varphi_n)$ and $\partial_{x_i} u = G_{2\gamma-1} * \varphi_i$. Let $x, \xi \in \mathbb{R}^n$. Then

$$\nabla_x u(x+\xi) = G_{2\gamma-1}(x+\xi-\cdot) * \varphi(\cdot), \qquad \nabla_x u(x) = G_{2\gamma-1}(x-\cdot) * \varphi(\cdot).$$

Recall that the convolution operator satisfies the following inequality:

$$\|\psi * \varphi\|_{L^p(\mathbb{R}^n)} \le \|\psi\|_{L^q(\mathbb{R}^n)} \|\varphi\|_{L^r(\mathbb{R}^n)},$$

where $p, q, r \geq 1$ and $1/p + 1 = 1/q + 1/r$ (see [27, Section 1.6]). In particular, for $q = 1$ we have $\|\psi * \varphi\|_{L^p} \leq \|\psi\|_{L^1}\|\varphi\|_{L^p}$. For the modulus of continuity of the Bessel potential function $G_\alpha, \alpha \in (0,1)$, the following estimate holds:

$$\|G_\alpha(\cdot + h) - G_\alpha(\cdot)\|_{L^1} \leq C_0|h|^\alpha,$$

for any $h \in \mathbb{R}^n$ (c.f., [67, Chapter 5.4, Proposition 7]). Let ξ_1, ξ_2 be bounded vector valued functions, i.e., $\xi_1, \xi_2 \in (L^\infty(\mathbb{R}^n))^n$. Then, for any $x \in \mathbb{R}^n$ and $\theta \in [0,1]$, we have

$$\begin{aligned}
& u(x+\xi_1(x)) - u(x+\xi_2(x)) - (\xi_1(x) - \xi_2(x)) \cdot \nabla_x u(x) \\
= \; & u(x+\xi_1(x)) - \nabla_x u(x) + \xi_1(x) \cdot \nabla_x u(x) \\
& -[u(x+\xi_2(x)) - \nabla_x u(x) - \xi_2(x) \cdot \nabla_x u(x)] \\
= \; & (\xi_1(x) - \xi_2(x)) \int_0^1 \nabla_x u(x+\xi_1(x)) - \nabla_x u(x) \mathrm{d}\theta \\
& + \int_0^1 \nabla_x u(x+\theta\xi_1(x)) - \nabla_x u(x+\theta\xi_2(x)) \mathrm{d}\theta \; .
\end{aligned}$$

Now,

$$\begin{aligned}
& \|Q(u,\xi_1) - Q(u,\xi_2)\|^p_{L^p(\mathbb{R}^n)} \\
= \; & \int_{\mathbb{R}^n} |u(x+\xi_1(x)) - u(x+\xi_2(x)) - (\xi_1(x) - \xi_2(x)) \cdot \nabla_x u(x)|^p \mathrm{d}x \\
\leq \; & \int_{\mathbb{R}^n} \left| (\xi_1(x) - \xi_2(x)) \int_0^1 \nabla_x u(x+\theta\xi_1(x)) - \nabla_x u(x) \mathrm{d}\theta \right|^p \mathrm{d}x \\
& + \int_{\mathbb{R}^n} \left| \xi_2(x) \int_0^1 \nabla_x u(x+\theta\xi_1(x)) - \nabla_x u(x+\theta\xi_2(x)) \mathrm{d}\theta \right|^p \mathrm{d}x \\
\leq \; & \|\xi_1 - \xi_2\|^p_\infty \int_0^1 \int_{\mathbb{R}^n} |\nabla_x u(x+\theta\xi_1(x)) - \nabla_x u(x)|^p dx d\theta \\
& + \|\xi_2\|^p_\infty \int_0^1 \int_{\mathbb{R}^n} |\nabla_x u(x+\theta\xi_1(x)) - \nabla_x u(x+\theta\xi_2(x)|^p dx d\theta \\
\leq \; & \|\xi_1 - \xi_2\|^p_\infty \int_0^1 \| \left(G_{2\gamma-1}(\cdot + \theta\xi_1) - G_{2\gamma-1}(\cdot)\right) * \varphi\|^p_{L^p} d\theta \\
& + \|\xi_2\|^p_\infty \int_0^1 \| \left(G_{2\gamma-1}(\cdot + \theta\xi_1) - G_{2\gamma-1}(\cdot + \theta\xi_2)\right) * \varphi\|^p_{L^p} d\theta \\
\leq \; & \|\xi_1 - \xi_2\|^p_\infty \int_0^1 \|G_{2\gamma-1}(\cdot + \theta\xi_1) - G_{2\gamma-1}(\cdot)\|^p_{L^1} d\theta \|\varphi\|^p_{L^p} \\
& + \|\xi_2\|^p_\infty \int_0^1 \|G_{2\gamma-1}(\cdot + \theta\xi_1) - G_{2\gamma-1}(\cdot + \theta\xi_2)\|^p_{L^1} d\theta \|\varphi\|^p_{L^p} \\
\leq \; & \left(\|\xi_1 - \xi_2\|^p_\infty \|\xi_1\|^{(2\gamma-1)p}_\infty + \|\xi_2\|^p_\infty \|\xi_1 - \xi_2\|^{(2\gamma-1)p}_\infty\right) C_0^p \|\nabla_x u\|^p_{X^{\gamma-1/2}} \\
\leq \; & \|\xi_1 - \xi_2\|^{(2\gamma-1)p}_\infty \left(\|\xi_1\|^p_\infty + \|\xi_2\|^{(2-2\gamma)p}_\infty \|\xi_1\|^{(2\gamma-1)p}_\infty + \|\xi_2\|^p_\infty\right) C_0^p \|\nabla_x u\|^p_{X^{\gamma-1/2}} \; .
\end{aligned}$$

By Young's inequality, we have $ab \leq \frac{a^\alpha}{\alpha} + \frac{b^\beta}{\beta}$ for any $a, b \geq 0$, and $\alpha, \beta > 1$ with $1/\alpha + 1/\beta = 1$ (c.f., [13]). Set $\alpha = 1/(2-2\gamma), \beta = 1/(2\gamma - 1)$. Then, $1/\alpha + 1/\beta = 1$ and we obtain $\|\xi_2\|_\infty^{(2-2\gamma)p}\|\xi_1\|_\infty^{(2\gamma-1)p} \leq (2-2\gamma)\|\xi_2\|_\infty^p + (2\gamma-1)\|\xi_1\|_\infty^p \leq 2\|\xi_2\|_\infty^p + \|\xi_1\|_\infty^p$. Therefore,

$$
\begin{aligned}
\|Q(u,\xi_1) - Q(u,\xi_2)\|^p_{L^p(\mathbb{R}^n)} &\leq 2\|\xi_1 - \xi_2\|_\infty^{(2\gamma-1)p} \left(\|\xi_1\|_\infty^p + \|\xi_2\|_\infty^p\right) C_0^p \|\nabla_x u\|^p_{X^{\gamma-1/2}} \\
&\leq 2C_0^p\|\xi_1 - \xi_2\|_\infty^{(2\gamma-1)p}(\|\xi_1\|_\infty + \|\xi_2\|_\infty)^p \|\nabla_x u\|^p_{X^{\gamma-1/2}}.
\end{aligned}
$$

Hence, the pointwise estimate holds with the constant $\hat{C} = 2^{1/p}C_0 > 0$. □

8.1.2 Proof of Proposition 2

Proof. The Lévy measure $\nu(\mathrm{d}z)$ is given by $\nu(\mathrm{d}z) = h(z)\mathrm{d}z$. Let us denote the auxiliary function $\tilde{h}(z) = |z|^\alpha h(z)$. Then, $0 \leq \tilde{h}(z) \leq C_0 e^{-D|z| - \mu|z|^2}$. Since $h(z) = |z|^{-\alpha}\tilde{h}(z) = h_1(z)h_2(z)$, where $h_1(z) = |z|^{-\beta}\tilde{h}(z)^{\frac{1}{2}}$ and $h_2(z) = |z|^{\beta-\alpha}\tilde{h}(z)^{\frac{1}{2}}$. Applying Proposition 2 with $\xi_1 = \xi, \xi_2 = 0$, and using the Hölder inequality, we obtain

$$
\begin{aligned}
\|f(u)\|^p_{L^p} &= \int_{\mathbb{R}^n} \left| \int_{\mathbb{R}^n} (u(x + \xi(x,z)) - u(x) - \xi(x,z) \cdot \nabla_x u(x)) h(z) \mathrm{d}z \right|^p \mathrm{d}x \\
&\leq \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} |u(x + \xi(x,z)) - u(x) - \xi(x,z) \cdot \nabla_x u(x)|^p \, h_1(z)^p \mathrm{d}z \\
&\qquad \times \left(\int_{\mathbb{R}^n} h_2(z)^q \mathrm{d}z \right)^{p/q} \mathrm{d}x \\
&= \int_{\mathbb{R}^n} \left(\int_{\mathbb{R}^n} |u(x + \xi(x,z)) - u(x) - \xi(x,z) \cdot \nabla_x u(x)|^p \, \mathrm{d}x \right) h_1(z)^p \mathrm{d}z \\
&\qquad \times \left(\int_{\mathbb{R}^n} h_2(z)^q \mathrm{d}z \right)^{p/q} \\
&\leq C_0^p \|\nabla_x u\|^p_{X^{\gamma-1/2}} \int_{\mathbb{R}^n} |\xi(x,z)|^{2\gamma p} |z|^{-\beta p} \tilde{h}(z)^{p/2} \mathrm{d}z \left(\int_{\mathbb{R}^n} h_2(z)^q \mathrm{d}z \right)^{p/q} \\
&\leq C_0^p \|\nabla_x u\|^p_{X^{\gamma-1/2}} \int_{\mathbb{R}^n} |z|^{(2\gamma\omega - \beta)p} \tilde{h}(z)^{p/2} \mathrm{d}z \left(\int_{\mathbb{R}^n} h_2(z)^q \mathrm{d}z \right)^{p/q}.
\end{aligned}
$$

Assuming $p, q \geq 1, 1/p + 1/q = 1$ are such that

$$
(2\gamma\omega - \beta)p > -n, \qquad (\beta - \alpha)q = (\beta - \alpha)\frac{p}{p-1} > -n,
$$

then, the integrals $\int_{\mathbb{R}^n} |z|^{(2\gamma\omega-\beta)p}\tilde{h}(z)^{p/2}\mathrm{d}z$ and $\int_{\mathbb{R}^n} h_2(z)^q \mathrm{d}z = \int_{\mathbb{R}^n} |z|^{(\beta-\alpha)q}\tilde{h}(z)^{q/2}\mathrm{d}z$ are finite, provided that the shape parameters satisfy: either $\mu > 0, D \in \mathbb{R}$, or $\mu = 0, D > D_0 \geq 0$. As $\gamma > (\alpha - n)/(2\omega)$, there exists $\beta > 1$ satisfying

$$
\alpha - n + n/p < \beta < 2\gamma\omega + n/p.
$$

Therefore, there exists $C_0 > 0$ such that $\|f(u)\|_{L^p} \le C_0 \|\nabla_x u\|_{X^{\gamma-1/2}}$. $\square$

8.1.3 Proof of Theorem 5

Proof. First, we outline the idea of the proof. The initial condition $u(0,\cdot) \notin X^\gamma$ because of two reasons. It is not smooth for $x = 0$, and it grows exponentially for $x \to \infty$ (call option) or $x \to -\infty$ (put option). The shifted function $U = u - u^{BS}$ satisfies $U(0,\cdot) \equiv 0$, and so the initial condition $U(0,\cdot)$ belongs to X^γ. On the other hand, the shift function u^{BS} enters the governing PIDE as it includes the term $f(u^{BS}(\tau,\cdot))$ in the right-hand side. Since $u^{BS}(0,x)$ is not sufficiently smooth for $x = 0$, the shift term $f(u^{BS}(\tau,\cdot))$ is singular for $\tau \to 0^+$. Following the ideas of [18], for the shift term $f(u^{BS}(\tau,\cdot))$, we can provide Hölder estimates which are sufficient for proving the main result of this theorem (c.f., [18, Lemma 4.1]). Furthermore, the exponential growth of the function u^{BS} will be overcome since $\tilde{f}(e^x) = 0$, where $\tilde{f}(u) = f(u) - \delta(\tau,\cdot)\partial_x u$, i.e.,

$$\tilde{f}(u)(x) = \int_{\mathbb{R}} \left(u(x+\xi) - u(x) - (e^\xi - 1)\partial_x u(x) \right) \nu(\mathrm{d}z).$$

Next, we present more details of the proof. The function u^{BS} solves the linear PDE (3.14). Thus, the difference $U = u - u^{BS}$ of a solution u to (3.13) and u^{BS} satisfies the PIDE with the right-hand side:

$$\begin{aligned}
\frac{\partial U}{\partial \tau} &= \frac{\sigma^2}{2}\frac{\partial^2 U}{\partial x^2} + \left(r - \frac{\sigma^2}{2} - \delta(\tau,x) \right) \frac{\partial U}{\partial x} + f(U) + f(u^{BS}) - \delta(\tau,x)\frac{\partial u^{BS}}{\partial x} \\
&= \frac{\sigma^2}{2}\frac{\partial^2 U}{\partial x^2} + f(U) + g(\tau,x,\partial_x U) + h(\tau,\cdot),
\end{aligned}$$

$U(0,x) = 0,\ x \in \mathbb{R}, \tau \in (0,T)$. Here $g(\tau,x,\partial_x U) = (r - \sigma^2/2 - \delta(\tau,x))\partial_x U$, and $h(\tau,\cdot) = \tilde{f}(u^{BS}(\tau,\cdot))$. According to Proposition 2, $f : X^\gamma \to X$ is a bounded linear mapping. Consequently, it is Lipschitz continuous, provided that $1/2 \le \gamma < 1$ and $\gamma > (\alpha - 1)/(2\omega)$. Clearly, $\tilde{f}(e^x) = 0$. Hence,

$$\tilde{f}(u^{BS}) = \tilde{f}(u^{BS} - Ke^{r\tau + x}), \quad \text{and} \quad \partial_\tau \tilde{f}(u^{BS}) = \tilde{f}(\partial_\tau(u^{BS} - Ke^{r\tau+x})).$$

Now, it follows from [18, Lemma 4.1] that the following estimate holds true:

$$\|h(\tau_1,\cdot) - h(\tau_2,\cdot)\|_{L^p} = \|\tilde{f}(u^{BS}(\tau_1,\cdot)) - \tilde{f}(u^{BS}(\tau_2,\cdot))\|_{L^p} \le C_0|\tau_1 - \tau_2|^{-\gamma + \frac{p+1}{2p}},$$

$$\|h(\tau,\cdot)\|_{L^p} = \|\tilde{f}(u^{BS})(\tau,\cdot))\|_{L^p} \le C_0|\tau^{-(2\gamma-1)\left(\frac{1}{2}-\frac{1}{2p}\right)},$$

for any $0 < \tau_1, \tau_2, \tau \le T$. The function $h : [0,T] \to X \equiv L^p(\mathbb{R})$ is $((p+1)/(2p) - \gamma)$-

Hölder continuous because $\gamma < \frac{p+1}{2p}$. Moreover,

$$\int_0^T \|h(\tau,\cdot)\|_{L^p} d\tau = \int_0^T \|\tilde{f}(u^{BS}(\tau,\cdot))\|_{L^p} d\tau \le C_0 \int_0^T \tau^{-(2\gamma-1)\left(\frac{1}{2}-\frac{1}{2p}\right)} d\tau < \infty,$$

because $(2\gamma-1)\left(\frac{1}{2}-\frac{1}{2p}\right) < 1$. We recall that the crucial part of the proof of [18, Lemma 4.1] was based on the estimates:

$$\|\tilde{f}(u^{BS}(\tau,\cdot))\|_{L^p} \le C_0 \|v(\tau,\cdot)\|_{X^{\gamma-1/2}}, \quad \text{and } \|\partial_\tau \tilde{f}(u^{BS}(\tau,\cdot))\|_{L^p} \le C_0 \|\partial_\tau v(\tau,\cdot)\|_{X^{\gamma-1/2}},$$

where $v(\tau,x) = \partial_x \left(u^{BS}(\tau,x) - Ke^{r\tau+x}\right) = Ke^{r\tau+x}(N(d_1(\tau,x))-1)$. This estimate is fulfilled because of Proposition 2 under the assumptions made on γ. The proof for the case of a put option is similar. The final estimate on the Hölder continuity of the mapping h follows from careful estimates of the solution u^{BS} derived in the proof of [18, Lemma 4.1]. The proof now follows from Theorem 4 and Proposition 2. □

8.2 Appendix B

In this section, we present the proofs of our main results, which are contained in the paper (see Ševčovič and Udeani [68]).

8.2.1 Proof of Theorem 6

Proof. Recall that $H = L^2(\mathbb{R}^d)$ and $V = H^1(\mathbb{R}^d)$, its dual space being $V' = H^{-1}(\mathbb{R}^d)$. Let the scalar products in V and V' be defined as follows:

$$(f,g)_V = (A^{1/2}f, A^{1/2}g)_H = (Af,g)_H, \; (f,g)_{V'} = (A^{-1/2}f, A^{-1/2}g)_H = (A^{-1}f,g)_H,$$

respectively. Let us define the operator $\mathcal{A}(\tau,\cdot) : V \to V'$ by

$$\langle \mathcal{A}(\tau,\varphi), \psi \rangle = (A^{-1}A\alpha(\cdot,\tau,\varphi), \psi)_H = (\alpha(\cdot,\tau,\varphi), \psi)_H.$$

Under the assumption made on the function α we can conclude that the mapping $\varphi \mapsto \alpha(\cdot,\tau,\varphi)$ maps V into V. Indeed, if $\varphi \in V$ and $\eta = \alpha(\cdot,\tau,\varphi)$ then $\eta(x) = \alpha(x,\tau,\varphi(x)) - \alpha(x,\tau,0) + \alpha(x,\tau,0)$ and so

$$|\eta(x)| \le (\max_\varphi \alpha'_\varphi(x,\tau,\varphi))|\varphi(x)| + |h(x,\tau)| \le L|\varphi(x)| + |h(x,\tau)|.$$

Thus, $\int_{\mathbb{R}^d} |\eta(x)|^2 dx \le 2\int_{\mathbb{R}^d} L^2|\varphi(x)|^2 + |h(x,\tau)|^2 dx \le 2L^2\|\varphi\|_H^2 + 2\|h(\cdot,\tau)\|_H^2$. Since

$\nabla\eta(x) = \nabla_x \alpha(x,\tau,\varphi(x)) + \alpha'_\varphi(x,\tau,\varphi(x))\nabla\varphi(x)$, we have

$$
\begin{aligned}
\|\eta\|_V^2 &= \int_{\mathbb{R}^d} |\eta(x)|^2 + |\nabla\eta(x)|^2 dx \\
&\le 2\int_{\mathbb{R}^d} L^2|\varphi(x)|^2 + |h(x,\tau)|^2 dx + 2\int_{\mathbb{R}^d} |p(x,\tau)|^2 + L_0^2|\varphi(x)|^2 dx + 2\int_{\mathbb{R}^d} L^2|\nabla\varphi(x)|^2 dx \\
&\le 2(L^2\|\varphi\|_V^2 + \|h(\cdot,\tau)\|_H^2 + \|p(\cdot,\tau)\|_H^2 + L_0^2\|\varphi\|_H^2) < \infty,
\end{aligned}
$$

because $p, h \in L^\infty((0,T);H)$. Consequently, $\eta \in V$, as claimed.

Next, we show that the operator $\mathcal{A}$ is monotone in the space V'. According to (4.17) we have $(\alpha(x,\tau,\varphi_1) - \alpha(x,\tau,\varphi_2))(\varphi_1 - \varphi_2) \ge \omega(\varphi_1 - \varphi_2)^2$, for any $\varphi_1, \varphi_2 \ge \varphi_{min}, x \in \mathbb{R}, \tau \in [0,T]$.

$$
\begin{aligned}
\langle \mathcal{A}(\tau,\varphi_1) - \mathcal{A}(\tau,\varphi_2), \varphi_1 - \varphi_2\rangle &= (\alpha(\cdot,\tau,\varphi_1) - \alpha(\cdot,\tau,\varphi_2), \varphi_1 - \varphi_2) \\
&= \int_{\mathbb{R}^d} (\alpha(x,\tau,\varphi_1(x)) - \alpha(x,\tau,\varphi_2(x)))(\varphi_1(x) - \varphi_2(x)) dx \\
&\ge \int_{\mathbb{R}^d} \omega|\varphi_1(x) - \varphi_2(x)|^2 dx = \omega\|\varphi_1 - \varphi_2\|_H^2.
\end{aligned}
$$

This implies that the operator $\mathcal{A}(\tau,\cdot)$ is strongly monotone.

For a given $\tilde\varphi \in \mathcal{H}$, we have $\hat f \in \mathcal{V}'$, where $\hat f(\tau) = g_0(\tau,\tilde\varphi(\cdot,\tau)) + \nabla\cdot \boldsymbol{g}_1(\tau,\tilde\varphi(\cdot,\tau))$, because $g_0, g_{1j} : [0,T]\times H \to H$ are globally Lipschitz continuous, $H \hookrightarrow V'$, and the operator ∇ maps H into V'. The hemicontinuity, boundedness, and coercivity of the operator $\mathcal{A}$ follows from the assumption that α is globally Lipschitz continuous and strictly increasing.

Applying Theorem 3, we deduce the existence of a unique solution $\varphi \in \mathcal{V}$ such that

$$
\partial_\tau\varphi + \mathcal{A}(\tau,\varphi) = \hat f(\tau), \qquad \varphi_0 \in H, \tag{8.1}
$$

where $\mathcal{A}(\tau,\varphi) = A\alpha(\cdot,\tau,\varphi)$. Next, we multiply (8.1) by A^{-1} to obtain

$$
\partial_\tau A^{-1}\varphi + \alpha(\cdot,\tau,\varphi) = f, \tag{8.2}
$$

where $f = f(\tau,\tilde\varphi) = A^{-1}\hat f(\tau)$. For $\tau \in [0,T]$, we denote $\tilde f(\tilde\varphi) = A^{-1/2}\hat f(\tau) = A^{-1/2}g_0(\tau,\tilde\varphi) + A^{-1/2}\sum_{j=1}^d \partial_{x_j} g_{1j}(\tau,\tilde\varphi)$. For the Fourier transform of $\tilde f$, we have

$$
\widehat{\tilde f(\tilde\varphi)}(\xi) = \frac{1}{(1+|\xi|^2)^{1/2}}\widehat{g_0(\tau,\tilde\varphi)}(\xi) + \sum_{j=1}^d \frac{(-i\xi_j)}{(1+|\xi|^2)^{1/2}}\widehat{g_{1j}(\tau,\tilde\varphi)}(\xi).
$$

Let $\beta > 0$ be the Lipschitz constant of the mappings $g_0, g_{1j}, j = 1,\cdots,d$. Using

Parseval's identity and Lipschitz continuity of g_0, g_{1j} in H, we obtain, for $\tilde{\varphi}_1, \tilde{\varphi}_2 \in \mathcal{H}$,

$$\begin{aligned}
\|\tilde{f}(\tilde{\varphi}_1) - \tilde{f}(\tilde{\varphi}_2)\|_H^2 = \|\widehat{\tilde{f}(\tilde{\varphi}_1)} - \widehat{\tilde{f}(\tilde{\varphi}_2)}\|_H^2 &= \int_{\mathbb{R}^d} |\widehat{\tilde{f}(\tilde{\varphi}_1)}(\xi) - \widehat{\tilde{f}(\tilde{\varphi}_2)}(\xi)|^2 d\xi \\
&\leq 2 \int_{\mathbb{R}^d} \frac{1}{1+|\xi|^2} |\widehat{g_0(\tau, \tilde{\varphi}_1)}(\xi) - \widehat{g_0(\tau, \tilde{\varphi}_2)}(\xi)|^2 \\
&\quad + \sum_{j=1}^{d} \frac{|\xi|^2}{1+|\xi|^2} |\widehat{g_{1j}(\tau, \tilde{\varphi}_1)}(\xi) - \widehat{g_{1j}(\tau, \tilde{\varphi}_2)}(\xi)|^2 d\xi \\
&\leq 2\|\widehat{g_0(\tau, \tilde{\varphi}_1)} - \widehat{g_0(\tau, \varphi_2)}\|_H^2 + 2\sum_{j=1}^{d} \|\widehat{g_{1j}(\tau, \tilde{\varphi}_1)} - \widehat{g_{1j}(\tau, \varphi_2)}\|_H^2 \\
&= 2\|g_0(\tau, \tilde{\varphi}_1) - g_0(\tau, \tilde{\varphi}_2)\|_H^2 + 2\sum_{j=1}^{d} \|g_{1j}(\tau, \tilde{\varphi}_1) - g_{1j}(\tau, \tilde{\varphi}_2)\|_H^2 \\
&\leq \tilde{\beta}^2 \|\tilde{\varphi}_1 - \tilde{\varphi}_2\|_H^2,
\end{aligned}$$

where $\tilde{\beta}^2 = 2(1+d)\beta^2$. Hence, we obtain

$$\|\tilde{f}(\tilde{\varphi}_1) - \tilde{f}(\tilde{\varphi}_2)\|_H \leq \tilde{\beta}\|\tilde{\varphi}_1 - \tilde{\varphi}_2\|_H. \tag{8.3}$$

Suppose $\varphi_1, \varphi_2 \in \mathcal{H}$ are such that $\varphi_1 = F(\tilde{\varphi}_1)$ and $\varphi_2 = F(\tilde{\varphi}_2)$. Here, the map $F : \mathcal{H} \to \mathcal{H}$ is defined by $\varphi = F(\tilde{\varphi})$, where φ is a solution to the Cauchy problem

$$\partial_\tau A^{-1}\varphi + \alpha(\cdot, \tau, \varphi) = f(\tau, \tilde{\varphi}), \qquad \varphi(0) = \varphi_0.$$

Letting $\varphi = \varphi_1 - \varphi_2 = F(\tilde{\varphi}_1) - F(\tilde{\varphi}_2)$, we obtain

$$\partial_\tau A^{-1}(\varphi_1 - \varphi_2) + \alpha(\cdot, \tau, \varphi_1) - \alpha(\cdot, \tau, \varphi_2) = f(\tilde{\varphi}_1) - f(\tilde{\varphi}_2). \tag{8.4}$$

Next, we multiply (8.4) by $\varphi_1 - \varphi_2$ and take the scalar product in the space H to obtain

$$\begin{aligned}
(\partial_\tau A^{-1}(\varphi_1 - \varphi_2), \varphi_1 - \varphi_2) \quad &+ \quad (\alpha(\cdot, \tau, \varphi_1) - \alpha(\cdot, \tau, \varphi_2), \varphi_1 - \varphi_2) \\
&= \quad (f(\tau, \tilde{\varphi}_1) - f(\tau, \tilde{\varphi}_2), \varphi_1 - \varphi_2).
\end{aligned} \tag{8.5}$$

Using (8.3) and the fact that $A^{-1/2}$ is self-adjoint in H, then (8.5) gives

$$\begin{aligned}
&\frac{1}{2}\frac{d}{d\tau}\|A^{-1/2}(\varphi_1 - \varphi_2)\|_H^2 + \omega\|\varphi_1 - \varphi_2\|_H^2 \\
&\quad \leq \langle f(\tau, \tilde{\varphi}_1) - f(\tau, \tilde{\varphi}_2), \varphi_1 - \varphi_2\rangle = \langle A^{1/2}(f(\tau, \tilde{\varphi}_1) - f(\tau, \tilde{\varphi}_2)), A^{-1/2}(\varphi_1 - \varphi_2)\rangle \\
&\quad \leq \|A^{1/2}(f(\tau, \tilde{\varphi}_1) - f(\tau, \tilde{\varphi}_2))\|_H \|\varphi_1 - \varphi_2\|_{V'} = \|\tilde{f}(\tilde{\varphi}_1) - \tilde{f}(\tilde{\varphi}_2)\|_H \|\varphi_1 - \varphi_2\|_{V'} \\
&\quad \leq \tilde{\beta}\|\tilde{\varphi}_1 - \tilde{\varphi}_2\|_H \|\varphi_1 - \varphi_2\|_{V'}.
\end{aligned}$$

This implies

$$\frac{1}{2}\frac{d}{d\tau}\|\varphi_1-\varphi_2\|_{V'}^2+\omega\|\varphi_1-\varphi_2\|_H^2\leq\tilde{\beta}\|\tilde{\varphi_1}-\tilde{\varphi_2}\|_H\|\varphi_1-\varphi_2\|_{V'}.$$

Then, integrating on a small time interval $[0,T]$ from 0 to t and noting that $\varphi_1(0)=\varphi_2(0)=\varphi_0$, we obtain

$$\begin{aligned}\frac{1}{2}\|\varphi_1(\tau)-\varphi_2(\tau)\|_{V'}^2+\omega\int_0^\tau\|\varphi_1(s)-\varphi_2(s)\|_H^2ds&\\ \leq\tilde{\beta}\int_0^\tau\|\tilde{\varphi_1}(s)-\tilde{\varphi_2}(s)\|_H\|\varphi_1(s)-\varphi_2(s)\|_{V'}ds&\\ \leq\tilde{\beta}\max_{\tau\in[0,T]}\|\varphi_1(\tau)-\varphi_2(\tau)\|_{V'}\int_0^T\|\tilde{\varphi_1}(\tau)-\tilde{\varphi_2}(\tau)\|_Hd\tau.&\end{aligned}$$

Taking the maximum over $\tau\in[0,T]$ and using the fact that for any $a,b\in\mathbb{R}$, $ab\leq\frac{1}{2}a^2+\frac{1}{2}b^2$, we obtain

$$\begin{aligned}\frac{1}{2}(\max_{\tau\in[0,T]}\|\varphi_1(\tau)-\varphi_2(\tau)\|_{V'})^2+\omega\int_0^T\|\varphi_1(\tau)-\varphi_2(\tau)\|_H^2d\tau&\\ \leq\tilde{\beta}\max_{\tau\in[0,T]}\|\varphi_1(\tau)-\varphi_2(\tau)\|_{V'}\int_0^T\|\tilde{\varphi_1}(\tau)-\tilde{\varphi_2}(\tau)\|_Hd\tau&\\ \leq\frac{1}{2}(\max_{\tau\in[0,T]}\|\varphi_1(\tau)-\varphi_2(\tau)\|_{V'})^2+\frac{\tilde{\beta}^2}{2}(\int_0^T\|\tilde{\varphi_1}(\tau)-\tilde{\varphi_2}(\tau)\|_Hd\tau)^2.&\end{aligned}$$

Using the Cauchy–Schwartz inequality, we obtain $\omega\int_0^T\|\varphi_1(\tau)-\varphi_2(\tau)\|_H^2d\tau\leq\frac{\tilde{\beta}^2}{2}\int_0^Td\tau\int_0^T\|\tilde{\varphi_1}(\tau)-\tilde{\varphi_2}(\tau)\|_H^2d\tau=\frac{\tilde{\beta}^2T}{2}\int_0^T\|\tilde{\varphi_1}(\tau)-\tilde{\varphi_2}(\tau)\|_H^2d\tau$. This implies that

$$\|F(\tilde{\varphi_1})-F(\tilde{\varphi_2})\|_{\mathcal{H}}^2\leq\frac{\tilde{\beta}^2T}{2\omega}\|\tilde{\varphi}_1-\tilde{\varphi}_2\|_{\mathcal{H}}^2.$$

Thus, for T sufficiently small such that $\frac{\tilde{\beta}^2T}{2\omega}<1$, the operator F is a contraction on the space $\mathcal{H}$; therefore by the Banach fixed point theorem, F has a unique fixed point in $\mathcal{H}$. It is worth noting that $\tilde{\beta}$ and ω are given such that they are independent of T. If $T>0$ is arbitrary, then we can apply a simple continuation argument. Indeed, if the solution exists in $(0,T_0)$ interval with $\frac{\tilde{\beta}^2T_0}{2\omega}<1$, then starting from the initial condition $\varphi_0=\varphi(T_0/2)$ we can continue the solution φ from the interval $(0,T_0)$ over the interval $(0,T_0)\cup(T_0/2,T_0/2+T_0)\equiv(0,3T_0/2)$. Continuing in this manner, we obtain the existence and uniqueness of a solution $\varphi\in\mathcal{H}$ defined on the time interval $(0,T)$.

Finally, the solution belongs to the space $\mathcal{V}$ because the right-hand side, i.e., the function $\hat{f}(\tau)=g_0(\tau,\varphi(\cdot,\tau))+\nabla\cdot\boldsymbol{g}_1(\tau,\varphi(\cdot,\tau))$ belongs to $\mathcal{V}'$. Applying Theorem 3, we conclude $\varphi \in \mathcal{V}$, as claimed.

$\square$

8.2.2 Proof of Theorem 7

Proof. Since $\hat{f} \in \mathcal{V}'$, where $\hat{f} = g_0 + \nabla \cdot \boldsymbol{g}_1$ and $\mathcal{A}(\tau, \varphi) \in \mathcal{V}'$, then $\partial_\tau \varphi \in \mathcal{V}'$. Therefore, for each $\varphi_0 \in H$, we have $\varphi \in W$ where W is the Banach space $W = \{\varphi, \varphi \in \mathcal{V}, \partial_\tau \varphi \in \mathcal{V}'\}$. According to [66, Proposition 1.2], we have $W \hookrightarrow C([0, T]; H)$. Hence, the unique solution φ to the Cauchy problem belongs to the space $C([0, T]; H)$, as claimed.

Next, we show that the unique solution satisfies a-priori energy estimate (4.15). Let φ be a unique solution to the Cauchy problem (4.14). Multiply (8.2) by φ and take the scalar product in H to obtain

$$(\partial_\tau A^{-1}\varphi, \varphi)_H + (\alpha(\cdot, \tau, \varphi), \varphi)_H = (A^{-1}g_0(\tau, \varphi) + A^{-1}\nabla \cdot \boldsymbol{g}_1(\tau, \varphi), \varphi). \tag{8.6}$$

Using the Lipschitz continuity of $g_0, \boldsymbol{g}_1$ and strong monotonicity of α, we obtain

$$\begin{aligned}
\frac{1}{2}\frac{d}{d\tau}\|\varphi\|_{V'}^2 + \omega\|\varphi\|_H^2 &= (\partial_\tau A^{-1}\varphi, \varphi) + \omega\|\varphi\|_H^2 \\
&\leq (\partial_\tau A^{-1}\varphi, \varphi) + (\alpha(\cdot, \varphi) - \alpha(\cdot, 0), \varphi) \\
&= (A^{-1}(g_0(\cdot, \varphi) + \nabla \cdot \boldsymbol{g}_1(\tau, \varphi)) - \alpha(\cdot, 0), \varphi) \\
&= (A^{-1}(g_0(\cdot, \varphi) - g_0(\cdot, 0) + \nabla \cdot \boldsymbol{g}_1(\cdot, \varphi) - \nabla \cdot \boldsymbol{g}_1(\cdot, 0)), \varphi) \\
&\quad + (A^{-1}(g_0(\cdot, 0) + \nabla \cdot \boldsymbol{g}_1(\cdot, 0)), \varphi) - (\alpha(\cdot, 0), \varphi) \\
&= (A^{-1/2}(g_0(\cdot, \varphi) - g_0(\cdot, 0) + \nabla \cdot \boldsymbol{g}_1(\cdot, \varphi) - \nabla \cdot \boldsymbol{g}_1(\cdot, 0)), A^{-1/2}\varphi) \\
&\quad + (A^{-1/2}(g_0(\cdot, 0) + \nabla \cdot \boldsymbol{g}_1(\cdot, 0)), A^{-1/2}\varphi) - (\alpha(\cdot, 0), \varphi) \\
&\leq \beta(1 + d)\|\varphi\|_H\|\varphi\|_{V'} + \|A^{-1/2}(g_0(\cdot, 0) + \nabla \cdot \boldsymbol{g}_1(\cdot, 0))\|_H\|\varphi\|_{V'} \\
&\quad + \|\alpha(\cdot, 0)\|_H\|\varphi\|_H \\
&\leq \frac{\omega}{4}\|\varphi\|_H^2 + \frac{\beta^2(1 + d)^2}{\omega}\|\varphi\|_{V'}^2 + \frac{1}{2}\|A^{-1/2}(g_0(\cdot, 0) + \nabla \cdot \boldsymbol{g}_1(\cdot, 0))\|_H^2 \\
&\quad + \frac{1}{2}\|\varphi\|_{V'}^2 + \frac{1}{\omega}\|\alpha(\cdot, 0)\|_H^2 + \frac{\omega}{4}\|\varphi\|_H^2.
\end{aligned}$$

Hence, there exist constants $C_0, C_1 > 0$ such that

$$\frac{d}{d\tau}\|\varphi\|_{V'}^2 + \omega\|\varphi\|_H^2 \leq C_1\|\varphi\|_{V'}^2 + C_0\big(\|g_0(\cdot, 0)\|_H^2 + \sum_{j=1}^{d}\|g_{1j}(\cdot, 0)\|_H^2 + \|\alpha(\cdot, 0)\|_H^2\big).$$

Solving the differential inequality $y'(\tau) \leq C_1 y(\tau) + r(\tau)$, where $y(\tau) = \|\varphi(\cdot, \tau)\|_{V'}^2$ and

$r(\tau) = C_0\big(\|g_0(\cdot,\tau,0)\|_H^2 + \sum_{j=1}^d \|g_{1j}(\cdot,\tau,0)\|_H^2 + \|\alpha(\cdot,\tau,0)\|_H^2\big)$, yields

$$y(\tau) \le e^{C_1 T}\big(y(0) + \int_0^T r(s)ds\big),$$

and the proof of the Theorem follows.

□

8.2.3 Proof of Theorem 8

Proof. Since $\sigma^2(\boldsymbol{\theta}) > 0$ and $\triangle$ is a compact set, there exist constants $0 < \omega \le L$ such that $0 < \omega \le \sigma^2(\boldsymbol{\theta}) \le L$ for all $\boldsymbol{\theta} \in \triangle$. It follows from Proposition 3 that

$$\omega|\varphi| \le |\alpha(x,\varphi) - \alpha(x,0)| \le L|\varphi|. \tag{8.7}$$

Since $\varphi_0, h \in L^\infty(\mathbb{R})$ and $h(x) = \alpha(x,0)$, we obtain $M := \sup_{x\in\mathbb{R}} |\alpha(x,\varphi_0(x))| < \infty$.

Let us define the shifted diffusion function by $\tilde{\alpha}(x,\varphi) = \alpha(x,\varphi) - \alpha(x,0)$. Notice that $\alpha(x,0) = \min_{\boldsymbol{\theta}\in\triangle} -\mu(x,\boldsymbol{\theta}) = h(x)$. Then equation (4.16) is equivalent to

$$\partial_\tau\varphi + A\tilde{\alpha}(\cdot,\varphi) = \tilde{\alpha}(\cdot,\varphi) + \partial_x^2 h - \partial_x\left(\alpha(\cdot,\varphi)\varphi\right),$$

where $A = I - \partial_x^2$.

Next, let $g_0(\varphi) = \tilde{\alpha}(\cdot,\varphi) + \partial_x^2 h$ and $g_1(\varphi) = -w(\alpha(\cdot,\varphi))\varphi$. Here, $w : \mathbb{R} \to \mathbb{R}$ is a suitable cut-off function

$$w(\alpha) = \begin{cases} \underline{\psi} e^{\lambda T}, & \text{if } \alpha \le \underline{\psi} e^{\lambda T}, \\ \alpha, & \text{if } \underline{\psi} e^{\lambda T} < \alpha < \overline{\psi} e^{\lambda T}, \\ \overline{\psi} e^{\lambda T}, & \text{if } \alpha \ge \overline{\psi} e^{\lambda T}, \end{cases}$$

where $\overline{\psi} = M, \underline{\psi} = -M$. Then, the functions $g_0, g_1 : H \to H$ are globally Lipschitz continuous.

Notice that the diffusion function $\tilde{\alpha}$ fulfills assumptions of Theorem 6 with $\tilde{h}(x) = \tilde{\alpha}(x,0) \equiv 0$. Now, applying Theorems 6 and 7, we obtain the existence and uniqueness of a solution $\varphi \in C([0,T];H) \cap L^2((0,T);V)$ to the Cauchy problem (4.14). The solution φ satisfies the point-wise estimate (4.15). Hence, $w(\alpha(x,\varphi(x,\tau))) = \alpha(x,\varphi(x,\tau))$ and φ is a solution to the Cauchy problem (4.16), as well.

Finally, from (8.7), we deduce the $L^\infty((0,T)\times\mathbb{R})$ estimate for the solution φ since $\sup_{x\in\mathbb{R}} |\alpha(x,\varphi(x,\tau))| \le Me^{\lambda\tau}$, where $\lambda = \sup_{x\in\mathbb{R}} p(x)$. Furthermore, $\varphi \in L^\infty((0,T)\times\mathbb{R})$, and

$$\sup_{x\in\mathbb{R},\tau\in[0,T]} |\varphi(x,\tau)| \le \omega^{-1}(Me^{\lambda T} + \max_{x\in\mathbb{R}} |h(x)|).$$

□

8.2.4 Proof of Proposition 3

Proof. Let us define $\alpha^{\boldsymbol{\theta}}(x,\tau,\varphi) := -\mu(x,t,\boldsymbol{\theta}) + \frac{\varphi}{2}\sigma(x,t,\boldsymbol{\theta})^2$, where $t = T-\tau$. Then

$$\alpha(x,\tau,\varphi) = \min_{\boldsymbol{\theta}\in\triangle} \alpha^{\boldsymbol{\theta}}(x,\tau,\varphi)\,.$$

For any given $\boldsymbol{\theta} \in \triangle$, the function $\alpha^{\boldsymbol{\theta}}(x,\tau,\varphi)$ is globally Lipschitz continuous in all variables. The minimal function α is therefore globally Lipschitz continuous as well. Moreover, the function $\alpha^{\boldsymbol{\theta}}(x,\tau,\varphi)$ satisfies the inequality (4.17) for any $\boldsymbol{\theta} \in \triangle$, and so does the minimal function α.

Next, we prove inequality (4.18). Let $x_1, x_2 \in \mathbb{R}^d$ such that $x_2 = x_1 + he^i$, where $e^i, i = 1,\cdots,d$, is the standard normal vector, i.e., $e^i = (0,0,...,0,1,0,...,0)^T$. We have that

$$\begin{aligned}
&\alpha^{\boldsymbol{\theta}}(x_1,\tau,\varphi) - \alpha^{\boldsymbol{\theta}}(x_2,\tau,\varphi) = -(\mu(x_1,\tau,\boldsymbol{\theta}) - \mu(x_2,\tau,\boldsymbol{\theta})) + \frac{\varphi}{2}(\sigma(x_1,\tau,\boldsymbol{\theta})^2 - \sigma(x_2,\tau,\boldsymbol{\theta})^2)\\
&= \int_0^h (-\partial_{x_i}\mu(x_1+\xi e^i,\tau,\boldsymbol{\theta}))d\xi + \int_0^h \frac{\varphi}{2}\partial_{x_i}\sigma^2(x_1+\xi e^i,\tau,\boldsymbol{\theta})d\xi\\
&\le \int_0^h |\partial_{x_i}\mu(x_1+\xi e^i,\tau,\boldsymbol{\theta}))|d\xi + \int_0^h \frac{|\varphi|}{2}|\partial_{x_i}\sigma^2(x_1+\xi e^i,\tau,\boldsymbol{\theta})|d\xi\\
&\le \max_{\boldsymbol{\theta}\in\triangle, 0\le\xi\le h} |\partial_{x_i}\mu(x_1+\xi e^i,\tau,\boldsymbol{\theta})|\, h + \max_{\boldsymbol{\theta}\in\triangle, x\in\mathbb{R}^d} |\partial_{x_i}\sigma^2(x,\tau,\theta)|\frac{|\varphi|}{2}h.
\end{aligned}$$

Hence,

$$\alpha^{\boldsymbol{\theta}}(x_1,\tau,\varphi) \le \alpha^{\boldsymbol{\theta}}(x_2,\tau,\varphi) + \max_{\boldsymbol{\theta}\in\triangle, 0\le\xi\le h} |\partial_{x_i}\mu(x_1+\xi e^i,\tau,\boldsymbol{\theta})|\, h + \max_{\boldsymbol{\theta}\in\triangle, x\in\mathbb{R}^d} |\partial_{x_i}\sigma^2(x,\tau,\theta)||\varphi|h.$$

We note that $x_2 - x_1 = he^i$ so that $|x_2 - x_1| = h$. Taking minimum over $\boldsymbol{\theta} \in \triangle$, we obtain

$$\alpha(x_1,\tau,\varphi) \le \alpha(x_2,\tau,\varphi) + \max_{\boldsymbol{\theta}\in\triangle, 0\le\xi\le h} |\partial_{x_i}\mu(x_1+\xi e^i,\tau,\boldsymbol{\theta})|\, h + \max_{\boldsymbol{\theta}\in\triangle, x\in\mathbb{R}^d} |\partial_{x_i}\sigma^2(x,\tau,\theta)||\varphi|h.$$

Exchanging the role of x_1 and x_2 and taking the limit as $x_2 \to x_1$, i.e., $h \to 0$, we obtain inequality (4.18), as stated.

$\square$

List of Figures

2.1 Sector $S_{a,\varphi}$ in the complex plane [61] 16

4.1 The path $\hat{\boldsymbol{\theta}}(\varphi)$ as a function of φ [61]. 41
4.2 a) A graph of the value function α, b) its second derivative $\alpha''(\varphi)$ for the portfolio consisting of the stocks index and bonds (c.f. [38]) for the convex compact decision set $\triangle$. The dotted line in a) corresponds to the discrete decision set $\hat{\triangle} = \{\boldsymbol{\theta}^1, \boldsymbol{\theta}^2, \boldsymbol{\theta}^3\} \subset \triangle$. Source: our computation is based on the method from [68]. 42
4.3 a) A graph of the value function $\alpha(\varphi)$, and b) the second derivative $\alpha''_\varphi(\varphi)$ corresponding to five stocks (BASF, Bayer, Degussa–Huls, FMC Scheringfrom) entering DAX30 index. Source: our computation is based on the method from [32, 68]. 42
4.4 The optimal vector $\boldsymbol{\theta} = (\theta_1, \cdots, \theta_n)^T$ as a function of φ for the German DAX30 index. Source: our computation is based on the method from [35, 68] 43
4.5 A solution $\varphi(x, \tau)$ for the DARA utility function with $a_0 = 9$, $a_1 = 8$, $x^* = 2$. Source: our computations based on the numerical method from [32, 68]. 43
4.6 a) Graphs of the solution φ and $\tilde{\varphi}$ corresponding to the decision sets $\triangle$ and $\tilde{\triangle}$, respectively 46

5.1 The Black-Scholes price at $t = 0$ 54
5.2 3D Black-Scholes surface for $\delta = 0.0025$ 55

6.1 Neural network architecture with two input vectors, three hidden layers (each with six neurons), and one output vector. 58

6.2 Simple architecture of physics informed neural network for solving partial differential equations [71] 61
6.3 Schematic of DeepONet 65
6.4 Simple architecture of physics informed DeepONet [60] 66
6.5 Comparision of a PI-DeepONet solution and the numerical solution obtained by the finite difference method (FDM numerical solution). The right-hand side represents the input function g [60]. 69

CHAPTER 9

References

[1] R. Abe and N. Ishimura. Existence of solutions for the nonlinear partial differential equation arising in the optimal investment problem. *Proc. Japan Acad., Ser. A.* **84**(1) (2008), 11–14.

[2] H. Abels and M. Kassmann. The Cauchy problem and the martingale problem for integro-differential operators with non-smooth kernels. *Osaka J. Math.* **46** (2009), 661–683.

[3] K. Aboodh, R. Farah, I. Almardy, F. Almostafa. Solution of partial integro-differential equations by using Aboodh and double Aboodh transforms methods. *Global Journal of Pure and Applied Mathematics.* **13** (2016), 4347–4360.

[4] I. Arregui, B. Salvador, D. Ševčovič, C. Vázquéz, C. Mathematical analysis of a nonlinear PDE model for European options with counterparty risk. *Comptes Rendus Mathematique.* **357** (2019), 252–257.

[5] I. Arregui, B. Salvador, D. Ševčovič, C. Vázquéz. Total value adjustment for European options with two stochastic factors. Mathematical model, analysis and numerical simulation. *Computers and Mathematics with Applications.* **76** (2018), 725–740.

[6] M. Avellaneda, A. Levy, A. Paras. Pricing and hedging derivative securities in markets with uncertain volatilities. *Applied Mathematical Finance.* **2** (1995), 73–88.

[7] V. Barbu. Nonlinear differential equations of monotone types in Banach spaces. *Springer Science & Business Media.* (2010).

[8] G. Barles, R. Buckdahn, E. Pardoux. Backward stochastic differential equations and integral-partial differential equations. *Stochastics: An International Journal of Probability and Stochastic Processes.* **60** (1997), 57–83.

[9] G. Barles, H. M.Soner. Option pricing with transaction costs and a nonlinear Black-Scholes equation. *Finance and Stochastics.* (1998).

[10] O .E. Barndorff–Nielsen and S. Z. Levendorskiĭ. Feller processes of normal inverse Gaussian type. *Quant. Finance.* **1** (2001), 318–331.

[11] A.U. Bello, M.T. Omojola, M.A. Onyido, M.O. Uba, and C.I. Udeani. New method for computing zeros of monotone maps in lebesgue spaces with applications to integral equations, fixed points, optimization, and variational inequality problems. *Acta Mathematica Universitatis Comenianae.* **91**(2) (2022).

[12] D. P. Bertsekas. Dynamic programming and stochastic control. *Academic Press, New York.* (1976).

[13] H. Brezis. Functional analysis, Sobolev spaces and partial differential equations. *Springer Science* & *Business Media.* (2010).

[14] M. Burzoni, V. Ignazio, A. M. Reppen, H.M. Soner. Viscosity solutions for controlled McKean–Vlasov jump-diffusions. *SIAM Journal on Control and Optimization.* **58** (2020), 1676–1699.

[15] T. Chen and H. Chen. Universal approximation to nonlinear operators by neural networks with arbitrary activation functions and its application to dynamical systems. *IEEE Transactions on Neural Networks.* **6**(4) (1995), 911–917.

[16] R. Cont and P. Tankov. Financial Modeling with Jump Processes. *Journal of Business.* (2004).

[17] R. Cont and E. Voltchkova. Integro-differential equations for option prices in exponential Lévy models. *Finance and Stochastics.* **9** (2005), 299–325.

[18] J. Cruz and D. Ševčovič. On solutions of a partial integro-differential equation in Bessel potential spaces with applications in option pricing models. *Jpn. J. Ind. Appl. Math.* **37** (2020), 697–721.

[19] J. Cruz and D. Ševčovič. Option pricing in illiquid markets with jumps. *Applied Mathematical Finance.* **25** (2018), 389–409.

[20] G. Deelstra, I. Diallo, and M. Vanmaele. Bounds for Asian basket options. *Journal of Computational and Applied Mathematics.* **218** (2008), 215–228.

[21] L. C. Evans. Partial differential equations. *American Mathematical Society.* **19** (2022).

[22] S. Federico, P. Gassiat, and F. Gozzi. Utility maximization with current utility on the wealth: regularity of solutions to the HJB equation. *Finance Stoch.* **19** (2015), 415–448.

[23] I. Florescu and M. C. Mariani. Solutions to integro-differential parabolic problems arising in the pricing of financial options in a Lévy market. *Electronic Journal of Differential Equations (EJDE)[electronic only.]* (2010), 1–10.

[24] R. Frey. Perfect option hedging for a large trader. *Finance Stochast.* **2** (1998), 305–325.

[25] R. Frey and P. Patie. Risk management for derivatives in illiquid markets: a simulation study. *Advances in finance and stochastics: Springer, Berlin.* (2002), pp. 137–159.

[26] R. Frey and A. Stremme. Market volatility and feedback effects from dynamic hedging. *Math. Finance.* **7** (1997), 351–374.

[27] D. Henry. Geometric theory of semilinear parabolic equations. *Lecture Notes in Mathematics: Springer-Verlag, Berlin-New York.* **840** (1981), pp. iv+348.

[28] H. Ishii. Viscosity solutions of nonlinear partial differential equations. *Sugaku Expositions.* **9** (1996), 135–152.

[29] N. Ishimura and D. Ševčovič. On traveling wave solutions to a Hamilton-Jacobi-Bellman equation with inequality constraints. *Japan J. Ind. Appl. Math.* **30**(1) (2013), 51–67.

[30] M. Jandačka and D. Ševčovič. On the risk-adjusted pricing-methodology-based valuation of vanilla options and explanation of the volatility smile. *J. Appl. Math.* (2005), pp. 235–258.

[31] A. D. Jagtap, E. Kharazmi, and G. E. Karniadakis. Conservative physics-informed neural networks on discrete domains for conservation laws: Applications to forward and inverse problems. *Computer Methods in Applied Mechanics and Engineering, Elsevier.* **365** (2020), 113028.

[32] C. Jose, G. Maria, D. Ševčovič, and C. I. Udeani. "Linear and nonlinear partial integro-differential equations arising from finance." In: understanding integro-differential equations, Eds: J. Vasundhara Devi, Z. Drici, F. A. McRae. *Nova Science Publishers, Inc., Hauppauge.* (2023), 191–256, ISBN: 979-8-89113-040-1.

[33] S. Justin and S. Konstantinos. DGM: A deep learning algorithm for solving partial differential equations. *Journal of computational physics, Elsevier.* **375** (2018), 1339–1364.

[34] E. Kharazmi, Z. Zhang, G. E. Karniadakis. Variational physics-informed neural networks for solving partial differential equations. *arXiv preprint arXiv:1912.00873.* (2019).

[35] S. Kilianová, and D. Ševčovič. A transformation method for solving the Hamilton–Jacobi–Bellman equation for a constrained dynamic stochastic optimal allocation problem. *ANZIAM Journal.* **55** (2013), 14–38.

[36] S. Kilianová, and D. Ševčovič. Expected utility maximization and conditional value-at-risk deviation-based sharpe ratio in dynamic stochastic portfolio optimization. *Kybernetika.* **54**(6) (2018), 1167–1183.

[37] S. Kilianová and D. Ševčovič. Dynamic intertemporal utility optimization by means of Riccati transformation of Hamilton–Jacobi– Bellman equation. *Japan Journal of Industrial and Applied Mathematics.* **36**(2) (2019), 497–517.

[38] S. Kilianová, I. Melicherčík, D. Ševčovič. Dynamic Accumulation Model for the Second Pillar of the Slovak Pension System, *Finance a uver - Czech Journal of Economics and Finance.* **56**(11–12) (2006), 506–521.

[39] S. Kilianová and M. Trnovská. Robust Portfolio Optimization via solution to the Hamilton–Jacobi–Bellman Equation. *Int. Journal of Computer Mathematics.* **93** (2016), 725–734.

[40] D. Klatte. On the Lipschitz behavior of optimal solutions in parametric problems of quadratic optimization and linear complementarity. *Optimization: A Journal of Mathematical Programming and Operations Research.* **16**(6) (1985), 819–831.

[41] S. Kou. A jump-diffusion model for option pricing. *Management Science* **48** (2002), 1086–1101.

[42] Y. K. Kwok. Mathematical models of financial derivatives. *Second ed.: Springer Finance, Springer, Berlin.* (2008), pp. xvi+530.

[43] H. E Leland. Option pricing and replication with transactions costs. *The journal of finance.* **40** (1985), 1283–1301.

[44] Lu, Lu, Jin, P. Pang, G. Zhang, Z. G. Karniadakis. Learning nonlinear operators via DeepONet based on the universal approximation theorem of operators. *Nature Machine Intelligence, Elsevier.* **3**(3) (2021), 218–229.

[45] Z. Macová and D. Ševčovič. Weakly nonlinear analysis of the Hamilton–Jacobi–Bellman equation arising from pension savings management. *Int. J. Numer. Anal. Model.* **7**(4) (2010), 619–638.

[46] D. B. Madan, P. Carr, and E. C. Chang. The Variance Gamma Process and Option Pricing. *European Finance Review.* **2** (1998), 79–105.

[47] R. C. Merton. Option pricing when underlying stock returns are discontinuous. *J. Financial Economics.* **3** (1976), 125–144.

[48] J. C. Meyer and D. J. Needham. Extended weak maximum principles for parabolic partial differential inequalities on unbounded domains. *Proc. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci.* **470** (2014), 20140079.

[49] R. Mikulevičius and H. Pragarauskas. On the Cauchy problem for certain integro-differential operators in Sobolev and Hölder spaces. *Lithuanian Mathematical Journal.* **32** (1992), 238–264.

[50] P. Milgrom and I. Segal. Envelope theorems for arbitrary choice sets. *Econometrica.* **70**(2) (2002), 583–601.

[51] M. A. Onyido, M. O. Uba, C. I. Udeani, and P.U. Nwokoro. A hybrid scheme for fixed points of a countable family of generalized nonexpansive-type maps and finite families of variational inequality and equilibrium problems with applications. *Carpathian Journal of Mathematics.* **39**(1) (2023), 281–292.

[52] M.A. Onyido, R. B. Salako, M.O. Uba, and C.I. Udeani. Dynamics of classical solutions of a two-stage structured population model with nonlocal dispersal. *Mathematics.* **11**(4) (2023), 925.

[53] G. Palatucci and T. Kuusi. Recent developments in nonlocal theory. *De Gruyter Open.* (2017).

[54] C. V. Pao and W. H. Ruan. Positive solutions of quasilieanr parabolic systems with Dirichlet boundary condition. *Journal of Differential Equation.* **248**(5) (2010), 1175–1211.

[55] T. Post, Y. Fang, and M. Kopa. Linear tests for DARA stochastic dominance. *Management Science.* **61** (2015), 1615–1629.

[56] M. H. Protter and H. F. Weinberger. Maximum principles in differential equations. *Springer Science & Business Media.* (2012).

[57] M. Raissi, P. Perdikaris, and G. E. Karniadakis. Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *Journal of Computational physics, Elsevier.* **378** (2019), 686–707.

[58] I. SenGupta, M. C. Mariani, P. Amster. Solutions to integro-differential problems arising on pricing options in a Lévy market. *Acta Appl. Math.* **118** (2012) 237–249.

[59] D. Ševčovič. Nonlinear parabolic equations arising in mathematical finance. *Novel Methods in Computational Finance, Springer.* **15** (2017), 3–15.

[60] D. Ševčovič and C. I. Udeani. Learning the solution operator of a nonlinear parabolic equation using physics-informed deep operator network. *ECMI2023 Proceedings, Springer.* (2023).

[61] D. Ševčovič and C. I. Udeani. Hamilton–Jacobi–Bellman equation arising from optimal portfolio selection problem. *ECMI2023 Proceedings, Springer.* (2023).

[62] D. Ševčovič and C. I. Udeani Multidimensional linear and nonlinear partial integro-differential equation in Bessel potential spaces with applications in option pricing. *Mathematics.* **9**(13) (2021) 1463.

[63] D. Ševčovič, B. Stehlíková and K. Mikula. Analytical and numerical methods for pricing financial derivatives. *Nova Science Publishers, Inc., Hauppauge.* (2011), pp. 1–309.

[64] D. Ševčovič and M. Žitňanská. Analysis of the nonlinear option pricing model under variable transaction costs. *Asia-Pac. Financ. Mark.* **23** (2016), 153–174.

[65] P. J. Schönbucher and P. Wilmott. The feedback effect of hedging in illiquid markets. *SIAM J. Appl. Math.* **61** (2000), 232–272.

[66] R. E. Showalter. Monotone operators in Banach space and nonlinear partial differential equations. *American Mathematical Soc.* **49** (2013).

[67] E. M. Stein. Singular integrals and differentiability properties of functions. *Princeton Mathematical Series: Princeton University Press, Princeton, N.J.* **30** (1970), pp. xiv+290.

[68] C. I. Udeani and D. Ševčovič. Application of maximal monotone operator method for solving Hamilton–Jacobi–Bellman equation arising from optimal portfolio selection problem. *Japan Journal of Industrial and Applied Mathematics, Springer.* **5** (2021), pp 1–21.

[69] R. G. Vickson. Stochastic dominance for decreasing absolute risk aversion. *Journal of Financial and Quantitative Analysis.* **10** (1975), 799–811.

[70] Wu et al. Blow-up of solutions for a semilieanr parabolic equation involving variable source and positive initial energy. *Applied Mathematics Letters.* **26**(5) (2013), 539–543.

[71] G. C. Yanan, X. Bainian, Liu G. Mei Solving partial differential equations using deep learning and physical ponstraints. *Applied Sciences.* **10** (2020), 5917. doi = 10.3390/app10175917.

[72] B. Yu. The deep Ritz method: a deep learning-based numerical algorithm for solving variational problems. *Communications in Mathematics and Statistics: Springer.* **6**(1) (2018), 1–12.

[73] S. Yüzbaşı. Improved Bessel collocation method for linear Volterra integro-differential equations with piecewise intervals and application of a Volterra population model. *Applied Mathematical Modeling.* **40** (2016), 5349–5363.

[74] Y. Zhu, N. Zabaras, P. Koutsourelakis, and P. Perdikaris. Physics-constrained deep learning for high-dimensional surrogate modeling and uncertainty quantification without labeled data. *Journal of Computational Physics, Elsevier.* **394** (2019), 56–81.