No d'ordre : 500/23

UNIVERSITÉ SULTAN MOULAY SLIMANE

Faculté des Sciences et Techniques

Béni Mellal

Centre d'Etudes Doctorales: Sciences et Techniques

Formation doctorale: Mathématiques et Physique Appliquées

THÈSE

Présentée par

OULMELK Abdessamad

Pour l'obtention du grade de

Docteur

Spécialité : Mathématiques Appliquées

Identification of potential parameters in a nonlinear time-fractional diffusion model

---

Soutenue le 21/03/2023 devant le jury:

Pr. Aziz OUHINOU,	Université Sultan Moulay Slimane, FST Béni Mellal,	Président
Pr. Abdelmoutalib METRANE,	Université Sultan Moulay Slimane, ENSA Khouribga,	Rapporteur
Pr. Idriss BOUTAAYAMOU,	Université Ibn Zohr, FP Ouarzazate,	Rapporteur
Pr. Rachid FAKHAR,	Université Sultan Moulay Slimane, FP Khouribga,	Rapporteur
Pr. Amine LAGHRIB,	Université Sultan Moulay Slimane, ENSA Khouribga,	Examinateur
Pr. Mourad NACHAOUI,	Université Sultan Moulay Slimane, FST Béni Mellal,	Examinateur
Pr. Aissam HADRI,	Université Ibn Zohr, FP Ouarzazate	Co-Directeur
Pr. Lekbir AFRAITES,	Univesrité Sultan Moulay Slimane, FST Béni Mellal	Directeur

***DEDICATION**:*

I dedicate this thesis to my parents for bringing me up, to my brothers Mohamed and Mustapha, and to my sisters. This work is my gift in return for their selfless love.

Acknowledgments

First, I would like to express my immense gratitude and sincere thanks to my supervisor, Professor **Lekbir AFRAITES**, for his continuous support of my doctoral studies and related research. His patience, motivation, and immense knowledge have been invaluable in guiding me throughout the research and writing of this thesis.

I would also like to thank my co-supervisor, Professor **Aissam HADRI**, for the trust and sympathy they have shown me during the years of my thesis. Their human and scientific qualities have been an invaluable contribution to me.

It is a pleasure to thank Professor **Aziz OUHINOU** for agreeing to be the president of the jury.

I would like to extend my appreciation to Professors **Idriss BOUTAAYAMOU**, **Abdelmoutalib METRANE** and **Rachid FAKHAR** for agreeing to be referees of this work.

In addition, I am grateful to Professors **Amine LAGHRIB** and **Mourad NACHAOUI** for accepting examinate this work and to be members of the jury.

Obviously, none of this would have been possible without the support of my family. I would like to thank my parents, my brothers, my sisters, and my friends.

Abstract

This thesis aims to study the inverse problem of identifying parameters in a nonlinear subdiffusion model from additional data. The nonlinear subdiffusion model involves a Caputo fractional derivative of order $\alpha \in (0, 1)$ in time. The first part explores the well-posedness of the direct problem; showing the existence, uniqueness, and regularity of its solution using Mittag-Leffler functions. Furthermore, the numerical approximation of the fractional equation will be determined using the finite difference method, and some numerical tests will be presented. The second part examines the inverse potential problem using regularized optimal control problem approach; establishing existence of optimal solution and showing uniqueness and stability with respect to data based on optimality conditions of considered functional and solution regularity of direct problem. Numerical results for certain examples are presented based on gradient descent approach. The next part is devoted to conceptual implementation of alternating direction multiplier method (ADMM) to solve optimal control problem subject to fractional diffusion equation constraint; firstly showing convexity of functional so as to apply ADMM and then establishing its convergence. To show effectiveness of proposed method, numerical experiments are presented. In the fourth part, an artificial neural network approach (ANN) for solving our inverse potential problem was proposed as a new and more widely used method. The last part is devoted to numerical comparison of all proposed methods, such as the gradient descent method, the ADMM method, and the ANN method, to show their effectiveness.

Keywords: Time-fractional diffusion equation, inverse problem, identification of parameters, optimal control problems, gradient descent algorithm, alternating direction multiplier method (ADMM) and artificial neural networks (ANN).

Résumé

Cette thèse vise à étudier le problème inverse de l'identification des paramètres dans un modèle de subdiffusion non linéaire à partir de données supplémentaires. Le modèle de subdiffusion non linéaire fait intervenir une dérivée fractionnaire de Caputo d'ordre $\alpha \in (0, 1)$ dans le temps. La première partie explore l'étude du problème direct. Dans ce cadre, nous montrons l'existence, l'unicité et la régularité de la solution du problème direct en utilisant les fonctions de Mittag-Leffler. En outre, l'approximation numérique de l'équation fractionnaire sera déterminée en utilisant la méthode des différences finies, et quelques tests numériques seront présentés. La deuxième partie examine le problème inverse en utilisant l'approche de contrôle optimal régularisé, où nous montrons l'existence du minimiseur et la stabilité de notre problème inverse à l'aide des conditions d'optimalité de la fonctionnelle coût considérée. Nous fournissons des résultats numériques pour certains exemples basés sur l'approche de descente du gradient. La troisième partie est consacrée à la mise en œuvre conceptuelle de la méthode du multiplicateur à direction alternée (ADMM) pour résoudre le problème de contrôle optimal. Nous montrons au début la convexité de la fonctionnelle pour pouvoir appliquer l'ADMM, ensuite, nous établissons sa convergence. Afin de montrer l'efficacité de la méthode proposée, nous présentons quelques expériences numériques. Quant à la quatrième partie, nous proposons une approche de réseau de neurones artificiels (ANN) pour résoudre notre problème inverse. Il s'agit d'une nouvelle méthode plus largement utilisée pour résoudre les problèmes inverses. La dernière partie est consacrée à la comparaison numérique de toutes les méthodes proposées, telles que la méthode de descente de gradient, la méthode ADMM et la méthode ANN, pour montrer leurs efficacités.

Mot clé: Equation de diffusion fractionnaire, problème inverse, identification des paramètres, problèmes de contrôle optimal, algorithme de descente du gradient, méthode du multiplicateur à direction alternée (ADMM) et méthode de réseau de neurones artificiels (ANN).

Contents

List of Figures

List of Tables

General introduction

The Non-Integer Order Calculus, usually known as Fractional Calculus, is the field of mathematics that aims to interpolate the classical derivatives and integrals which generalize them for any orders, not necessarily integer orders. More precisely, it studies the generalization of the derivation and integration of integer and non-integer orders (fractional). The theory of fractional derivation is a subject almost as old as classical calculus as we know it today; its origins go back to the end of the 17th century, the time when Isaac Newton and Gottfried Wilhelm Leibniz developed the foundations of differential and integral calculus [50]. Fractional calculus has become one of the most developed fields of mathematical analysis, especially after the first specialized conference organized by B. Ross on "Partial Calculus and Its Applications" and also have provided in [59] a brief historical review of the evolution of fractional calculus. Many applications of various kinds of fractional differential equations have become a target of specialists due to both theoretical and practical reasons [7]. It has gone through rapid development and has been revealed as a powerful tool in the modeling of certain phenomena in several sciences such as Physics, Chemistry, Biology, Engineering, and Finance, especially during the past three decades.

In the literature, numerous definitions of fractional derivatives have been offered, such as the Riemann-Liouville, the Caputo, the Grünwald-Letnikov, and the Riesz-Feller [57]. Although they are different, they are all related to each other. The Riemann-Liouville played an important role for its application in pure mathematics, while Caputo was introduced to respond to applied problems. Indeed, M. Caputo was the first to give applications of fractional calculus to mechanics, especially to linear models of viscoelasticity [10, 11]. Caputo derivatives allow the use of physically interpretable initial conditions, which is not permitted by the Riemann-Liouville. The Grünwald-Letnikov is the most obvious approach to define fractional derivatives. It is mainly used for numerical approximation of fractional derivatives. However, dealing with fractional derivatives as a limit of fractional-order difference is not convenient due to its mathematical complexity. Therefore, some

studies use the Grünwald-Letnikov numerically, but try to solve the initial problem with other definitions.

The fractional diffusion equation is obtained from the standard diffusion equation by replacing the first-order derivative with a fractional derivative of order α. Such equations have proved particularly useful in the context of anomalous diffusion, which is characterized by an asymptotic long-time behavior of the mean square displacement of the form $\langle x^2(t)\rangle \sim t^\alpha,\ t \longrightarrow \infty$. They have received great attention from many researchers working in different disciplines of science and technology. In other words, some recent publications show the importance of fractional diffusion equations in the mathematical modeling of many real-world phenomena, due to their ability to describe anomalous diffusion, which are known as sub-diffusion or super-diffusion phenomena [8, 9, 46, 63]. Moreover, these equations have many successful applications in physics, biology, and finance [45, 58]. In the literature, there are many theoretical and numerical studies on fractional diffusion equations with distinct kinds of fractional derivatives, such as the Riemann-Liouville fractional derivative and the Caputo fractional derivative. For instance, in [42], the authors show the maximum principle of the solution, while [60] gave some existence and uniqueness results. For numerical methods, finite element methods and finite difference approximations for time-fractional partial differential equations have been provided in [23, 26, 38].

However, in real situations, the data of the fractional diffusion equation, such as the initial data, the diffusion coefficient, the source term, or the order fractional derivative, are not all known, which leads to the study of a fractional inverse problem. For this reason, additional measurement data is required to deal with this type of problem. At this stage, numerous contributions have been introduced to solve various fractional inverse problems. For the inverse source problems, many results concerning this problem have been given in [41, 52, 60, 70, 75]. Also, several theoretical and numerical studies on the inverse coefficient problems have been presented in [27, 65, 67, 69]. Regarding other types of inverse problems, we cite [25, 60, 68].

The main objective of our thesis is to study an inverse potential problem for a nonlinear time-fractional diffusion equation in a bounded domain. Let Ω be a bounded domain in $\mathbb{R}$ with boundary $\partial\Omega$. We consider an initial-boundary value problem for the nonlinear

time-fractional diffusion equation in the form:

$$\begin{cases} D_t^\alpha u(x,t) - \Delta u(x,t) + p(x)f(u) &= 0, & (x,t) \in \Omega \times (0,T), \\ u(x,0) &= \phi(x), & x \in \Omega, \\ u(x,t) &= 0, & x \in \partial\Omega \quad t \in (0,T), \end{cases} \tag{1}$$

where D_t^α is the Caputo fractional derivative of order $0 < \alpha < 1$. The fractional derivative $D_t^\alpha u$ recovers the usual first order derivative as the order $\alpha \longrightarrow 1^-$ for sufficiently smooth functions u. The functions f is a smooth function and ϕ is the initial data. The time-fractional equation (1) is a direct problem if all elements such as the fractional order α, potential p, nonlinear term f, and initial ϕ are provided. In the case where one of the parameters is unknown, we talk about a fractional inverse problem. The fractional inverse problem treated in our thesis is to determine the potential p in time-fractional equation (1) from the following additional data:

$$u(x,T) = g(x), \qquad x \in \Omega, \tag{2}$$

where g is the noisy data.

There are many works on the mathematical treatment of inverse problems (1)-(2) when $\alpha = 1$ and $f(u) = u$. In the case of $\alpha = 1$, this corresponds to parabolic equations rewritten as follows:

$$\partial_t u - \Delta u + p(x)f(u) = 0, \qquad (x,t) \in \Omega \times (0,T). \tag{3}$$

The inverse problem for the recovery of p in (3) has been extensively studied by several authors. The existence of p is obtained by a fixed point theorem in [72], as well as the existence, uniqueness, and stability results are established on the basis of the optimal control method in [78]. Further, in the linear case $f(u) = u$, the inverse problem (1)-(2) of recovering p has been studied in [29] and [46], where they proved that the inverse problem is locally Lipschitz for small terminal time, under certain conditions on the initial data.

We will need to make several assumptions about the data, denoted (**H**), in order to study our inverse problem, which we outline in the lines that follow:

$(\mathbf{H}_1)$: The function f is Lipschitz, $f(0) = 0$.

$(\mathbf{H}_2)$: The function f is convex, differentiated, $f' > 0$ and f' is Lipschitz.

$(\mathbf{H}_3)$: The function $g \in L^2(\Omega)$.

$(\mathbf{H}_4)$: $\phi(x)$ is in $L^2(\Omega)$.

Motivated by the aforementioned studies, this thesis proposes a theoretical study and different numerical methods to solve the inverse problem (1)-(2). This thesis uses and compares the several tools bellow:

- ***Optimal control problem:*** The purpose of optimal control is to influence the behavior of a dynamic system to achieve a desired objective. Optimal control has a wide variety of applications where dynamics can be optimally controlled, such as in aerospace, aeronautics, chemical plants, mechanical systems, finance and economics. It can also be used to solve inverse problems, where the goal is to determine the unknown in an equation from additional conditions. An important application that we will study in this work is to transform the inverse problem into an optimal control problem, in which we determine the unknown parameter p in a time-fractional PDE (1) from additional conditions.

- ***Gradient computation and gradient descent:*** Gradient descent is an iterative approach that finds the minimum of a function. It is an optimization algorithm that finds the parameters or coefficients of a function for which the function has a minimum value. This approach is also a useful way to solve control problems numerically, requiring the computation of the gradient for the regular cost function.

- ***ADMM:*** The Alternating Direction Method of Multipliers (ADMM) is an algorithm developed to provide a general framework that can be applied to a variety of convex optimization problems. It can split a large problem into a series of subproblems, each of which is easy to solve. ADMM can be viewed as an attempt to combine the benefits of dual decomposition and augmented Lagrangian methods for constrained optimization; it turns out to be equivalent, fast, making very few assumptions about the properties of the objective function, and having relations to many other algorithms. In practice, it converges quickly, with reasonable precision, and is easy to implement.

In the literature, there are several works using the ADMM method to solve some control problems. For example, the authors of [64] applied ADMM to parabolic optimal control problems with control constraints and demonstrated the convergence of the algorithm using an inexactness criterion. In [18] the ADMM numerical method was used to solve linear parabolic state-constrained optimal control problems. Also, the ADMM method has been applied to the optimization problem constrained by a stationary Maxwell system in [20]. To the best of our knowledge in the literature, this is the first time that the idea of the ADMM has been explored for solving the inverse problem of identifying the parameters of a nonlinear time-fractional PDE model.

- ***Artificial neural network:*** Artificial Neural Networks (ANNs) are currently the most widely-used Machine Learning techniques. These Neural Networks were first developed in the 1970s, but due to recent increases in computing power, they have become extremely popular and are now found almost everywhere. Recently, they have been widely used to solve inverse problems; this method is based on replacing the unknown parameter with a neural network. Therefore, instead of identifying the unknown parameter in a model, we estimate the weights and biases as neural network parameters.

  In the literature, neural network technique for PDEs was early proposed by Lagaris et al.[33, 34], and has been used to a vast number of studies on PDE approximation problem. Specifically, the approach was developed in various follow-up studies for the high order differential equations [44] and partial differential equations [2]. Then they applied it to represent the solution of PDEs implicitly, and the boundary constraints are enforced through collocation. The approaches of these techniques primarily depend on the universal approximation feature of neural network, which was demonstrated in the pioneering work [12] for one single hidden layer, and later expanded and improved in [21].

  In real-world research, solving inverse problems is one of the great applications by artificial neural networks. Recently, there has been a lot of interest in Artificial Neural Networks (ANN) methods for PDE inverse problems. It is verified to be effective for various problems, including fractional ADEs [55], stochastic problems

[74] and the mechanism of physics in [15]and so on. As a matter of fac, the artificial neural networks method for determining the nonlinearity of the monodomain model is shown in [32]. A novel artificial neural network method for solving Cauchy inverse problems is proposed.

We describe the main challenges and innovations of this thesis.

The main difficulties are:

1. We treat a nonlinear fractional model, which is challenging to examine,because the fractional derivative and nonlinear term make studying existence, uniqueness, and stability more difficult.

2. The optimal control problem subject to both fractional diffusion equation (FDE) constraints and additional constraints on the control variables is generally challenging, both from a theoretical analysis and algorithmic design perspective. To answer this, the ADMM approach can be directly applied to such an optimal control problem. However, the convergence results associated with this approach are more difficult.

3. Solving inverse potential problems using neural networks is much more difficult, because of the limited literature in this new field.

The following are the most significant contributions made by this thesis:

1. We first address the question of the existence, uniqueness, and regularity of solutions for the direct problem and we present the reformulation of the inverse problem into an optimal control one.

2. We study the existence of the minimizer and the stability of the inverse problem using optimal conditions.

3. To apply the ADMM technique, we need to prove the convexity of the optimal control problem. To do so, we show the continuity of the observation operator and its differentiability. Then, we compute the first derivative of the cost function with respect to the potential parameter p, which helps us to demonstrate the convexity.

4. In the literature, many convergence results for the ADMM method applied to parabolic optimal control problems have been discussed [18, 64]. In this thesis,

we prove the convergence of the proposed ADMM method applied to our inverse potential problem.

5. An artificial neural network method is proposed for solving fractional inverse problem.

6. The numerical simulations of all proposed methods are presented for cases with and without noise, and a comparison between them is established.

Outline of the thesis

The thesis is organized as follows:

In **Chapter** 1, we provide a mathematical background that serves as a foundation for understanding the successive chapters. In particular, we give some fundamental results on Sobolev spaces, vector-valued Lebesgue spaces, and inequalities on Sobolev spaces. We also present the definitions and some properties of fractional derivatives of order $\alpha \in (0, 1)$ as a significant factor in this model. The Mittag-Leffler functions and their properties play an important role in our study, so we discuss, especially their importance in proving the well-posedness of the direct problem.

Chapter 2 is devoted to study the existence, uniqueness and regularity of the solution for direct problem (1), which will be needed later in the identification methods. To overcome the difficulties posed by the nonlinearity of the model, we opt for the representation of the solution for nonlinear fractional diffusion problem by means of two-parameters $(\alpha,\ \beta)$ Mittag-Leffler functions $E_{\alpha,\beta}(z)$ as well as the eigenvalues and eigenfunctions of the elliptic operator. We also propose the numerical approximation of the direct problem by the finite difference method, and present some numerical tests. The results of this chapter are published in [51].

In **Chapter** 3, we reformulate the inverse problem into an optimal control one using the least squares method. First, we will establish the existence of the optimal solution. Second, we compute the necessary optimality conditions using the adjoint problem and study the main result concerning the stability and uniqueness of the minimizer. Finely, we give the numerical approximation of the optimal control problem and some numerical tests based on the gradient descent algorithm. The results of this chapter are published in [51].

In **Chapter** 4, We study the inverse problem from another observation data defined in the given Ω_1 subdomain of Ω. First, we reformulate it into a Least-Squares optimization problem. Then, we prove the convexity of the considered cost function by using its first derivative. To solve this problem numerically, we adapt the Alternating Direction Method of Multipliers (ADMM) and establish its convergence. To show the effectiveness of the proposed method, we present some numerical experiments. The results of this chapter are published in [53].

The **Chapter** 5 deals with the identification of the parameter p in the subdiffusion model (1) using artificial neural networks. To do so, we first give the formulas and properties of neural networks. We then show that the optimal control problem parametrized by a neural network is well-posed and demonstrate the existence of the solution to the control problem, providing a mathematical analysis and the derivation of optimal conditions. Moreover, various numerical examples show that our method is effective.

In **Chapter** 6, we present numerical comparison between the proposed methods to demonstrate the efficiency of each algorithm with regular and singular tests. These results will be provided in cases with and without noise to prove the stability of each algorithm.

Finally, we will conclude this thesis with a conclusion and some perspectives.

Chapter 1

Preliminary results

In this chapter, we recall several definitions, basic concepts, notations, and elementary results that are used throughout this thesis. In particular, we give some fundamental results on Sobolev spaces, vector-valued Lebesgue spaces, and inequalities on Sobolev spaces. We also present the definitions and some properties of fractional derivatives of order $\alpha \in (0,1)$ as a significant factor in this model. The Mittag-Leffler functions and their properties play an important role in this study, which we will also discuss, especially its importance in proving the well-posedness of the direct problem.

1.1 Notations and definitions

In this section, we introduce some notations that will be used throughout the thesis. Let $\Omega \subset \mathbb{R}^d$ ($d \in \mathbb{N}^\star$) be an open bounded domain with piecewise smooth boundary $\partial\Omega$,

$$I = (0,T), \quad Q = \Omega \times I,$$

and let's φ be an functions defined from Ω into $\mathbb{R}$ and ψ be an functions defined from Ω into $\mathbb{R}^d$.

- We use the symbol $\mathcal{O}$ to denote a domain which may stand for Ω, I, Q or $\mathbb{R}$. Let $L^2(\mathcal{O})$ be the space of measurable functions whose square is Lebesgue integrable in $\mathcal{O}$. The inner product of $L^2(\mathcal{O})$ is defined by

$$(u,v)_{\mathcal{O}} = \int_{\mathcal{O}} uv \, \mathrm{d}\mathcal{O}, \quad \forall u, v \in L^2(\mathcal{O})$$

- We denote by $\langle ., . \rangle$ the scalar product in $\mathbb{R}^d$ defined by:

$$\langle x, y \rangle = \sum_{i=1}^{d} x_i y_i \quad \text{for all} \quad x, y \in \mathbb{R}^d,$$

where the associated norm is: $|x| = \langle x, x \rangle^{\frac{1}{2}}$.

- The gradient of φ is represented by:

$$\nabla \varphi = \begin{pmatrix} \frac{\partial \varphi}{\partial x_1} \\ \vdots \\ \frac{\partial \varphi}{\partial x_d} \end{pmatrix}.$$

- The Laplacian of φ is written as follows:

$$\Delta \varphi = \sum_{i=1}^{d} \frac{\partial^2 \varphi}{\partial x_i^2}.$$

- The divergence of ψ is denoted by:

$$\nabla \cdot \psi = \sum_{i=1}^{d} \frac{\partial \psi_i}{\partial x_i}.$$

- We denote the normal derivative of φ by:

$$\frac{\partial \varphi}{\partial \nu} = \nabla \varphi \cdot \nu,$$

where ν is a direction that is outwardly normal on Ω.

- Let $\alpha = (\alpha_1, \ldots, \alpha_d)$ is d-dimensional multi-index in $\mathbb{N}^d$ of length $|\alpha| = \alpha_1 + \cdots + \alpha_d$, the $|\alpha|$-order partial derivative of function φ noted $D^\alpha \varphi$, is defined as follows:

$$D^\alpha \varphi(x) = \frac{\partial^{|\alpha|} \varphi(x)}{\partial x_1^{\alpha_1} \cdots \partial x_d^{\alpha_d}}.$$

- The Lebesgue measure of Ω is denoted by the symbol $|\Omega|$, while the dimensional Lebesgue measure of $\partial\Omega$ is denoted by the symbol $|\partial\Omega|$.

1.2 Functional spaces

In this section, we recall some definitions and properties of some functional spaces. In particular, we present some definitions of Sobolev spaces, then we introduce some Sobolev injections and some Sobolev inequalities. To do this, we start by defining Lebesgue spaces. We refer the reader, for example, to [3, 13, 39], for more details about these functional spaces and their properties.

1.2.1 The spaces $L^p(\Omega)$

Let $p \in [1, +\infty[$, we define the space $L^p(\Omega)$ as the space of measurable functions whose p^{th} power is integrable, in the Lebesgue sense. The space $L^p(\Omega)$ is equipped with the norm

$$\|\varphi\|_{p,\Omega} = \left(\int_\Omega |\varphi|^p dx \right)^{\frac{1}{p}},$$

is a Banach space. In particular, for $p = 2$, $L^2(\Omega)$ is the Hilbert space equipped with the norm

$$\|\varphi\|_{\Omega} = \left(\int_\Omega |\varphi|^2 dx \right)^{\frac{1}{2}},$$

associated with the scalar product

$$(\varphi, \psi)_\Omega = \int_\Omega \varphi(x)\psi(x)dx.$$

$L^\infty(\Omega)$ is the space of measurable functions φ, endowed with the norm

$$\|\varphi\|_{L^\infty(\Omega)} = \inf_{\substack{M \subset \Omega \\ mesM=0}} \sup_{x \in \Omega \setminus M} |\varphi(x)| \equiv ess \sup_{x \in \Omega} |\varphi(x)| \tag{1.1}$$

is finite. $L^\infty(\Omega)$ equipped with the norm (1.1) is a Banach space.

Prior to discussing the Sobolev spaces, we provide some results on the spaces of functions of class $C^m(\Omega)$, for $m \in \mathbb{N}$.

1.2.2 The spaces $C^m(\Omega)$, $C^m(\bar{\Omega})$, $D(\Omega)$ and $D(\bar{\Omega})$

- The space of continuous functions on Ω is denoted by $C^0(\bar{\Omega})$ and has the following norm:

$$\|\varphi\|_\infty = \sup_{x\in\Omega} |\varphi(x)|.$$

- $C^m(\Omega)$ is the space of functions φ, which have partial derivatives $D^\alpha\varphi$ continuous on Ω, for any $\alpha \in \mathbb{N}^d$, such that $|\alpha| \leq m$.

- We note by $C_0^m(\Omega)$ the subspace of $C^m(\Omega)$ of functions that vanish at infinity.

- $C^m(\bar{\Omega})$ is a subspace of $C^m(\Omega)$ of functions φ such that $D^\alpha\varphi$ is continuous on $\bar{\Omega}$, for any $\alpha \in \mathbb{N}^d$, such that $|\alpha| \leq m$. It has the following norm:

$$\|\varphi\|_{C^m(\bar{\Omega})} = \max_{0\leq|\alpha|\leq m} \sup_{x\in\bar{\Omega}} |D^\alpha\varphi(x)|\,. \tag{1.2}$$

- For $0 < \lambda \leq 1$, we denote by $C^{m,\lambda}(\bar{\Omega})$ the subspace of $C^m(\bar{\Omega})$, of functions φ whose derivatives $D^\alpha\varphi$ are Hölder continuous of exponent λ, for any $\alpha \in \mathbb{N}^d$, such that $|\alpha| = m$; i.e. there exists $c > 0$, such that

$$|D^\alpha\varphi(x) - D^\alpha\varphi(y)| \leq c|x-y|^\lambda \qquad \forall x, y \in \Omega. \tag{1.3}$$

  In $C^{m,\lambda}(\bar{\Omega})$, we consider the following norm

$$\|\varphi\|_{C^{m,\lambda}(\bar{\Omega})} = \|\varphi\|_{C^m(\bar{\Omega})} + \max_{|\alpha|=m} \sup_{\substack{x,y\in\Omega \\ x\neq y}} \frac{|D^\alpha\varphi(x) - D^\alpha\varphi(y)|}{|x-y|^\lambda}. \tag{1.4}$$

- The space of indefinitely differentiable functions in Ω with compact support is denoted by $D(\Omega)$, whose the topological dual $D'(\Omega)$ is the space of distributions on Omega.

- $D(\bar{\Omega})$ is the space of restrictions to Ω of functions in $D(\mathbb{R}^d)$.

The next section includes definitions and properties of Sobolev spaces.

1.2.3 The Sobolev spaces

Let m be an integer number, we define the Sobolev space $H^m(\Omega)$ by

$$H^m(\Omega) = \left\{\varphi \in L^2(\Omega)/D^\alpha\varphi \in L^2(\Omega), \text{ for any } |\alpha| \leq m\right\} \tag{1.5}$$

where D^α is the derivative in the distributions sense of order $|\alpha|$ of the function φ.

The Hilbert space $H^m(\Omega)$ is equipped with the following norm:

$$\|\varphi\|_{m,\Omega} = \left(\sum_{|\alpha|\leq m} \|D^\alpha\varphi\|^2_{0,\Omega}\, dx\right)^{\frac{1}{2}}. \tag{1.6}$$

For any real $s = m + \sigma$ with $m \in \mathbb{N}$ and $0 < \sigma < 1$, we define $H^s(\Omega)$ as being the space of functions $\varphi \in H^m(\Omega)$ which verify:

$$\int\int_{\Omega\times\Omega} \frac{|D^\alpha\varphi(x) - D^\alpha\varphi(y)|^2}{|x-y|^{d+2\sigma}} dxdy < +\infty \quad \text{for} \quad |\alpha| = m. \tag{1.7}$$

$H^s(\Omega)$ equipped with the following norm

$$\|\varphi\|_{s,\Omega} = \left(\|\varphi\|^2_{m,\Omega} + \sum_{|\alpha|=d} \int\int_{\Omega\times\Omega} \frac{|D^\alpha\varphi(x) - D^\alpha\varphi(y)|^2}{|x-y|^{d+2\sigma}} dx\, dy\right)^{\frac{1}{2}} \tag{1.8}$$

is a Banach space.

The space $H^1_0(\Omega)$ is defined as the adherence of $D(\Omega)$ in $H^1(\Omega)$. If the boundary $\partial\Omega$ is smooth, $H^1_0(\Omega)$ is defined as follows

$$H^1_0(\Omega) = \left\{\varphi \in H^1(\Omega) \;/\; \gamma(\varphi) = \varphi_{|\partial\Omega} = 0\right\} \tag{1.9}$$

where γ denotes the trace operator which is a linear and continuous application of $H^1(\Omega)$ in $L^2(\partial\Omega)$. The space $H^1_0(\Omega)$ equipped with the following norm

$$|\varphi|_{1,\Omega} = \|\nabla\varphi\|_{0,\Omega} \tag{1.10}$$

is a Hilbert space, with topological dual is $H^{-1}(\Omega)$.

We note by Γ_0 a non-zero measure part of $\partial\Omega$, the space

$$H^1_{\Gamma_0}(\Omega) = \left\{\varphi \in H^1(\Omega) \ / \ \varphi_{|\Gamma_0} = 0\right\} \tag{1.11}$$

equipped with the norm (1.10) is also a Hilbert space.

1.2.4 Embeddings of Sobolev spaces

We now give some embeddings of Sobolev spaces, which will be used later. For more details, See [39, 40]. The following continuous embedding of the Sobolev spaces is now valid.

Theorem 1.1. *Let Ω be a domain with Lipschitz boundary. The following embeddings are continuous:*

i) If $m < \frac{d}{2}$, then $H^m(\Omega) \hookrightarrow L^{p^*}(\Omega)$, where $p^* = \frac{2d}{d-2m}$.

ii) $H^s(\Omega) \hookrightarrow L^{p^*}(\Omega)$ for all p^* verifying $\frac{1}{p^*} = \frac{1}{2} - \frac{s}{2}$.

iii) $H^1(\Omega) \hookrightarrow H^{1-\delta}(\Omega)$, for all $0 < \delta < \frac{1}{2}$.

iv) If k is an integer, such that $0 \leq k < m - \frac{d}{2} < k + 1$, then $H^m(\Omega) \hookrightarrow C^{k,\alpha}(\overline{\Omega})$, for $\alpha = m - \frac{d}{2} - k$.

We will now quote the general result concerning compact embeddings of Sobolev spaces.

Theorem 1.2. *Let Ω be a domain with Lipschitz boundary. Then the following embeddings are compact:*

i) If $m < \frac{d}{2}$, then $H^m(\Omega) \hookrightarrow L^q(\Omega)$ for all $q < p^* = \frac{2d}{d-2m}$.

ii) $H^1(\Omega) \hookrightarrow H^{1-\delta}(\Omega)$, for all $0 < \delta < \frac{1}{2}$.

iii) If k is an integer value, such that $0 \leq k < m - \frac{d}{2} < k + 1$, then $H^m(\Omega) \hookrightarrow C^{k,\beta}(\overline{\Omega})$ for all $\beta < \alpha$.

1.3 Vector-valued spaces

In this section, we present some properties of Lebesgue spaces and vector-valued Sobolev spaces. The latter play an important role in the study of evolutionary problems (see for example [14, 73]). Let $p \in [1, +\infty[$, I be an interval of $\mathbb{R}$ and E be a Banach space equipped with the norm $\|.\|_E$.

1.3.1 The spaces $L^p(I, E)$

Definition 1.1. *The definition of the space $L^p(I, E)$ is:*

$$L^p(I, E) = \left\{ u : I \longrightarrow E \text{ mesurable } / \int_I \|u(t)\|_E^p dt < +\infty \right\}.$$

It is equipped with the norm

$$\|u\|_{L^p(I,E)} = \left(\int_I \|u(t)\|_E^p dt \right)^{\frac{1}{p}}.$$

The definition of the space $L^\infty(I, E)$ if $p = +\infty$ is

$$L^\infty(I, E) = \left\{ u : I \longrightarrow E \text{ mesurable } / \exists C > 0, \ \|u(t)\|_E \leq C \quad \text{for almost all } t \in I \right\},$$

it is equipped with the norm

$$\|u\|_{L^\infty(I,E)} = \inf \left\{ C > 0 \ / \ \|u(t)\|_E \leq C \ \text{ for almost all } t \in I \right\}.$$

We then have the following results (see for example [14, 73]).

Proposition 1.1. *For all p, q and $r \in [1, +\infty[$, we have*

i) Let u be a measurable function defined from I into E, then $u \in L^p(I, E)$ if and only if there exists a function $v \in L^p(I)$, such that $\|u(t)\|_E \leq v$ for almost all $t \in I$.

ii) In the event that E is continuously embedding into F, then $L^p(I, E)$ is similarly continuously embedding into $L^p(I, F)$.

iii) If p, q et r are real numbers such that $p \leq r \leq q$, then $L^p(I, E) \cap L^q(I, E) \subset L^r(I, E)$. Moreover if I is bounded, we have $L^p(I, E)$ continuously embedding into $L^q(I, E)$

for all $q \leq p$.

Proposition 1.2. *For all* $p \in [1, +\infty[$, *we have*

i) $L^p(I, E)$ *is a Banach space.*

ii) The space $\mathcal{D}(I, E)$ *is dense in* $L^p(I, E)$, *where* $\mathcal{D}(I, E)$ *is the space of functions defined from* I *into* E, *of class* $\mathcal{C}^\infty$ *and with compact support in* I.

Theorem 1.3. *Assume that the interval* I *is open, then for all* $p \in]1, +\infty[$, *we have*

i) If E *is reflexive, then the space* $L^p(I, E)$ *is reflexive and its dual space is* $L^q(I, E^{'})$, *where* $\frac{1}{q} + \frac{1}{p} = 1$ *and* $E^{'}$ *is the dual of* E.

ii) If E *is separable, then the space* $L^p(I, E)$ *is separable.*

In the next subsections, we shall introduce vector-valued Sobolev spaces.

1.3.2 The spaces $W^{1,p}(I, E)$

Definition 1.2. *The definition of the space* $W^{1,p}(I, E)$ *is:*

$$W^{1,p}(I, E) = \{u \in L^p(I, E) \ / \ u' \in L^p(I, E)\},$$

where u' *is the first derive of* u. *It is equipped with the norm*

$$\|u\|_{W^{1,p}(I,E)} = \|u\|_{L^p(I,E)} + \|u'\|_{L^p(I,E)}.$$

Remark 1.1. *Generally speaking, the following space is often used to solve evolutionary problems*

$$W^{1,(p,q)}(I, E, F) = \{u \in L^p(I, E) \ / \ u' \in L^q(I, F)\},$$

which is equipped with the norm $\|u\|_{W^{1,p}(I,E)} = \|u\|_{L^p(I,E)} + \|u'\|_{L^q(I,F)}$, *where the spaces* E *and* F *are Banach spaces and* $\frac{1}{q} + \frac{1}{p} = 1$. *In particular, in most cases* F *is considered as the topological dual of the space* E. *More precisely, one often uses the space* $W(I)$ *defined by*

$$W(I) = \left\{u \in L^p(I, E) : u' \in L^q(I, E^{'})\right\},$$

where $E^{'}$ *is the dual of* E.

Now, we are in position to state the following properties (see for example [14, 73]).

Proposition 1.3. *For all $p \in [1, +\infty[$, we have*

i) $W^{1,p}(I, E)$ *is Banach space.*

ii) If E *is separable then* $W^{1,p}(I, E)$ *is separable.*

iii) If E *is separable and reflexive then* $W^{1,p}(I, E)$ *is reflexive.*

iv) $\mathcal{C}^\infty(\bar{I}, E)$ *is dense in* $W^{1,p}(\bar{I}, E)$*, where* $\mathcal{C}^\infty(\bar{I}, E)$ *is the set of functions of class* $\mathcal{C}^\infty$ *defined from* $\bar{I}$ *into* E*.*

We also have the following result (See [14]).

1.3.3 Embeddings into spaces $L^p(I, E)$

The best known compactness results on vector-valued spaces are those of Aubin [6] and Simon [62]. The first one assumes the reflexivity of the spaces, while Simon generalizes the result to the case of non-reflexivity. These results are very important, in particular when studying certain partial differential equations.

Lemma 1.1 (Aubin-Lions)**.** *Let $p \in]1, +\infty[$ and let V, E and F be three Banach spaces such that $V \hookrightarrow E \hookrightarrow F$ (continuously embedded). if A is a bounded part in $W^{1,p}(I; F)$, then A is relatively compact in $\mathcal{C}(\bar{I}; F)$.*

In the case of the space $W = \left\{u \in L^p\left(\bar{I}; E\right) \mid u' \in L^q\left(\bar{I}; F\right)\right\}$, the following lemma gives a similar result known as the Aubin-Simon lemma, the proof can be found for example in [6].

Lemma 1.2 (Aubin-Lions)**.** *Let E, V and F be three Banach spaces with $E \subseteq V \subseteq F$. Suppose that E is compactly embedded in V and that V is continuously embedded in F. For $1 \leq p, q \leq \infty$, let*

$$W = \left\{u \in L^p\left(\bar{I}; E\right) \mid u' \in L^q\left(\bar{I}; F\right)\right\}.$$

(i) If $p < \infty$ then the embedding of W into $L^p(\bar{I}; V)$ is compact.

(ii) If $p = \infty$ and $q > 1$ then the embedding of W into $C(\bar{I}; V)$ is compact.

1.4 Useful functions and fractional derivatives

In this section, we present the definitions and some properties of the Gamma and Mittag-Lafller functions. These functions are crucial in both the theory of fractional-order differential equations and its applications.

1.4.1 The Gamma function

One of the basic tools of fractional calculus is the Gamma function, which naturally extends the factorial to the positive real numbers (and even to the complex numbers with positive real parts).

Definition 1.3. *For $z \in \mathbb{C}$, such that $\mathrm{Re}(z) > 0$ the Gamma function $\Gamma(z)$ is defined by the integral*

$$\Gamma(z) = \int_0^\infty t^{z-1} e^{-t} dt,$$

with $\Gamma(1) = 1$, $\Gamma(0^+) = +\infty, \Gamma(x)$ is a monotonic and strictly decreasing function for any real $0 < x < 1$.

We will present some propositions related to the Gamma function that we can use in the next chapters (see for example [31, 57]).

Propositions 1.1. *For all $\mathrm{Re}(z) > 0$ and for all $n \in \mathbb{N}$, we have*

i $\Gamma(n+1) = n!$.

ii $\Gamma(z+1) = z\Gamma(z) = z!$.

iii $\Gamma(z)\Gamma(z-1) = \dfrac{\pi}{\sin(\pi z)}$.

iv $\Gamma\left(n + \frac{1}{2}\right) = \frac{(2n)!}{2^{2n} n!}\sqrt{\pi}$.

1.4.2 Classical Mittag-Leffler functions

In partial differential equations of fractional order, the Mittag-Leffler function plays a role similar to that of the exponential function in the case of integer differential equations. It is also widely used in the search for solutions of fractional order differential equations, this function was introduced by G.M. Mittag-Leffler [57, 47]. In this part, we present the two

most important derivatives of fractional calculus: in the Riemann-Liouville sense and in the Caputo sense, including some of their properties and the relationship between these two derivatives.

Definition 1.4. *For $z \in \mathbb{C}$ and α is a strictly positive real number, we define the classical Mittag-Leffler function $E_\alpha(z)$ by the following series development:*

$$E_\alpha(z) = \sum_{k=0}^{\infty} \frac{z^k}{\Gamma(\alpha k + 1)}.$$

For all $\alpha > 0$ and $\beta > 0$, the generalized Mittag-Leffler function $E_{\alpha,\beta}(z)$ can be defined by two parameters as:

$$E_{\alpha,\beta}(z) = \sum_{k=0}^{\infty} \frac{z^k}{\Gamma(\alpha k + \beta)}.$$

Taking $\alpha = 1$ the exponential function is recovered, $E_\alpha(z) = e^z$ and for $\beta = 1$ the classical Mittag-Leffler function is recovered, i.e. $E_{\alpha,1}(z) = E_\alpha(z)$.

By virtue of Definition 1.4, the following relations are obtained in the following form [30].

Theorem 1.4. *For all $\alpha > 0$ and $\beta > 0$, we have*

$$E_{\alpha,\beta}(z) = \frac{1}{\Gamma(\beta)} + zE_{\alpha,\alpha+\beta}(z)$$

and

$$E_{\alpha,\beta}(z) = \beta E_{\alpha,\beta+1}(z) + \alpha z \frac{d}{dz} E_{\alpha,\beta+1}(z).$$

Lemma 1.3. *Let $\alpha \in (0,2), \beta \in \mathbb{R}$, and $\mu \in (\alpha\pi/2, \min(\pi, \alpha\pi))$, and $N \in \mathbb{N}$. Then for $|\arg(z)| \leq \mu$ with $|z| \to \infty$, we obtain*

$$E_{\alpha,\beta}(z) = \frac{1}{\alpha} z^{\frac{1-\beta}{\alpha}} e^{z^{\frac{1}{\alpha}}} - \sum_{k=1}^{N} \frac{1}{\Gamma(\beta - \alpha k)} \frac{1}{z^k} + O\left(\frac{1}{z^{N+1}}\right),$$

and for $\mu \leq |\arg(z)| \leq \pi$ with $|z| \to \infty$

$$E_{\alpha,\beta}(z) = -\sum_{k=1}^{N} \frac{1}{\Gamma(\beta - \alpha k)} \frac{1}{z^k} + O\left(\frac{1}{z^{N+1}}\right).$$

In particular, for x real and positive, $E_{\alpha,\beta}(-x)$ decays to zero linearly as $x \to \infty$.

The usable estimate that follows is a straightforward corollary of the Lemma 1.3.

Corollary 1.1. *Let $0 < \alpha < 2, \beta \in \mathbb{R}$, and $\pi\alpha/2 < \mu < \min(\pi, \pi\alpha)$. Then the following estimates hold*

$$|E_{\alpha,\beta}(z)| \leq c_1(1+|z|)^{\frac{1-\beta}{\alpha}} e^{\Re\left(z^{\frac{1}{\alpha}}\right)} + \frac{c_2}{1+|z|} \quad |\arg(z)| \leq \mu,$$

and

$$|E_{\alpha,\beta}(z)| \leq \frac{c}{1+|z|} \quad \mu \leq |\arg(z)| \leq \pi.$$

We will also need the standard formulas for recursion and derivative of the Mittag-Laffler function and in particular the following lemma.

Lemma 1.4. *([60]). For $\lambda > 0, \alpha > 0$ and $n \in \mathbb{N}^*$, we have*

$$\frac{d^n}{dt^n} E_{\alpha,1}\left(-\lambda t^\alpha\right) = -\lambda t^{\alpha-n} E_{\alpha,\alpha-n+1}\left(-\lambda t^\alpha\right), t > 0,$$

In particular, if we set $n = 1$, then there holds

$$\frac{d}{dt} E_{\alpha,1}\left(-\lambda t^\alpha\right) = -\lambda t^{\alpha-1} E_{\alpha,\alpha}\left(-\lambda t^\alpha\right), t > 0.$$

1.4.3 Fractional derivatives

Fractional calculus is a branch of analysis whose study relates to integration and derivation operators of non-integer order. There are several definitions of fractional derivatives; unfortunately, not all of them are equivalent. In this part, we present the definitions of Riemann-Liouville and Caputo, which are the most used, and give some of their associated properties([1, 10, 56]).

Definition 1.5. *The Riemann-Liouville left-sided and right-sided fractional derivatives ${}^R D_t^\alpha u$ and ${}^R D_T^\alpha u$ of order $0 < \alpha < 1$, are defined by*

$${}^R D_t^\alpha u(t) = \frac{1}{\Gamma(1-\alpha)} \frac{d}{dt} \int_0^t \frac{u(s)}{(t-s)^\alpha} ds, \quad \forall t \in [0, T]$$

and

$${}^R D_T^\alpha u(t) = -\frac{1}{\Gamma(1-\alpha)} \frac{d}{dt} \int_t^T \frac{u(s)}{(s-t)^\alpha} ds, \quad \forall t \in [0, T].$$

Definition 1.6. *The Caputo left-sided and right-sided fractional derivatives $D_t^\alpha u$ and*

$D_T^\alpha u$ of order $0 < \alpha < 1$, are defined by

$$D_t^\alpha u(t) = \frac{1}{\Gamma(1-\alpha)} \int_0^t \frac{u'(s)}{(t-s)^\alpha} ds, \quad \forall t \in [0, T]$$

and

$$D_T^\alpha u(t) = -\frac{1}{\Gamma(1-\alpha)} \int_t^T \frac{u'(s)}{(s-t)^\alpha} ds, \quad \forall t \in [0, T].$$

One of the remarkable properties of the Riemann-Liouville fractional derivative is given in the following lemma.

Lemma 1.5. *For all positive real α, if $u \in H^\alpha(I), \quad v \in C^\infty(I)$, then*

$$\left({}^R D_t^\alpha u(t), v(t)\right)_I = \left(u(t), {}^R D_T^\alpha v(t)\right)_I.$$

Since $C_0^\infty(I)$ is dense in $H_0^{\frac{\alpha}{2}}(I)$, the above inequality remains true for all $v \in H_0^{\frac{\alpha}{2}}(I)$.

Lemma 1.6. *If $0 < \alpha < 2, \alpha \neq 1, u, v \in H_0^{\frac{\alpha}{2}}(I)$, then*

$$\left({}^R D_t^\alpha u(t), v(t)\right)_I = \left({}^R D_t^{\frac{\alpha}{2}} u(t), {}^R D_T^{\frac{\alpha}{2}} v(t)\right)_I,$$
$$\left({}^R D_T^\alpha u(t), v(t)\right)_I = \left({}^R D_T^{\frac{\alpha}{2}} u(t), {}^R D_t^{\frac{\alpha}{2}} v(t)\right)_I.$$

The following theorem establishes the link between the fractional derivative in the Caputo sense and the fractional derivative in the Riemann-Liouville sense. (For more details, see [66],[71]).

Theorem 1.5. *For $0 < \alpha < 1, u \in H^1(I)$, we have*

$$D_T^\alpha u(t) = {}^R D_T^{\frac{\alpha}{2}} u(t) - \frac{u(T)}{\Gamma(1-\alpha)(T-t)^\alpha}$$

and

$$D_t^\alpha u(t) = {}^R D_t^\alpha u(t) - \frac{u(0)t^{-\alpha}}{\Gamma(1-\alpha)}.$$

Lemma 1.7. . *For $0 < \alpha < 1$, $u \in H^1(I)$ and $v \in H_0^{\frac{\alpha}{2}}(I)$, then*

$$(D_T^\alpha u(t), v(t))_I = \left({}^R D_T^{\frac{\alpha}{2}} u(t), {}^R D_t^{\frac{\alpha}{2}} v(t)\right)_I - \left(\frac{u(T)}{\Gamma(1-\alpha)(T-t)^\alpha}, v(t)\right)_I,$$

and

$$(D_t^\alpha u(t), v(t))_I = \left({}^R D_t^{\frac{\alpha}{2}} u(t), {}^R D_T^{\frac{\alpha}{2}} v(t)\right)_I - \left(\frac{u(0)t^{-\alpha}}{\Gamma(1-\alpha)}, v(t)\right)_I.$$

Remark 1.2.

- *If* $u(T) = 0$ *one will have* $D_T^\alpha u(t) = {}^R D_T^\alpha u(t)$.
- *If* $u(0) = 0$ *one will have* $D_t^\alpha u(t) = {}^R D_t^\alpha u(t)$.
- *The main advantage of the Caputo approach is that the initial conditions of fractional differential equations in Caputo derivatives sense accept the same form as for integer differential equations.*

The following lemma about the composition between D_t^α and the fractional integral I_t^α is presented in [61].

Lemma 1.8. *Define the Riemann-Liouville* α*-th order integral* I_t^α *as*

$$I_t^\alpha u = \frac{1}{\Gamma(\alpha)} \int_0^t (t-\tau)^{\alpha-1} u(\tau) \mathrm{d}\tau.$$

For $0 < \alpha < 1, u(t), D_t^\alpha u \in C[0,T]$*, we have*

$$\left(D_t^\alpha \circ I_t^\alpha u\right)(t) = u(t), \quad \left(I_t^\alpha \circ D_t^\alpha u\right)(t) = u(t) - u(0), t \in [0,T].$$

Chapter 2

Study of the direct problem

Recently, mathematicians have shown a special interest in fractional diffusion equations, resulting in plenty of research works that have been carried out. This has made a valuable contribution ranging from the qualitative theory of the solutions of fractional diffusion equations, such as existence, uniqueness, stability and controllability to the numerical analysis. Speaking in this context, the well-posedness of the direct problem is important in many applications such as optimization and numerical analysis. In this chapter, we present a theoretical framework for the direct problem, which consists of showing the existence, uniqueness and regularity of the solution, as well as a numerical framework dedicated to the numerical approximation of the fractional diffusion equation.

2.1 Well-posedness of the direct problem

In this section, we study the existence, uniqueness and regularity results of the solution for the direct problem (1), and we establish some a prior estimates on the solution. Which will be needed later in the identification methods. To overcome the difficulties posed by the nonlinearity of the model, we derive a solution representation to nonlinear fractional diffusion problem using the two-parameter Mittag-Leffler functions $E_{\alpha,\beta}(z)$, as well as the eigenvalues and eigenfunctions of the elliptic operator $\mathbb{A} = -\Delta$. The asymptotic behavior of these functions plays a very important role in the interpretation of the solution of various problems of physics connected with fractional reaction, fractional relaxation, fractional diffusion, fractional reaction-diffusion and so forth in complex systems.

In this study, we consider the parameter p is bounded such that

$$0 < c_1 \leq p \leq c_2 \text{ .a.e in } \Omega,$$

where c_1 and c_2 are a positive constants.

2.1.1 The solution representation and regularity results for direct problem

In this part, we provide the regularity results for the direct problem (1), to show this result, we need some background results, such as those concerning the Mittag-Leffler functions that were established in the preliminary chapter 1, some estimates of the solution operators, and the version of the Gronwall inequality for the fractional operator case showed in [30, 26].

Firstly, we reformulate the direct problem (1) into the following operator form.

$$\begin{cases} D_t^{\alpha} u(t) + \mathbb{A}u(t) &= -pf(u(t)), \quad t \in I, \\ u(0) &= \phi(x). \end{cases} \tag{2.1}$$

Specifically, we define a symmetric uniformly elliptic operator $\mathbb{A} = -\Delta$ in $L^2(\Omega)$ by

$$\begin{cases} \mathbb{A}u(x) &= -\Delta u(x), \qquad in \quad \Omega, \\ \mathcal{D}(\mathbb{A}) &= H_0^1(\Omega) \cap H^2(\Omega). \end{cases} \tag{2.2}$$

Since $\mathbb{A}$ a symmetric uniformly elliptic operator, the spectrum of $\mathbb{A}$ is entirely composed of eigenvalues and counting according to the multiplicities. Then $\mathbb{A}$ possesses eigenvalues $0 < \lambda_1 < \lambda_2 < .... < \lambda_j < ...$ Let φ_j be the $L^2(\Omega)$ orthonormal eigenfunctions corresponding to λ_j, satisfies $\mathbb{A}\varphi_j = \lambda_j \varphi_j$ for all $\varphi_j \in \mathcal{D}(\mathbb{A})$. The notation $\mathcal{D}(\mathbb{A}^{\frac{s}{2}})$ denotes the domain of the fractional power $\mathbb{A}^s = (-\Delta)^s$, with the graph norm $\|.\|_{\mathcal{D}(\mathbb{A}^s)}$, given by

$$\|v\|^2_{\mathcal{D}(\mathbb{A}^s)} = \sum_{j=0}^{\infty} \lambda_j^s |(v, \varphi_j)|^2 = \|\mathbb{A}^s v\|^2_{L^2(\Omega)}.$$

Next we introduce a space $\bar{H}^s(\Omega)$ by

$$\bar{H}^s(\Omega) = \{v \in L^2(\Omega); \sum_{j=0}^{\infty} \lambda_j^s |(v, \varphi_j)|^2 < \infty\}$$

which is a Hilbert space [30], with the inner product

$$\langle u, v \rangle_{\bar{H}^s(\Omega)} = (\mathbb{A}^{\frac{s}{2}} u, \mathbb{A}^{\frac{s}{2}} v),$$

where $\mathbb{A}^{\frac{s}{2}}v = \sum_{j=0}^{\infty} \lambda_j (\mathbb{A}^{\frac{s}{2}}v, \varphi_j)\varphi_j$. Then, we can define the norm of $\bar{H}^s(\Omega)$ as follows

$$\|v\|^2_{\bar{H}^s(\Omega)} = \sum_{j=0}^{\infty} \lambda_j^s |(v, \varphi_j)|^2 = \sum_{j=0}^{\infty} |(v, \lambda_j^{\frac{s}{2}}\varphi_j)|^2 = \sum_{j=0}^{\infty} (v, \mathbb{A}^{\frac{s}{2}}\varphi_j)^2 = \sum_{j=0}^{\infty} (\mathbb{A}^{\frac{s}{2}}v, \varphi_j)^2 = \|\mathbb{A}^{\frac{s}{2}}v\|^2_{L^2(\Omega)}. \tag{2.3}$$

We have $\bar{H}^s(\Omega) \subseteq H^s(\Omega)$ for $s \in [0, 2]$, since $\bar{H}^s(\Omega) \subseteq L^2(\Omega)$, identifying the dual $(L^2(\Omega))'$ with itself, we have $\bar{H}^s(\Omega) \subseteq L^2(\Omega) \subseteq \bar{H}^{-s}(\Omega)$ where $\bar{H}^{-s}(\Omega)$ is the dual space of $\bar{H}^s(\Omega)$.

We prove that for suitably smooth ϕ and any $p \in L^2(\Omega)$, problem (2.1) has a unique classical solution $u = u(p) \in C(\bar{I}; \bar{H}^2(\Omega)) \cap C^1(]0, T]; L^2(\Omega))$. By viewing $pf(u(t))$ as the non-linear inhomogeneous term and applying Duhamel's principle, we deduce that then solution representation $u(t)$ to the nonlinear problem (1) satisfies (see, [28]).

$$\begin{aligned} u(t) &= \sum_{j=1}^{\infty} E_\alpha(-\lambda_j t^\alpha)(\phi, \varphi_j)\varphi_j + \sum_{j=1}^{\infty} \int_0^t (t-s)^{\alpha-1} E_{\alpha,\alpha}(-\lambda_j (t-s)^\alpha)(-pf(u(s)), \varphi_j)\varphi_j ds \\ &= \mathbb{E}(t)\phi + \int_0^t \bar{\mathbb{E}}(t-s)(-pf(u(s))ds, \qquad 0 < t < T, \end{aligned} \tag{2.4}$$

where the solution operators $\mathbb{E}$ and $\bar{\mathbb{E}}$ are defined by, respectively

$$\mathbb{E}(t)v = \sum_{j=1}^{\infty} E_\alpha(-\lambda_j t^\alpha)(v, \varphi_j)\varphi_j \tag{2.5}$$

and

$$\bar{\mathbb{E}}(t)v = \sum_{j=1}^{\infty} t^{\alpha-1} E_{\alpha,\alpha}(-\lambda_j t^\alpha)(v, \varphi_j)\varphi_j \tag{2.6}$$

Remark 2.1. *Many of the results for parabolic equations carry over to fractional order operators although there are significant exceptions. In the case of the non-homogeneous parabolic equation $u_t + \mathbb{A}u = F$ defined on $\Omega \times [0, \infty)$ with $u(x, 0) = \phi$, the evolution operator $\mathbb{E}(t)$ is the sequel to the semigroup operator $e^{t\mathbb{A}}\phi$.*

The following useful identities is a direct corollary of Lemma 1.4.

Corollary 2.1. *The following identities hold*

$$\frac{d}{dt}\mathbb{E}(t)v = \bar{\mathbb{E}}(t)(-\mathbb{A}v). \tag{2.7}$$

Proof. We can be seen from the definitions of the solution operators (2.5) and (2.6)

$$\frac{d}{dt}\mathbb{E}(t)v = \sum_{j=1}^{\infty} \frac{d}{dt} E_\alpha(-\lambda_j t^\alpha)(v, \varphi_j)\varphi_j. \tag{2.8}$$

From the Lemma 1.4, we have

$$\frac{d}{dt}E_\alpha(-\lambda_j t^\alpha) = -\lambda_j t^{\alpha-1}E_{\alpha,\alpha}(-\lambda_j t^\alpha). \tag{2.9}$$

Combining (2.8) and (2.8), we obtain

$$\begin{aligned}\frac{d}{dt}\mathbb{E}(t)v &= \sum_{j=1}^{\infty} -\lambda_j t^{\alpha-1}E_{\alpha,\alpha}(-\lambda_j t^\alpha)(v,\varphi_j)\varphi_j \\ &= \sum_{j=1}^{\infty} t^{\alpha-1}E_{\alpha,\alpha}(-\lambda_j t^\alpha)(-\lambda_j v,\varphi_j)\varphi_j \\ &= \bar{\mathbb{E}}(t)(-\mathbb{A}v).\end{aligned}$$

This shows also the identity. □

The following lemmas given the stability estimates on the solution operators $\mathbb{E}(t)$ and $\bar{\mathbb{E}}(t)$ defined in (2.5) and (2.6) respectively. (For more detail, see [24]. Chapter 6.)

Lemma 2.1. *For the solution opretors $\mathbb{E}(t)$ and $\bar{\mathbb{E}}(t)$ defined in (2.5) and (2.6) respectively, for any $t > 0$, we have*

$$\|\mathbb{E}(t)v\|_{\bar{H}^m(\Omega)} \leq ct^{\left(\frac{n-m}{2}\right)\alpha}\|v\|_{\bar{H}^n(\Omega)}, \tag{2.10}$$

$$\|\bar{\mathbb{E}}(t)v\|_{\bar{H}^m(\Omega)} \leq ct^{-1+\left(1+\frac{n-m}{2}\right)\alpha}\|v\|_{\bar{H}^n(\Omega)}, \tag{2.11}$$

where $0 \leq n \leq 2$ and $n \leq m \leq n+2$ and the constant c depends only on α and $m-n$.

Gronwall's inequalities are a basic tool for differential equations, the following lemma presents the convolution version of Gronwall's inequality in the fractional operator case. (See [24]. Chapter 6.)

Lemma 2.2. *Suppose $b \geq 0$, $\beta > 0$ and $a(t)$ is a nonnegative function locally integrable on $0 \leq t < T$ and suppose $u(t)$ is a nonnegative function locally integrable on $0 \leq t < T$ with*

$$u(t) \leq a(t) + \frac{b}{\Gamma(\beta)}\int_0^t s^{\beta-1}u(t-s)ds, \qquad 0 \leq t < T.$$

Then

$$u(t) \leq a(t) + b\int_0^t (t-s)^{\beta-1} E_{\alpha,\alpha}(b(t-s)^{\beta}) a(s) ds, \qquad 0 \leq t < T.$$

Now we discuss the existence, uniqueness, and regularity for the solutions to direct problem (1). These results are needed in the theoretical and numerical analysis . The main result of this chapter is the following theorem

Theorem 2.1. *Let $\phi \in \bar{H}^2(\Omega)$ and $f(\phi) \in L^2(\Omega)$ and under assumption $(\boldsymbol{H}_1) - (\boldsymbol{H}_2)$. Then problem (1) has a unique solution u such that*

$$u \in C(\bar{I}; \bar{H}^2(\Omega)) \cap C^1(]0,T]; L^2(\Omega)), \;\; D_t^{\alpha} u \in C(\bar{I}; L^2(\Omega)), \tag{2.12}$$

$$\frac{du}{dt}(t) \in L^2(\Omega), \text{ and } \|\frac{du}{dt}(t)\|_{L^2(\Omega)} \leq ct^{\alpha-1} \text{ for } t \in (0,T], \tag{2.13}$$

$$\text{There exists a constant } C \text{ such that } \|u\|_{C(\bar{I};\bar{H}^2(\Omega))} \leq C. \tag{2.14}$$

Proof. For the unique existence of u, it suffices to discuss the integral equation

$$u(t) = \mathbb{E}(t)\phi + \int_0^t \bar{\mathbb{E}}(t-s)(-pf(u(s)))ds, \qquad 0 < t < T. \tag{2.15}$$

First we estimate $\dfrac{du}{dt}(t)$, which is given by

$$\frac{du}{dt}(t) = \frac{d}{dt}\left(\mathbb{E}(t)\phi\right) + \frac{d}{dt}\left(\int_0^t \bar{\mathbb{E}}(s)(-pf(u(t-s)))ds\right),$$

using the result (2.7) in the first term, and the definition of the standard derivative $\frac{d}{dt}$ in the second, we obtain

$$\frac{du}{dt}(t) = \bar{\mathbb{E}}(t)(-\mathbb{A}\phi - pf(\phi)) + \int_0^t \bar{\mathbb{E}}(s)(-pf'(u(t-s))\frac{du}{dt}(t-s))ds.$$

Now, we will show the estimation of $\dfrac{du}{dt}(t)$ in $L^2(\Omega)$, actually

$$\|\frac{du}{dt}(t)\|_{L^2(\Omega)} \leq \|\bar{\mathbb{E}}(t)(-\mathbb{A}\phi - pf(\phi))\|_{L^2(\Omega)} + \int_0^t \|\bar{\mathbb{E}}(s)(-pf'(u(t-s))\frac{du}{dt}(t-s))\|_{L^2(\Omega)} ds.$$

For any $0 < t \leq T$, from the Lemma 2.1, using the inequality (2.11) with $m = n = 0$, we get

$$\|\bar{\mathbb{E}}(t)v\|_{L^2(\Omega)} \leq ct^{\alpha-1}\|v\|_{L^2(\Omega)}.$$

Based on this result, we obtain

$$\|\frac{du}{dt}(t)\|_{L^2(\Omega)} \leq ct^{\alpha-1}\|\mathbb{A}\phi + pf(\phi)\|_{L^2(\Omega)} + c\int_0^t s^{\alpha-1}\|pf'(u(t-s))\frac{du}{dt}(t-s)\|_{L^2(\Omega)}ds.$$

The hypothesis $\mathbf{H}_2$ gives that f is bounded, the parameter $p \in \mathcal{U}_{ad}$ also bounded , then we have

$$\begin{aligned}\|\frac{du}{dt}(t)\|_{L^2(\Omega)} &\leq ct^{\alpha-1}\|\mathbb{A}\phi + pf(\phi)\|_{L^2(\Omega)} + c\int_0^t s^{\alpha-1}\|pf'(u(t-s))\|_{L^\infty(\Omega)}\|\frac{du}{dt}(t-s)\|_{L^2(\Omega)}ds.\\ &\leq ct^{\alpha-1}\|\mathbb{A}\phi + pf(\phi)\|_{L^2(\Omega)} + c\int_0^t s^{\alpha-1}\|\frac{du}{dt}(t-s)\|_{L^2(\Omega)}ds.\end{aligned}$$

We apply Gronwell's inequality (lemma 2.2) on the last inequality, we obtain

$$\begin{aligned}\|\frac{du}{dt}(t)\|_{L^2(\Omega)} &\leq ct^{\alpha-1}\|\mathbb{A}\phi + pf(\phi)\|_{L^2(\Omega)} + b\int_0^t (t-s)^{\alpha-1}E_{\alpha,\alpha}(b(t-s)^\alpha)cs^{\alpha-1}\|\mathbb{A}\phi + pf(\phi)\|_{L^2(\Omega)}ds\\ &\leq ct^{\alpha-1}\|\mathbb{A}\phi + pf(\phi)\|_{L^2(\Omega)} + bc\|\mathbb{A}\phi + pf(\phi)\|_{L^2(\Omega)}\int_0^t (t-s)^{\alpha-1}s^{\alpha-1}E_{\alpha,\alpha}(b(t-s)^\alpha)ds,\end{aligned}$$

where $b = c\Gamma(\alpha)$. The estimation of the Mittag-Leffler function in the Corollary 1.1 gives

$$\begin{aligned}\|\frac{du}{dt}(t)\|_{L^2(\Omega)} &\leq ct^{\alpha-1}\left(\|\mathbb{A}\phi\|_{L^2(\Omega)} + \|pf(\phi)\|_{L^2(\Omega)}\right)\\ &\leq ct^{\alpha-1}\left(\|\phi\|_{\bar{H}^2(\Omega)} + \|f(\phi)\|_{L^2(\Omega)}\right),\end{aligned}$$

which implies that $u \in C^1(]0,T];L^2(\Omega))$ and thus

$$\|D_t^\alpha u\|_{L^2(\Omega)} \leq \frac{1}{\Gamma(\alpha+1)}\int_0^t (t-s)^{-\alpha}\|\frac{du}{dt}(s)\|_{L^2(\Omega)}ds.$$

Using the estimate of $\dfrac{du}{dt}$ above, we obtain

$$\|D_t^\alpha u\|_{L^2(\Omega)} \leq \frac{c}{\Gamma(\alpha+1)}\int_0^t (t-s)^{-\alpha}s^{\alpha-1}\left(\|\phi\|_{\bar{H}^2(\Omega)} + \|f(\phi)\|_{L^2(\Omega)}\right)ds.$$

From the definition of Caputo's fractional derivatives above, we have

$$\|D_t^\alpha u\|_{L^2(\Omega)} \leq cD_t^\alpha(t^\alpha)\left(\|\phi\|_{\bar{H}^2(\Omega)} + \|f(\phi)\|_{L^2(\Omega)}\right),$$

the fact that $D_t^\alpha(t^\alpha) = \Gamma(\alpha+1)$, we get

$$\|D_t^\alpha u\|_{L^2(\Omega)} \leq c\Gamma(\alpha+1)\left(\|\phi\|_{\bar{H}^2(\Omega)} + \|f(\phi)\|_{L^2(\Omega)}\right).$$

i.e $D_t^\alpha u \in C(\bar{I}; L^2(\Omega))$.
It remains to show that $u \in C(\bar{I}; \bar{H}^\sigma(\Omega))$ with $\sigma \in [0,2]$. We will estimate $\mathbb{A}^{\frac{\sigma}{2}}u(t)$, which is given by

$$\begin{aligned}\mathbb{A}^{\frac{\sigma}{2}}u(t) &= \mathbb{A}^{\frac{\sigma}{2}}\mathbb{E}(t)\phi + \int_0^t \mathbb{A}^{\frac{\sigma}{2}}\bar{\mathbb{E}}(t-s)(-pf(u(s)))ds \\ &= \mathbb{A}^{\frac{\sigma}{2}}\mathbb{E}(t)\phi + \int_0^t \mathbb{A}^{\frac{\sigma}{2}}\bar{\mathbb{E}}(t-s)(-pf(u(t)))ds + \int_0^t \mathbb{A}^{\frac{\sigma}{2}}\bar{\mathbb{E}}(t-s)(-p\left[f(u(s)) - f(u(t))\right])ds,\end{aligned}$$

in view of the identity (2.7), we get

$$\mathbb{A}^{\frac{\sigma}{2}}u(t) = \underbrace{\mathbb{A}^{\frac{\sigma}{2}}\mathbb{E}(t)\phi}_{I_1} + \underbrace{\mathbb{A}^{\frac{\sigma}{2}-1}(\mathbb{E}(t)-I)(-pf(u(t)))}_{I_2} + \underbrace{\int_0^t \mathbb{A}^{\frac{\sigma}{2}}\bar{\mathbb{E}}(t-s)(-p\left[f(u(s)) - f(u(t))\right])ds}_{I_3}.$$

Clearly, by using the identity (2.10) of the Lemma 2.1 under parameters $m = \sigma$ and $n = 0$, we obtain

$$\|I_1\|_{L^2(\Omega)} = \|\mathbb{A}^{\frac{\sigma}{2}}\mathbb{E}(t)\phi\|_{L^2(\Omega)} = \|\mathbb{E}(t)\phi\|_{\bar{H}^\sigma(\Omega)} \leq cT^{\frac{\sigma}{2}}\|\phi\|_{L^2(\Omega)}.$$

By the direct application of this lemma 2.1 for I_2, we have

$$\|I_2\|_{L^2(\Omega)} \leq c\|pf(u)\|_{L^2(\Omega)} \leq c\|u\|_{L^2(\Omega)}.$$

The fact that f is a Lipschitz function (the assumption $(\mathbf{H}_2)$), we have

$$\begin{aligned}\|pf(u(s)) - pf(u(t))\|_{L^2(\Omega)} \leq c\|u(s)-u(t)\|_{L^2(\Omega)} &= c\|\int_t^s u_\tau(\tau)d\tau\|_{L^2(\Omega)} \\ &\leq c\int_s^t \|u_\tau(\tau)\|_{L^2(\Omega)}d\tau.\end{aligned}$$

Using the estimate of u_τ above, we obtain

$$\begin{aligned}\|pf(u(s)) - pf(u(t))\|_{L^2(\Omega)} &\leq c\int_s^t \tau^{\alpha-1}d\tau(\|f(\phi)\|_{L^2(\Omega)} + \|\phi\|_{\bar{H}^2(\Omega)}) \\ &\leq c(t^\alpha - s^\alpha)(\|f(\phi)\|_{L^2(\Omega)} + \|\phi\|_{\bar{H}^2(\Omega)}),\end{aligned}$$

based on this calculation, we have

$$\begin{aligned}\|I_3\|_{L^2(\Omega)} &\leq c(\|f(\phi)\|_{L^2(\Omega)} + \|\phi\|_{\bar{H}^2(\Omega)}).\int_0^t \|\mathbb{A}^{\frac{\sigma}{2}}\bar{\mathbb{E}}(t-s)\|_{L^2(\Omega)}(t^\alpha - s^\alpha)ds \\ &\leq c(\|f(\phi)\|_{L^2(\Omega)} + \|\phi\|_{\bar{H}^2(\Omega)}).\int_0^t \|\bar{\mathbb{E}}(t-s)\|_{\bar{H}^\sigma(\Omega)}(t^\alpha - s^\alpha)ds.\end{aligned}$$

Hence, by Lemma 2.1 with $m = \sigma$ and $n = 0$, we obtain

$$\|I_3\|_{L^2(\Omega)} \leq c(\|f(\phi)\|_{L^2(\Omega)} + \|\phi\|_{\bar{H}^2(\Omega)}).$$

Therefore, we have $\mathbb{A}^{\frac{\sigma}{2}}u(t) \in C(\bar{I}; L^2(\Omega))$ then $u \in C(\bar{I}; \bar{H}^\sigma(\Omega))$.
This completes the proof of the theorem. □

2.1.2 Uniqueness of solution for the direct problem

This part will be devoted to the study of the uniqueness of the solution of the time-fractional diffusion problem (1), on the basis of this study, we are thus led to we use $H^s(\Omega)$ to denote the usual Sobolev spaces, $\forall s > 0$, whose norm are denoted by $\|.\|_{s,\Omega}$ (see [3]).
We denote by the space $B^s(Q) = H^s(I; L^2(\Omega)) \cap L^2(I; H_0^1(\Omega))$, $\forall s > 0$, equipped by the norm :

$$\|v\|_{B^s(Q)} = \left(\|v\|^2_{H^s(I;L^2(\Omega))} + \|v\|^2_{L^2(I;H_0^1(\Omega))}\right)^{\frac{1}{2}}.$$

We start by giving the following Riemann Liouville weak formulation of the problem (1)

$$\begin{cases} \text{Find} \quad u \in B^{\frac{\alpha}{2}}(Q) \\ \left({}^R D_t^{\frac{\alpha}{2}} u, {}^R D_T^{\frac{\alpha}{2}} v\right)_Q + (\nabla u, \nabla v)_Q + (p(x)f(u), v)_Q = \left(\dfrac{\phi(x)t^{-\alpha}}{\Gamma(1-\alpha)}, v\right)_Q, \ \forall v \in B^{\frac{\alpha}{2}}(Q). \end{cases} \tag{2.16}$$

From the weak variational (2.16), we can then conclude the following estimate

Lemma 2.3. *If u is a solution of (2.16) and under assumptions $(\boldsymbol{H}_1)-(\boldsymbol{H}_2)$, we have the following estimation*

$$\|u\|_{B^{\frac{\alpha}{2}}(Q)} \leq \frac{c\gamma_t}{\cos(\frac{\alpha\pi}{2})}, \tag{2.17}$$

where $\gamma_t = \|\phi(x)\|_\Omega \|t^{-\alpha}\|_{L^q(I)}$ with $q = \frac{2}{1+\alpha}$.

Proof. Taking $v = u$ in the weak formulation (2.16), we obtain

$$({}^R D_t^{\frac{\alpha}{2}} u, {}^R D_T^{\frac{\alpha}{2}} u)_Q + (\nabla u, \nabla u)_Q + (pf(u), u)_Q = (\frac{\phi(x)t^{-\alpha}}{\Gamma(1-\alpha)}, u)_Q. \tag{2.18}$$

Using the following result (see [36])

$$\begin{aligned}({}^R D_t^{\frac{\alpha}{2}} u, {}^R D_T^{\frac{\alpha}{2}} u)_Q + (\nabla u, \nabla u)_Q &\geq \cos(\frac{\alpha\pi}{2}) \left(({}^R D_t^{\frac{\alpha}{2}} u, {}^R D_t^{\frac{\alpha}{2}} u)_Q + (\nabla u, \nabla u)_Q \right) \\ &\geq \cos(\frac{\alpha\pi}{2}) \|u\|^2_{B^{\frac{\alpha}{2}}(Q)}, \ \forall u \in B^{\frac{\alpha}{2}}(Q). \end{aligned} \tag{2.19}$$

By applying mean value theorem, the equation (2.18) becomes as follow

$$\cos(\alpha\pi/2) \ \|u\|^2_{B^{\frac{\alpha}{2}}(Q)} + (pf'(u-\theta u), u^2)_Q \leq (\frac{\phi(x)t^{-\alpha}}{\Gamma(1-\alpha)}, u)_Q,$$

where $0 < \theta < 1$. Based on the assumption $(\mathbf{H}_2)$, we obtain

$$\|u\|^2_{B^{\frac{\alpha}{2}}(Q)} \leq \frac{1}{\cos(\alpha\pi/2)} \left| (\frac{\phi(x)t^{-\alpha}}{\Gamma(1-\alpha)}, u)_Q \right|. \tag{2.20}$$

By direct calculation, we get

$$\left| (\frac{\phi(x)t^{-\alpha}}{\Gamma(1-\alpha)}, u)_Q \right| \leq \frac{c}{\cos(\alpha\pi/2)} \|\phi(x)\|_\Omega \|t^{-\alpha}\|_{L^q(I)} \|u\|_{L^{q'}(I,L^2(\Omega))}. \tag{2.21}$$

where

$$q = 2/(1+\alpha), \quad q' = 2/(1-\alpha)$$

Furthermore, by the Embedding Theorem [43], we know that

$$H^{\frac{\alpha}{2}}(I) \hookrightarrow L^{q'}(I).$$

Thus

$$\|u\|_{L^{q'}(I;L^2(\Omega))} \leq \|u\|_{H^{\frac{\alpha}{2}}(I;L^2(\Omega))} \leq \|u\|_{B^{\frac{\alpha}{2}}(Q)}, \tag{2.22}$$

Combining the equations (2.20), (2.21) and (2.22), we obtain

$$\|u\|^2_{B^{\frac{\alpha}{2}}(Q)} \leq \frac{c}{\cos(\alpha\pi/2)}\|\phi(x)\|_{\Omega}\|t^{-\alpha}\|_{L^q(I)}\|u\|_{B^{\frac{\alpha}{2}}(Q)}, \tag{2.23}$$

the proof is hence completed. □

The following theorem shows the existence of unique solution.

Theorem 2.2. *Under the assumptions* (***H***)*, the problem* (2.16) *admits a unique solution in* $L^2(Q)$.

Proof. The existence is given by the above subsection by mean Mettag-Leffler functions. For the uniqueness of the solution, we assume that u_1 and u_2 are two distinct solutions of the problem (2.16). By subtracting the weak formulations associated to the solutions u_1 and u_2, we obtain

$$({}^RD_t^{\frac{\alpha}{2}}(u_1-u_2),{}^R D_T^{\frac{\alpha}{2}}v)_Q + (\nabla(u_1-u_2),\nabla v)_Q + \int_Q p(x)[f(u_1)-f(u_2)]v dxdt = 0.$$

By taking $v = u_1 - u_2$, we obtain

$$({}^RD_t^{\frac{\alpha}{2}}(u_1-u_2),{}^R D_T^{\frac{\alpha}{2}}(u_1-u_2))_Q+(\nabla(u_1-u_2),\nabla(u_1-u_2))_Q+\int_Q p(x)[f(u_1)-f(u_2)](u_1-u_2)dxdt = 0.$$

We use the assumption ($\mathbf{H}_1$) and inequality (2.17), we get

$$\|u_1-u_2\|_{B^{\frac{\alpha}{2}}(Q)} \leq 0.$$

This achieves the proof of the theorem. □

2.2 Numerical approximation

In this section, we introduce the basic ideas for the approximate solution of the time fractional diffusion equations by the finite differences method. Finding the approximate solution of the direct problem (1) requires the evaluation of Caputo's right-sided fractional derivative. In this regard, the authors in [54] derive the numerical integration technique known as the trapezoidal rule to approximate the Caputo derivative. However, their approximation formula involved a symbolic variable known as the tau parameter that

needed to be assigned appropriate values during the computation of the derivatives. Consequently, the involvement of this symbolic variable leads to the involvement of symbolic computation, which increases the time taken to evaluate the derivative. This approach is obviously inefficient in terms of time. We address this issue in this work by replacing the tau parameter with a numerical implementation developed in [48, 77, 49], which eventually frees from any symbolic computation.

We anticipate a more accurate approximation of the direct problem using the Newton-Raphson method. We first approximate a linear time-fractional equation using the approximation of the fractional derivative.

Define $t_n = n\tau, n = 0, 1, 2, \cdots, N,\ x_i = ih,\ i = 0, 1, 2, \cdots, M$, where $\tau = \frac{T}{N+1}$ and $h = \frac{L}{M+1}$ are time and space steps, respectively. We seek in each of these points an approximate value, noted u_i^n be the numerical approximation to $u(x_i, t_n)$. For approximating the second-order space derivative in equation (1) we used a symmetric second difference quotient in space at level $t = t_n$. The time fractional derivative term can be approximated by the following scheme:

$$
\begin{aligned}
D_t^\alpha u(x_i, t_n) &= \frac{1}{\Gamma(1-\alpha)} \int_0^{t_n} \frac{u'(x_i, s)}{(t_n - s)^\alpha} ds \\
&= \frac{1}{\Gamma(1-\alpha)} \sum_{j=1}^{n} \int_{(j-1)k}^{jk} \frac{u'(x_i, s)}{(t_n - s)^\alpha} ds \\
&= \frac{1}{\Gamma(1-\alpha)} \sum_{j=1}^{n} \int_{(j-1)k}^{jk} \frac{u_i^j - u_i^{j-1}}{k} (t_n - s)^{-\alpha} ds \\
&= \frac{1}{\Gamma(1-\alpha)} \sum_{j=1}^{n} \left(\frac{u_i^j - u_i^{j-1}}{k} \right) \int_{(j-1)k}^{jk} (t_n - s)^{-\alpha} ds \\
&= \frac{1}{\Gamma(1-\alpha)} \sum_{j=1}^{n} \left(\frac{u_i^j - u_i^{j-1}}{k} \right) \left[\frac{-(t_n - s)^{1-\alpha}}{1-\alpha} \right]_{(j-1)k}^{jk} ds \\
&= \frac{1}{(1-\alpha)\Gamma(1-\alpha)} \sum_{j=1}^{n} \left(\frac{u_i^j - u_i^{j-1}}{k} \right) \left((n-j+1)k)^{1-\alpha} - ((n-j)k)^{1-\alpha} \right) \\
&= \frac{\tau^{-\alpha}}{\Gamma(2-\alpha)} \sum_{j=1}^{n} \left(u_i^j - u_i^{j-1} \right) \left((n-j+1)^{1-\alpha} - (n-j)^{1-\alpha} \right).
\end{aligned}
$$

We use the change of variable $j \longleftarrow n - j + 1$ and, set the shifting indices as follows

$$
b_j = (j)^{1-\alpha} - (j-1)^{1-\alpha}, \qquad b_1 = 1.
$$

Hence, the first-order approximation method for the computation of Caputo's fractional derivative is then given by the expression

$$D_t^\alpha u(x_i,t_n) = \frac{\tau^{-\alpha}}{\Gamma(2-\alpha)} \sum_{j=1}^{n} \left(u_i^{n-j+1} - u_i^{n-j}\right) b_j. \tag{2.24}$$

Remark 2.2. *The quadrature formula (2.24) does not provide the values of the time fractional derivative at $t = 0$, which are not required by the implicit finite difference scheme that follows.*

Using (2.24), the restriction of the exact solution to the grid points centered at $(x_i,t_n) = (ih, n\tau)$, in linear equation, satisfies

$$D_t^\alpha u(x_i,t_n) = \Delta u(x_i,t_n) - p(x_i)u(x_i,t_n).$$

The initial-boundary value conditions of problem (1) are approximate as follows

$$u_i^0 = u(x_i,0) = \phi(x_i), \quad u_0^n = u(0,t_n) = 0 \text{ et } u_1^n = u(1,t_n) = 0.$$

From the results obtained above, we have

$$\begin{cases} D_t^\alpha u(x_i,t_n) \simeq \dfrac{\tau^{-\alpha}}{\Gamma(2-\alpha)} \displaystyle\sum_{j=1}^{n} \left(u_i^{n-j+1} - u_i^{n-j}\right) b_j \\ \Delta u(x_i,t_n) \simeq \dfrac{u_{i+1}^n - 2u_i^n + u_{i-1}^n}{h^2} \\ u(x_i,t_n) \simeq u_i^n, \end{cases} \tag{2.25}$$

Then, we can approximate the equation (2.2) as follows

$$\frac{k^{-\alpha}}{\Gamma(2-\alpha)} \sum_{j=1}^{n} \left(u_i^{n-j+1} - u_i^{n-j}\right) b_j = \frac{u_{i+1}^n - 2u_i^n + u_{i-1}^n}{h^2} - p_i u_i^n.$$

To facilitate expressions, we take

$$\mu = \frac{k^\alpha}{h^2}, \quad r = \mu\Gamma(2-\alpha) \text{ et } r_1 = k^\alpha\Gamma(2-\alpha).$$

Then

$$\begin{aligned}
\sum_{j=1}^{n}\left(u_i^{n-j+1}-u_i^{n-j}\right)b_j &= ru_{i+1}^n-2ru_i^n+ru_{i-1}^n-r_1p_iu_i^n\\
u_i^n-u_i^{n-1}+\sum_{j=2}^{n}\left(u_i^{n-j+1}-u_i^{n-j}\right)b_j &= ru_{i+1}^n-2ru_i^n+ru_{i-1}^n-r_1p_iu_i^n\\
-u_i^{n-1}+\sum_{j=2}^{n}\left(u_i^{n-j+1}-u_i^{n-j}\right)b_j &= ru_{i+1}^n-(1+2r+r_1p_i)u_i^n+ru_{i-1}^n\\
-u_i^{n-1}+\sum_{j=2}^{n}u_i^{n-j+1}b_j-\sum_{j=2}^{n}u_i^{n-j}b_j &= ru_{i+1}^n-(1+2r+r_1p_i)u_i^n+ru_{i-1}^n\\
-u_i^{n-1}+u_i^{n-1}b_2+\sum_{j=3}^{n}u_i^{n-j+1}b_j-\sum_{j=2}^{n}u_i^{n-j}b_j &= ru_{i+1}^n-(1+2r+r_1p_i)u_i^n+ru_{i-1}^n,
\end{aligned}$$

yield

$$\begin{aligned}
-u_i^{n-1}+u_i^{n-1}b_2+\sum_{j=2}^{n-1}u_i^{n-j}b_{j+1}-\sum_{j=2}^{n-1}u_i^{n-j}b_j-u_i^0b_n &= ru_{i+1}^n-(1+2r+r_1p_i)u_i^n+ru_{i-1}^n\\
(b_2-b_1)u_i^{n-1}+\sum_{j=2}^{n-1}u_i^{n-j}\left(b_{j+1}-b_j\right)-u_i^0b_n &= ru_{i+1}^n-(1+2r+r_1p_i)u_i^n+ru_{i-1}^n\\
(b_1-b_2)u_i^{n-1}+\sum_{j=2}^{n-1}u_i^{n-j}\left(b_j-b_{j+1}\right)+u_i^0b_n &= -ru_{i+1}^n+(1+2r+r_1p_i)u_i^n-ru_{i-1}^n.
\end{aligned}$$

Hence, for $n=1$, we obtain:

$$ru_{i+1}^1+(1+2r+r_1p_i)u_i^1-ru_{i-1}^1=b_2u_i^0. \tag{2.26}$$

For $n>1$, we obtain

$$-ru_{i+1}^n+(1+2r+r_1p_i)u_i^n-ru_{i-1}^n=(1-b_1)\,u_i^{n-1}+\sum_{j=2}^{n-1}u_i^{n-j}\left(b_j-b_{j+1}\right)+b_nu_i^0.$$

We define

$$c_j=b_j-b_{j+1}\quad\text{ with }\quad c_1=1-b_2,$$

These notations give

$$\begin{aligned}
c_1u_i^{n-1}+\sum_{j=2}^{n-1}u_i^{n-j}c_j+u_i^0b_n &= -ru_{i+1}^n+(1+2r+r_1p_i)u_i^n-ru_{i-1}^n\\
c_1u_i^{n-1}+c_2u_i^{n-2}+...+c_{n-1}u_i^1+b_nu_i^0 &= -ru_{i+1}^n+(1+2r+r_1p_i)u_i^n-ru_{i-1}^n,
\end{aligned}\tag{2.27}$$

where $i = 1, 2, \cdots, M$. The equations (2.26) and (2.27) can be written as

$$\begin{cases} \mathbf{u}^0 = \phi \\ A\mathbf{u}^1 = b_n \mathbf{u}^0, \\ A\mathbf{u}^n = c_1 \mathbf{u}^{n-1} + c_2 \mathbf{u}^{n-2} + \cdots + c_n \mathbf{u}^1 + b_n \mathbf{u}^0, n > 0 \end{cases} \tag{2.28}$$

where

$$A = \begin{bmatrix} 1+2r+r_1p_1 & -r & 0 & \cdots & 0 & 0 \\ -r & 1+2r+r_1p_2 & -r & \cdots & 0 & 0 \\ 0 & -r & 1+2r+r_1p_3 & \cdots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & 1+2r+r_1p_{M-2} & -r \\ 0 & 0 & 0 & \ldots & -r & 1+2r+r_1p_{M-1} \end{bmatrix}$$

$$\mathbf{u}^n = \begin{bmatrix} u_1^n \\ u_2^n \\ \vdots \\ u_{N-1}^n \end{bmatrix}, \phi = \begin{bmatrix} \phi(x_1) \\ \phi(x_2) \\ \vdots \\ \phi(x_{M-1}) \end{bmatrix}$$

Remark 2.3. *In equation (2.28) we see that A is strictly diagonally dominant with positive diagonal terms and non positive off-diagonal terms. Then, the equation (2.28) can be solved.*

To solve numerically the nonlinear equation (1). we first discretize it by finite differences method

$$\frac{k^{-\alpha}}{\Gamma(2-\alpha)} \sum_{j=1}^{n} \left(u_i^{n-j+1} - u_i^{n-j} \right) b_j - \frac{u_{i+1}^n - 2u_i^n + u_{i-1}^n}{h^2} + p_i f_i^n(u_i^n) = 0, \quad \forall i = 0, ...M. \tag{2.29}$$

Starting from this system (2.29), we construct a sequence of values $\{u_i^n\}_{i=0,M}$ such that:

$$\mathcal{K}_i \left(u_0^n, u_1^n, ..., u_M^n \right) = 0 \qquad \forall i = 0, ..., M. \text{ and } n = 0, ..., N. \tag{2.30}$$

Note that the Newton-Raphson method [5] is one of the standard methods used to solve non-linear equations. If the sequence $\{u_i^n\}_{i=0,M}$ converges, it converges to a solution of (2.30) and thus to the solution of the nonlinear equation (1).

To find the roots of the nonlinear equations (2.30), we can use the Newton-Raphson method, which consists in constructing the following iterative fixed point sequence:

$$\left[u_i^{n+1}\right] = [u_i^n] - [\mathbf{J}^n]_{i,j}^{-1} \left[\mathcal{K}_j(u_0^n, ..., u_N^n)\right], \tag{2.31}$$

where $\mathbf{J}^n$ is the Jacobian matrix of the functions $\mathcal{K}_i$: $\mathbf{J}_{i,j}^n = \frac{\partial \mathcal{K}_i}{\partial u_j}(u_0^n..u_N^n)$ computed at the iteration n.

2.3 Numerical experiments

The numerical simulations presented are in one-dimension case, and we set $T = 1$, $\Omega = (0, 1)$, consider the number of discretization points in space $M = 81$ and time $N = 91$ and select the nonlinear term f that satisfies the assumptions $(\mathbf{H}_1)$ and $(\mathbf{H}_2)$ as defined below:

$$f(x) = e^x - 1.$$

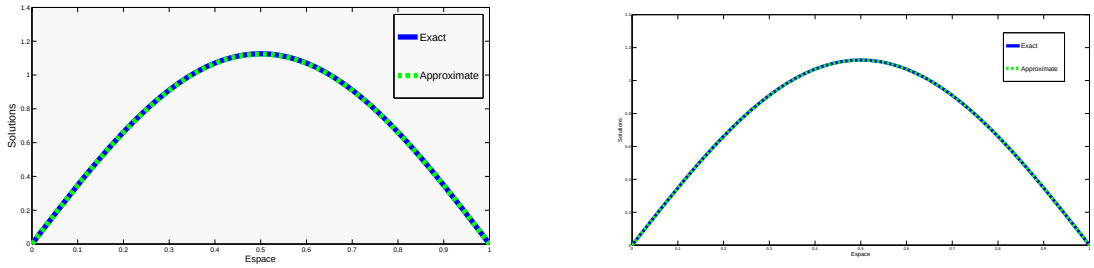
In this numerical experiment, we give the solution to the direct problem by $u(x, t) = t^2 \sin(\pi x) + \sin(\pi x)$, the order of the fractional derivative to be $\alpha = 0.5$ and the initial data $u_0 = \sin(\pi x)$. Using the finite difference method developed above, the direct problem (1) is solved. The numerical solution is compared with the exact solution in Figure 2.1 for tow following Examples

Exemple 2.1. *The first example is regular supposed as follows* $p(x) = \pi^2 \sin(\pi x)$.

Exemple 2.2. *In the second example, we test a non-smooth one with a peak. Let the parameter given by the following expression*

$$p(x) = \begin{cases} x & 0 \leq x \leq \frac{1}{2} \\ 1 - x & \frac{1}{2} \leq x \leq 1 \end{cases} \tag{2.32}$$

Figure 2.1 shows the exact and approximate solution of the direct problem with $\alpha = 0.5$, and prove that the impact made by the order of the fractional derivative.



(a) Example 2.1. (b) Example 2.2.

Figure 2.1: The comparison of the exact and approximate solutions with $\alpha = 0.5$ for test example 2.1 and 2.2.

Table 2.1: Numerical values of the exact solution and approximate solutions for different order α of equation (1) with $x = 0.5$.

t	$\alpha = 0.2$	$\alpha = 0.6$	$\alpha = 0.8$	exact solution
0.1	0.8660	0.8660	0.8660	0.8660
0.2	0.8969	0.8939	0.8926	0.8947
0.3	0.9243	0.9243	0.9251	0.9207
0.4	0.9690	0.9699	0.9717	0.9640
0.5	1.0314	1.0326	1.0349	1.0246

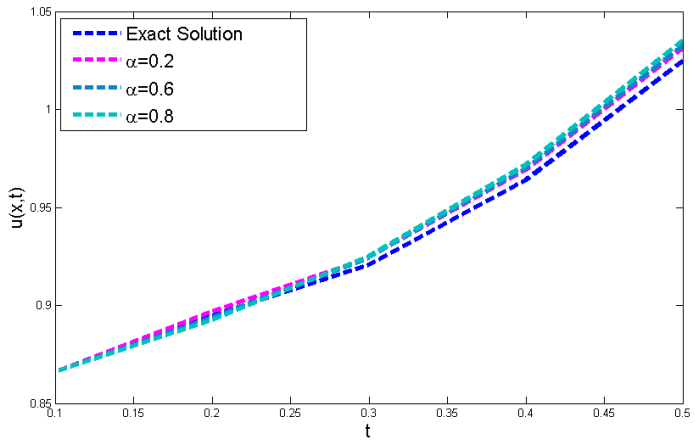


Figure 2.2: The behaviour of the exact solution and the approximate solutions of equation (1) with $x = 0.5$.

Chapter 3

Optimal control problem

In this chapter, we propose a regularized optimal control approach to identify the unknown parameter $p(x)$ in the time-fractional diffusion equation (1) from given data (2). First, we reformulate the inverse problem (1)-(2) into a regularized optimal control problem using the least squares method. Therefore, we establish the existence of the solution for control model. Then, we compute the necessary optimality conditions using the adjoint problem and study the main result concerning the stability and uniqueness of the minimizer. Finally, after calculating the gradient of the cost functional, we give the gradient descent algorithm and some numerical tests.

The formulation of our inverse problem into a regularized optimal control problem yields the following

$$J(p^*) = \arg\min_{p \in \mathcal{U}_{ad}} J(p), \tag{3.1}$$

with

$$J(p) = \frac{1}{2}\int_\Omega |u(x,T;p) - g(x)|^2 dx + \frac{\beta}{2}\int_\Omega |\nabla p|^2 dx, \tag{3.2}$$

and $\mathcal{U}_{ad}$ is a set of admissible parameters under the following definition

$$\mathcal{U}_{ad} = \Big\{ p \in L^2(\Omega) : 0 < c_1 \leq p \leq c_2 \text{ .a.e in } \Omega \text{ and } \|\nabla p\|_{L^2(\Omega)} \leq c_3 \Big\},$$

where c_1, c_2 and c_3 are a positives constants, $u(x,t;p)$ is the solution of the problem (1) for a given potential $p \in \mathcal{U}_{ad}$ and β is a regularization parameter.

3.1 Existence of minimizer

The result of the existence of the solution for the problem (3.1)-(3.2) in the following theorem.

Theorem 3.1. *Under the assumptions $\boldsymbol{H}_1$, there exists at least one solution for the op-*

timization problem (3.1)-(3.2).

Proof. Clearly, $\inf J(p)$ is finite over the admissible set $\mathcal{U}_{ad}$, and thus there exists a minimizing sequence $p_n \in \mathcal{U}_{ad}$ such that

$$\lim_{n\longrightarrow\infty} J(p_n) = \min_{p\in\mathcal{U}_{ad}} J(p).$$

By definition of admissible set $\mathcal{U}_{ad}$, the boundedness of the sequence $p_n \in \mathcal{U}_{ad}$ implies that there exists a subsequence, still denoted by p_n, and $p^* \in \mathcal{U}_{ad}$ such that

$$p_n \rightharpoonup p^* \text{ weakly in } H^1(\Omega). \tag{3.3}$$

Let $u_n = u(p_n)$ be the solution in $B^{\frac{\alpha}{2}}(Q)$ of the equation (1), corresponding to p_n, satisfying

$$({}^R D_t^{\frac{\alpha}{2}} u_n, {}^R D_T^{\frac{\alpha}{2}} v)_Q + (\nabla u_n, \nabla v)_Q + (p_n(x) f(u_n), v)_Q = (\frac{\phi(x) t^{-\alpha}}{\Gamma(1-\alpha)}, v)_Q, \ \forall v \in B^{\frac{\alpha}{2}}(Q). \tag{3.4}$$

Using Lemma 2.3, we have $\|u_n\|_{B^{\frac{\alpha}{2}}(Q)}$ is uniformly bounded independent of n. Then we can extract a subsequence denoted again u_n, and some $u^* \in B^{\frac{\alpha}{2}}(Q)$ such that

$$u_n \rightharpoonup u^* \text{ weakly in } B^{\frac{\alpha}{2}}(Q). \tag{3.5}$$

Thanks to the compact embedding of $B^{\frac{\alpha}{2}}(Q)$ in $L^2(Q)$ (see,[35]), we can extract a subsequence denoted again by u_n, such that

$$u_n \longrightarrow u^* \text{ in } L^2(Q). \tag{3.6}$$

By using the assumption $(\mathbf{H}_1)$, we obtain

$$f(u_n) \longrightarrow f(u^*) \text{ in } L^2(Q). \tag{3.7}$$

Next we prove that $u^* = u(p^*)$ be the solution of (1) corresponding to p^*. For every $v \in B^{\frac{\alpha}{2}}(Q)$ we have

$$\int_Q p_n(x) f(u_n) v dx dt = \int_Q p_n \left(f(u_n) - f(u^*)\right) v dx dt + \int_Q \left(p_n - p^*\right) f(u^*) v dx dt + \int_Q p^* f(u^*) v dx dt.$$

Now the strong convergences (3.6)-(3.7) and the weak convergence of p_n in $L^2(\Omega)$, (5.11) yield that

$$\lim_{n\longrightarrow\infty}({}^RD_t^{\frac{\alpha}{2}}u_n,{}^R\,D_T^{\frac{\alpha}{2}}v)_Q=({}^RD_t^{\frac{\alpha}{2}}u^*,{}^R\,D_T^{\frac{\alpha}{2}}v)_Q,$$
$$\lim_{n\longrightarrow\infty}(\nabla u_n,\nabla v)_Q=(\nabla u^*,\nabla v)_Q,$$
$$\lim_{n\longrightarrow\infty}\int_Q p_n(x)f(u_n)vdxdt=\int_Q p^*f(u^*)vdxdt.$$

Letting n tend to infinity in equation (3.4), we conclude that the function u^* satisfies

$$({}^RD_t^{\frac{\alpha}{2}}u^*,{}^R\,D_T^{\frac{\alpha}{2}}v)_Q+(\nabla u^*,\nabla v)_Q+(p^*(x)f(u^*),v)_Q=(\frac{\phi(x)t^{-\alpha}}{\Gamma(1-\alpha)},v)_Q. \tag{3.8}$$

this gives $u^*=u(p^*)$ by the definition of $u(p^*)$.

Moreover, by using the Theorem 2.1, the solution u_n is bounded in $\in C(\bar{I};\bar{H}^2(\Omega))$. From the continuous embedding $C(\bar{I};\bar{H}^2(\Omega))$ of into $L^\infty(I;L^2(\Omega))$, then u_n is bounded in $L^\infty(I;L^2(\Omega))$, we obtain

$$u_n\rightharpoonup u^*\text{ weakly in }L^\infty(I;L^2(\Omega)). \tag{3.9}$$

and since we have the

$$p_n\rightharpoonup p^*\text{ weakly in }H^1(\Omega), \tag{3.10}$$

we obtain the following

$$\begin{aligned}J(p^*)&=\frac{1}{2}\int_\Omega(u(x,T;p^*)-g(x))^2dx+\frac{\beta}{2}\int_\Omega(\nabla p^*)^2dx,\\&\leq\lim_{n\longrightarrow\infty}\inf\left(\frac{1}{2}\int_\Omega(u(x,T;p_n)-g(x))^2dx+\frac{\beta}{2}\int_\Omega(\nabla p_n)^2dx\right),\\&\leq\lim_{n\longrightarrow\infty}\inf J(p_n),\\&=\inf_{p\in\mathcal{U}_{ad}}J(p).\end{aligned}$$

Therefore p^* is a minimizer of the functional J. □

3.2 Necessary optimality condition

In this section, we establish the necessary condition of our optimal control problem that allows us to establish a stability result in the following section.

Theorem 3.2. *Let p be the solution of the optimal control problem (3.1). Then there exists a unique adjoint state v, and (u, v, p) satisfying the following system:*

$$\begin{cases} D_t^\alpha u(x,t) - \Delta u(x,t) + p(x)f(u) &= 0, & (x,t) \in Q, \\ u(x,0) &= \phi(x), & x \in \Omega, \\ u(x,t) &= 0, & x \in \partial\Omega \quad t \in I. \end{cases} \tag{3.11}$$

$$\begin{cases} D_T^\alpha v(x,t) - \Delta v(x,t) + pf'(u)v(x,t) &= F(x,t) & (x,t) \in Q, \\ v(x,T) &= 0 & x \in \Omega, \\ v(x,t) &= 0 & x \in \partial\Omega \quad t \in I, \end{cases} \tag{3.12}$$

where $F(x,t) = (u(x,t) - g(x))\delta_T(t)$. Moreover

$$\int_0^T \int_\Omega (p-h)f(u)v dx dt - \beta \int_\Omega \nabla p.\nabla(p-h)dx \geq 0, \tag{3.13}$$

for any $h \in \mathcal{U}_{ad}$.

Proof. For any $h \in \mathcal{U}_{ad}$, we consider $p_\gamma = (1-\gamma)p + \gamma h \in \mathcal{U}_{ad}$, $\forall \gamma \in [0,1]$. Then

$$J(p_\gamma) = \frac{1}{2}\int_\Omega (u(x,T;p_\gamma) - g(x))^2 dx + \frac{\beta}{2}\int_\Omega (\nabla p_\gamma)^2 dx.$$

Let u_γ be the solution of problem (3.11) with given $p = p_\gamma$. Since p is an optimal solution, then

$$\frac{dJ_\gamma}{d\gamma} \mid_{\gamma=0} = \int_\Omega (u(x,T;p) - g(x))\frac{\partial u_\gamma}{\partial \gamma} \mid_{\gamma=0} dx + \beta \int_\Omega \nabla p.\nabla(h-p)dx \geq 0. \tag{3.14}$$

Let $\tilde{u}_\gamma = \dfrac{\partial u_\gamma}{\partial \gamma}$, direct calculations lead to the following system:

$$\begin{cases} D_t^\alpha \tilde{u}_\gamma(x,t) - \Delta \tilde{u}_\gamma(x,t) + p_\gamma f'(u_\gamma)\tilde{u}_\gamma(x,t) &= -(h-p)f(u_\gamma), & (x,t) \in Q, \\ \tilde{u}_\gamma(x,0) &= 0, & x \in \Omega, \\ \tilde{u}_\gamma(0,t) = \tilde{u}_\gamma(1,t) &= 0, & t \in I. \end{cases} \tag{3.15}$$

Let $u^1 = \tilde{u}_\gamma$ at $\gamma = 0$. Then u^1 satisfies the following sensitive problem

$$\begin{cases} D_t^\alpha u^1(x,t) - \Delta u^1(x,t) + pf'(u)u^1(x,t) &= -(h-p)f(u), \quad (x,t) \in Q, \\ u^1(x,0) &= 0, \qquad\qquad x \in \Omega, \\ u^1(x,t) &= 0, \qquad\qquad x \in \partial\Omega \quad t \in I. \end{cases} \tag{3.16}$$

From (3.14), we have

$$\int_\Omega (u(x,T;p) - g(x))u^1 dx + \beta \int_\Omega \nabla p.\nabla(h-p)dx \geq 0. \tag{3.17}$$

Multiplying a function $v \in B^{\frac{\alpha}{2}}(Q)$ on both sides of the first equation in (3.16) and integrating over Q, we obtain

$$\int_0^T \int_\Omega -(h-q)f(u)v dxdt = \int_0^T \int_\Omega \quad D_t^\alpha u^1 v \, dxdt - \int_0^T \int_\Omega \Delta u^1 v dxdt + \int_0^T \int_\Omega p(x)f'(u)u^1 v dxdt,$$

which means

$$\int_0^T \int_\Omega -(h-p)f(u)v \, dxdt = (D_t^\alpha u^1, v)_Q - (\Delta u^1, v)_Q + (p(x)f'(u)u^1, v)_Q. \tag{3.18}$$

By Lemma 1.7, we have

$$\begin{aligned} (D_t^\alpha u^1, v)_Q &= \left({}^R D_t^{\frac{\alpha}{2}} u^1, {}^R D_T^{\frac{\alpha}{2}} v\right)_Q - \left(\frac{u^1(x,0)t^{-\alpha}}{\Gamma(1-\alpha)}, v\right)_Q \\ &= \left(u^1, {}^R D_T^\alpha v\right)_Q \\ &= (u^1, D_T^\alpha v)_Q + \left(u^1, \frac{v(x,T)}{\Gamma(1-\alpha)(T-t)^\alpha}\right)_Q. \end{aligned} \tag{3.19}$$

Using the Green formula and the boundary conditions of v and u^1, we obtain :

$$\left(\Delta u^1, v\right)_Q = \left(u^1, \Delta v\right)_Q. \tag{3.20}$$

Therefore, from (3.18), (3.19) and (3.20) we have

$$\begin{aligned} \int_0^T \int_\Omega -(h-p)f(u)v dxdt = \left(u^1, D_T^\alpha v\right)_Q - \left(u^1, \Delta v\right)_Q + (p(x)f'(u)v, u^1)_Q \\ + \left(u^1, \frac{v(x,T)}{\Gamma(1-\alpha)(T-t)^\alpha}\right)_Q \end{aligned}$$

Supposing that v is the solution of the adjoint problem (3.12), we obtain

$$\int_Q -(h-q)f(u)v(x,t)dxdt = \int_\Omega [u(x,T)-g(x)]u^1(x,T)dx. \tag{3.21}$$

Combining (3.17) and (3.21), one can easily obtain that

$$\int_0^T \int_\Omega (p-h)f(u)vdxdt - \beta \int_\Omega \nabla p.\nabla(p-h)dx \geq 0.$$

This completes the proof. □

3.3 Uniqueness and stability

In this section, we will study the stability and uniqueness of our optimal control problem (3.1). First of all, we consider the weak formulation of problem (3.12) as follows: For $F \in B^{\frac{\alpha}{2}}(Q)'$,

$$\begin{cases} \text{Find} \quad v \in B^{\frac{\alpha}{2}}(Q) \quad \text{such that} \\ \left({}^R D_T^{\frac{\alpha}{2}} v, {}^R D_t^{\frac{\alpha}{2}} \varphi\right)_Q + (\nabla v, \nabla \varphi)_Q + (p(x)f'(u)v, \varphi)_Q = (F,\varphi)_{(B^{\frac{\alpha}{2}}(Q)', B^{\frac{\alpha}{2}}(Q))} \\ \forall \varphi \in B^{\frac{\alpha}{2}}(Q). \end{cases} \tag{3.22}$$

In preparation for this section, we need to prove the well-posedness of the weak problem (3.22), which is presented in the following theorem. In sequel, we denote by $C > 0$ a generic constant.

Theorem 3.3. *For all $0 < \alpha < 1$ and $F \in B^{\frac{\alpha}{2}}(Q)'$, problem (3.22) admits a unique solution. Furthermore, if v is its solution, then it holds*

$$\|v\|_{B^{\frac{\alpha}{2}}(Q)} \leq C\|u(x,T)-g(x)\|_{L^\infty(\Omega)},$$

where C is independent of p.

Proof. The existence and uniqueness of the solution is guaranteed by the well-known Lax-Milgram lemma (see [66]). It consists in verifying the continuity and coercivity of the bilinear form and the continuity linear form. From this, we have

$$\|v\|_{B^{\frac{\alpha}{2}}(Q)} \leq (F,v)_Q \leq C\|u(x,T)-g(x)\|_{L^2(\Omega)}\|v\|_{L^2(Q)},$$

then employing Theorem 2.1 yields

$$\|v\|_{B^{\frac{\alpha}{2}}(Q)} \leq C\|u(x,T)-g(x)\|_{L^\infty(\Omega)}.$$

□

In the following theorem, we give the regularity of the solution for the adjoint problem through a change of variable to transform the problem at final time to a initial one.

Theorem 3.4. *Let f satisfies the assumptions (**H**). Then the solution v to problem (3.12) belongs to $C(\bar{I};\bar{H}^2(\Omega))$.*

Proof. First of all, we take $\tau = T-t$, then the adjoint problem (3.12) is reduced to

$$\begin{cases} D_\tau^\alpha v(x,\tau) - \Delta v(x,\tau) + pf'(u(x,T-\tau))v(x,\tau) = F(x,\tau) & (x,\tau)\in Q, \\ v(x,0) = 0 & x\in\Omega, \\ v(x,\tau) = 0 & x\in\partial\Omega \quad \tau\in I, \end{cases} \tag{3.23}$$

where $F(x,\tau) = (u(x,T-\tau)-g(x))\delta_0(\tau)$.
To prove this result, it suffices to discuss the integral equation

$$\begin{aligned} v(\tau) &= \int_0^\tau \bar{\mathbb{E}}(\tau-s)(F(x,s)ds + \int_0^\tau \bar{\mathbb{E}}(\tau-s)(-pf'(u(T-s))v(s))ds, \qquad 0<\tau<T \\ &= \bar{\mathbb{E}}(\tau)(u(x,T)-g(x)) + \int_0^\tau \bar{\mathbb{E}}(\tau-s)(-pf'(u(T-s))v(s))ds. \end{aligned}$$

Let $\sigma\in[0,1]$ one can estimate $\mathbb{A}^{\frac{\sigma}{2}}v(\tau)$

$$\mathbb{A}^{\frac{\sigma}{2}}v(\tau) = \mathbb{A}^{\frac{\sigma}{2}}\bar{\mathbb{E}}(\tau)(u(x,T)-g(x)) + \int_0^\tau \mathbb{A}^{\frac{\sigma}{2}}\bar{\mathbb{E}}(\tau-s)(-pf'(u(T-s))v(s))ds.$$

Then, we have

$$\begin{aligned} \|\mathbb{A}^{\frac{\sigma}{2}}v(\tau)\|_{L^2(\Omega)} &\leq \|\mathbb{A}^{\frac{\sigma}{2}}\bar{\mathbb{E}}(\tau)(u(x,T)-g(x))\|_{L^2(\Omega)} + \int_0^\tau \|\mathbb{A}^{\frac{\sigma}{2}}\bar{\mathbb{E}}(\tau-s)(-pf'(u(T-s))v(s))\|_{L^2(\Omega)}ds \\ &\leq \|\bar{\mathbb{E}}(\tau)(u(x,T)-g(x))\|_{\bar{H}^\sigma(\Omega)} + \int_0^\tau \|\bar{\mathbb{E}}(\tau-s)(-pf'(u(T-s))v(s))\|_{\bar{H}^\sigma(\Omega)}ds. \end{aligned}$$

By the direct application of the Lemma 2.1, we have

$$\begin{aligned}\|\mathbb{A}^{\frac{\sigma}{2}}v(\tau)\|_{L^2(\Omega)} &\leq C\tau^{\alpha-1}\|u(x,T)-g(x)\|_{\bar{H}^\sigma(\Omega)} + C\int_0^\tau s^{\alpha-1}\|pf'(u(T-s))v(s))\|_{\bar{H}^\sigma(\Omega)}ds \\ &\leq C\tau^{\alpha-1}\|u(x,T)-g(x)\|_{\bar{H}^\sigma(\Omega)} + C\int_0^\tau s^{\alpha-1}\|v(s)\|_{\bar{H}^\sigma(\Omega)}ds.\end{aligned}$$

It is easy to deduce that $\|\mathbb{A}^{\frac{\sigma}{2}}v(\tau)\|_{L^2(\Omega)} = \|v(\tau))\|_{\bar{H}^\sigma(\Omega)}$. From the Lemma 2.1, we obtain

$$\begin{aligned}\|v(\tau))\|_{\bar{H}^\sigma(\Omega)} \leq& C\tau^{\alpha-1}\|u(x,T)-g(x)\|_{\bar{H}^\sigma(\Omega)} \\ &+ Cb\int_0^\tau s^{\alpha-1}(\tau-s)^{\alpha-1}E_{\alpha,\alpha}(b(\tau-s)^\alpha)\|u(x,T-\tau)-g(x)\|_{\bar{H}^\sigma(\Omega)}ds \\ \leq& C\tau^{\alpha-1}\|u(x,T)-g(x)\|_{\bar{H}^\sigma(\Omega)}.\end{aligned}$$

For any $T > \tau_0 > 0$, $\tau \in [\tau_0, T]$, we have

$$\|v(\tau))\|_{\bar{H}^\sigma(\Omega)} \leq C\tau_0^{\alpha-1}\|u(x,T)-g(x)\|_{\bar{H}^\sigma(\Omega)}.$$

Therefore, we have $\mathbb{A}^{\frac{\sigma}{2}}v \in C(\bar{I}; L^2(\Omega))$ then $v \in C(\bar{I}; \bar{H}^\sigma(\Omega))$.
This completes the proof of the theorem. □

Now we are in position to prove the stability of the optimal control model, which means to show the difference between p_1 and p_2 remains the same or less for the difference between g_1 and g_2. For this, Assuming that $g_1(x)$ and $g_2(x)$ are two given functions which satisfy the condition (2). Let $p_1(x)$ and $p_2(x)$ be the minimizers of the optimal control problem (3.1) corresponding to $g_1(x)$ and $g_2(x)$ respectively. Let also:

- $\{u_i, i = 1, 2\}$ be solutions of the problem (3.11) associated to $p = p_i$ with $i = 1, 2$ respectively.

- $\{v_i, i = 1, 2\}$ be solutions of the problem (3.12) associated to $p = p_i$ with $i = 1, 2$ respectively.

By taking

$$U = u_1 - u_2, \quad V = v_1 - v_2, \text{ and } P = p_1 - p_2,$$

where U solution

$$\begin{cases} D_t^\alpha U - \Delta U + p_1(f(u_1) - f(u_2)) &= -Pf(u_2), & (x,t) \in Q, \\ U(x,0) &= 0, & x \in \Omega, \\ U(x,t) &= 0, & x \in \partial\Omega \quad t \in I, \end{cases} \tag{3.24}$$

and V satisfy

$$\begin{cases} D_T^\alpha V - \Delta V + p_1 f'(u_1)V &= -v_2 f'(u_1)P - p_2 v_2(f'(u_1) - f'(u_2)) + \mathbf{F}(x,t), \quad (x,t) \in Q, \\ V(x,T) &= 0, \quad x \in \Omega, \\ V(x,t) &= 0, \quad x \in \partial\Omega \quad t \in I, \end{cases} \tag{3.25}$$

where $\mathbf{F}(x,t) = (U(x,t) - (g_1 - g_2))\,\delta_T(t)$.

Before giving the theorem of the stability, we give some useful that will used in the proof of the main results. The following lemma show an estimation on the solution U of the equation (3.24).

Lemma 3.1. *For U solution of the equation (3.24), we have the following estimation:*

$$\int_0^T \int_\Omega U^2 dxdt \leq C(\max_{x\in\Omega} \mid P \mid)^2 \int_0^T \int_\Omega \mid u_2 \mid^2 dxdt. \tag{3.26}$$

where C is independent of T and Ω.

Proof. Under assumption $(\mathbf{H}_1) - (\mathbf{H}_2)$ and applying the mean value theorem, we have

$$\begin{aligned} f(u_1) - f(u_2) &= f'(u_2 + \theta(u_1 - u_2))(u_1 - u_2) \\ &= f'(u_2 + \theta U)U, \end{aligned}$$

where $0 < \theta < 1$. Then the equation (3.24) can be rewritten as follow

$$\begin{cases} D_t^\alpha U - \Delta U + p_1 f'(u_2 + \theta U)U &= -Pf(u_2), & (x,t) \in Q, \\ U(x,0) &= 0, & x \in \Omega, \\ U(x,t) &= 0, & x \in \partial\Omega \quad t \in I. \end{cases} \tag{3.27}$$

Multiplying the function U on both sides of the first equation in (3.27) and integrating

over Q, we have

$$(D_t^{\alpha}U, U)_Q - (\Delta U, U)_Q + (p_1 f'(u_2 + \theta U)U, U)_Q = (-Pf(u_2), U)_Q.$$

We apply the lemma 1.7 and Green formula, we obtain

$$({}^R D_t^{\frac{\alpha}{2}}U, {}^R D_T^{\frac{\alpha}{2}}U)_Q + (\nabla U, \nabla U)_Q + (p_1 f'(u_2 + \theta U), U^2)_Q = (-Pf(u_2), U)_Q,$$

Using the fact that f' and p_1 are positive, then we obtain

$$\|U\|^2_{L^2(I;H_0^1(\Omega))} \leq \int_0^T \int_\Omega \mid Pf(u_2)U \mid dxdt.$$

The assumptions $(\mathbf{H}_1) - (\mathbf{H}_2)$ on the function f imply that

$$\|U\|^2_{L^2(I;L^2(\Omega))} \leq C \int_0^T \int_\Omega \mid P \mid\mid u_2 \mid\mid U \mid dxdt.$$

We use the Young's inequality, we obtain

$$\int_0^T \int_\Omega \mid U \mid^2 dxdt \leq \frac{1}{2} \int_0^T \int_\Omega \mid U \mid^2 dxdt + \frac{C}{2} \int_0^T \int_\Omega \mid P\ u_2 \mid^2 dxdt.$$

hence

$$\int_0^T \int_\Omega U^2 dxdt \leq C(\max_{x\in\Omega} \mid P \mid)^2 \int_0^T \int_\Omega \mid u_2 \mid^2 dxdt.$$

This completes the proof of the Lemma. □

Lemma 3.2. *For equation (3.25) we have the following estimate:*

$$\int_0^T \int_\Omega \mid V \mid^2 dxdt \leq C(\max_{x\in\Omega} \mid P \mid)^2 \int_0^T \int_\Omega \left(\mid v_2 \mid^2 + \mid u_2 \mid^2 \right) dxdt + C \int_\Omega \mid g_1 - g_2 \mid^2 dx. \tag{3.28}$$

where C is independent of T and Ω.

Proof. Multiplying the function V on both sides of the first equation in (3.25) and integrating over Q, we have

$$(D_T^{\alpha}V, V)_Q - (\Delta V, V)_Q + (p_1 f'(u_1)V, V)_Q = (-v_2 f'(u_1)P - p_2 v_2 (f'(u_1) - f'(u_2)), V)_Q + (\mathbf{F}, V)_Q.$$

we apply the Lemma 1.7 and Green formula, we obtain

$$({}^{R}D_{T}^{\frac{\alpha}{2}}V, {}^{R}D_{t}^{\frac{\alpha}{2}}V)_Q + (\nabla V, \nabla V)_Q + (p_1 f'(u_1), V^2)_Q = (-v_2 f'(u_1)P - p_2 v_2 (f'(u_1) - f'(u_2)), V)_Q + (\mathbf{F}, V)_Q.$$

Using the fact that f' and p_1 are positive, then we obtain

$$\|V\|^2_{L^2(I;H_0^1(\Omega))} \leq \int_0^T \int_\Omega \mid [v_2 f'(u_1)P + p_2 v_2 (f'(u_1) - f'(u_2))]V \mid dxdt + \int_0^T \int_\Omega \mid \mathbf{F}V \mid dxdt,$$

the assumptions $(\mathbf{H}_1) - (\mathbf{H}_2)$ on the function f imply that

$$\|V\|^2_{L^2(I;H_0^1(\Omega))} \leq C \int_0^T \int_\Omega \mid [v_2 f'(u_1)P + p_2 v_2 (f'(u_1) - f'(u_2))]V \mid dxdt + C \int_0^T \int_\Omega \mid UV\delta_T(t) \mid dxdt + C \int_0^T \int_\Omega \mid (g_1 - g_2)V\delta_T(t) \mid dxdt.$$

We use the Young's inequality and the fact $V(x,T) = 0$ in Ω, we obtain

$$\|V\|^2_{L^2(I;H_0^1(\Omega))} \leq \frac{1}{2} \int_0^T \int_\Omega \mid V \mid^2 dxdt + C \int_0^T \int_\Omega \left(\mid v_2 f'(u_1)P \mid^2 + \mid p_2 v_2 (f'(u_1) - f'(u_2) \mid^2 \right) dxdt + C \int_\Omega \mid g_1 - g_2 \mid^2 dxdt.$$

This yields

$$\int_0^T \int_\Omega \mid V \mid^2 dxdt \leq C (\max_{x\in\Omega} \mid P \mid)^2 \int_0^T \int_\Omega \mid v_2 f'(u_1) \mid^2 dxdt + C \int_0^T \int_\Omega \mid p_2 v_2 (f'(u_1) - f'(u_2) \mid^2 dxdt + C \int_\Omega \mid g_1 - g_2 \mid^2 dxdt$$

Note from $(\mathbf{H}_1)$ we have f is Lipschitz, then we have f' and f'' is bounded in $L^\infty(I; L^\infty(\Omega))$ and the fact p_2 is bounded. From Theorem 3.4 and Lemma 3.1, we obtain

$$\int_0^T \int_\Omega \mid p_2 v_2 (f'(u_1) - f'(u_2) \mid^2 dxdt \leq C \int_0^T \int_\Omega \mid f''(u_1 + \theta U) \mid^2 \mid U \mid^2 dxdt \leq C (\max_{x\in\Omega} \mid P \mid)^2 \int_0^T \int_\Omega \mid u_2 \mid^2 dxdt.$$

and

$$\int_0^T \int_\Omega \mid v_2 f'(u_1) \mid^2 dxdt \leq C \int_0^T \int_\Omega \mid v_2 \mid^2 dxdt \tag{3.29}$$

Then we conclude

$$\int_0^T \int_\Omega \mid V \mid^2 dxdt \leq C(\max_{x\in\Omega} \mid P \mid)^2 \int_0^T \int_\Omega \left(\mid v_2 \mid^2 + \mid u_2 \mid^2 \right) dxdt + C \int_\Omega \mid g_1 - g_2 \mid^2 dxdt.$$

Hence the proof of the lemma. □

Now, we give the main contribution concerned the stability result of our inverse problem.

Theorem 3.5. *Let p_1 and p_2 be the minimizers of the optimal control problem (3.1) corresponding to g_1 and g_2 respectively. If there exists a point $x_0 \in \Omega$ such that $p_1(x_0) = p_2(x_0)$, then for $T << 1$ we have the following estimation:*

$$\max_{x\in\Omega} \mid P \mid \leq C \|g_1 - g_2\|_{L^2(\Omega)}.$$

Proof. By taking $h = p_2$ when $p = p_1$ and $h = p_1$ when $p = p_2$ in (3.13), we have

$$\int_0^T \int_\Omega (p_1 - p_2) f(u_1) v_1 dxdt - \beta \int_\Omega \nabla p_1 . \nabla (p_1 - p_2) dx \geq 0, \tag{3.30}$$

$$\int_0^T \int_\Omega (p_2 - p_1) f(u_2) v_2 dxdt - \beta \int_\Omega \nabla p_2 . \nabla (p_2 - p_1) dx \geq 0, \tag{3.31}$$

where $\{u_i, v_i; i = 1, 2\}$ are solutions of problems (3.11) − (3.12) associated to $\{p = p_i; i = 1, 2\}$, respectively.
by combining (3.30) and (3.31), we obtain

$$\beta \int_\Omega |\nabla P|^2 dx = \beta \int_\Omega \mid \nabla (p_1 - p_2) \mid^2 dx \leq \int_0^T \int_\Omega P \left[(f(u_1) - f(u_2)) v_1 + f(u_2) V \right] dxdt. \tag{3.32}$$

We have also

$$\mid P(x) \mid \leq \mid P(x_0) \mid + \int_{x_0}^x |\nabla P| dx \leq \mid P(x_0) \mid + \int_\Omega |\nabla P| dx.$$

From the assumption of Theorem 3.5, there exists a point $x_0 \in \Omega$ such that $P(x_0) =$

$p_1(x_0) - p_2(x_0) = 0$ and by applying the Cauchy–Schwarz inequality, we obtain

$$(\max_{x\in\Omega} \mid P(x) \mid)^2 \leq mes(\Omega) \int_\Omega \mid \nabla P \mid^2 dx, \tag{3.33}$$

since we have $mes(\Omega) = 1$, then we have

$$(\max_{x\in\Omega} \mid P(x) \mid)^2 \leq \int_\Omega \mid \nabla P \mid^2 dx. \tag{3.34}$$

From the inequalities (3.32) – (3.34), we obtain

$$(\max_{x\in\Omega} \mid P(x) \mid)^2 \leq \frac{1}{\beta} \int_0^T \int_\Omega P\left[(f(u_1) - f(u_2))v_1 + f(u_2)V\right] dxdt. \tag{3.35}$$

Using Cauchy-Schwarz inequality and the assumptions on the function f imply that

$$\begin{aligned}(\max_{x\in\Omega} \mid P(x) \mid)^2 &\leq \frac{1}{\beta} \int_0^T \int_\Omega P\left[(f(u_1) - f(u_2))v_1 + f(u_2)V\right] dxdt. \\ &\leq \frac{C}{\beta} \max_{x\in\Omega} \mid P(x) \mid \int_0^T \int_\Omega \mid Uv_1 \mid + \mid u_2V \mid dxdt \\ &\leq \frac{C}{\beta} \max_{x\in\Omega} \mid P(x) \mid \left[\|U\|_{L^2(Q)}\|v_1\|_{L^2(Q)} + \|u_2\|_{L^2(Q)}\|V\|_{L^2(Q)}\right].\end{aligned}$$

Now, we use the Lemma 3.1 and lemma 3.2, we obtain

$$\begin{aligned}(\max_{x\in\Omega} \mid P(x) \mid)^2 \leq& \frac{C}{\beta}(\max_{x\in\Omega} \mid P(x) \mid)^2\|u_2\|_{L^2(Q)}\|v_1\|_{L^2(Q)} + \frac{C}{\beta}(\max_{x\in\Omega} \mid P(x) \mid)^2\|u_2\|_{L^2(Q)}\|v_2\|_{L^2(Q)} \\ &+ \frac{C}{\beta}(\max_{x\in\Omega} \mid P(x) \mid)^2\|u_2\|_{L^2(Q)}\|u_2\|_{L^2(Q)} + \frac{C}{\beta} \max_{x\in\Omega} \mid P(x) \mid \|u_2\|_{L^2(Q)}\|g_1 - g_2\|_{L^2(\Omega)}.\end{aligned}$$

From the theorems 2.1, there exists a constant C such that

$$\|u_2\|_{L^2(Q)} \leq C\sqrt{T}, \tag{3.36}$$

then, we have

$$\begin{aligned}(\max_{x\in\Omega} \mid P(x) \mid)^2 \leq& \frac{C\sqrt{T}}{\beta}(\max_{x\in\Omega} \mid P(x) \mid)^2 \left(\|v_1\|_{L^2(Q)} + \|v_2\|_{L^2(Q)} + \|u_2\|_{L^2(Q)}\right) \\ &+ \frac{C\sqrt{T}}{\beta} \max_{x\in\Omega} \mid P(x) \mid \|g_1 - g_2\|_{L^2(\Omega)}.\end{aligned} \tag{3.37}$$

Choosing $T << 1$ such that

$$\frac{C\sqrt{T}}{\beta} < 1. \tag{3.38}$$

Combining (3.37), (3.38) and the theorem 3.4, one can easily get

$$\max_{x\in\Omega} \mid P(x) \mid \leq C\|g_1 - g_2\|_{L^2(\Omega)}. \tag{3.39}$$

This completes the proof of Theorem 3.5. □

3.4 Gradient of functional and descent gradient algorithm

In this section, we compute the gradient of the cost function using the adjoint state and we give the descent gradient algorithm. First, we establish the differentiability of the solution $u(p)$ with respect to p. To this end, we take $q = h - p$ in (3.16) and the sensitivity problem u^1 becomes

$$\begin{cases} D_t^\alpha u^1(x,t) - \Delta u^1(x,t) + pf'(u)u^1(x,t) &= -qf(u), \quad (x,t) \in Q, \\ u^1(x,0) &= 0, \qquad x \in \Omega \\ u^1(x,t) &= 0, \qquad x \in \partial\Omega \quad t \in I \end{cases} \tag{3.40}$$

Lemma 3.3. *For any $p \in \mathcal{U}_{ad}$ the solution $u(p)$ is differentiable with respect to p in the sense that*

$$\frac{\|u(p+q) - u(p) - u^1\|_{B^{\frac{\alpha}{2}}(Q)}}{\|q\|_{L^\infty(\Omega)}} \longrightarrow 0 \text{ as } q \longrightarrow 0.$$

Proof. The function $w \equiv u(p+q) - u(p) - u^1$ satisfies

$$\begin{cases} D_t^\alpha w - \Delta w + p\left(f(u(p+q)) - f(u(p)) - f'(u(p))u^1\right) &= -q\left(f(u(p+q)) - f(u(p))\right), \quad in \quad Q \\ w(x,0) &= 0, \quad x \in \Omega \\ w(x,t) &= 0, \quad x \in \partial\Omega \quad t \in I \end{cases} \tag{3.41}$$

From the convexity of the function f implies that

$$f'(u(p))\left[u(p+q) - u(p)\right] \leq f(u(p+q)) - f(u(p)). \tag{3.42}$$

Thus

$$p\left(f(u(p+q)) - f(u(p)) - f'(u(p))u^1\right) \geq pf'(u(p))w. \tag{3.43}$$

Multiplying a function $w \in B^{\frac{\alpha}{2}}(Q)$ on both sides of the first equation in (3.41), integrating

over Q and using (3.43), we obtain

$$({}^{R}D_t^{\frac{\alpha}{2}}w, {}^{R}D_T^{\frac{\alpha}{2}}w)_Q + (\nabla w, \nabla w)_Q + (pf'(u(p)), w^2)_Q \leq \left(-q\left(f(u(p+q)) - f(u(p))\right), w\right)_Q . \tag{3.44}$$

Using the assumptions $\mathbf{H}_1$ and $\mathbf{H}_2$ of the function f, we obtain

$$({}^{R}D_t^{\frac{\alpha}{2}}w, {}^{R}D_T^{\frac{\alpha}{2}}w)_Q + (\nabla w, \nabla w)_Q \leq C\int_Q |q|\ |u(p+q) - u(p)|\ |w| dxdt. \tag{3.45}$$

Moreover, we have for all $w \in B^{\frac{\alpha}{2}}(Q)$

$$({}^{R}D_t^{\frac{\alpha}{2}}w, {}^{R}D_T^{\frac{\alpha}{2}}w)_Q + (\nabla w, \nabla w)_Q \geq C\left(({}^{R}D_t^{\frac{\alpha}{2}}w, {}^{R}D_t^{\frac{\alpha}{2}}w)_Q + (\nabla w, \nabla w)_Q\right) \geq C\|w\|^2_{B^{\frac{\alpha}{2}}(Q)}. \tag{3.46}$$

By noting (3.45) and (3.46), we have

$$\|w\|^2_{B^{\frac{\alpha}{2}}(Q)} \leq C\int_Q |q|\ |u(p+q) - u(p)|\ |w| dxdt. \tag{3.47}$$

Now the Cauchy-Schwarz inequality, the fact that $B^{\frac{\alpha}{2}}(Q) \hookrightarrow$ in $L^2(Q)$ and $\mathcal{U}_{ad}$ include in $L^\infty(\Omega) \cap H^1(\Omega)$, implies that

$$\|w\|_{B^{\frac{\alpha}{2}}(Q)} \leq C\|q\|_{L^\infty(\Omega)}\|u(p+q) - u(p)\|_{B^{\frac{\alpha}{2}}(Q)}.$$

Similarly, we can show that

$$\|u(p+q) - u(p)\|_{B^{\frac{\alpha}{2}}(Q)} \leq C\|q\|_{L^\infty(\Omega)}\|u(p+q)\|_{B^{\frac{\alpha}{2}}(Q)}.$$

However, the Lemma 2.3 indicates that, $\|u(p+q)\|_{B^{\frac{\alpha}{2}}(Q)}$ is uniformly bounded. Thus it follows directly that

$$\frac{\|u(p+q) - u(p) - u^1\|_{B^{\frac{\alpha}{2}}(Q)}}{\|q\|_{L^\infty(\Omega)}} \longrightarrow 0 \text{ as } q \longrightarrow 0.$$

This completes the proof of the lemma. □

Thanks to the sensitivity problem (3.40) of the state, we are able to compute in next the gradient of the cost function (3.2).

Proposition 3.1. *The gradient of the cost function J is given by*

$$J'(p) = -\int_I f(u)v \ dt - \beta \Delta p, \tag{3.48}$$

where v is the solution of the adjoint problem

$$\begin{cases} D_T^\alpha v(x,t) - \Delta v(x,t) + pf'(u)v(x,t) &= F(x,t), \quad (x,t) \in Q, \\ v(x,T) &= 0, \qquad x \in \Omega, \\ v(x,t) &= 0, \qquad x \in \partial\Omega \quad t \in I, \end{cases} \tag{3.49}$$

with

$$F(x,t) = (u(x,t) - g(x))\delta_T(t)$$

in which $\delta_T(.)$ is the Dirac function.

Proof. From (3.2), we observe that

$$J'(p)[q] = \lim_{\gamma \longrightarrow 0} \frac{J(p+\gamma q) - J(p)}{\gamma} = \int_\Omega [u(x,T) - g(x)]u^1(x,T)dx - \beta \int_\Omega \Delta p[q]dx. \tag{3.50}$$

where u^1 is solution of the sensibility problem (3.40). Employing the same techniques used in the proof of the theorem 3.2, we deduce

$$\begin{aligned} \int_Q -qf(u)v(x,t)dxdt &= \int_Q [u(x,T) - g(x)]\delta_T(t)u^1(x,t)dx \\ &= \int_\Omega [u(x,T) - g(x)]u^1(x,T)dx. \end{aligned}$$

Using formula (3.50), we obtain

$$J'(p)[q] = \int_Q -qf(u)v(x,t)dxdt - \beta \int_\Omega \Delta p[q]dx,$$

which implies

$$< J'(p), q > = < -\int_I f(u)v \ dt, q > -\beta < \Delta p, q >,$$

thus we conclude that

$$J'(p) = -\int_I f(u)vdt - \beta \ \Delta p.$$

This completes the proof of Proposition. □

Now, we present the algorithm descent for optimal control problem (3.1).

Algorithm 1 Gradient descent algorithm

Begin

Input: the error tol, the initial condition u_0, the data g, the function f, α , β and ς^0.

Initialize: $p^0 \in \mathcal{U}_{ad}$ and set $k = 0$.

Repeat :

1. Compute u^k by solving the state problem (1).
2. Compute v^k by solving the adjoint problem (4.35).
3. Compute $J'(p^k)$ by (3.48).
4. **Update** $p^{k+1} = p^k - \varsigma^k J'(p^k)$ with ς^k is computed by Armijo linear search.

$k = k + 1$

Until $\| J'(p^{k+1})\| \leq tol$.

End

3.5 Numerical approximation and experiments

In this section, we present detailed numerical results to demonstrate the effectiveness of the descent gradient method for solving our inverse problem. Initially, we demonstrate a computationally efficient finite difference approximation to solve the optimal control problem, so that we can approximate both the direct and adjoint equations. To demonstrate the stability of the problem and the efficiency of the algorithm, we show numerical results with comments in the cases with and without noise on the data. Thus, in several examples of fractional derivative of order α, the numerical errors for different noise levels are computed.

3.5.1 Numerical approximation

In this section, we focus on approximating the problem in optimal control (3.1). Thus, the optimization problem is solved using a Gradient descent method presented in the algorithm (1). Recall that the state problem is approximated by the finite difference method using Newton's scheme in Chapter 2.

We shall now try to approximate the adjoint problem by finite differences method

using the explicit schema. The adjoint problem obtained in this study is recalled by :

$$\begin{cases} D_T^{\alpha} v(x,t) - \Delta v(x,t) + pf(u(x,t))v(x,t) &= (u(x,T) - g(x))\delta(t-T), \quad (x,t) \in Q, \\ v(x,T) &= 0, \qquad\qquad x \in \Omega, \\ v(0,t) = v(1,t) &= 0, \qquad\qquad t \in I. \end{cases} \tag{3.51}$$

Define $t_n = n\tau, n = 0, 1, 2, \cdots, N,\ x_i = ih,\ i = 0, 1, 2, \cdots, M$, where $\tau = \frac{T}{N+1}$ and $h = \frac{L}{M+1}$ are time and space steps, respectively. We seek in each of these points an approximate value, noted v_i^n be the numerical approximation to $v(x_i, t_n)$. The approximate values at the mesh points at the edge of the domain and at $t = T$ are given by the exact value -given of the function v:

$$v_0^n = v(0, t_n) = 0 \text{ for all } n = 0, 1, ..., N$$

$$v_1^n = v(1, t_n) = 0 \text{ for all } n = 0, 1, ..., N$$

$$v_i^T = v(x_i, T) = 0 \text{ for all } i = 0, 1, ..., M + 1$$

This leaves $N \times M$ unknowns to be determined, the v_i^n for $1 \leq n \leq N$ and for $1 \leq i \leq M$. For the vertices that is remains, the Caputo left-sided fractional derivative term in time given in the above equation (3.51) can be approximated by the following scheme

$$\begin{aligned} D_T^{\alpha} v(x_i, t_n) &= \frac{-1}{\Gamma(1-\alpha)} \int_{t_n}^{T} \frac{v'(x_i, s)}{(s - t_n)^{\alpha}} ds \\ &= \frac{-1}{\Gamma(1-\alpha)} \sum_{j=n+1}^{N} \int_{(j-1)k}^{jk} \frac{v'(x_i, s)}{(s - t_n)^{\alpha}} ds \\ &= \frac{-1}{\Gamma(1-\alpha)} \sum_{j=n+1}^{N} \int_{(j-1)k}^{jk} \frac{v_i^j - v_i^{j-1}}{k} (s - t_n)^{-\alpha} ds \\ &= \frac{-1}{\Gamma(1-\alpha)} \sum_{j=n+1}^{N} \left(\frac{v_i^j - v_i^{j-1}}{k} \right) \int_{(j-1)k}^{jk} (s - t_n)^{-\alpha} ds \\ &= \frac{-1}{\Gamma(1-\alpha)} \sum_{j=n+1}^{N} \left(\frac{v_i^j - v_i^{j-1}}{k} \right) \left[\frac{(s - t_n)^{1-\alpha}}{1-\alpha} \right]_{(j-1)k}^{jk} \\ &= \frac{-1}{(1-\alpha)\Gamma(1-\alpha)} \sum_{j=n+1}^{N} \left(\frac{v_i^j - v_i^{j-1}}{k} \right) \left((j-n)k)^{1-\alpha} - ((j-1-n)k)^{1-\alpha} \right) \\ &= \frac{-k^{-\alpha}}{\Gamma(2-\alpha)} \sum_{j=n+1}^{N} \left(v_i^j - v_i^{j-1} \right) \left((j-n)^{1-\alpha} - (j-1-n)^{1-\alpha} \right). \end{aligned}$$

Taking $j \longleftarrow j-n+1$, we obtain

$$ {}_t\partial_T^\alpha v(x_i,t_n) = \frac{-k^{-\alpha}}{\Gamma(2-\alpha)} \sum_{j=1}^{N-n} \left(v_i^{n+j} - u_i^{n+j-1}\right)\left((j)^{1-\alpha} - (j-1)^{1-\alpha}\right). $$

We will first calculate

$$ D_T^\alpha v(x_i,t_n) = v_{xx}(x_i,t_n) - p(x_i)f'(u(x_i,t_n))v(x_i,t_n) + (u(x_i,t_N) - g(x_i))\delta(t_n - t_N). $$

From what is done above, we summarize that

$$ \begin{cases} D_T^\alpha v(x_i,t_n) \simeq \dfrac{-k^{-\alpha}}{\Gamma(2-\alpha)} \displaystyle\sum_{j=1}^{N-n} \left(v_i^{n+j} - v_i^{n+j-1}\right)\left((j)^{1-\alpha} - (j-1)^{1-\alpha}\right) \\ v_{xx}(x_i,t_n) \simeq \dfrac{v_{i+1}^n - 2v_i^n + v_{i-1}^n}{h^2} \\ v(x_i,t_n) \simeq v_i^n \end{cases} \tag{3.52} $$

Thus

$$ \frac{-k^{-\alpha}}{\Gamma(2-\alpha)} \sum_{j=1}^{N-n} \left(v_i^{n+j} - v_i^{n+j-1}\right)\left((j)^{1-\alpha} - (j-1)^{1-\alpha}\right) = \frac{v_{i+1}^n - 2v_i^n + v_{i-1}^n}{h^2} - p_i v_i^n + (u_i^N - g_i). $$

For ease of calculation, we set

$$ \mu = \frac{k^\alpha}{h^2}, \quad r = \mu\Gamma(2-\alpha) \text{ et } r_1 = k^\alpha\Gamma(2-\alpha). $$

We suppose

$$ b_j = (j+1)^{1-\alpha} - (j)^{1-\alpha}, \qquad b_0 = 1, $$

yield

$$ \begin{aligned} -\sum_{j=1}^{N-n} \left(v_i^{n+j} - v_i^{n+j-1}\right) b_{j-1} &= r v_{i+1}^n - 2r v_i^n + r v_{i-1}^n - r_1 p_i v_i^n + r_1(u_i^N - g_i) \\ -v_i^{n+1} + v_i^n - \sum_{j=2}^{N-n} \left(v_i^{n+j} - v_i^{n+j-1}\right) b_{j-1} &= r v_{i+1}^n - 2r v_i^n + r v_{i-1}^n - r_1 p_i v_i^n + r_1(u_i^N - g_i). \end{aligned} $$

For $n = N-1$, we give

$$v_i^N = -rv_{i+1}^N + (1+2r+r_1p_i)v_i^{N-1} - rv_{i-1}^{N-1} - r_1(u_i^N - g_i),$$

the matrix form of this equation can then be written as

$$\begin{cases} \mathbf{v}^N \text{ is gevin} \\ A\mathbf{v}^{N-1} = \mathbf{v}^N + r_1\left(\mathbf{u}^N - g)\right), \end{cases} \tag{3.53}$$

For $n < N-1$, we take

$$c_j = b_{j-1} - b_j \quad \text{with} \quad c_1 = 1 - b_1.$$

This it helps us to show that

$$\begin{aligned}
-v_i^{n+1} + v_i^n - \sum_{j=2}^{N-n}\left(v_i^{n+j} - v_i^{n+j-1}\right)b_{j-1} &= rv_{i+1}^n - 2rv_i^n + rv_{i-1}^n - r_1p_iv_i^n + r_1(u_i^N - g_i) \\
v_i^{n+1} + \sum_{j=2}^{N-n}\left(v_i^{n+j} - v_i^{n+j-1}\right)b_{j-1} &= -rv_{i+1}^n + (1+2r+r_1p_i)v_i^n - rv_{i-1}^n - r_1(u_i^N - g_i) \\
v_i^{n+1} + \sum_{j=2}^{N-n} v_i^{n+j}b_{j-1} - \sum_{j=2}^{N-n} v_i^{n+j-1}b_{j-1} &= -rv_{i+1}^n + (1+2r+r_1p_i)v_i^n - rv_{i-1}^n - r_1(u_i^N - g_i) \\
v_i^{n+1} + \sum_{j=2}^{N-n} v_i^{n+j}b_{j-1} - \sum_{j=1}^{N-n-1} v_i^{n+j}b_j &= -rv_{i+1}^n + (1+2r+r_1p_i)v_i^n - rv_{i-1}^n - r_1(u_i^N - g_i) \\
v_i^{n+1} - b_1v_i^{n+1} + \sum_{j=2}^{N-n-1} v_i^{n+j}(b_{j-1} - b_j) + b_{N-n}v_i^N &= -rv_{i+1}^n + (1+2r+r_1p_i)v_i^n - rv_{i-1}^n - r_1(u_i^N - g_i) \\
(1-b_1)v_i^{n+1} + \sum_{j=2}^{N-n-1} v_i^{n+j}c_j + b_{N-n}v_i^N &= -rv_{i+1}^n + (1+2r+r_1p_i)v_i^n - rv_{i-1}^n - r_1(u_i^N - g_i),
\end{aligned}$$

yield

$$c_1\mathbf{v}^{n+1} + c_2\mathbf{v}^{n+2} + ... + c_{N-n-1}\mathbf{v}^{N-1} + b_{N-n}\mathbf{v}^N = A\mathbf{v}^n - r_1(\mathbf{u}^N - g). \tag{3.54}$$

The adjoint equation can then be written in the following matrix form

$$\begin{cases} \mathbf{u}^N \text{ given} \\ \mathbf{v}^N \text{ given} \\ A\mathbf{v}^{N-1} = \mathbf{v}^N + r_1\left(\mathbf{u}^N - g)\right) \\ A\mathbf{v}^n = c_1\mathbf{v}^{n+1} + c_2\mathbf{v}^{n+2} + ... + c_{N-n-1}\mathbf{v}^{N-1} + b_{N-n}\mathbf{v}^N + r_1(\mathbf{u}^N - g), \end{cases} \tag{3.55}$$

where

$$A = \begin{bmatrix} 1+2r+r_1p_1 & -r & 0 & \cdots & 0 & 0 \\ -r & 1+2r+r_1p_2 & -r & \cdots & 0 & 0 \\ 0 & -r & 1+2r+r_1p_3 & \cdots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & 1+2r+r_1p_{M-2} & -r \\ 0 & 0 & 0 & \ldots & -r & 1+2r+r_1p_{M-1} \end{bmatrix},$$

and

$$\mathbf{v}^n = \begin{bmatrix} v_1^n \\ v_2^n \\ \vdots \\ v_M^n \end{bmatrix}. \tag{3.56}$$

For the discretization of the optimal control problem, we define the family of discrete admissible parameters :

$$\mathcal{U}_{ad}^h = \left\{ p_h \in L^2(\Omega) : \ 0 < c_1 < p_h < c_2 \ \text{ and } \ \|\nabla_h p_h\|_{L^2(\Omega)} < c_3, \ \forall i = 1, ..., M \right\},$$

and we approximate the discrete cost functional by:

$$J_h(p_h) = \frac{1}{2} \sum_{i=1}^{M} \int_{(i-1)h}^{ih} |u_i^N - g_i|^2 dx + \frac{\beta}{2} \sum_{i=1}^{M} \int_{(i-1)h}^{ih} |\nabla_h p_h|^2 dx, \tag{3.57}$$

where $u_i^N \approx u(x_i, T)$ is the approximated solution of direct problem (2.31) at point x_i and ∇_h is the numerical approximation of ∇. We thus state the discrete optimal control problem control problem as follows:

$$\text{Find } \bar{p}_h \in \mathcal{U}_{ad}^h \text{ such that } J_h(\bar{p}_h) = \min_{p_h \in \mathcal{U}_{ad}^h} J_h(p_h), \tag{3.58}$$

In order to determine a solution for the minimisation problem (3.58), we use the gradient descent method presented in the algorithm 1 with the gradient of the cost functional is given by

$$J_h'(p_h) = -\sum_{n=1}^{N} \int_{(n-1)\tau}^{n\tau} f(u_i^n) v_i^n d\tau - \beta \ \Delta_h p_h,$$

where u_i^n and v_i^n are the approximated solutions of the direct problem (2.31) and the adjoint problem (3.51) respectively, and Δ_h is the numerical approximation of Δ.

3.5.2 Numerical experiments

The numerical simulations presented in our work are in one-dimension case, and we set $T << 1$, $\Omega = (0, 1)$. The grid points on $[0, 1]$ and $[0, T]$ are both 51 while the resolution of the direct problem is by finite difference method.

We build the synthetic data g by fixing the unknown parameter p, selecting the nonlinear term f, and setting the initial conditions ϕ, then solving the direct problem (1), and extracting the measurement g. We add a random perturbation with a fixed amplitude, ε, to the data for the tests with noisy data, i.e.

$$g^{\varepsilon} = g(\varepsilon(2rand(size(g)) - 1) + 1).$$

To show the accuracy of the numerical solution for proposed method, we compute the approximate relative error denoted by

$$e_k = \frac{\|p_k(x) - p(x)\|_{L^2(\Omega}}{\|p(x)\|_{L^2(\Omega}}.$$

where $p_k(x)$ is the parameter reconstructed at the kth iteration and $p(x)$ is the exact solution. The approximate error in the observation data g and the residual at the kth iteration are gives as follows

$$E_k = \|u_k(x, T) - g\|_{L^2(\Omega)} \qquad E_k^{\varepsilon} = \|u_k(x, T) - g^{\varepsilon}\|_{L^2(\Omega)}.$$

The implementation also requires a stopping criteria. In our numerical experiments, it was difficult to find criteria which works well on both synthetic and measured test data. Therefore, in our case, the iterative process was interactively monitored and stopped when the change in the reconstruction obtained from an iteration step was no longer noticeable.

Along the reconstruction, we select the nonlinear term f that satisfies the assumptions $(\mathbf{H}_1)$ and $(\mathbf{H}_2)$ as defined below:

$$f(x) = e^x - 1.$$

Now, we present some examples in the one-dimensional case to verify the effectiveness for gradient descent algorithm, where first four examples are smooth and last two are non smooth. We have added these discontinuous examples to show the power and the limitations of the proposed method.

In this numerical experiment, we give the solution to the direct problem by $u(x, t) = t^2 \sin(\pi x) + \sin(\pi x)$, the order of the fractional derivative to be $\alpha = 0.5$ and the regularity

parameter. We identify the parameter p for $u(x,t)$ the following exact solution of problem (1), where the initial data $u_0 = \sin(\pi x)$. The final data $u(x,T)$ are obtained by solving the direct problem (1).

Exemple 3.1. *The first example is regular supposed as follows* $p(x) = \pi^2 \sin(\pi x)$.

Exemple 3.2. *The second example is synthetic, we assume* u_0 *supplied in the above and we estimate the unknown parameter* p *as follows*

$$p(x) = 2 + \sin(2\pi x).$$

The observation data g *are obtained by solving the direct problem* (1).

Exemple 3.3. *Take* $p(x) = e^{4x}$ *and the data* g *is obtained by solving the direct problem* (1) *by using the finite difference method as the exact input data.*

Exemple 3.4. *This example is synthetic, we consider* $\phi(x)$ *and* α *given in the above and we suppose the unknown parameter as follows*

$$p(x) = e^{-2x} \sin(5\pi x) + 1.5.$$

The data g *is obtained by solving the direct problem* (1) *by the finite difference method.*

Exemple 3.5. *In the fated example,we test a non-smooth one with a peak. Let the parameter given by the following expression*

$$p(x) = \begin{cases} x & 0 \leq x \leq \frac{1}{2} \\ 1 - x & \frac{1}{2} \leq x \leq 1 \end{cases}$$

The data g *is obtained by solving the direct problem* (1) *by a finite difference method as the exact input data.*

Exemple 3.6. *In the last example, we consider a discontinuous potential as follow:*

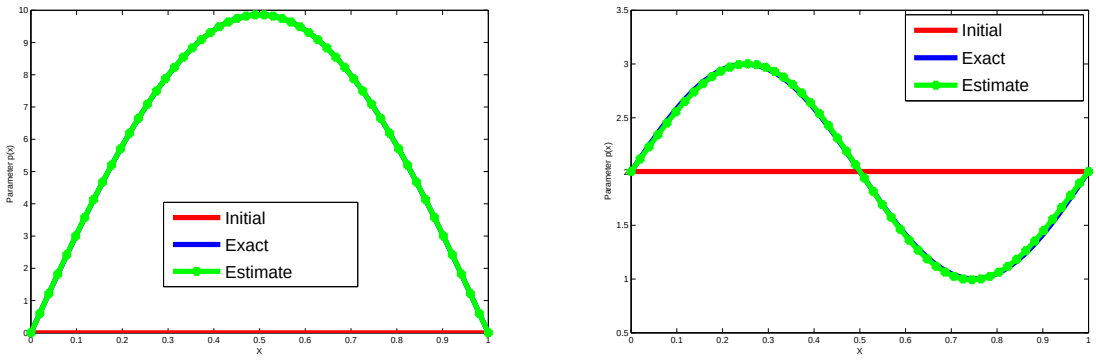
$$p(x) = \begin{cases} 0 & 0 \leq x \leq \frac{1}{3} \\ 1 & \frac{1}{3} < x \leq \frac{2}{3} \\ 0 & \frac{2}{3} < x \leq 1 \end{cases}.$$

The data g is obtained by solving the direct problem (1) by a finite difference method as the exact input data.

Now, we present the numerical results obtained for the above examples using the Gradient Descent method in the cases with and without noise on the data to show the stability of the problem as well as the effectiveness of the method. In the various examples of fractional derivative order α, the numerical errors for different noise levels are determined.

3.5.2.1 Identification results without noise

in this paragraph, we present the numerical results of the identification of the parameter p without noise for the examples above. We start with the regular examples like Example 3.1, Example 3.2, Example 3.3, Example 3.4 where the result presented in Figure 3.1 and 3.2. We observe that in these regular cases, the results obtained almost coincide with the exact parameter p. On the other hand, for complex Parameter such as Example 3.5 and Example 3.6, the method does not construct in an exact way the singularities of the parameter Figure 3.3. This is due to the type of functional and regularization used in our optimal control problem which are of type L^2 and to capture these singularities we must use a minimization of type L^1.



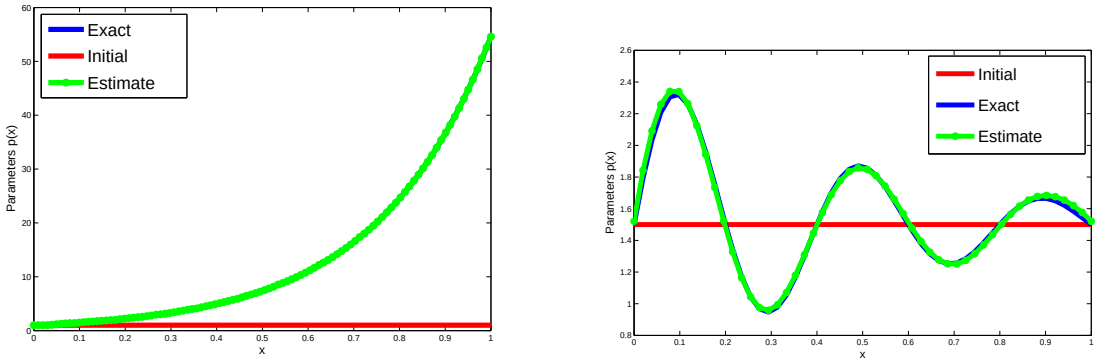
(a) Example 3.1 without noise.

(b) Example 3.2 without noise.

Figure 3.1: The numerical results of descent algorithm for example 3.1 and 3.2.

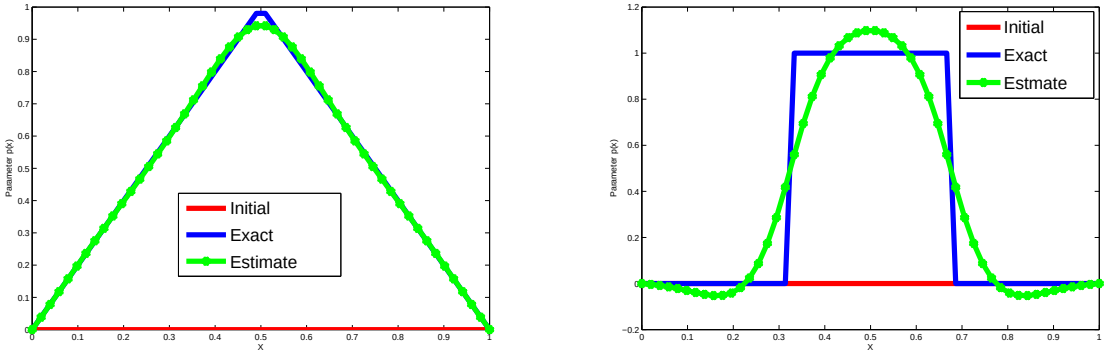
3.5.2.2 Identification results with noise

In this section, we will see the stability the proposed method, we present the numerical results obtained by noisy data. Figure 3.4 and Figure 3.5 shows the results obtained with noise for the regular cases and Figure 3.6 presents the results obtained with noise for the complex cases. The numerical tests with noise show that the proposed method is stable with respect to the noise, however, it cannot capture the singularity of the parameters.



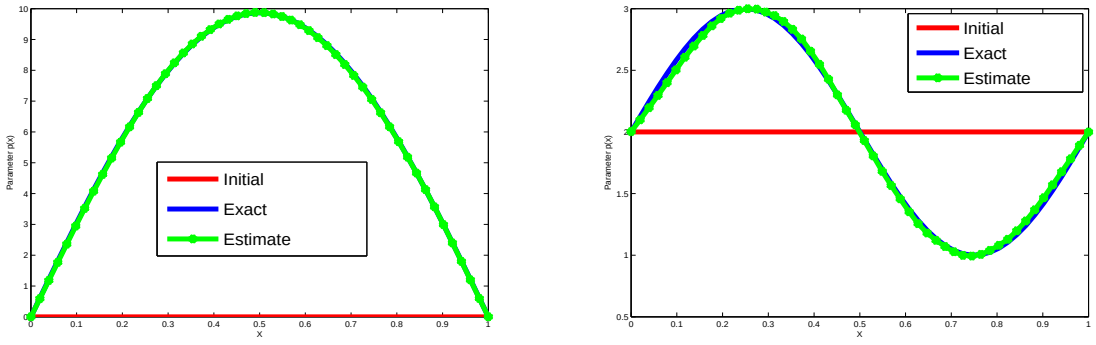
(a) Example 3.3 without noise. (b) Example 3.4 without noise.

Figure 3.2: The numerical results of descent algorithm for example 3.3 and 3.4.



(a) Example 3.5 without noise. (b) Example 3.6 without noise.

Figure 3.3: The numerical results of descent algorithm for examples 3.5 and 3.6.

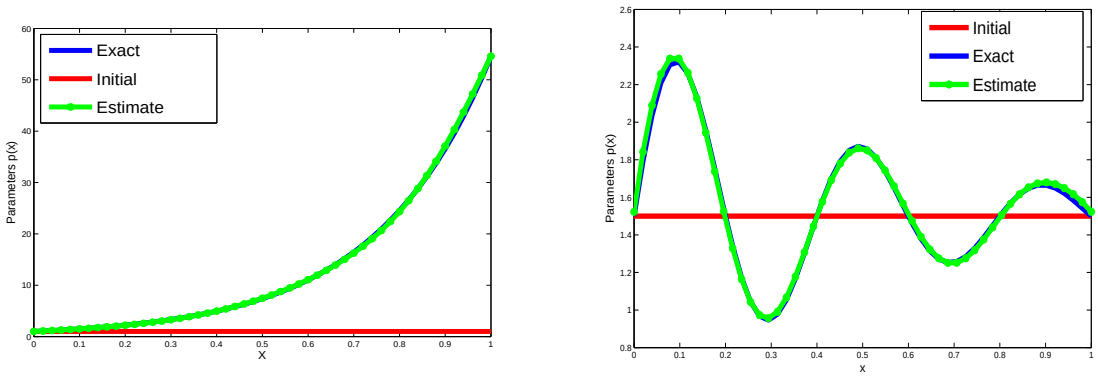


(a) Example 3.1 with 5% noise. (b) Example 3.2 with 5% noise.

Figure 3.4: The numerical results of descent algorithm with noise for example 3.1 and 3.2.

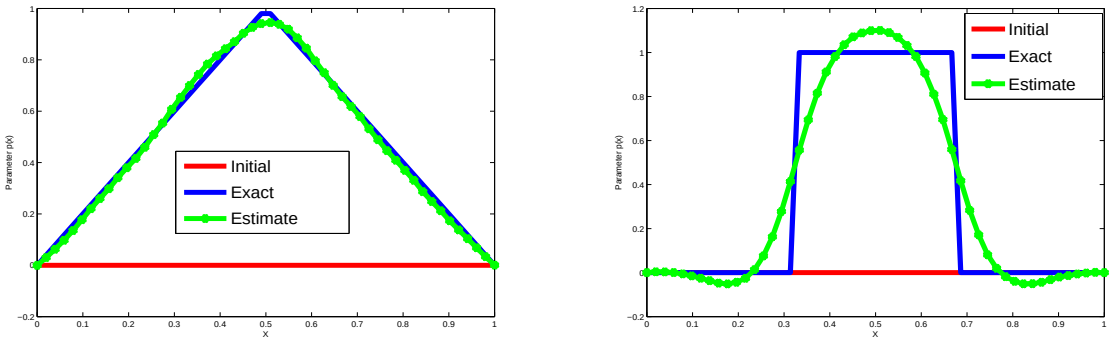
3.5.2.3 Relative error

In this paragraph, we show the results of the relative error e_k with respect to increasing the number of iterations. As well as the impact of the fractional order α. In Figure 3.7, we notice



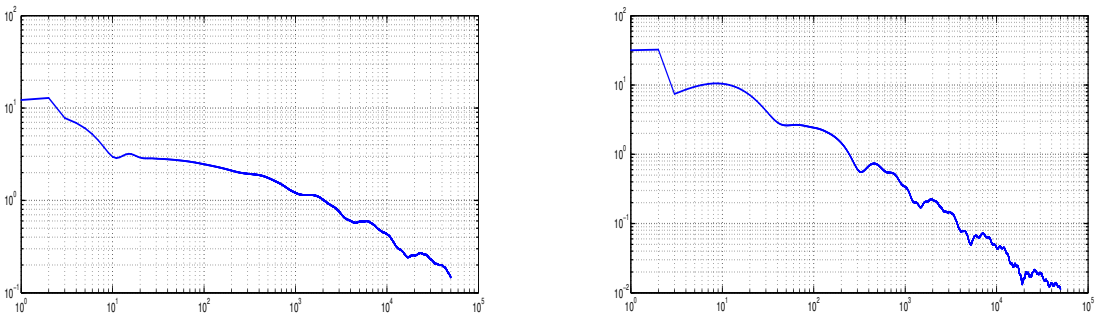
(a) Example 3.3 with 5% noise. (b) Example 3.4 with 5% noise.

Figure 3.5: The numerical results of descent algorithm with noise for example 3.3 and 3.4.



(a) Example 3.5 with 5% noise. (b) Example 3.6 with 5% noise.

Figure 3.6: The numerical results of descent algorithm with noise for examples 3.5 and 3.6.



(a) Error e_k for Example 3.1 (b) Error e_k for Example 3.5

Figure 3.7: Relative errors for Example 3.1 and 3.5.

that in both regular and singular cases, the relative error decreases in a fast way and, at a certain stage, becomes flat. But for the non-smooth case, a singular perturbation appears at a given

moment. In the Table 3.1, we present the numerical errors e_k of the Example 3.1 for different

Table 3.1: Numerical results for e_k with different ε for Example 3.1

ε	0	0.001	0.01	0.10	0.20
$\alpha = 0.3$	0.0313	0.0497	0.0621	0.0874	0.923
$\alpha = 0.7$	0.0236	0.0241	0.0305	0.516	0.743

noise values ε and different α. We can notice that the relative error e decreases when the noise level in the data decreases.

3.6 Conclusion

In this chapter, we have studied an inverse problem of identifying the parameter governed by a nonlinear time-fractional diffusion equation using the regularized optimal control method. The existence and uniqueness results of the solution of the direct problem are proved by means the Mittag-Leffler functions. Then, we reformulate our inverse problem to an optimal control problem, the existence of the minimizer is proved. The gradient descent algorithm associated with the adjoint problem is used to solve the control problem. We set the necessary conditions to show the stability of our inverse problem. In the end, we present some numerical tests to show the efficiency of the proposed algorithm

Chapter 4

ADMM for solving inverse potential problem

In this chapter, we treat the inverse problem of the identification the potential parameter p in problem (1) from new observation defined in the given Ω_1 subset of Ω:

$$u(x,t) = g(x,t) \text{ .a.e } x \in \Omega_1 \text{ and } t \in I,$$

where g is the noisy data and Ω_1 is subset of Ω.

One of the most commonly used methods to solve inverse problems is optimization method. To this end, we reformulate our inverse problem into the following Least-Squares optimal control problem.

$$J(\bar{p}) = \min_{p \in \mathcal{U}_{ad}} J(p) \tag{4.1}$$

where

$$J(p) = \frac{1}{2}\int_0^T \int_{\Omega_1} |u(x,t;p) - g(x,t)|^2 dxdt + \frac{\beta}{2}\int_\Omega |p|^2 dx \tag{4.2}$$

with $u(x,t;p)$ is the solution of the nonlinear time-fractional diffusion equation (1) and the admissible set $\mathcal{U}_{ad}$ is defined by

$$\mathcal{U}_{ad} = \Big\{ p \in L^2(\Omega) : 0 < c_1 \leq p \leq c_2 \ \ a.e\ \Omega \Big\}.$$

The constant $\beta > 0$ is a regularization parameter, c_1 and c_2 are positive constants.

In order to implement ADMM efficiently, we first establish the convexity of the optimal control problem using the solution estimate of the derived problem, and then, we propose an inexactness criterion that is independent of the mesh size of the involved discretization. Theoretically, we provide the convergence of the developed algorithm to a stationary point. Numerical examples in one-dimensional case are provided to show the effectiveness of this proposed method.

4.1 Convexity of the optimization problem

In this section, we show the convexity of the optimization problem in order to apply the ADMM method. The proof of the convexity is based on differentiability, which we obtained by Lipschitz continuity and the variational formulation. We cite in this context several works using the same approach [19, 4, 76] where they show the convexity of the object functional from the estimates of the first and higher derivative solutions. We now show the continuity of the observation operator and its differentiability with respect to the parameter p in order to establish the convexity of the cost functional.

We use $H^s(\Omega)$ and $H_0^s(\Omega)$ to denote the usual Sobolev spaces, $\forall s > 0$, whose norm are denoted by $\|.\|_{s,\Omega}$ (see [3]). We denote by the space $B^s(Q) = H^s(I; L^2(\Omega)) \cap L^2(I; H_0^1(\Omega))$, $\forall s > 0$, equipped by the norm :

$$\|v\|_{B^s(Q)} = \left(\|v\|^2_{H^s(I;L^2(\Omega))} + \|v\|^2_{L^2(I;H_0^1(\Omega))}\right)^{\frac{1}{2}}.$$

We start by giving the following Riemann Liouville weak formulation of the problem (1) which given as follow

$$\begin{cases} \text{Find} \quad u \in B^{\frac{\alpha}{2}}(Q) \\ \left({}^R D_t^{\frac{\alpha}{2}} u, {}^R D_T^{\frac{\alpha}{2}} v\right)_Q + (\nabla u, \nabla v)_Q + (p(x) f(u), v)_Q = \left(\dfrac{\phi(x) t^{-\alpha}}{\Gamma(1-\alpha)}, v\right)_Q \quad \forall v \in B^{\frac{\alpha}{2}}(Q). \end{cases} \tag{4.3}$$

4.1.1 Continuity of the observation operators

In order to prove the continuity, we define the following operator S from $\mathcal{U}_{ad}$ to $B^{\frac{\alpha}{2}}(Q)$: $p \in \mathcal{U}_{ad} \longmapsto S(p) = u(x,t;p) \in B^{\frac{\alpha}{2}}(Q)$, where u is the solution of variational problem (4.3). In the following proposition, we give the continuity of the operator S.

Proposition 4.1. *Under the assumption $(\boldsymbol{H}_2)$, the operator S is continuous in $\mathcal{U}_{ad}$ and we have :*

$$\|S(p_1) - S(p_2)\|_{B^{\frac{\alpha}{2}}(Q)} \leq \frac{c\sqrt{T}}{\cos(\alpha\pi/2)} \|p_1 - p_2\|_{L^2(\Omega)}. \qquad \forall (p_1, p_2) \in \mathcal{U}_{ad}^2 \tag{4.4}$$

Proof. Let $u_1 = S(p_1)$ and $u_2 = S(p_2)$ are solutions of (4.3) for p_1 and p_2 respectively. By taking $w = u_1 - u_2$, then w satisfies the following equation

$$\begin{cases} D_t^\alpha w(x,t) - \Delta w(x,t) + p_1(f(u_1) - f(u_2)) &= (p_2 - p_1) f(u_2), & (x,t) \in Q, \\ w(x,0) &= 0, & x \in \Omega, \\ w(x,t) &= 0, & x \in \partial\Omega \; t \in I. \end{cases}$$

Note that there exists $u_1 < \xi < u_2$ such that

$$f(u_1) - f(u_2) = f'(\xi)(u_1 - u_2) = f'(\xi)w,$$

what's given w satisfies the following equation

$$\begin{cases} D_t^\alpha w(x,t) - \Delta w(x,t) + p_1 f'(\xi) w &= (p_2 - p_1) f(u_2), \quad (x,t) \in Q, \\ w(x,0) &= 0, \quad x \in \Omega, \\ w(x,t) &= 0, \quad x \in \partial\Omega \quad t \in I, \end{cases} \tag{4.5}$$

we give its variational formulation as follows

$$\begin{cases} \left({}^R D_t^{\frac{\alpha}{2}} w, {}^R D_T^{\frac{\alpha}{2}} v \right)_Q + (\nabla w, \nabla v)_Q + (p_1(x) f'(\xi) w, v)_Q = ((p_2 - p_1) f(u_2), v)_Q, \\ \forall v \in B^{\frac{\alpha}{2}}(Q). \end{cases} \tag{4.6}$$

Set $v = w$ and using the result (3.46), we obtain

$$\cos(\alpha\pi/2) \ \|w\|^2_{B^{\frac{\alpha}{2}}(Q)} + (p_1 f'(\xi), w^2)_Q \leq (-(p_1 - p_2) f(u_2), w)_Q. \tag{4.7}$$

Based on the assumption $(\mathbf{H}_2)$, using the Theorem 2.1, the solution u belongs to $C(\bar{I}; \bar{H}^2(\Omega))$ and Cauchy-Schwartz inequality, we obtain

$$\|w\|^2_{B^{\frac{\alpha}{2}}(Q)} \leq \frac{c\sqrt{T}}{\cos(\alpha\pi/2)} \|p_1 - p_2\|_{L^2(\Omega)} \|w\|_{L^2(Q)}. \tag{4.8}$$

This completes the proof of the proposition. □

4.1.2 Differentiability of the operator S

Given that p in $\mathcal{U}_{ad}$, for all h sufficiently small such as $p + h \in \mathcal{U}_{ad}$ and $S(p + h) \in B^{\frac{\alpha}{2}}(Q)$ is well-defined. Let $w^1 = S(p + h) - S(p)$, then by definition of S, w^1 satisfies the following equation

$$\begin{cases} D_t^\alpha w^1(x,t) - \Delta w^1(x,t) + p[f(u(p+h) - f(u(p))] &= -h f(u(p+h)), \quad (x,t) \in Q, \\ w^1(x,0) &= 0, \quad x \in \Omega, \\ w^1(x,t) &= 0, \quad x \in \partial\Omega \ t \in I, \end{cases} \tag{4.9}$$

we give its variational formulation as follows: Find $w^1 \in B^{\frac{\alpha}{2}}(Q)$ such that

$$\begin{cases} \left({}^R D_t^{\frac{\alpha}{2}} w^1, {}^R D_T^{\frac{\alpha}{2}} v\right)_Q + (\nabla w^1, \nabla v)_Q + (p[f(u(p+h) - f(u(p))], v)_Q = (-hf(u(p+h)), v)_Q, \\ \qquad\qquad \forall v \in B^{\frac{\alpha}{2}}(Q). \end{cases} \tag{4.10}$$

Now, we can study the differentiability of S and deduce the form of $DS(p)$.

Proposition 4.2. *Under the assumption $(\boldsymbol{H}_1) - (\boldsymbol{H}_3)$. For any $p \in \mathcal{U}_{ad}$ the function $S(p)$ is differentiable with respect to p, if $u' = DS(p)h$ is the unique solution to the following sensitivity equation*

$$\begin{cases} D_t^\alpha u'(x,t) - \Delta u'(x,t) + pf'(u)u' &= -hf(u), & (x,t) \in Q, \\ u'(x,0) &= 0, & x \in \Omega, \\ u'(x,t) &= 0, & x \in \partial\Omega \quad t \in I, \end{cases} \tag{4.11}$$

its weak variational formulation is given by: Find $u' \in B^{\frac{\alpha}{2}}(Q)$ such that

$$\left({}^R D_t^{\frac{\alpha}{2}} u', {}^R D_T^{\frac{\alpha}{2}} v\right)_Q + (\nabla u', \nabla v)_Q + (pf'(u)u', v)_Q = (-hf(u), v)_Q\,; \ \forall v \in B^{\frac{\alpha}{2}}(Q). \tag{4.12}$$

with $u = S(p)$. Then we have

$$\|DS(p)\|_{B^{\frac{\alpha}{2}}(Q)} \leq \frac{c\sqrt{T}}{\cos(\alpha\pi/2)^2}. \tag{4.13}$$

Proof. Firstly, we combine (4.10) and (4.12) to get

$$\begin{aligned} \left({}^R D_t^{\frac{\alpha}{2}}(w^1 - u'), {}^R D_T^{\frac{\alpha}{2}} v\right)_Q + \left(\nabla(w^1 - u'), \nabla v\right)_Q + (p[f(u(p+h) - f(u(p))) - f'(u)u'], v)_Q \\ = (-h[f(u(p+h)) - f(u(p))], v)_Q \end{aligned} \tag{4.14}$$

for all $v \in B^{\frac{\alpha}{2}}(Q)$. The convexity of the function f implies that

$$f'(u(p))\,[u(p+h) - u(p)] \leq f(u(p+h)) - f(u(p)). \tag{4.15}$$

Thus

$$p\left[f(u(p+h)) - f(u(p)) - f'(u(p))u^1\right] \geq pf'(u(p))(w^1 - u'). \tag{4.16}$$

Then, the equation (4.14) will be as follows

$$({}^{R}D_{t}^{\frac{\alpha}{2}}(w^1-u'),{}^{R}D_{T}^{\frac{\alpha}{2}}v)_Q+(\nabla(w^1-u'),\nabla v)_Q+(pf'(u(p))(w^1-u'),v)_Q \\ \leq \left(-h\left(f(u(p+h))-f(u(p))\right),v\right)_Q. \tag{4.17}$$

By taking $v=w^1-u'$, we obtain

$$\left({}^{R}D_{t}^{\frac{\alpha}{2}}(w^1-u'),{}^{R}D_{T}^{\frac{\alpha}{2}}(w^1-u')\right)_Q+\left(\nabla(w^1-u'),\nabla(w^1-u')\right)_Q \\ +\left(pf'(u(p))(w^1-u'),(w^1-u')\right)_Q\leq\left(-h\left(f(u(p+h))-f(u(p))\right),(w^1-u')\right)_Q.$$

Based on the assumptions $(\mathbf{H}_1)-(\mathbf{H}_3)$, we applying ellipticity (3.46) and Cauchy-Schwartz inequality, we obtain

$$\cos(\alpha\pi/2)\ \|w^1-u'\|^2_{B^{\frac{\alpha}{2}}(Q)}+(pf'(u(p)),(w^1-u')^2)_Q \\ \leq c\|h\|_{L^2(\Omega)}\|u(p+h)-u(p)\|_{L^2(I,L^\infty(\Omega))}\|w^1-u'\|_{L^2(Q)}.$$

Using the fact f' is positive and the continuous embedding $L^2(\bar{I},H^1(\Omega))$ of into $L^2(\bar{I},L^\infty(\Omega))$, we obtain

$$\cos(\alpha\pi/2)\ \|w^1-u'\|_{B^{\frac{\alpha}{2}}(Q)}\leq c\|h\|_{L^2(\Omega)}\|w^1\|_{B^{\frac{\alpha}{2}}(Q)}, \tag{4.18}$$

Appealing to Proposition 4.1 again for (4.9), we have

$$\|w^1\|_{B^{\frac{\alpha}{2}}(Q)}=\|S(p+h)-S(p)\|_{B^{\frac{\alpha}{2}}(Q)}\leq\frac{c\sqrt{T}}{\cos(\alpha\pi/2)^2}\|h\|_{L^2(\Omega)}. \tag{4.19}$$

From (4.18) and (4.19), we have

$$\|w^1-u'\|_{B^{\frac{\alpha}{2}}(Q)}\leq\frac{c\sqrt{T}}{\cos(\alpha\pi/2)^2}\|h\|^2_{L^2(\Omega)}, \tag{4.20}$$

which gives

$$\frac{\|S(p+h)-S(p)-u'\|_{B^{\frac{\alpha}{2}}(Q)}}{\|h\|_{L^2(\Omega)}}\leq\frac{c\sqrt{T}}{\cos(\alpha\pi/2)^2}\|h\|_{L^2(\Omega)}. \tag{4.21}$$

Thus, S is differentiable at p and $DS(p)h=u'$.

Also, let $v=u'$ in (4.12), and use ellipticity (3.46), the hypotheses $\mathbf{H}$ as well as Cauchy-Schwartz

inequality again to obtain

$$\cos(\alpha\pi/2)\ \|u'\|^2_{B^{\frac{\alpha}{2}}(Q)} + (pf'(\xi),(u')^2)_Q \leq c\|h\|_{L^2(\Omega)}\|u\|_{L^2(I,L^\infty(\Omega))}\|u'\|_{L^2(Q)}.$$

Then

$$\|u'\|_{B^{\frac{\alpha}{2}}(Q)} \leq \frac{c\sqrt{T}}{\cos(\alpha\pi/2)}\|h\|_{L^2(\Omega)}.$$

which shows that $DS(p)$ is uniformly bounded. □

4.1.3 Convexity of the cost function

Now, we can prove that the functional cost is convex and that its minimizer in $\mathcal{U}_{ad}$ is unique. For any $p \in \mathcal{U}_{ad}$, and h perturbation, we have

$$J'(p)h = \int_0^T \int_{\Omega_1} (u(p) - g)\, u' dxdt + \beta \int_\Omega phdx, \tag{4.22}$$

where $u' = DS(p)(h)$ is the unique solution to sensitive problem (4.11).
The result of convexity of the functional cost is established in the following proposition

Proposition 4.3. *For any $\beta > 0$ and under assumption $(\boldsymbol{H}_4)$, the functional $J(p)$ is strictly convex.*

Proof. Let p and $p + h$ in $\mathcal{U}_{ad}$, then

$$\begin{aligned} J(p+h) - J(p) - J'(p)h = \\ &\frac{1}{2}\int_0^T \int_{\Omega_1} |u(p+h) - g|^2 dxdt + \frac{\beta}{2}\int_\Omega |p+h|^2 dx - \frac{1}{2}\int_0^T \int_{\Omega_1} |u(p) - g|^2 dxdt \\ &- \frac{\beta}{2}\int_\Omega |p|^2 dx - \int_0^T \int_{\Omega_1} (u(p) - g)\, u' dxdt - \beta \int_\Omega phdx. \end{aligned}$$

By direct calculation, we obtain

$$\begin{aligned} J(p+h) - J(p) - J'(p)h =& \frac{1}{2}\int_0^T \int_{\Omega_1} (u(p+h) + u(p) - 2g)\,(u(p+h) - u(p))\, dxdt \\ &- \int_0^T \int_{\Omega_1} (u(p) - g)\, u' dxdt + \frac{\beta}{2}\|h\|^2_{L^2(\Omega)}. \end{aligned}$$

It is easy to see that the last term is non-negative, it remains to show that the difference between

the first terms is positive, then

$$J(p+h)-J(p)-J'(p)h = \underbrace{\frac{1}{2}\int_0^T\int_{\Omega_1}(u(p)-g)\left[u(p+h)-u(p)-u'\right]dxdt}_{I_1}$$
$$+\underbrace{\frac{1}{2}\int_0^T\int_{\Omega_1}(u(p+h)-g)\left[u(p+h)-u(p)-u'\right]dxdt}_{I_2}$$
$$+\underbrace{\frac{1}{2}\int_0^T\int_{\Omega_1}(u(p+h)-u(p))\,u'dxdt}_{I_3}+\frac{\beta}{2}\|h\|^2_{L^2(\Omega)}.$$

Applying the Cauchy-Schwartz inequality, we have

$$|I_1|\leq\frac{1}{2}\|u(p)-g\|_{L^2(Q)}\|u(p+h)-u(p)-u'\|_{L^2(Q)}$$

and

$$|I_2|\leq\frac{1}{2}\|u(p+h)-g\|_{L^2(Q)}\|u(p+h)-u(p)-u'\|_{L^2(Q)}.$$

Using the Theorem 2.1, the solution u belongs to $C(\bar{I};\bar{H}^2(\Omega))$ and from the inequality (4.21), we obtain

$$|I_1|\leq\frac{cT}{2\cos(\alpha\pi/2)^2}\|h\|^2_{L^2(\Omega)} \tag{4.23}$$

and

$$|I_2|\leq\frac{cT}{2\cos(\alpha\pi/2)^2}\|h\|^2_{L^2(\Omega)}. \tag{4.24}$$

In the same way, we have

$$|I_3|\leq\frac{1}{2}\|u(p+h)-u(p)\|_{L^2(Q)}\|u'\|_{L^2(Q)}.$$

From the Proposition 4.1 and the Proposition 5.1, we get

$$|I_3|\leq\frac{cT}{2\cos(\alpha\pi/2)^3}\|h\|^2_{L^2(\Omega)}. \tag{4.25}$$

Combining (4.23), (4.24) and (4.25) one can easily get

$$J(p+h)-J(p)-J'(p)h>\left(\frac{\beta}{2}-\frac{cT}{2\cos(\alpha\pi/2)^3}\right)\|h\|^2_{L^2(\Omega)}.$$

Choosing T such that

$$T<\frac{\cos(\alpha\pi/2)^3\beta}{c}.$$

Thus $J(p)$ is strictly convex. □

4.2 Implementation of ADMM method

In this section, we will try to apply the ADMM directly to the optimal control problem, constrained by a time-fractional diffusion equation (1)-(4.1). We first reformulate (1)-(4.1) as a convex optimization problem with a separable structure. To that end, an auxiliary variable z satisfying $p = z$ is introduced, and the problem (1)-(4.1) can be written as follows:

$$\begin{cases} \min\limits_{(p,z)\in L^2(\Omega)\times L^2(\Omega)} \tilde{J}(p,u) + I_{\mathcal{U}_{ad}}(z), \\ s.t \qquad p = z, \end{cases} \tag{4.26}$$

where $I_{\mathcal{U}_{ad}}(z) = \begin{cases} 0 \text{ if } & z \in \mathcal{U}_{ad} \\ \infty & \text{otherwise} \end{cases}$ is the indicator function of the admissible set $\mathcal{U}_{ad}$ which is a non-empty convex closed set of $L^2(\Omega)$ and

$$\tilde{J}(p,u) = \int_0^T \frac{\eta}{2} \int_{\Omega_1} |u(x,t;p) - g(x)|^2 dxdt + \frac{1}{2}\int_\Omega |p|^2 dx, \text{ with } \eta = \frac{1}{\beta}.$$

In the following this work, The notations $(.,.)$ and $\|.\|$ denotes the canonical inner product and the operator norm on $L^2(\Omega)$,respectively.

Let us define the augmented Lagrangian function of (4.26) as

$$L_r(p,z;\lambda) = \tilde{J}(p,u) + I_{\mathcal{U}_{ad}}(z) - (\lambda, p-z) + \frac{r}{2}\|p-z\|^2, \tag{4.27}$$

where $\lambda \in L^2(\Omega)$ is the Lagrange multiplier associated the constraint $p = z$, and $r > 0$ is a penalty parameter. Throughout this work, to simplify the discussion, we consider that the penalty parameter is fixed. Then, by applying the ADMM to (4.27), we immediately obtain the algorithm 2.

Now, we develop the resulting subproblems (a) and (b) in Algorithm 2, and discuss the implementation details of Algorithm 2. For simplicity of notation we denote the spaces $L^2(\Omega)$ by U. Let us recall that the augmented Lagrangian functional defined in (4.27) is given by

$$L_r(p,z;\lambda) = \tilde{J}(p,u) + I_{\mathcal{U}_{ad}}(z) - (\lambda, p-z) + \frac{r}{2}\|p-z\|^2,$$

Algorithm 2 An ADMM algorithm for the solution of problem (1)-(4.1).

- **Input** The order of the fractional derivative $\alpha \in (0,1)$, $T > 0$, β, r , the initial condition ϕ and the functions $f, g \in \mathbf{L}^2(\mathbf{\Omega})$.
- **Initialisation** Set initial values $\{z^0, \lambda^0\} \in L^2(\Omega) \times L^2(\Omega)$.
- **Repeat**

  For $k > 0$, $\{z^k, \lambda^k\} \longrightarrow p^{k+1} \longrightarrow z^{k+1} \longrightarrow \lambda^{k+1}$ via the solution and computation of

$$\begin{cases} p^{k+1} &= \arg \min\limits_{p \in L^2(\Omega)} L_r(p, z^k; \lambda^k), & (a) \\ z^{k+1} &= \arg \min\limits_{z \in L^2(\Omega)} L_r(p^{k+1}, z; \lambda^k), & (b) \\ \lambda^{k+1} &= \lambda^k - r(p^{k+1} - z^{k+1}), & (c) \end{cases}$$

  end

and the resulting subproblems (a) and (b) are

$$p^{k+1} = \arg \min_{p \in L^2(\Omega)} L_r(p, z^k; \lambda^k),$$
$$z^{k+1} = \arg \min_{z \in L^2(\Omega)} L_r(p^{k+1}, z; \lambda^k),$$

respectively.

4.2.1 Solution of the z-subproblem

In this subsection, we focus on the resolution of z-subproblem (b) by minimizing the augmented Lagrangian functional L_r with respect to z. Since, $\tilde{J}$ is independent of z, the z-subproblem (b) becomes

$$z^{k+1} = \underset{z \in L^2(\Omega)}{argmin} \left\{ I_{\mathcal{U}_{ad}}(z) - (\lambda^k, p^{k+1} - z) + \frac{r}{2}\|p^{k+1} - z\|^2 \right\}. \tag{4.28}$$

To solve this subproblem, we apply the proximal operator defined by

$$\mathbf{prox}_{\tau \mathrm{G}}(f) = arg \min_{z} \frac{\|z - f\|^2}{2\tau} + G(z).$$

Through this definition, it appears that there is a congruence between the subproblem and the proximal operator definition with $\tau = \frac{1}{r}$, $G(z) = I_{\mathcal{U}_{ad}}(z) - (\lambda^k, p^{k+1} - z)$ and $f = p^{k+1}$.
By calculation direct, we have

$$z^{k+1} = P_{\mathcal{U}_{ad}}\left(p^{k+1} - \frac{\lambda^k}{r}\right), \tag{4.29}$$

where $P_{\mathcal{U}_{ad}}(.)$ denotes the projection into the admissible set $\mathcal{U}_{ad}$, then

$$z^{k+1} = \max\left(c_1, \min\left(p^{k+1} - \frac{\lambda^k}{r}, c_2\right)\right). \tag{4.30}$$

4.2.2 Solution of the p-subproblem

In this subsection, we look for the solution of the first p-subproblem (a) of the Algorithm 2 where we minimize the augmented Lagrangian functional L_r with respect to p. Notice that

$$L_r(p, z^k; \lambda^k) = \tilde{J}(p, u) - (\lambda^k, p - z^k) + \frac{r}{2}\|p - z^k\|^2,$$

which implies that the problem (a) is equivalent to the following unconstrained optimal control problem:

$$\begin{cases} p^{k+1} \in U, \\ j_k(p^{k+1}) \leq j_k(p), \quad \forall p \in U, \end{cases} \tag{4.31}$$

where

$$j_k(p) = \tilde{J}(p, u) - (\lambda^k, p - z^k) + \frac{r}{2}\|p - z^k\|^2.$$

Let $Dj_k(p)$ be the first order differential of $j_k(.)$ at p, then the unique solution of problem (4.31) is characterized by

$$Dj_k(p^{k+1}) = 0.$$

To compute $Dj_k(p)$, we consider δp as a perturbation of $p \in U$, then

$$Dj_k(p)(\delta p) = \nabla \tilde{J}(p, u)(\delta p) - (\lambda^k, \delta p) + r(p - z^k, \delta p), \tag{4.32}$$

with

$$\nabla \tilde{J}(p) = -\eta \int_0^T f(u) v dt + p, \tag{4.33}$$

where v is the solution of the adjoint problem associated to u which is a solution of the state problem and they are obtained from the successive solution of the following two fractional equations.

$$\begin{cases} D_t^\alpha u(x,t) - u_{xx}(x,t) + p(x) f(u) &= 0, & (x,t) \in Q, \\ u(x,0) &= \phi(x), & x \in \Omega, \\ u(0,t) = u(1,t) &= 0, & t \in I. \end{cases} \tag{4.34}$$

and

$$\begin{cases} D_T^{\alpha} v(x,t) - v_{xx}(x,t) + p(x) f'(u) v &= (u-g)\chi_{\Omega_1}(x)\ , & (x,t) \in Q, \\ v(x,T) &= 0, & x \in \Omega, \\ v(0,t) = v(1,t) &= 0, & t \in I. \end{cases} \tag{4.35}$$

Combining (4.32) and (4.33), we obtain

$$Dj_k(p) = -\eta \int_0^T f(u) v dt + (1+r)p - \lambda^k - r z^k. \tag{4.36}$$

Theorem 4.1. *Let p^{k+1} be the unique solution of the subproblem (a). Then, it is characterized by the following optimality condition*

$$Dj_k(p^{k+1}) = -\eta \int_0^T f(u^{k+1}) v^{k+1} dt + (1+r)p^{k+1} - \lambda^k - r z^k = 0, \tag{4.37}$$

where u^{k+1} is a solution of the state problem (4.34) for p^{k+1} and v^{k+1} is a solution of the adjoint problem (4.35) associated to u^{k+1}.

4.2.3 Inexactness criterion

In this subsection, we first take a closer look at the solutions of subproblems $(a)-(c)$, and then propose an inexactness criterion for solving subproblem (a) iteratively. Thus, we propose an inexact version of ADMM with two-layer nested iterations. It is worth noting that the optimality condition of the p-subproblem (a) can be characterized as (4.37). Since the p-subproblem (a) is strongly convex, the above necessary condition is also sufficient. As a result, if $p \in U$ satisfies $Dj_k(p) = 0$, then p is the unique solution to the p-subproblem (a). To propose an inexactness criterion, we define $e_k(p)$ as

$$e_k(p) = -\eta \int_0^T f(u) v dt + (1+r)p - \lambda^k - r z^k, \tag{4.38}$$

with v is the solution of adjoint problem associated with p.

It is easy to notice that $e_k(p) = Dj_k(p)$ and p is the solution of the p-subproblem (a) at the $(k+1)$-th iteration if and only if $e_k(p^{k+1}) = 0$. Therefore, we can use $e_k(p)$ as the residual for the p-subproblem (a). With the help of $e_k(p)$, we propose the following inexactness criterion. For a given arbitrary constant σ satisfying

$$0 < \sigma < \frac{\sqrt{2}}{\sqrt{2}+\sqrt{r}} \in (0,1), \tag{4.39}$$

we compute p^{k+1} such that

$$\|e_k(p^{k+1})\| \leq \sigma \|e_k(p^k)\|. \tag{4.40}$$

To solve p-subproblem (a), we propose a numerical algorithm based on the DG method. The choice of this method is justified by the fast convergence towards an optimal solution. It is widely used as an optimization algorithm for classification problems. The DG Algorithm is given as follows.

Algorithm 3 Descent Gradient for p-subproblem (a).

Begin

k is an iteration of ADMM algorithm

Input: η, r,λ^k, ς_0 ,$p_0^k = p^k$, $u_0^k = u^k$, $v_0^k = v^k$.

Compute $J_0^k = -\eta \int_0^T f(u_0^k) v_0^k dt + (1+r)p_0^k - \lambda^k - rz^k$, $e_k(p_0^k) = J_0^k$.

Repeat

Compute u_m^k by solving the state problem (4.34).

Compute v_m^k by solving the adjoint problem (4.35).

Compute $J_m^k = -\eta \int_0^T f(u_m^k) v_m^k dt + (1+r)p_m^k - \lambda^k - rz^k$.

Update $p_{m+1}^k = p_m^k - \varsigma_m J_m^k$ with ς_m is computed by Armijo line search.

$m = m + 1$

Until $\|e_k(p_{m+1}^k)\| \leq \sigma \|e_k(p_m^k)\|$

Output $p^{k+1} = p_m^k$.

End

4.3 Convergence analysis

In this section, we prove the convergence of Algorithm 2. First of all, notice that if the augmented Lagrangian functional L_r in (4.27) has a saddle-point $(p, z; \lambda) \in (U \times U) \times U$, then it follows from, e.g., [16, 17], that the sequence $\{p^k\}$ generated by Algorithm 2 converges strongly in U to the solution of the problem (4.26). Although there are many works in the literature studying the convergence of the ADMM [18, 64]. We hence present the complete analysis for the convergence of Algorithm 2. We start from some known preliminaries. We denote $w \in W := U \times U \times U$, $y \in V := U \times U$ and the function $F(w)$ as follows:

$$w = \begin{pmatrix} p \\ z \\ \lambda \end{pmatrix}, \quad y = \begin{pmatrix} z \\ \lambda \end{pmatrix}, \text{ and } F(w) = \begin{pmatrix} D\tilde{J}(p) - \lambda \\ \lambda \\ p - z \end{pmatrix}, \tag{4.41}$$

where $D\tilde{J}(p)$ is the first-order derivative of $\tilde{J}$, we also define the norm

$$\|y\|_{\Lambda} = \sqrt{(y, \Lambda y)} = \sqrt{r\|z\|^2 + \frac{1}{r}\|\lambda\|^2}, \qquad \forall y \in V, \tag{4.42}$$

which is induced by the matrix operator

$$\Lambda = \begin{pmatrix} rI & 0 \\ 0 & \frac{1}{r}I \end{pmatrix},$$

with these notations, it is easy to see that the problem (4.26) can be characterized as the following variational inequality

$$VI(W, F) : \begin{cases} \text{Find } w^* = (\ p^*, z^*, \lambda^*\ )^T \in W^* \text{ such that} \\ I_{\mathcal{U}_{ad}}(z) - I_{\mathcal{U}_{ad}}(z^*) + (w - w^*, F(w^*)) \geq 0, \quad \forall w \in W. \end{cases} \tag{4.43}$$

We denote by W^* the solution set of the variational inequality (4.43).

Since $\tilde{J}$ is strongly convex, then

$$\|p - q\|^2 \leq \left(p - q, D\tilde{J}(p) - D\tilde{J}(q)\right), \qquad \forall p, q \in U. \tag{4.44}$$

4.3.1 Optimality conditions

In the Algorithm 2, the p-subproblem (a) is solved inexactly subject to the inexactness criterion (4.38). But, the z-subproblem (b) and the λ-subproblem (c) are be solved exactly. Now, we consider $w^{k+1} = (p^{k+1}, z^{k+1}, \lambda^{k+1})^T$ a sequence generated by the Algorithm 2, the first order optimality conditions can be expressed as follows:

$$\begin{cases} D_p L_r(p^{k+1}, z^k, \lambda^k) = e_k(p^{k+1}), \\ I_{\mathcal{U}_{ad}}(z) - I_{\mathcal{U}_{ad}}(z^{k+1}) + (z - z^{k+1}, \lambda^k - r(p^{k+1} - z^{k+1})) \geq 0, \quad \forall z \in U \\ \lambda^{k+1} = \lambda^k - r(p^{k+1} - z^{k+1}), \end{cases} \tag{4.45}$$

where $D_p L_r(p^{k+1}, z^k, \lambda^k)$ is the first-order partial derivative of $L_r(p, z, \lambda)$ with respect to p at $(p^{k+1}, z^k, \lambda^k)^T$. To show the convergence of the Algorithm 2, the analysis of the residual $e_k(p^{k+1})$ is of great interest. It follows from (4.38) and (4.40) that

$$\|e_k(p^{k+1})\| \leq \sigma\|e_k(p^k)\| \leq \sigma\|e_{k-1}(p^k)\| + \sigma\|r(z^{k-1} - z^k) + \lambda^{k-1} - \lambda^k\|.$$

Using the definition of $\|.\|$, we have

$$\begin{aligned}\|e_k(p^{k+1})\| \leq & \sigma\|e_{k-1}(p^k)\| \\ & + \sigma\left(\|r(z^{k-1}-z^k)\|^2 + \|\lambda^{k-1}-\lambda^k\|^2 + 2r(z^{k-1}-z^k, \lambda^{k-1}-\lambda^k)\right)^{\frac{1}{2}}.\end{aligned} \tag{4.46}$$

Furthermore, it results from (4.45) that

$$I_{\mathcal{U}_{ad}}(z^k) - I_{\mathcal{U}_{ad}}(z^{k+1}) + (z^k - z^{k+1}, \lambda^k - r(p^{k+1}-z^{k+1})) \geq 0, \quad \forall z^k \in U,$$

and

$$I_{\mathcal{U}_{ad}}(z^{k+1}) - I_{\mathcal{U}_{ad}}(z^k) + (z^{k+1} - z^k, \lambda^{k-1} - r(p^k - z^k)) \geq 0, \quad \forall z^{k+1} \in U.$$

We make the sum of two above expressions and using $\lambda^{k+1} = \lambda^k - r(p^{k+1} - z^{k+1})$, we obtain

$$(z^{k+1} - z^k, \lambda^{k+1} - \lambda^k) \leq 0. \tag{4.47}$$

From (4.46) and (4.47), we have

$$\begin{aligned}\|e_k(p^{k+1})\| &\leq \sigma\|e_{k-1}(p^k)\| + \sigma\|r(z^{k-1}-z^k) + \lambda^{k-1} - \lambda^k\| \\ &\leq \sigma\|e_{k-1}(p^k)\| + \sigma\left(\|r(z^{k-1}-z^k)\|^2 + \|\lambda^{k-1}-\lambda^k\|^2\right)^{\frac{1}{2}} \\ &= \sigma\|e_{k-1}(p^k)\| + \sigma.\sqrt{r}\|v^k - v^{k-1}\|_\Lambda.\end{aligned}$$

In the same way we finally obtain

$$\|e_k(p^{k+1})\| \leq \sum_{i=0}^{k-1} \sigma^{k-i}\sqrt{r}\|v^i - v^{i+1}\|_\Lambda + \sigma^k\|e_0(p^1)\|. \tag{4.48}$$

Moreover, recall that the parameter σ controlling the accuracy in (4.40) is restricted by the condition (4.39). Hence,

$$0 < \frac{r\sigma^2}{2(1-\sigma)} = \left(\frac{\sigma}{2(1-\sigma)}\right)\left(\frac{r\sigma}{2(1-\sigma)}\right) < 1,$$

and obviously there exists a $\mu > 0$ such that

$$0 < \left(\frac{\mu\sigma}{2(1-\sigma)}\right) < 1, \quad \text{and } 0 < \left(\frac{r\sigma}{2\mu(1-\sigma)}\right) < 1. \tag{4.49}$$

This relationship will be often used in the coming analysis.

4.3.2 Convergence

Now we prove the convergence of the sequence generated by the Algorithm 2. To simplify the notation, let us introduce an auxiliary variable

$$\bar{w}^k = \begin{pmatrix} \bar{p}^k \\ \bar{z}^k \\ \bar{\lambda}^k \end{pmatrix} = \begin{pmatrix} p^{k+1} \\ z^{k+1} \\ \lambda^k - r(p^{k+1} - z^k) \end{pmatrix}. \tag{4.50}$$

This notation is just to simplify the notation in our analysis. Next, we will show a results that will be useful for the continuation of our study.

The first result is to analyse that the point $\bar{w}^k$ defined in (4.50) is different from the solution w^* of (4.43) and how to quantify this difference by the iterations generated by the Algorithm 2.

Lemma 4.1. *Let $\{w^k\} = \{(p^k, z^k, \lambda^k)^T\}$ be the sequence generated by Algorithm 2 and $\{\bar{w}^k\} = \{(\bar{p}^k, \bar{z}^k, \bar{\lambda}^k)^T\}$ defined by (4.50). Then , for all $w \in W$, we have*

$$\begin{aligned} I_{\mathcal{U}_{ad}}(\bar{z}^k) - I_{\mathcal{U}_{ad}}(z) + (\bar{w}^k - w, F(\bar{w}^k)) \leq & \frac{1}{2}\left(\|v^k - v\|_\Lambda^2 - \|v^{k+1} - v\|_\Lambda^2 - \|v^k - v^{k+1}\|_\Lambda^2\right) \\ & + \left(p^{k+1} - p, D_p L_r(p^{k+1}, z^k, \lambda^k)\right). \end{aligned} \tag{4.51}$$

Proof. First, we construct the expression of $D_p L_r(p^{k+1}, z^k, \lambda^k)$ as

$$D_p L_r(p^{k+1}, z^k, \lambda^k) = D\tilde{J}(p^{k+1}) - (\lambda^k - r(p^{k+1} - z^k)) = D\tilde{J}(p^{k+1}) - \bar{\lambda}^k. \tag{4.52}$$

Based on the above equality, for all $w \in W$, we obtain

$$\begin{aligned} R :=& I_{\mathcal{U}_{ad}}(\bar{z}^k) - I_{\mathcal{U}_{ad}}(z) + (\bar{w}^k - w, F(\bar{w}^k)) \\ =& I_{\mathcal{U}_{ad}}(\bar{z}^k) - I_{\mathcal{U}_{ad}}(z) + (\bar{p}^k - p, D\tilde{J}(p^{k+1}) - \bar{\lambda}^k) + (\bar{z}^k - z, \bar{\lambda}^k) + (\bar{\lambda}^k - \lambda, \bar{p}^k - \bar{z}^k). \end{aligned}$$

By direct calculation, we get

$$\begin{aligned} R =& I_{\mathcal{U}_{ad}}(z^{k+1}) - I_{\mathcal{U}_{ad}}(z) + (p^{k+1} - p, D_p L_r(p^{k+1}, z^k, \lambda^k)) \\ & + (z^{k+1} - z, \lambda^k - r(p^{k+1} - z^{k+1})) + r(z^{k+1} - z, z^k - z^{k+1}) \\ & + \frac{1}{r}(\bar{\lambda}^k - \lambda, \lambda^k - \lambda^{k+1}). \end{aligned}$$

From (4.45), we have

$$\begin{aligned} R \leq & (p^{k+1} - p, D_p L_r(p^{k+1}, z^k, \lambda^k)) + r(z^{k+1} - z, z^k - z^{k+1}) \\ & + \frac{1}{r}(\bar{\lambda}^k - \lambda^{k+1}, \lambda^k - \lambda^{k+1}) + \frac{1}{r}(\lambda^{k+1} - \lambda, \lambda^k - \lambda^{k+1}). \end{aligned}$$

By applying the elementary equation

$$(c - a, b - c) = \frac{1}{2}\left(-\|a - c\|^2 + \|a - b\|^2 - \|b - c\|^2\right), \tag{4.53}$$

and using the expressions for $\bar{\lambda}^k$ and λ^{k+1}, we have

$$\begin{aligned} R \leq & (p^{k+1} - p, D_p L_r(p^{k+1}, z^k, \lambda^k)) + \frac{r}{2}\left(-\|z - z^{k+1}\|^2 + \|z - z^k\|^2 - \|z^k - z^{k+1}\|^2\right) \\ & + (z^k - z^{k+1}, \lambda^k - \lambda^{k+1}) + \frac{1}{2r}\left(-\|\lambda - \lambda^{k+1}\|^2 + \|\lambda - \lambda^k\|^2 - \|\lambda^k - \lambda^{k+1}\|^2\right). \end{aligned}$$

From (4.47), we obtain

$$\begin{aligned} R \leq & (p^{k+1} - p, D_p L_r(p^{k+1}, z^k, \lambda^k)) + \frac{r}{2}\left(-\|z - z^{k+1}\|^2 + \|z - z^k\|^2 - \|z^k - z^{k+1}\|^2\right) \\ & \frac{1}{2r}\left(-\|\lambda - \lambda^{k+1}\|^2 + \|\lambda - \lambda^k\|^2 - \|\lambda^k - \lambda^{k+1}\|^2\right). \end{aligned}$$

Using the definition of Λ-norm in (4.42), The last result can be rewritten as

$$R \leq \left(p^{k+1} - p, D_p L_r(p^{k+1}, z^k, \lambda^k)\right) + \frac{1}{2}\left(\|v^k - v\|_\Lambda^2 - \|v - v^{k+1}\|_\Lambda^2 - \|v^k - v^{k+1}\|_\Lambda^2\right).$$

The proof is complete. □

The difference between the variational inequality reformulation (4.43) and the inequality (4.51) reflects the difference between the solution point w^* and $\bar{w}^k$. The first three terms on the right-hand side of (4.51) are quadratic, and they are easy to manipulate on different indicators using algebraic operations, but it is not clear how the last cross term can be controlled for the final purpose of proving the convergence of the sequence $\bar{w}^k$. We investigate this term in detail and demonstrate that the sum of these crossover terms over K iterations can also be bounded by some quadratic terms. This result is summarized in the following Lemma

Lemma 4.2. *Let $\{w^k\} = \{(p^k, z^k, \lambda^k)^T\}$ be the sequence generated by Algorithm 2. For any*

integer $K > 0$, $\forall p \in U$, *we have*

$$\begin{aligned}\sum_{k=1}^{K}\left(p^{k+1}-p, D_pL_r(p^{k+1},z^k,\lambda^k)\right) \leq & \frac{\mu}{2}\sum_{k=1}^{K}\frac{\sigma}{1-\sigma}\|p^{k+1}-p\|^2 \\ & + \frac{1}{2\mu}\sum_{k=1}^{K}\frac{\sigma}{1-\sigma}r\|v^i-v^{i+1}\|_\Lambda^2 \\ & + \frac{1}{2\mu}\frac{\sigma}{1-\sigma}\left[\|e_0(p^1)\|+\sqrt{r}\|v^0-v^1\|_\Lambda\right]^2.\end{aligned} \tag{4.54}$$

Proof. We start the calculation by using the Cauchy-Schwartz and $D_pL_r(p^{k+1}) = e_k(p^{k+1})$, we have

$$\sum_{k=1}^{K}\left(p^{k+1}-p, D_pL_r(p^{k+1},z^k,\lambda^k)\right) \leq \sum_{k=1}^{K}\|p^{k+1}-p\|\|e_k(p^{k+1})\|.$$

Using (4.55) for any $k \geq 1$, we obtain

$$\|e_k(p^{k+1})\| \leq \sum_{i=0}^{k-1}\sigma^{k-i}\sqrt{r}\|v^i-v^{i+1}\|_\Lambda + \sigma^k\|e_0(p^1)\|,$$

then we have

$$\begin{aligned}\sum_{k=1}^{K}\|p^{k+1}-p\|\|e_k(p^{k+1})\| \leq & \sum_{k=1}^{K}\sum_{i=1}^{k-1}\sigma^{k-i}\sqrt{r}\|p^{k+1}-p\|\|v^i-v^{i+1}\|_\Lambda \\ & + \sum_{k=1}^{K}\sigma^k\|p^{k+1}-p\|\left[\|e_0(p^1)\|+\sqrt{r}\|v^0-v^1\|_\Lambda\right].\end{aligned}$$

According to Young's inequality, we get

$$\begin{aligned}&\sum_{k=1}^{K}\|p^{k+1}-p\|\|e_k(p^{k+1})\| \\ &\quad\leq \frac{\mu}{2}\sum_{k=1}^{K}\sum_{i=1}^{k-1}\sigma^{k-i}\|p^{k+1}-p\|^2 + \frac{1}{2\mu}\sum_{k=1}^{K}\sum_{i=1}^{k-1}\sigma^{k-i}r\|v^i-v^{i+1}\|_\Lambda^2 \\ &\qquad + \frac{\mu}{2}\sum_{k=1}^{K}\sigma^k\|p^{k+1}-p\|^2 + \frac{1}{2\mu}\sum_{k=1}^{K}\sigma^k\left[\|e_0(p^1)\|+\sqrt{r}\|v^0-v^1\|_\Lambda\right]^2 \\ &\quad\leq \frac{\mu}{2}\sum_{k=1}^{K}\sum_{i=0}^{k-1}\sigma^{k-i}\|p^{k+1}-p\|^2 + \frac{1}{2\mu}\sum_{k=1}^{K}\sum_{i=1}^{k-1}\sigma^{k-i}r\|v^i-v^{i+1}\|_\Lambda^2 \\ &\qquad + \frac{1}{2\mu}\sum_{k=1}^{K}\sigma^k\left[\|e_0(p^1)\|+\sqrt{r}\|v^0-v^1\|_\Lambda\right]^2.\end{aligned}$$

For any $\mu > 0$ satisfying (4.49), we obtain

$$\begin{aligned}\sum_{k=1}^{K} \|p^{k+1} - p\| \|e_k(p^{k+1})\| \leq & \frac{\mu}{2} \sum_{k=1}^{K} \frac{\sigma - \sigma^{k+1}}{1-\sigma} \|p^{k+1} - p\|^2 \\ & + \frac{1}{2\mu} \sum_{i=1}^{K-1} \frac{\sigma - \sigma^{K-i+1}}{1-\sigma} r \|v^i - v^{i+1}\|_\Lambda^2 \\ & + \frac{1}{2\mu} \frac{\sigma - \sigma^{K+1}}{1-\sigma} \left[\|e_0(p^1)\| + \sqrt{r} \|v^0 - v^1\|_\Lambda \right]^2 \\ \leq & \frac{\mu}{2} \sum_{k=1}^{K} \frac{\sigma}{1-\sigma} \|p^{k+1} - p\|^2 + \frac{1}{2\mu} \sum_{i=1}^{K-1} \frac{\sigma}{1-\sigma} r \|v^i - v^{i+1}\|_\Lambda^2 \\ & + \frac{1}{2\mu} \frac{\sigma}{1-\sigma} \left[\|e_0(p^1)\| + \sqrt{r} \|v^0 - v^1\|_\Lambda \right]^2 .\end{aligned}$$

The proof is complete. □

The convergence of the proposed in Algorithm 2 is established in the following theorem

Theorem 4.2. *Let $\{w^*\} = \{(p^*, z^*, \lambda^*)^T\}$ be the solution of the variational inequality (4.43) and $\{w^k\} = \{(p^k, z^k, \lambda^k)^T\}$ be the sequence generated by the Algorithm 2. Then, we have the following assertions*

1. $\|e_k(p^{k+1})\| \longrightarrow 0$, $\|z^k - z^{k+1}\| \longrightarrow 0$ *and* $\|p^{k+1} - z^{k+1}\| \longrightarrow 0$,

2. $p^k \longrightarrow p^*$, $z^k \longrightarrow z^*$ *and* $\lambda^k \longrightarrow \lambda^*$ *strongly in* U.

Proof. First, recall the definition of $F(w)$ in (4.41). We have that

$$\begin{aligned}\left(w - \bar{w}^k, F(w) - F(\bar{w}^k)\right) = & \left(p - \bar{p}^k, D\tilde{J}(p) - D\tilde{J}(\bar{p}^k) - \lambda + \bar{\lambda}^k\right) + \left(z - \bar{z}^k, \lambda - \bar{\lambda}^k\right) \\ & + \left(\lambda - \bar{\lambda}^k, p - \bar{p}^k - z + \bar{z}^k\right) \\ = & \left(p - \bar{p}^k, D\tilde{J}(p) - D\tilde{J}(\bar{p}^k)\right).\end{aligned}$$

From the inequality (4.44), we get

$$\left(w - \bar{w}^k, F(w) - F(\bar{w}^k)\right) = \left(p - \bar{p}^k, D\tilde{J}(p) - D\tilde{J}(\bar{p}^k)\right) \geq \|p - \bar{p}^k\|^2 = \|p - p^{k+1}\|^2. \quad (4.55)$$

Then, using the results (4.51) established in Lemma 4.1, we obtain

$$\begin{aligned}
RR :=& \sum_{k=1}^{K} \left\{ I_{\mathcal{U}_{ad}}(\bar{z}^k) - I_{\mathcal{U}_{ad}}(z) + (\bar{w}^k - w, F(w)) \right\} \\
=& \sum_{k=1}^{K} \left\{ I_{\mathcal{U}_{ad}}(\bar{z}^k) - I_{\mathcal{U}_{ad}}(z) + (\bar{w}^k - w, F(\bar{w}^k)) + (\bar{w}^k - w, F(w) - F(\bar{w}^k)) \right\} \\
\leq& \frac{1}{2} \left(\|v^1 - v\|_\Lambda^2 - \|v^{K+1} - v\|_\Lambda^2 \right) - \sum_{k=1}^{K} \|v^k - v^{k+1}\|_\Lambda^2 \\
&+ \sum_{k=1}^{K} \left\{ \left(p^{k+1} - p, D_p L_r(p^{k+1}, z^k, \lambda^k) \right) - (w - \bar{w}^k, F(w) - F(\bar{w}^k)) \right\}.
\end{aligned}$$

From the result (4.55), we have

$$\begin{aligned}
RR \leq& \frac{1}{2} \left(\|v^1 - v\|_\Lambda^2 - \|v^{K+1} - v\|_\Lambda^2 \right) - \frac{1}{2} \sum_{k=1}^{K} \|v^k - v^{k+1}\|_\Lambda^2 \\
&+ \sum_{k=1}^{K} \left\{ \left(p^{k+1} - p, D_p L_r(p^{k+1}, z^k, \lambda^k) \right) \right\} - \sum_{k=1}^{K} \|p^{k+1} - p\|^2.
\end{aligned}$$

According to the result (4.54) established in Lemma 4.2, we get

$$\begin{aligned}
RR \leq& \frac{1}{2} \left(\|v^1 - v\|_\Lambda^2 - \|v^{K+1} - v\|_\Lambda^2 \right) \\
&+ \sum_{k=1}^{K-1} \left(\frac{r}{2\mu} \frac{\sigma}{1-\sigma} - \frac{1}{2} \right) \|v^k - v^{k+1}\|_\Lambda^2 - \frac{1}{2} \|v^K - v^{K+1}\|_\Lambda^2 \\
&+ \sum_{k=1}^{K} \left(\frac{\mu}{2} \frac{\sigma}{1-\sigma} - 1 \right) \|p^{k+1} - p\|^2 + \frac{1}{2\mu} \frac{\sigma}{1-\sigma} \left[\|e_0(p^1)\| + \sqrt{r} \|v^0 - v^1\|_\Lambda \right]^2.
\end{aligned}$$

Setting $\{w^*\} = \{(p^*, z^*, \lambda^*)^T\}$ as the solution of the variational inequality (4.43), we have

$$I_{\mathcal{U}_{ad}}(\bar{z}^k) - I_{\mathcal{U}_{ad}}(z^*) + (\bar{w}^k - w^*, F(w^*)) \geq 0, \quad \forall \bar{w}^k \in W, \text{ for all } k \geq 1.$$

Taking $w = w^*$, together with the above property, for any integer $K > 1$, we obtain

$$\begin{aligned}
&\sum_{k=1}^{K-1} \left(\frac{1}{2} - \frac{r}{2\mu} \frac{\sigma}{1-\sigma} \right) \|v^k - v^{k+1}\|_\Lambda^2 + \sum_{k=1}^{K} \left(1 - \frac{\mu}{2} \frac{\sigma}{1-\sigma} \right) \|p^{k+1} - p^*\|^2 \\
&\leq \frac{1}{2} \left(\|v^1 - v^*\|_\Lambda^2 - \|v^{K+1} - v^*\|_\Lambda^2 \right) - \frac{1}{2} \|v^K - v^{K+1}\|^2 \\
&+ \frac{1}{2\mu} \frac{\sigma}{1-\sigma} \left[\|e_0(p^1)\| + \sqrt{r} \|v^0 - v^1\|_\Lambda \right]^2.
\end{aligned} \tag{4.56}$$

According to the results obtained in (4.49), we have

$$1-\left(\frac{\mu\sigma}{2(1-\sigma)}\right)>0, \quad \text{and } 1-\left(\frac{r\sigma}{2\mu(1-\sigma)}\right)>0.$$

Then, from the inequality (4.56), we conclude that

$$\|p^{k+1}-p^*\|\longrightarrow 0, \text{ and } \|v^k-v^{k+1}\|_\Lambda\longrightarrow 0 \text{ as } k\longrightarrow+\infty. \tag{4.57}$$

By definition of the norm $\|.\|_\Lambda$, we conclude that

$$\|z^{k+1}-z^k\|\longrightarrow 0, \text{ and } \|\lambda^{k+1}-\lambda^k\|\longrightarrow 0 \text{ as } k\longrightarrow+\infty. \tag{4.58}$$

Then, from the definition of λ^{k+1} in (4.45), we have $\|p^{k+1}-z^{k+1}\|=\frac{1}{r}\|\lambda^{k+1}-\lambda^k\|$, which implies that

$$\|p^{k+1}-z^{k+1}\|\longrightarrow 0 \text{ as } k\longrightarrow+\infty. \tag{4.59}$$

Furthermore, for any $\varepsilon>0$, there exists a k_0 such that for all $k\geq k_0$, we have

$$\|v^k-v^{k+1}\|_\Lambda^2<\varepsilon \text{ and } \sigma^k<\varepsilon.$$

Then, for all $k\geq k_0$, it follows from (4.55) that

$$\begin{aligned}
\|e_k(p^{k+1})\| &\leq \sum_{i=0}^{k_0-1}\sigma^{k-i}\sqrt{r}\|v^i-v^{i+1}\|_\Lambda+\sum_{i=k_0}^{k-1}\sigma^{k-i}\sqrt{r}\|v^i-v^{i+1}\|_\Lambda+\sigma^k\|e_0(p^1)\| \\
&\leq \left(\sqrt{r}\max_{0\leq i\leq k_0-1}\|v^i-v^{i+1}\|_\Lambda\sum_{i=0}^{k_0-1}\sigma^{k-k_0-i}\right)\sigma^{k_0}+\sigma^k\|e_0(p^1)\| \\
&+\left(\sqrt{r}\max_{k_0\leq i\leq k-1}\|v^i-v^{i+1}\|_\Lambda\sum_{i=k_0}^{k-1}\sigma^{k-i}\right) \\
&\leq \varepsilon.\left[\sqrt{r}\max_{0\leq i\leq k_0-1}\|v^i-v^{i+1}\|_\Lambda\sum_{i=0}^{k_0-1}\sigma^{k-k_0-i}+\sqrt{r}\sum_{i=k_0}^{k-1}\sigma^{k-i}+\|e_0(p^1)\|\right],
\end{aligned}$$

which means that

$$\|e_k(p^{k+1})\|\longrightarrow 0.$$

From (4.57), we know that $p^k\longrightarrow p^*$ strongly in U. In addition, from (4.59), we have $\|p^{k+1}-z^{k+1}\|\longrightarrow 0$, and $z^k\longrightarrow z^*$ strongly in U. From (4.43), it is easy to verify that $\lambda^*=D\tilde{J}(p^*)$.

On the other hand, one has:

$$\lambda^k = D\tilde{J}(p^{k+1}) + r(p^{k+1} - z^k) - e_k(p^{k+1}).$$

Then

$$\begin{aligned}\lambda^k - \lambda^* &= D\tilde{J}(p^{k+1}) - D\tilde{J}(p^*) + r(p^{k+1} - z^k) - e_k(p^{k+1}) \\ &= D\tilde{J}(p^{k+1}) - D\tilde{J}(p^*) + r(p^{k+1} - p^k) + r(p^k - z^k) - e_k(p^{k+1}).\end{aligned}$$

Recalls that $p^k \longrightarrow p^*$ strongly in U, $p^k \longrightarrow z^k$, $e_k(p^{k+1}) \longrightarrow 0$ and $D\tilde{J}$ is Lipschitz continuous, we have

$$\lambda^k \longrightarrow \lambda^* \text{ strongly in } U.$$

Hence the result of Theorem 4.2. □

4.4 Numerical results

In this section, we first give the implementation of the Algorithm 2 and the numerical tests of some examples in one dimension.

4.4.1 ADMM-DG algorithm

In this subsection, we discuss how to perform the proposed method in order to specify that the Agorithm 2 is concrete algorithm for the optimal control problem (4.1). Indeed, the subproblems $(a)-(b)$ are typical unconstrained optimal control problems and various numerical methods from the literature can be applied. In our study, we always propose the descent-gradient method to solve the p-subproblem (a) iteratively, its details are presented in Algorithm 3. We compute the z-subproblem (b) using the projection operator (4.30).

Now, with these discussions, the Algorithm 2 can be specified as a nested iterative ADMM-DG scheme for the optimal control problem (4.1). In the following, the instructions are given in the Algorithm 4.

4.4.2 Numerical experiments

In this subsection, we present the numerical results for four examples to show the effectiveness of the ADMM method. The numerical simulations presented are in one-dimension case, and we

Algorithm 4 An ADMM-DG nested iterative scheme for the optimal control problem (4.1).

Begin
Input: The data $T < 1$, $0 < \alpha < 1$, $r > 0$, $\beta > 0$, ϕ, the nonlinear term f, and $0 < \sigma < \frac{\sqrt{2}}{\sqrt{2}+\sqrt{r}}$.
Initialisation p^0, z^0, λ^0,
for $k > 0$ **do**
Compute p^{k+1} by the DG method in Algorithm 3
Compute z^{k+1} by (4.30)
Update the Lagrange multiplier

$$\lambda^{k+1} = \lambda^k - r(p^{k+1} - z^{k+1})$$

$k = k + 1$
end for
End

set $T < 1$, $\Omega = (0, 1)$. The grid points on $[0, 1]$ and $[0, T]$ are both 51 while the resolution of the direct problem is by finite difference method (see [48, 77]).

We build the synthetic data g by fixing the unknown parameter p, selecting the nonlinear term f, and setting the initial conditions ϕ, then solving the direct problem (1), and extracting the measurement g. We add a random perturbation with a fixed amplitude, δ, to the data for the tests with noisy data, i.e.

$$g^\delta = g(\delta(2rand(size(g)) - 1) + 1).$$

To show the accuracy of numerical solutions, we compute the relative error denoted by

$$E_k = \frac{\|p^k - p\|_{L^2(\Omega)}}{\|p\|_{L^2(\Omega)}}, \tag{4.60}$$

where p^k is the potential term reconstructed at the k-th iteration, and p is the exact solution of the inverse problem.

Along the reconstruction, we select the nonlinear term f that satisfies the assumptions $(\mathbf{H}_1)$ and $(\mathbf{H}_2)$ as defined below:

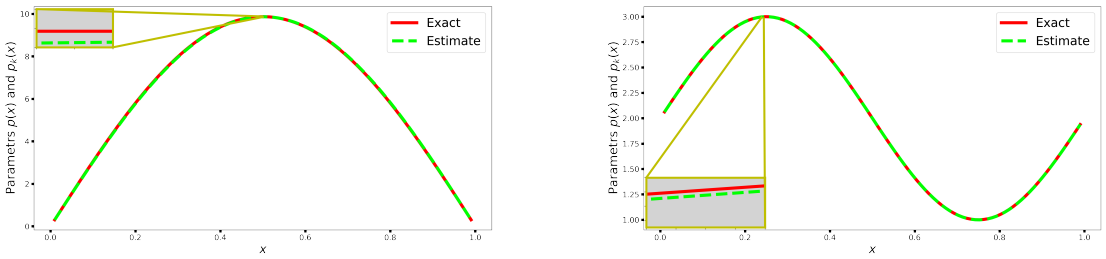
$$f(x) = e^x - 1.$$

According to (4.39), we choose $\sigma = 0.99\frac{\sqrt{2}}{\sqrt{2}+\sqrt{r}}$ for the constant σ in the inexactness criterion (4.40), because larger values of σ indicate that the criterion is looser and thus less computation is required for solving the subproblems. In our computation, we always set $T = 0.5$, $\beta = 10^{-6}$ and $\max^{iter} = 50$. Keep in this section the examples that we defined in the chapter 3. The

numerical results for these examples by using the discrepancy principle for various noise levels are shown in the below figures.

4.4.2.1 Identification results without noise

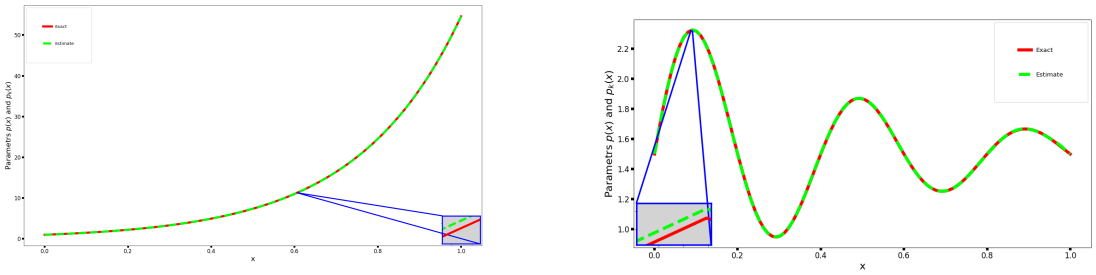
We now present the results of identifying the p-parameter for the previous four examples. We start with the regular Examples like 3.1, 3.2, 3.3 and 3.4 where the results are presented in Figure 4.1. We observe that in these regular cases, the results obtained almost coincide with the exact parameter p. On the other hand, for a complex parameter like examples 3.5 and 3.6, the method does not build exactly the singularities of the parameter but it is approach very well more than those obtained by the usual methods like (gradient method and its variant) by using only a functional of type L^2, which is an advantage. We think, it is due to the projection appearing in the method and the Lagrangian multiple. We have to add that the p chosen in the examples 3.5 and 3.6 their treatment requires TV regularization (see [19, 22]). Figure 4.3 shows the end result.



(a) Example 3.1 without noise

(b) Example 3.2 without noise

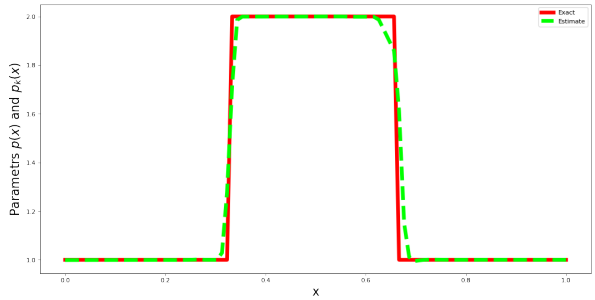
Figure 4.1: The numerical results of the ADMM algorithm for Examples 3.1 and 3.2 without noise.



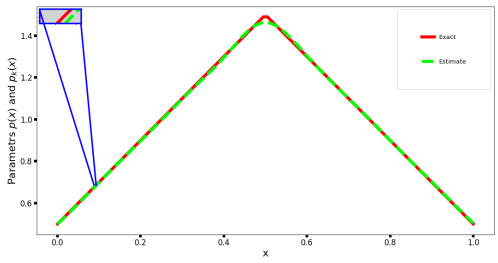
(a) Example 3.3 without noise

(b) Example 3.4 without noise

Figure 4.2: The numerical results of the ADMM algorithm for Examples 3.3 and 3.4 without noise.



(a) Example 3.5 without noise

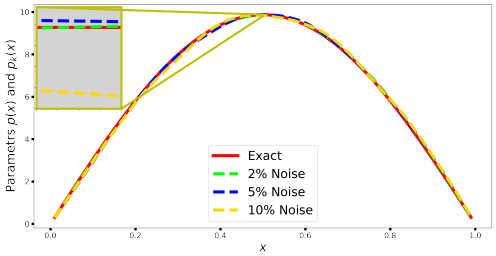


(b) Example 3.6 without noise

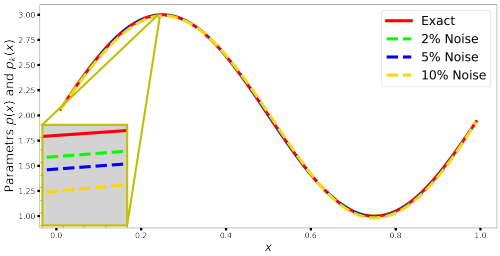
Figure 4.3: The numerical results of the ADMM algorithm for Examples 3.3 and 3.4 without noise.

4.4.2.2 Identification results with noise

In this subsection, we present the numerical results obtained by noisy data. Figure 4.4 and Figure 4.5 shows the results obtained with noise for the regular case, and Figure 4.6 for complex cases. The numerical tests with noise show that the ADMM method is stable with respect to

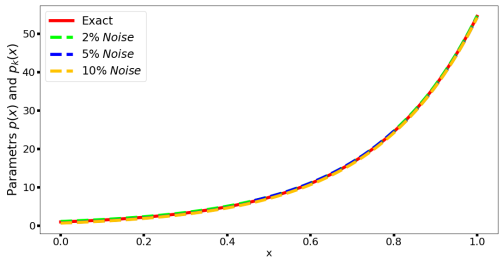


(a) Example 3.1 with noise

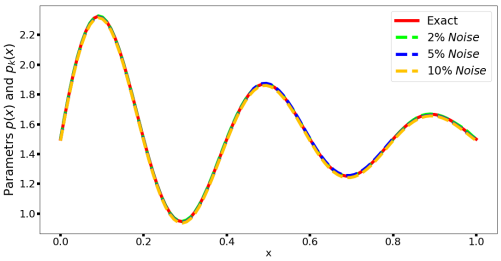


(b) Example 3.2 with noise

Figure 4.4: The numerical results of the ADMM algorithm for Examples 3.1 and 3.2 with noise.

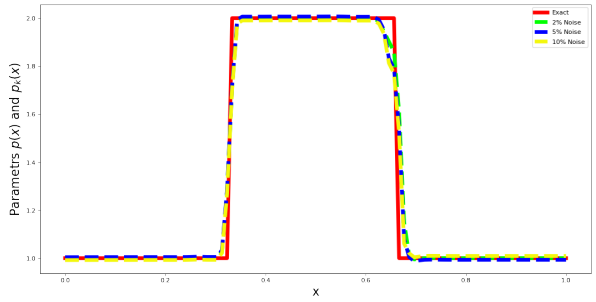


(a) Example 3.3 with noise

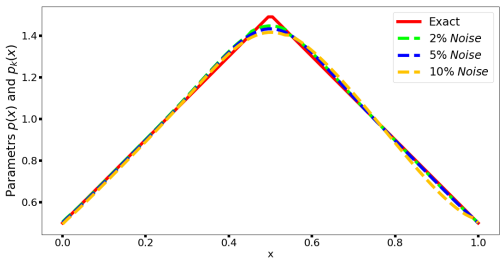


(b) Example 3.4 with noise

Figure 4.5: The numerical results of the ADMM algorithm for Examples 3.3 and 3.4 with noise.



(a) Example 3.5 with noise



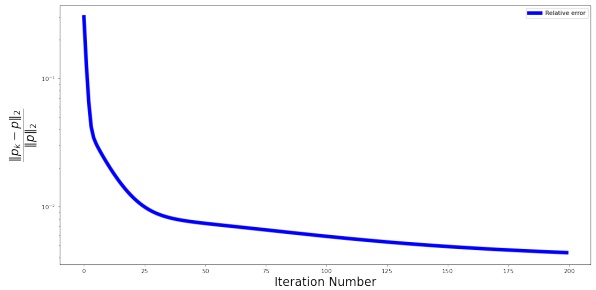
(b) Example 3.6 with noise

Figure 4.6: The numerical results of the ADMM algorithm for Exemples 3.5 and 3.6 with noise.

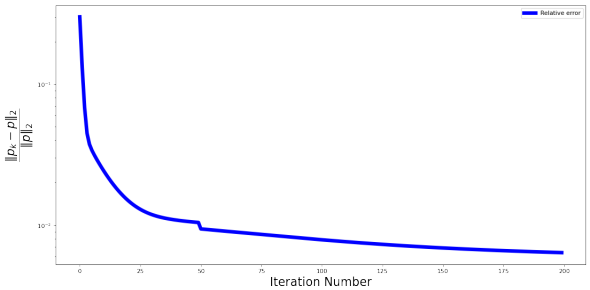
the noise. However, it cannot capture the singularity of the parameter, but it approaches it.

4.4.2.3 Relative error

In this paragraph, we show the results of the relative error e_k with respect to increasing the number of iterations. As well as the impact of the fractional order α. In Figure 4.7, we notice



(a) Relative error for Example 3.2



(b) Relative error for Example 3.4

Figure 4.7: The relative error of p for Exemples 3.2 and 3.4.

that in both regular and singular cases, the relative error decreases in a fast way and, at a certain stage, becomes flat. But for the non-smooth case, a singular perturbation appears at a given moment. In the Table 4.1, we present the numerical errors e_k of the Example 3.4 for different

Table 4.1: Numerical results for e_k with different ε for Example 3.4

ε	0	0.001	0.01	0.10	0.20
$\alpha = 0.3$	$3.13.10^{-2}$	$4.97.10^{-2}$	$6.21.10^{-2}$	$8.03.10^{-2}$	$9.13.10^{-2}$
$\alpha = 0.5$	$3.14.10^{-2}$	$4.97.10^{-2}$	$6.21.10^{-2}$	$8.74.10^{-2}$	$9.23.10^{-2}$
$\alpha = 0.7$	$2.36.10^{-2}$	$2.41.10^{-2}$	$4.05.10^{-2}$	$6.18.10^{-2}$	$7.43.10^{-2}$
$\alpha = 0.9$	$2.39.10^{-2}$	$2.43.10^{-2}$	$3.92.10^{-2}$	$4.91.10^{-2}$	$6.94.10^{-2}$

noise values ε and different α. We can notice that the relative error e_k decreases when the noise level in the data decreases.

4.5 Conclusion

In this chapter, we investigate an inverse space-dependent parameter problem for a time-fractional diffusion equation. The existence, uniqueness and regularity of the solution for the direct problem are all proved by the Mettag-Lafler function. To study our inverse parameter problem, we reformulate it into a least squares optimization problem. Then, we establish the existence of the optimal solution and we prove the convexity of the considered cost function by using its first derivative. Numerically, we adapt a recent method in the literature known as the alternating direction multiplier method (ADMM) to our optimal control problem. We then establish the convergence of the developed algorithm to a stationary point. Finally, some numerical examples are provided to show that the ADMM method is effective and stable for solving the inverse parameter problem.

Chapter 5

Artificial neural networks for identifying the potential parameter

---

In this chapter, we propose an artificial neural network approach to identify the parameter p in non-linear subdiffusion model (1) from additional data (2). Instead of determining the parameter in the fractional diffusion model, we design the model itself, in the form of an artificial neural network. The relevance of this approach relies on the approximation capability of neural networks. Then, the potential p is replaced by in the form of a neural network, namely a function whose expression is given by

$$p_L(x) = \begin{cases} W_2(\rho(W_1(x))) & \text{if } L = 2, \\ W_L(\rho(W_{L-1}(\rho(...\rho(W_1(x)))...) & \text{if } L \geq 3, \end{cases} \tag{5.1}$$

where ρ is called the *activation function*. The simplest neural network is a single layer of perception that mathematically performs a linear operation followed by a nonlinear composition applied to its input space

$$p_1(x) = A_1 x + b_1$$
$$p_2(x) = A_2(\rho(p_1(x))) + b_2.$$

Deep neural networks are multiple layers stacked together within some architecture

$$p_l = A_l(\rho(p_{l-1}(x))) + b_l,$$

where l indicates the index of a layer, A_l is a weight matrix of dimension $n_{l-l} \times n_l$ ($n_l :=$ *number of neurons in the layer* l), b_l is a bias vector of dimension n_l and p_l is a vector with n_l components.

The artificial neural networks approach for inverse potential problem (1)-(3.1) can be for-

mulated as follows:

$$\begin{cases} \min\limits_{W} J(p_L) = \dfrac{1}{2}\displaystyle\int_{\Omega} |u(x,T;p_L) - g(x)|^2 dx \\ \text{subject to } u \text{ solution of the problem (1) paraeterized with } W. \end{cases} \tag{5.2}$$

We show that the optimal control problem parametrized by ANN (5.2) is well-posed, Also, we demonstrate the existence of the solution to the control problem and provide a mathematical analysis and the derivation of optimal conditions.

5.1 Properties of neural networks

This section defines neural networks, introduces the basic configuration as well as other notations, and lists the basic elements needed for the neural network constructions considered throughout the work. There is a plethora of neural network architectures and activation functions in the literature. We restrict ourselves here to the activation function assumed to be globally Lipschitz and consider the following general network architecture with L layers.

$$p_L(x, W_1, W_2, ..., W_L) = \begin{cases} W_1(x) & \text{if } L = 1, \\ W_2(\rho(W_1(x))) & \text{if } L = 2, \\ W_L(\rho(W_{L-1}(\rho(...\rho(W_1(x)))...) & \text{if } L \geq 3, \end{cases} \tag{5.3}$$

where, for $l \in \{1, 2, ..., L\}$, $W_l : \mathbb{R}^{n_{l-1}} \to \mathbb{R}^{n_l}$; $W_l(x) = A_l x + b_l$ are the associated affine transformations with matrices $A_l \in \mathbb{R}^{n_{l-1} \times n_l}$ and vectors $b_l \in \mathbb{R}^{n_l}$ and the activation function ρ. The resulting affine spaces are denoted by $Aff(n_{l-1}, n_l,\ \mathbb{R})$. We denote a space of the family $(W_l)_{1 \leq l \leq L}$ by

$$\mathbf{W} = (W_1, W_2, ..., W_L) \in \mathcal{W} := \prod_{l=1}^{L} Aff(n_{l-1}, n_l,\ \mathbb{R}).$$

The space $\mathcal{W}$ is endowed with the Euclidean norm given by

$$\|\mathbf{W}\|^2_{\mathcal{W}} = \sum_{l=1}^{L} \left(|A_l|^2_{\mathbb{R}^{n_{l-1} \times n_l}} + |b_l|^2_{\mathbb{R}^{n_l}} \right) = \sum_{l=1}^{L} \left(trace(A_l^T A_l) + |b_l|^2_{\mathbb{R}^{n_l}} \right).$$

Observe that for all $1 \leq l \leq L-1$ the following formula holds:

$$p_{l+1}(x, W_1, W_2, ..., W_{l+1}) = p_{l+1} \circ \rho(p_l(x, W_1, W_2, ..., W_l)). \tag{5.4}$$

By induction we can verify that the derivative of p_L with respect to x satisfies the following identity.

$$\nabla_x p_L(x, W_1, W_2, ..., W_L) = A_L \prod_{l=1}^{L-1} \rho' \left(p_l(x, W_1, W_2, ..., W_l)\right) A_l, \tag{5.5}$$

with ρ' exists in $L^\infty(\mathbb{R})$ if the activation function ρ is assumed globally Lipschitz on $\mathbb{R}$. The symbol $\prod_{l=1}^{L-1}$ indicates that the product is calculated iteratively by multiplying to the left the different matrices, when the index l increases: $\prod_{l=1}^{L} B_l = B_L B_{L-1} ... B_1$. Here again the terms of type $\rho'(p_l)$ are vectors of $\mathbb{R}^{n_l}$ that have to be understood coordinate-wise. The vector/matrix product $\rho'(p_l(x, W_1, W_2, ..., W_l))A_l$ is a matrix, and has to be understood as $(\rho' A)_i = \rho'_i A_{ij}$. Then, from the mean-value theorem, we get the following Lipschitz estimate:

$$|p_L(x_1, \mathbf{W}) - p_L(x_2, \mathbf{W})| \leq \left(\prod_{l=1}^{L} |A_l|_{\mathbb{R}^{n_{l-1} \times n_l}} \right) \|\rho'\|_\infty^{\mathcal{L}w} |x_1 - x_2|, \tag{5.6}$$

where $\mathcal{L}w = \sum_{l=1}^{L-1} n_l$.

Proposition 5.1. *When p_L satisfies $p_L(0) = 0$ and the activation function ρ is globally Lipschitz on $\mathbb{R}$, we have*

$$\|p_L\|_\infty \leq C C_\rho^{\mathcal{L}W} \|\mathbf{W}\|_{\mathcal{W}}^L. \tag{5.7}$$

Proof. See [37]. □

The derivative of p_L with respect to the weights W_l is given by the following formula, for all $1 \leq l \leq L$, for all $\tilde{W} = \tilde{A}x + \tilde{b}$:

$$\frac{\partial p_L}{\partial W_l}(x, \mathbf{W}).\tilde{W} = \left(\prod_{k=l}^{L-1} A_{k+1} \rho'(p_k(x, W_1, W_2, ..., W_k)) \right) \left(\tilde{W} \circ \rho(p_{l-1}(x, W_1, W_2, ..., W_{l-1})) \right), \tag{5.8}$$

$$\frac{\partial p_L}{\partial W_L}(x, \mathbf{W}).\tilde{W} = \left(\tilde{W} \circ \rho(p_{L-1}(x, W_1, W_2, ..., W_{L-1})). \right) \tag{5.9}$$

5.2 Optimal control problem

In the section, we are interested in the following abstract optimal control problem

$$\begin{cases} \min\limits_{\mathbf{W}\in\mathcal{W}} J(u,\mathbf{W}) = \dfrac{1}{2}\displaystyle\int_\Omega |u(x,T;\mathbf{W}) - g(x)|^2 dx \\ \text{subject to } (u,\mathbf{W}) \text{ solution of the problem (1) paraeterized with (5.3) and } \|\mathbf{W}\|^2_{\mathcal{W}} \leq C. \end{cases} \tag{5.10}$$

5.2.1 Existence of minimizer

The result of the existence of the solution for the problem (5.10) is introduced in the following theorem.

Theorem 5.1. *There exists at least one solution for the optimization problem* (5.10).

Proof. Clearly, $\min_{\mathbf{W}\in\mathcal{W}} J(u,\mathbf{W})$ is bounded and thus there exists a minimizing sequence $(u_n,\mathbf{W}_n)$ with $u_n = u(x,T;\mathbf{W}_n)$ such that

$$\lim_{n\longrightarrow\infty} J(u_n,\mathbf{W}_n) = \min_{\mathbf{W}\in\mathcal{W}} J(u,\mathbf{W}).$$

Since $\|\mathbf{W}_n\|_{\mathcal{W}}$ is bounded with respect to n, implies that there exists a subsequence, still denoted by $\mathbf{W}_n$, and $\mathbf{W}^* \in \mathcal{W}$ such that

$$\mathbf{W}_n \rightharpoonup \mathbf{W}^* \text{ weakly in } \mathcal{W}. \tag{5.11}$$

The convergence of $(\mathbf{W}_n)_n$ is strong because $\mathcal{W}$ is finite-dimensional. Further, the bounds provided by Theorem 2.1 for the of u_n in space $C([0,T];\bar{H}^2(\Omega))$, we can extract a subsequence denoted again u_n, and some $u^* \in C([0,T];\bar{H}^2(\Omega))$ such that

$$u_n \rightharpoonup u^* \text{ weakly in } C([0,T];\bar{H}^2(\Omega)). \tag{5.12}$$

Moreover, the embedding $C([0,T];\bar{H}^2(\Omega)) \hookrightarrow L^\infty([0,T];H^2(\Omega))$ is compact. The sequences u_n converges strongly to u^* in $C([0,T];H^2(\Omega))$. On the other hand, the functional J is weakly lower semi-continuous, we obtain

$$J(u^*,\mathbf{W}^*) \leq \liminf_{n\longrightarrow\infty} J(u_n,\mathbf{W}_n) = \inf_{\mathbf{W}\in\mathcal{W}} J(u,\mathbf{W}). \tag{5.13}$$

□

5.2.2 Optimality conditions

In this section, the study will be carried out in the general case whose activation functions ρ are only Lipschitz which is not sufficient to derive the optimality conditions with standard tools. For this purpose, we consider an approximation ρ_ε of the activation function ρ. We assume that it satisfies the following set of assumptions:

(A1) For all $\varepsilon > 0$, the function ρ_ε is of class C^1 on $\mathbb{R}$.

(A2) The sequence $(\rho_\varepsilon)_\varepsilon$ converges uniformly to ρ in $C(\mathbb{R}, \mathbb{R})$.

Example: The ReLU activation function. The so-called ReLU (Rectified Linear Units) activation function $\rho : x \longrightarrow \max(0; x)$ can be approximated by

$$\rho_\varepsilon(x) \begin{cases} 0 & \text{if } x \leq -\varepsilon \\ \frac{1}{\varepsilon}(x+\varepsilon)^2 & \text{if } -\varepsilon \leq x \leq \varepsilon \\ x & \text{if } \varepsilon \leq x. \end{cases} \tag{5.14}$$

We can verify that ρ_ε is of class C^1 on $\mathbb{R}$, and Lipschitz with constant less than or equal to 1 independent of ε, and $(\rho_\varepsilon)_\varepsilon$ converges uniformly towards ρ in $C(\mathbb{R}, \mathbb{R})$.

Then, the optimal control problem (5.10) can be reformulated as follows

$$\begin{cases} \min\limits_{\mathbf{W}\in\mathcal{W}} J(u, \mathbf{W}) = \dfrac{1}{2}\displaystyle\int_\Omega |u(x, T; \mathbf{W}) - g(x)|^2 dx \\ \text{subject to } (u, \mathbf{W}) \text{ solution of the problem (1) with } \rho_\varepsilon \text{ as activation function and } \|\mathbf{W}\|^2_{\mathcal{W}} \leq C. \end{cases} \tag{5.15}$$

where u solution of state problem (1), We give our weak formulation

$$\begin{cases} \text{Find} \quad u \in B^{\frac{\alpha}{2}}(Q) \quad \text{such that} \\ \left({}^R D_t^{\frac{\alpha}{2}} u, {}^R D_T^{\frac{\alpha}{2}} v\right)_Q + (\nabla u, \nabla v)_Q + \left(p_L^{(\varepsilon)}(x) f(u), v\right)_Q = \left(\dfrac{\phi(x) t^{-\alpha}}{\Gamma(1-\alpha)}, v\right)_Q \\ \forall v \in B^{\frac{\alpha}{2}}(Q). \end{cases} \tag{5.16}$$

where $p_L^{(\varepsilon)}$ is the neural network with the regularized ρ_ε as activation function.

Now, we will find the optimal conditions needed for the above regularized problem. Hence, a solution to problem (5.15) can be seen as a saddle-point of the following Lagrangian functional:

$$\begin{aligned} \mathbb{J}(u, v, \mathbf{W}, \lambda) =& J(u, \mathbf{W}) - \left({}^R D_t^{\frac{\alpha}{2}} u, {}^R D_T^{\frac{\alpha}{2}} v\right)_Q - (\nabla u, \nabla v)_Q - \left(p_L^{(\varepsilon)}(x) f(u), v\right)_Q \\ &+ \left(\frac{\phi(x) t^{-\alpha}}{\Gamma(1-\alpha)}, v\right)_Q - \lambda(\|\mathbf{W}\|^2 - C). \end{aligned}$$

The adjoint problem is obtained by setting the derivative of $\mathbb{J}(u, v, \mathbf{W})$ with respect to u equal to zero. First, taking the derivative with respect to u, we obtain the following adjoint problem associated with the regularized problem (5.15)

$$\begin{cases} D_T^{\alpha} v(x,t) - \Delta v(x,t) + p_L^{(\varepsilon)} f'(u) v(x,t) &= (u(x,t) - g(x))\delta_T(t) \quad (x,t) \in Q \\ v(x,T) &= 0 \quad x \in \Omega \\ v(x,t) &= 0 \quad x \in \partial\Omega \quad t \in I \end{cases} \tag{5.17}$$

where $\delta(.)$ is the Dirac function.

Proposition 5.2. *Let $\bar{\mathbf{W}} = \left(\bar{W}_1, \bar{W}_2, ..., \bar{W}_L\right) \in \mathcal{W}$ be an optimal solution of regularized problem (5.15). Then there exists $\bar{\lambda} \geq 0$ such that for all $1 \leq l \leq L$, we have*

$$\int_0^T \int_\Omega \frac{\partial p_L^{(\varepsilon)}(x, \bar{\mathbf{W}})}{\partial W_l} f(\bar{u}) \bar{v} dx dt - 2\bar{\lambda}\bar{W}_l = 0, \tag{5.18}$$

and where $\bar{u}$ is state solution of (1) with $\bar{\mathbf{W}}$ as data, and where $\bar{v}$ is adjoint solution of (5.17) with $\bar{\mathbf{W}}$ as data

Proof. We assume additionally that u are continuously differentiable, we present the calculate of the first-order derivatives of $\mathbb{J}$ with respect to $\mathbf{W}$, in direction $\tilde{\mathbf{W}}$, we have

$$\begin{aligned} \nabla_{\mathbf{W}} \mathbb{J}(u, v, \mathbf{W}).\tilde{\mathbf{W}} =& \nabla_{\mathbf{W}} J(u, \mathbf{W}).\tilde{\mathbf{W}} - \left({}^R D_t^{\frac{\alpha}{2}} u'.\tilde{\mathbf{W}}, {}^R D_T^{\frac{\alpha}{2}} v\right)_Q - \left(\nabla u'.\tilde{\mathbf{W}}, \nabla v\right)_Q \\ & - \left(p_L^{(\varepsilon)}(x) f'(u) u'.\tilde{\mathbf{W}}, v\right)_Q - \left(\nabla_{\mathbf{W}} p_L^{(\varepsilon)}.\tilde{\mathbf{W}} f(u), v\right)_Q - 2\lambda \left(\mathbf{W}, \tilde{\mathbf{W}}\right)_{\mathcal{W}}. \end{aligned}$$

From the definition of the function J, its derivative with respect to $\mathbf{W}$ is written as follows

$$\nabla_{\mathbf{W}} J(u, \mathbf{W}).\tilde{\mathbf{W}} = \left(u(x, T; \mathbf{W}) - g(x), u'(x, T; \mathbf{W}).\tilde{\mathbf{W}}\right)_\Omega. \tag{5.19}$$

Using Lemma 1.7 and v is the solution of adjoint problem (5.17), we obtain

$$\begin{aligned} \nabla_{\mathbf{W}} \mathbb{J}(u, v, \mathbf{W}).\tilde{\mathbf{W}} &= \left(u(x, T; \mathbf{W}) - g(x), u'.\tilde{\mathbf{W}}\right)_\Omega - \left(u'.\tilde{\mathbf{W}}, D_T^{\alpha} v\right)_Q + \left(u'.\tilde{\mathbf{W}}, \Delta v\right)_Q \\ & - \left(u'.\tilde{\mathbf{W}}, p_L^{(\varepsilon)}(x) f'(u) v\right)_Q - \left(\nabla_{\mathbf{W}} p_L^{(\varepsilon)}.\tilde{\mathbf{W}} f(u), v\right)_Q - 2\lambda \left(\mathbf{W}, \tilde{\mathbf{W}}\right)_{\mathcal{W}}. \\ &= \left(u(x, T; \mathbf{W}) - g(x), u'.\tilde{\mathbf{W}}\right)_\Omega - \left(u'.\tilde{\mathbf{W}}, D_T^{\alpha} v - \Delta v + p_L^{(\varepsilon)}(x) f'(u) v\right)_Q \\ & - \left(\nabla_{\mathbf{W}} p_L^{(\varepsilon)}.\tilde{\mathbf{W}} f(u), v\right)_Q - 2\lambda \left(\mathbf{W}, \tilde{\mathbf{W}}\right)_{\mathcal{W}}. \end{aligned}$$

Since v is the solution of the adjoint problem 5.17 means that

$$\nabla_{\mathbf{W}}\mathbb{J}(u,v,\mathbf{W}).\tilde{\mathbf{W}} = \left(u(x,T;\mathbf{W}) - g(x), u'.\tilde{\mathbf{W}}\right)_{\Omega} - \left(u'.\tilde{\mathbf{W}}, (u(x,t;\mathbf{W}) - g(x))\delta_T(t)\right)_{Q} - \left(\nabla_{\mathbf{W}} p_L^{(\varepsilon)}.\tilde{\mathbf{W}} f(u), v\right)_{Q} - 2\lambda\left(\mathbf{W},\tilde{\mathbf{W}}\right)_{\mathcal{W}}.$$

consequently

$$\begin{aligned}\nabla_{\mathbf{W}}\mathbb{J}(u,v,\mathbf{W}).\tilde{\mathbf{W}} &= -\left(\nabla_{\mathbf{W}} p_L^{(\varepsilon)}.\tilde{\mathbf{W}} f(u), v\right)_{Q} - 2\lambda\left(\mathbf{W},\tilde{\mathbf{W}}\right)_{\mathcal{W}} \\ &= -\left(\left(\nabla_{\mathbf{W}} p_L^{(\varepsilon)} f(u), v\right)_{Q}, \tilde{\mathbf{W}}\right)_{\mathcal{W}} - 2\lambda\left(\mathbf{W},\tilde{\mathbf{W}}\right)_{\mathcal{W}}.\end{aligned}$$

Let us recall that $\mathbf{W} = (W_1, W_2, ..., W_L)$, $\tilde{\mathbf{W}} = (\tilde{W}_1, \tilde{W}_2, ..., \tilde{W}_L)$
and $\nabla_{\mathbf{W}} p_L^{(\varepsilon)} = \left(\nabla_{W_1} p_L^{(\varepsilon)}, \nabla_{W_2} p_L^{(\varepsilon)}, ..., \nabla_{W_L} p_L^{(\varepsilon)}\right)$.
Thus the proof is complete. □

5.2.3 Training algorithm

The gradient descent (GD) method is the original algorithm for neural networks. It may be expressed as follows in this approximation: The main formulas for weights W are shown as following, and formula for bias b is similar to it:

$$\begin{aligned}\mathbf{W}^{k+1} &= \mathbf{W}^k - \eta\nabla_{\mathbf{W}}\mathbb{J}(u,v,\mathbf{W}^k), \\ \mathbf{b}^{k+1} &= \mathbf{b}^k - \eta\nabla_{\mathbf{b}}\mathbb{J}(u,v,\mathbf{W}^k).\end{aligned}$$

where k is the iterations, $\mathbf{W}^k$ is the main formulas for weights and $\mathbf{b}^k$ is the formula for bias. To summarize, the ANN algorithm to solve the inverse problem is constructed as follows:

5.3 Numerical experiments

In this section, we present detailed numerical results to demonstrate the ANN method for inverse potential problem (1)-(2). First, different identification examples are presented to verify the accuracy of this method.

The structure of ANN is chosen with layers $[d, 10, 10, 10, 10, 1]$, where d is the dimension of input data. By setting initial w and b randomly in $(-1, 1)$. In each hidden layer, we use the ReLU activation function. Note that the input data here depends explicitly on the spatial variable x,

Algorithm 5 ANN algorithm algorithm to solve our inverse problem

Begin
Input: Given the tolerance tol, the initial condition u_0, the data g, x_{data}, f, α, β,η , max_{itr},

the number of hidden layers L;
Initialize :
Initialize weights $\mathbf{W}_l$, bias $\mathbf{b}_l$ and other parameters;
Initialize the structure of neural networks;
set $k = 0$.
while $k < max_{it}$ **do**
Compute $p_k = Model_{(\mathbf{W}^k,\mathbf{b}^k)}(x_{data})$
Compute u_k by solving the state problem (1).
Compute v_k by solving the state problem (5.17).
$\mathbf{W}^{k+1} = \mathbf{W}^k - \eta\nabla_{\mathbf{W}}\mathbb{J}(u, v, \mathbf{W}^k)$,
$\mathbf{b}^{k+1} = \mathbf{b}^k - \eta\nabla_{\mathbf{b}}\mathbb{J}(u, v, \mathbf{W}^k)$.
$p_{k+1} = Model_{(\mathbf{W}^{k+1},\mathbf{b}^{k+1})}(x_{data})$
Check convergence : **if** $\dfrac{\|p_{k+1} - p_k\|_{L^2(\Omega)}}{\|p_k\|_{L^2(\Omega)}} \leq tol$ stop.
Set $k = k + 1$.
end while
return p.
End

both the input and output layers have only one neuron, respectively. By setting the unknown

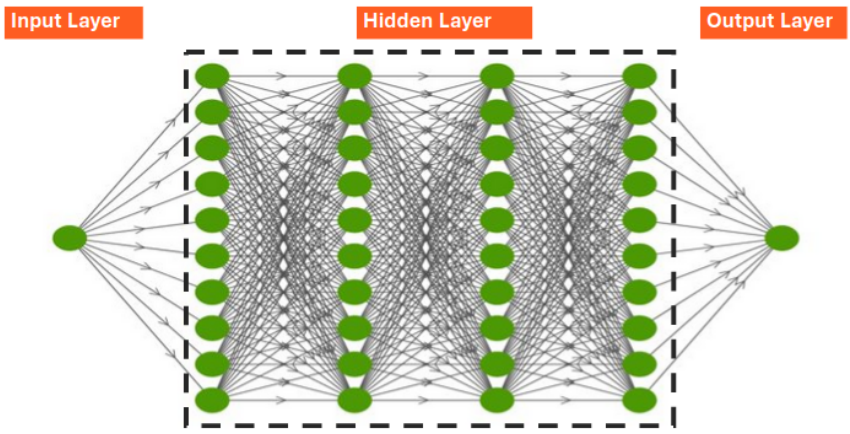


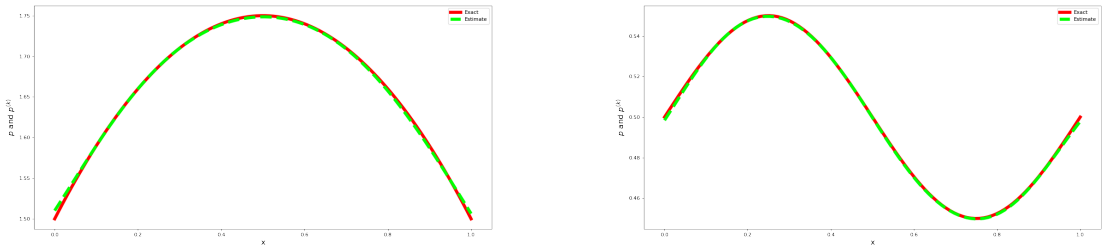
Figure 5.1: The ANN structure

coefficient p and selecting the initial condition $\phi(x)$ and the derivative order α, we construct the synthetic data $g(x)$ by solve the direct problem (1) for $p(x) = p_L$, and at the conclusion $g(x) = u(x, T)$, we extract the measurement $g(x)$ as the solution $u(x, T)$. In this numerical experiment, we assume that the solution of the direct problem to be $u(x, t) = t^2 \sin(\pi x)+\sin(\pi x)$,

the order of the fractional derivative to be $\alpha = 0.5$ and the regularity parameter $\beta = 10^{-6}$. Always keep in this section the examples we defined in the chapter 3. The numerical results for these examples using the divergence principle for different noise levels are shown in the figures below.

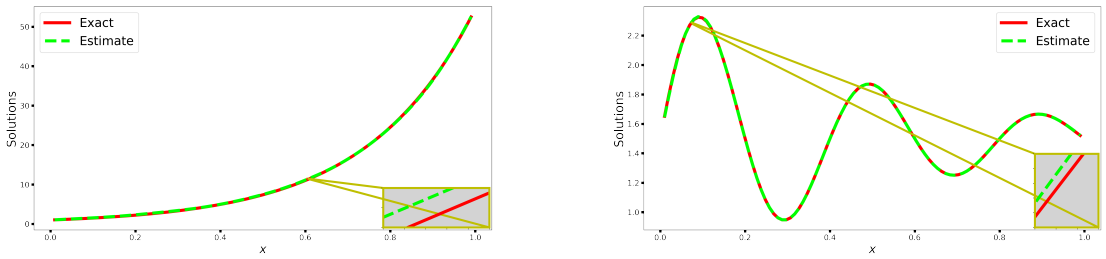
5.3.0.1 Identification results without noise

We now provide the identification results of the potential p for the four different examples. We classify the above examples into two types; regulars like Examples 3.1, Examples 3.2,Examples 3.3 and 3.4 and discontinuous ones like Examples 3.5 and 3.6. The result of the first type is presented in Figures 5.2and 5.3, and the second type is presented in Figure 5.4. We note that,



(a) Exemple 3.1 (b) Exemple 3.2

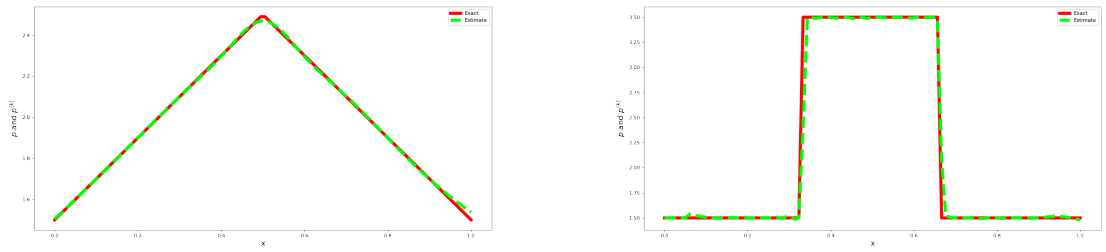
Figure 5.2: The numerical output of the ANN algorithm for Examples 3.1 and 3.2 without noise



(a) Exemple 3.3 (b) Exemple 3.4

Figure 5.3: The numerical output of the ANN algorithm for Examples 3.3 and 3.4 without noise

the results obtained in the regular cases (Figures 5.2 and 5.3) nearly exactly match the exact potential. On the other side, the method does not precisely construct the coefficient singularities for discontinuous coefficients. Figure 5.4.

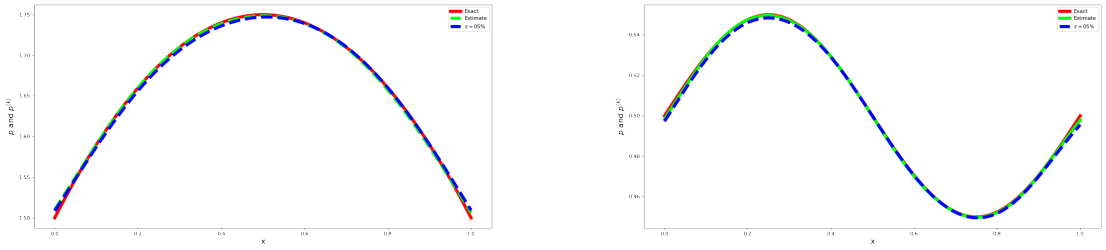


(a) Exemple 3.5 (b) Exemple 3.6

Figure 5.4: The numerical output of the ANN algorithm for Examples 3.5 and 3.6 without noise.

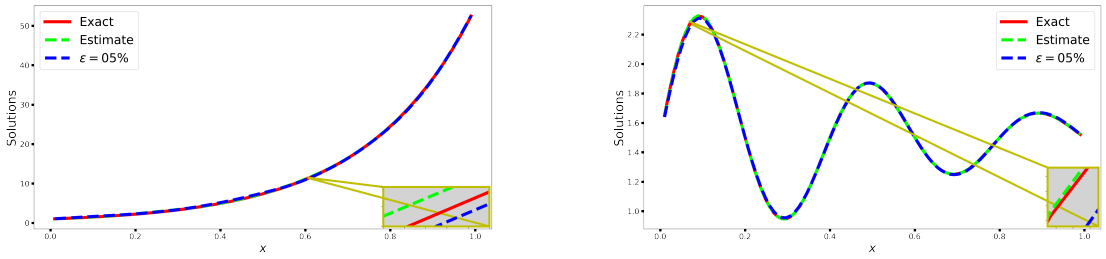
5.3.0.2 Identification results with noise

In this paragraph, we give the results of numerical tests obtained for noise level 5%. Figure 5.5 and Figure 5.6 presents the numerical outputs for the regular examples when noise level and Figure 5.7 shows the numerical outputs with noise level for the discontinuous examples. The



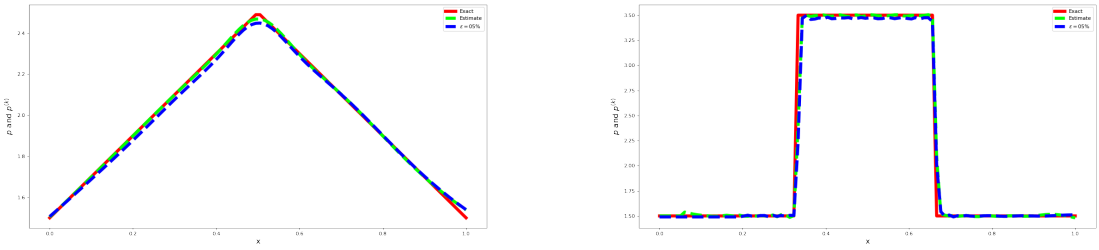
(a) Exemple 3.1 (b) Exemple 3.2

Figure 5.5: The numerical output of the ANN algorithm for Examples 3.1 and 3.2 with noise



(a) Exemple 3.3 (b) Exemple 3.4

Figure 5.6: The numerical output of the ANN algorithm for Examples 3.3 and 3.4 with noise



(a) Exemple 3.5 (b) Exemple 3.6

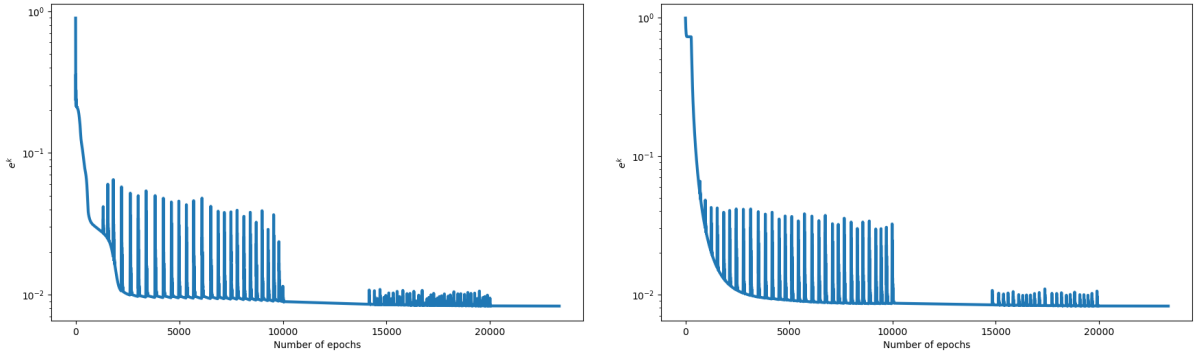
Figure 5.7: The numerical output of the ANN algorithm for Examples 3.5 and 3.6 with noise

numerical outputs obtained with noise in Figures 5.5, 5.6 and 5.7 demonstrate that while the ANN algorithm is stable with regard to noise, it is unable to capture the coefficient discontinuity.

We can deduce from the above numerical examples that the ANN method for the fractional inverse problem with many hidden layers is a stable method on various aspects, including good pose with noise on the measure and insensitivity on the singular area.

5.3.0.3 Relative error

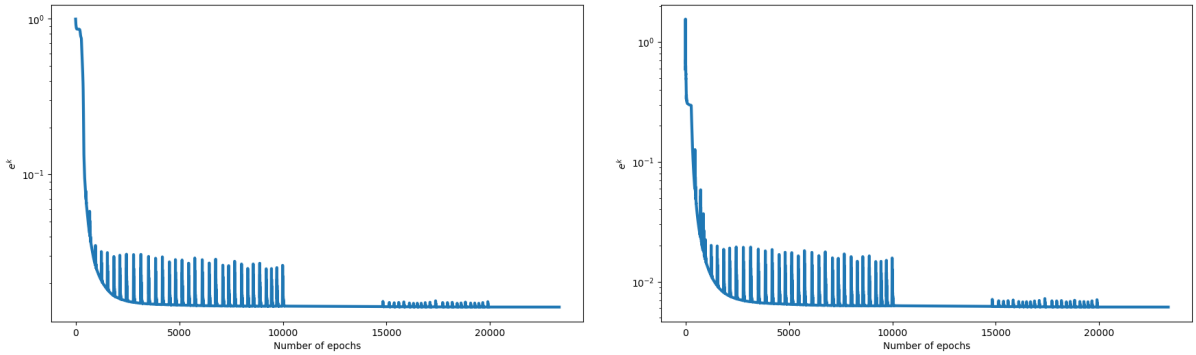
In this paragraph, we show the results of the relative error e_k with respect to increasing the number of iterations. As well as the impact of the fractional order α. In the Figures 5.8



(a) Exemple 3.3 (b) Exemple 3.4

Figure 5.8: The numerical results of the relative error e_k for the Examples 3.3 and 3.4.

and 5.9,we present the result of the relative errors e_k of four examples above with respect to increasing the number of iterations. We can notice that the relative errors tend towards the neighbourhood of zero in an exponential way. The Table 5.1 shows the numerical results of the relative errors e_k for the Example 3.2 according to the different noise values ε with respect to the different α. We can notice that the relative error e_k decreases when the noise level in the



(a) Exemple 3.5 (b) Exemple 3.6

Figure 5.9: The numerical results of the relative error for the Examples 3.5 and 3.6.

Table 5.1: Numerical results for e_k with different ε for Example 3.1

$\alpha \backslash \varepsilon$	$\varepsilon = 0$	$\varepsilon = 0.001$	$\varepsilon = 0.01$	$\varepsilon = 0.10$	$\varepsilon = 0.20$
$\alpha = 0.3$	$2.312.10^{-2}$	$3.178.10^{-2}$	$4.231.10^{-2}$	$5.421.10^{-2}$	$7.901.10^{-2}$
$\alpha = 0.5$	$2.271.10^{-2}$	$2.953.10^{-2}$	$3.284.10^{-2}$	$5.092.10^{-2}$	$7.312.10^{-2}$
$\alpha = 0.7$	$2.109.10^{-2}$	$2.641.10^{-2}$	$3.017.10^{-2}$	$4.467.10^{-2}$	$7.031.10^{-2}$
$\alpha = 0.9$	$1.319.10^{-2}$	$1.923.10^{-2}$	$2.594.10^{-2}$	$4.158.10^{-2}$	$6.964.10^{-2}$

data decreases. It is also clear to us how much the impacts of the fractal order α, as it has a small impact on the obtained results.

As can be observed, there is a good agreement between the current result and the reference solution. The aforementioned findings demonstrate that our technique is effective for solving time-fractional inverse problems, and we also learn more about our method's stability.

5.4 Conclusion

In this chapter, we proposed an artificial neural network approach to identify the parameter in a nonlinear subdiffusion model from additional data. Instead of determining the parameter in the temporal fractional diffusion model, we design the model itself, in the form of an artificial neural network. The choice of this approach constitutes a change in the nature of neural network approximation, which leads us to the development of parameter identification techniques. We reformulate this inverse problem as an optimal control problem. We show that the latter is well posed. We also proved the existence of the solution to the control problem and provided a mathematical analysis and derivation of the optimal conditions. In the end, we presented some numerical tests to show the efficiency of the proposed algorithm.

Chapter 6

Numerical comparison

In this section, we attempt to make a numerical comparison between the methods we proposed above to solve our inverse problem through the optimization approach. We studied the inverse problem through additional conditions on the data; the first is a condition in the final time, according to which we transformed our inverse problem into a control problem, and to study it we proposed two methods, namely the GD method in Chapter 3 and the ANN method in Chapter 5. The second is a condition in a set in the domain. This also transformed the inverse problem into an optimal control problem, and to study it we proposed the ADMM method in Chapter 4. However, in order to make an effective comparison, we will apply the three proposed methods to the following optimal control problem:

$$\text{Find } \bar{p}(x) \in \mathcal{U}_{ad} \text{ such that } J(\bar{p}) = \min_{p \in \mathcal{U}_{ad}} J(p), \tag{6.1}$$

with

$$J(p) = \frac{1}{2} \int_{\Omega} |u(x, T; p) - g(x)|^2 dx. \tag{6.2}$$

Obviously, the additional condition we choose in this case is $u(x, T) = g(x)$.

To compare, we present numerical simulations performed in one dimensional, and we set $T < 1$, $\Omega = (0, 1)$. The grid points on $[0, 1]$ and $[0, T]$ are respectively 51 and 70 while the resolution of the direct problem is by finite difference method.

We build the synthetic data g by fixing the unknown parameter p, selecting the nonlinear term f, and setting the initial conditions ϕ, then solving the direct problem (1), and extracting the measurement g. We add a random perturbation with a fixed amplitude, δ, to the data for the tests with noisy data, i.e.

$$g^{\delta} = g(\delta(2rand(size(g)) - 1) + 1).$$

The implementation also requires a stopping criteria. In our numerical experiments, it was difficult to find criteria which works well on both synthetic and measured test data. Therefore,

in our case, the iterative process was interactively monitored and stopped when the change in the reconstruction obtained from an iteration step was no longer noticeable.

Along the reconstruction, we select the nonlinear term f that satisfies the assumptions $(\mathbf{H}_1)$ and $(\mathbf{H}_2)$ as defined below:

$$f(x) = e^x - 1.$$

In this numerical experiment, we give the solution to the direct problem by $u(x,t) = t^2 \sin(\pi x) + \sin(\pi x)$, the order of the fractional derivative to be $\alpha = 0.5$. We identify the parameter p for $u(x,t)$ the following exact solution of problem (1), where the initial data $u_0 = \sin(\pi x)$.

Remark 6.1. *It is difficult to seek a suitable regularization parameter, thus the comparisons of minimum reconstruction errors will eliminate the effects resulting from utilizing different β. However, the regularization parameter in our situation was mainly chosen by trial and error.*

In the ADMM algorithm, we choose the constant $\sigma = 0.99\frac{\sqrt{2}}{\sqrt{2}+\sqrt{r}}$ (4.39) to check the inexactness criterion (4.40), because larger values of σ indicate that the criterion is looser and thus less computation is required to solve the subproblems. In our computation we always set $\max^{iter} = 50$. The final data $g = u(.,T)$ are obtained by solving the direct problem (1).

In the ANN algorithm, we chose the ANN structure with layers $[d, 10, 10, 10, 10, 1]$, where d is the dimension of the input data. By randomly setting the initial w and b in $(-1, 1)$. In each hidden layer we use the ReLU activation function. Note that the input data here explicitly depends on the spatial variable x, both the input and output layers have only one neuron each. By setting the exact potential p and choosing the initial condition $\phi(x)$ and the derivative order α, we construct the synthetic data $g(x)$ by solving the direct problem (1) for $p(x) = p_L$, and at the conclusion $g(x) = u(x,T)$, we extract the measurement $g(x)$ as the solution $u(x,T)$.

Now, we present some examples in the one-dimensional case to verify the effectiveness for each algorithm, where first two examples are smooth and last two are non smooth. We have added these discontinuous examples to show the power and the limitations of the proposed methods.

Exemple 6.1. *The first example is regular supposed as follows $p(x) = \pi^2 \sin(\pi x)$.*

Exemple 6.2. *The second example is synthetic, we assume u_0 supplied in the above and we estimate the unknown parameter p as follows*

$$p(x) = 2 + \sin(2\pi x). \tag{6.3}$$

The observation data g are obtained by solving the direct problem (1).

Exemple 6.3. *Take $p(x) = e^{4x}$ and the data g is obtained by solving the direct problem (1) by using the finite difference method as the exact input data.*

Exemple 6.4. *This example is synthetic, we consider $\phi(x)$ and α given in the above and we suppose the unknown parameter as follows*

$$p(x) = e^{-2x}\sin(5\pi x) + 1.5.$$

The data g is obtained by solving the direct problem (1) by the finite difference method.

Exemple 6.5. *In the fated example,we test a non-smooth one with a peak. Let the parameter given by the following expression*

$$p(x) = \begin{cases} x & 0 \le x \le \frac{1}{2} \\ 1 - x & \frac{1}{2} \le x \le 1 \end{cases} \tag{6.4}$$

The data g is obtained by solving the direct problem 1 by a finite difference method as the exact input data.

Exemple 6.6. *In the last example, we consider a discontinuous potential as follow:*

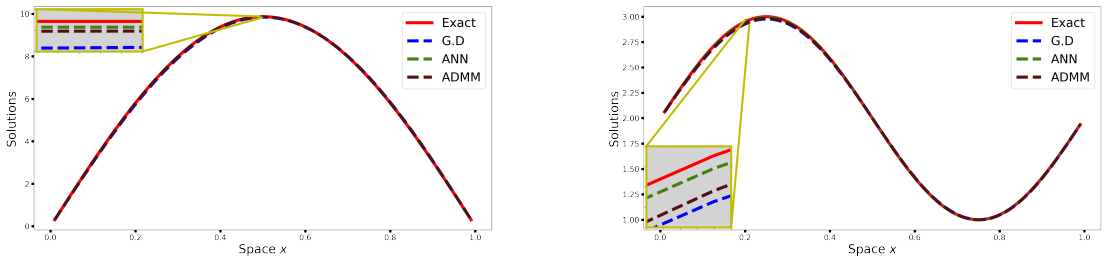
$$p(x) = \begin{cases} 0 & 0 \le x \le \frac{1}{3} \\ 1 & \frac{1}{3} < x \le \frac{2}{3} \\ 0 & \frac{2}{3} < x \le 1 \end{cases}. \tag{6.5}$$

The data g is obtained by solving the direct problem 1 by a finite difference method as the exact input data.

6.1 Identification results without noise

This part is devoted to the comparison without noise between the gradient descent, ANN and ADMM methods for all the above examples. Figures 6.1 and 6.2 show that the three methods have more or less identical performances for regular parameters. But the approximate solutions by ANN method is more close than the exact solutions. For the discontinuous cases presented

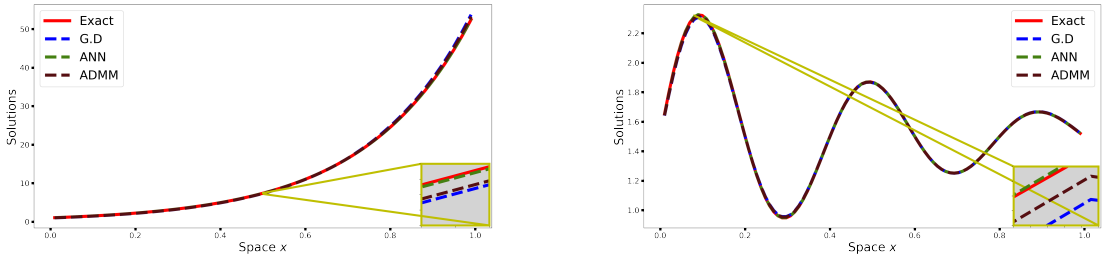
in Figure 6.3, however, the ANN method performs better than the ADMM method, which also performs better than the gradient descent method. In terms of efficiency, it is clear to us from the results obtained in all cases that the ANN method is more efficient than the other methods.



(a) Example 6.1 without noise

(b) Example 6.2 without noise

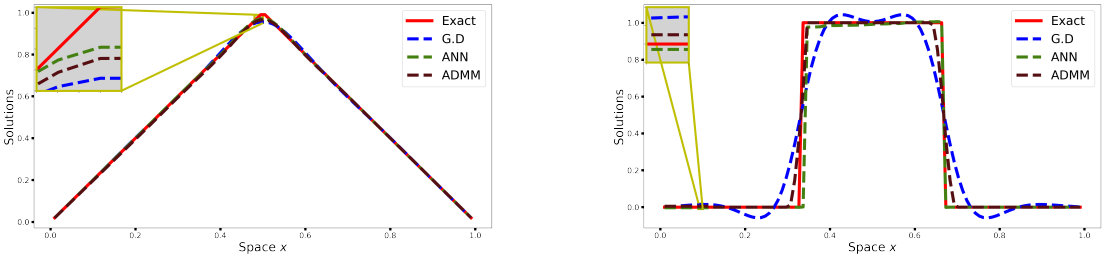
Figure 6.1: The comparison of the numerical results obtained with the different proposed methods for Examples 6.1 and 6.2 without noise.



(a) Example 6.3 without noise

(b) Example 6.4 without noise

Figure 6.2: The comparison of the numerical results obtained with the different proposed methods for Examples 6.3 and 6.4 without noise.



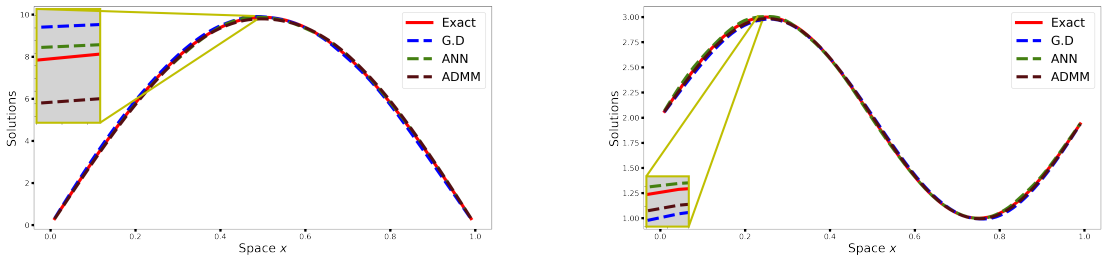
(a) Example 6.5 without noise

(b) Example 6.6 without noise

Figure 6.3: The comparison of the numerical results obtained with the different proposed methods for Examples 6.3 and 6.4 without noise.

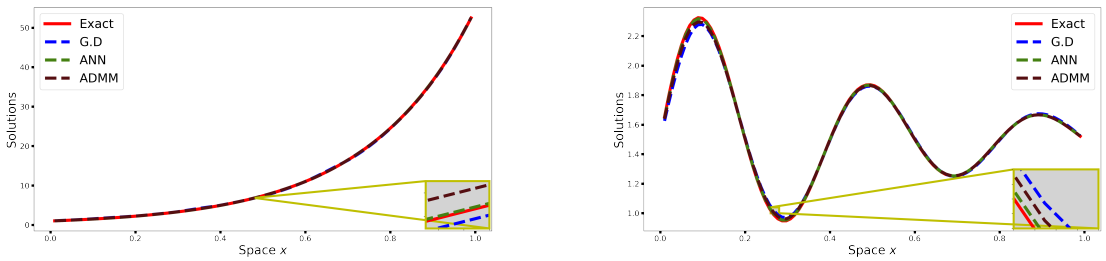
6.2 Identification results with noise

The numerical tests with noise show that all three methods are generally stable with respect to noise. But we can conclude through the numerical results obtained in Figures 6.4, 6.5 and 6.6 that the ANN method is more stable than the ADMM method, which more stable than the GD method. However, it cannot capture the singularity of the parameter, but it approaches it. However, the three methods cannot capture the singularity of the parameter, but the ANN method comes closer to it than the ADMM method which comes closer to it than the GD method.



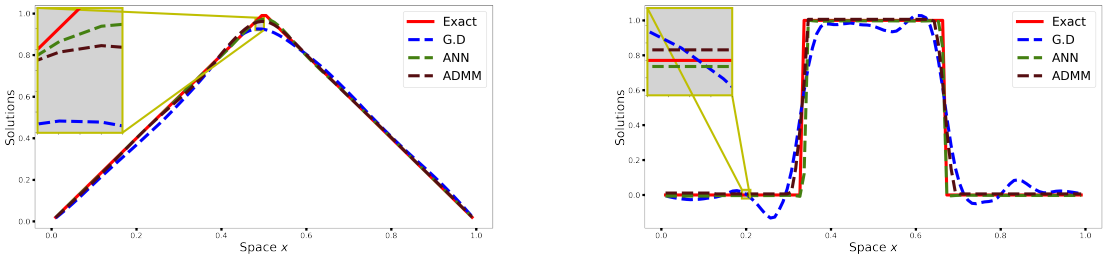
(a) Example 6.1 with noise (b) Example 6.2 with noise

Figure 6.4: The comparison of the numerical results obtained with the different proposed methods for Examples 6.1 and 6.2 with noise.



(a) Example 6.3 with noise (b) Example 6.4 with noise

Figure 6.5: The comparison of the numerical results obtained with the different proposed methods for Examples 6.3 and 6.4 with noise.



(a) Example 6.5 with noise

(b) Example 6.6 with noise

Figure 6.6: The comparison of the numerical results obtained with the different proposed methods for Exemples 6.5 and 6.6 with noise.

Conclusion and perspectives

In this thesis, we have studied the inverse problem of identifying potential parameters in a nonlinear subdiffusion model from additional conditions. The nonlinear subdiffusion model involves a Caputo fractional derivative of order $\alpha \in (0, 1)$ in time. In the first part, we explored the well-posedness of the direct problem using Mittag-Leffler functions. The second part investigated the inverse problem based on the regularized optimal control problem approach. We established the existence of the optimal solution and shown the stability with respect to the data of our inverse problem based on the optimality conditions of the considered functional. The third part is devoted to the conceptual implementation of the Alternating Direction Multiplier Method (ADMM) to solve the optimal control problem, showing the convexity of the functional to apply the ADMM and establishing the convergence of this proposed numerical method. The fourth part proposed the Artificial Neural Network (ANN) method to solve the inverse problem. This approach is based on the replacement of the unknown parameter by a neural network, which is identified by reformulating it into a control problem involving a neural network. This study started by giving some theoretical results on the ANN method, showing the well-posedness of the control problem, computing the optimal conditions with respect to the network weights and biases, and proving the ANN algorithm. The last part presented some numerical experiments to demonstrate and compare the effectiveness of the proposed algorithms, such as the gradient descent algorithm, ANN algorithm, and the ADMM algorithm.

We have to add that it is possible to extend the present technique to more general cases, such as more complex diffusion parameters, but under the condition that the diffusion parameter is bounded and non-degenerate or under appropriate general conditions. As future work, we will extend our study to the case where the minimization of type L^1. In order, to capture the singularities of the parameter p for the case non-smooth or even the capture the case of p measure.

Bibliography

[1] J. T. A. Kilbas, H. Srivastava. *Theory and Applications of Fractional Differential Equations*, volume 204. Elsevier Science, New York, 2006. 31

[2] L. P. Aarts and P. Van Der Veer. Neural network method for solving partial differential equations. *Neural Processing Letters*, 14(3):261–271, 2001. 16

[3] R. A. Adams and J. Fournier. *Sobolev Spaces Academic Press*, volume 41. 1975. 22, 41, 78

[4] L. Afraites and A. Oulmelk. Identification of robin coefficient in elliptic problem by a coupled complex boundary method. In *International Conference on Numerical Analysis and Optimization Days*, pages 71–86. Springer, 2021. 78

[5] R. P. Agarwal and D. O'Regan. Picard's method of successive approximations. *An Introduction to Ordinary Differential Equations*, pages 53–60, 2008. 47

[6] J.-P. Aubin. Analyse mathematique-un theoreme de compacite. *Comptes Rendus Hebdomadaires Des Seances De L Academie Des Sciences*, 256(24):5042, 1963. 28

[7] D. Baleanu, Z. B. Güvenç, J. T. Machado, et al. *New trends in nanotechnology and fractional calculus applications*, volume 10. Springer, 2010. 12

[8] D. A. Benson, S. W. Wheatcraft, and M. M. Meerschaert. The fractional-order governing equation of lévy motion. *Water resources research*, 36(6):1413–1423, 2000. 13

[9] B. Berkowitz, H. Scher, and S. E. Silliman. Anomalous transport in laboratory-scale, heterogeneous porous media. *Water Resources Research*, 36(1):149–158, 2000. 13

[10] M. Caputo. Linear models of dissipation whose q is almost frequency independent—ii. *Geophysical Journal International*, 13(5):529–539, 1967. 12, 31

[11] M. Caputo. Elasticity and dissipation zanichelli, 1969. 12

[12] G. Cybenko. Approximation by superpositions of a sigmoidal function. *Mathematics of control, signals and systems*, 2(4):303–314, 1989. 16

[13] R. Dautray and J.-L. Lions. Analyse mathématique et calcul numérique pour les sciences et les techniques. *Collection du Commissariat a l'Energie Atomique. Serie Scientifique*, 1985. 22

[14] R. Dautray and J.-L. Lions. *Mathematical analysis and numerical methods for science and technology: volume 1 physical origins and classical methods*. Springer Science & Business Media, 2012. 26, 28

[15] M. Dissanayake and N. Phan-Thien. Neural-network-based approximations for solving partial differential equations. *communications in Numerical Methods in Engineering*, 10(3):195–201, 1994. 17

[16] M. Fortin and R. Glowinski. *Augmented Lagrangian methods: applications to the numerical solution of boundary-value problems*. Elsevier, 2000. 88

[17] R. Glowinski. *Lectures on numerical methods for non-linear variational problems*. Springer Science & Business Media, 2008. 88

[18] R. Glowinski, Y. Song, and X. Yuan. An admm numerical approach to linear parabolic state constrained optimal control problems. *Numerische Mathematik*, 144(4):931–966, 2020. 16, 17, 88

[19] M. S. Gockenbach and A. A. Khan. An abstract framework for elliptic inverse problems: Part 1. an output least-squares approach. *Mathematics and mechanics of solids*, 12(3):259–276, 2007. 78, 99

[20] Y. Hao, H. Song, X. Wang, and K. Zhang. An alternating direction method of multipliers for the optimization problem constrained with a stationary maxwell system. *Commun. Comput. Phys.*, 24(5):1435–1454, 2018. 16

[21] K. Hornik, M. Stinchcombe, and H. White. Multilayer feedforward networks are universal approximators. *Neural networks*, 2(5):359–366, 1989. 16

[22] B. Jadamba, A. A. Khan, G. Rus, M. Sama, and B. Winkler. A new convex inversion framework for parameter identification in saddle point problems with an application to the elasticity imaging inverse problem of predicting tumor location. *SIAM Journal on Applied Mathematics*, 74(5):1486–1510, 2014. 99

[23] Y. Jiang and J. Ma. High-order finite element methods for time-fractional partial differential equations. *J. Comput. Appl.Math*, 235(11):3285–3290, 2011. 13

[24] B. Jin et al. *Fractional differential equations*. Springer, 2021. 37

[25] B. Jin and Y. Kian. Recovery of a distributed order fractional derivative in an unknown medium. *arXiv preprint arXiv:2207.12929*, 2022. 13

[26] B. Jin, B. Li, and Z. Zhou. Numerical analysis of nonlinear subdiffusion equations. *SIAM Journal on Numerical Analysis*, 56(1):1–23, 2018. 13, 35

[27] B. Jin and W. Rundell. An inverse problem for a one-dimensional time-fractional diffusion problem. *Inverse Problems*, 28(7):075010, jun 2012. 13

[28] B. Jin and W. Rundell. A tutorial on inverse problems for anomalous diffusion processes. *Inverse problems*, 31(3):035003, 2015. 36

[29] B. Jin and Z. Zhou. An inverse potential problem for subdiffusion: stability and reconstruction. *Inverse Problems*, 37(1):015006, 2020. 14

[30] B. Kaltenbacher and W. Rundell. On the identification of a nonlinear term in a reaction–diffusion equation. *Inverse Problems*, 35(11):115007, 2019. 30, 35

[31] A. Kilbas. *Theory and applications of fractional differential equations*. 29

[32] K. Kunisch et al. Design of the monodomain model by artificial neural networks. *arXiv e-prints*, pages arXiv–2107, 2021. 17

[33] I. E. Lagaris, A. Likas, and D. I. Fotiadis. Artificial neural networks for solving ordinary and partial differential equations. *IEEE transactions on neural networks*, 9(5):987–1000, 1998. 16

[34] I. E. Lagaris, A. C. Likas, and D. G. Papageorgiou. Neural-network methods for boundary value problems with irregular boundaries. *IEEE Transactions on Neural Networks*, 11(5):1041–1049, 2000. 16

[35] L. Li and J. G. Liu. Some compactaness criteria for weak solutions of time fractional pdes. *SIAM J. MATH. ANAL*, 40(5):3963–3995, 2018. 51

[36] X. Li and C. Xu. Existence and uniqueness of the weak solution of the space-time fractional diffusion equation and a spectral method approximation. *Communications in Computational Physics*, 8(5):1016, 2010. 42

[37] Y. Li and X. Hu. An artificial neural network approximation for cauchy inverse problems. *arXiv preprint arXiv:2001.01413*, 2020. 105

[38] Y. Lin and C. Xu. Finite difference/spectral approximations for the time-fractional diffusion equation. *J. Comput. Phys*, 225(2):1533–1552, 2007. 13

[39] J. L. Lions. Problemes aux limites non homogenes et applications. 1968. 22, 25

[40] J.-L. Lions. Quelques méthodes de résolution de problemes aux limites non linéaires. 1969. 25

[41] Y. Liu. Strong maximum principle for multi-term time-fractional diffusion equations and its application to an inverse source problem. *Computers and Mathematics with Applications*, 73(1):96–108, 2017. 13

[42] Y. Luchko. Maximum principle for the generalized time-fractional diffusion equation. *J. Math. Anal. Appl*, 351(1):218–223, 2009. 13

[43] Y.-K. Ma, P. Prakash, and A. Deiveegan. Optimization method for determining the source term in fractional diffusion equation. *Mathematics and Computers in Simulation*, 155:168–176, 2019. 42

[44] A. Malek and R. S. Beidokhti. Numerical solution for high order differential equations using a hybrid neural network—optimization method. *Applied Mathematics and Computation*, 183(1):260–271, 2006. 16

[45] R. Metzler and J. Klafter. The random walk's guide to anomalous diffusion: A fractional dynamics approach. *Phys.Rep*, 339(1):1–77, 2000. 13

[46] L. Miller and M. Yamamoto. Coefficient inverse problem for a fractional diffusion equation. *Inverse Problems*, 29(7):1–8, 2013. 13, 14

[47] G. Mittag-Leffler. Sur la représentation analytique d'une branche uniforme d'une fonction homogène. *Acta math*, 29:101–182, 1905. 29

[48] D. A. Murio. Implicit finite difference approximation for time fractional diffusion equations. *Computers & Mathematics with Applications*, 56(4):1138–1145, 2008. 44, 98

[49] Z. Odibat. Approximations of fractional integrals and caputo fractional derivatives. *Applied Mathematics and Computation*, 178(2):527–533, 2006. 44

[50] K. Oldham and J. Spanier. *The fractional calculus theory and applications of differentiation and integration to arbitrary order*. Elsevier, 1974. 12

[51] A. Oulmelk, L. Afraites, and A. Hadri. An inverse problem of identifying the coefficient in a nonlinear time-fractional diffusion equation. *Computational and Applied Mathematics*, 42(1):65, 2023. 18

[52] A. Oulmelk, L. Afraites, A. Hadri, and M. Nachaoui. An optimal control approach for determining the source term in fractional diffusion equation by different cost functionals. *Applied Numerical Mathematics*, 181:647–664, 2022. 13

[53] A. Oulmelk, M. Srati, L. Afraites, and A. Hadri. Implementation of the admm approach to constrained optimal control problem with a nonlinear time-fractional diffusion equation. *Discrete and Continuous Dynamical Systems - Series S*, 11 2022. 19

[54] M. Pakdaman, A. Ahmadian, S. Effati, S. Salahshour, and D. Baleanu. Solving differential equations of fractional order using an optimization technique based on training artificial neural network. *Applied Mathematics and Computation*, 293:81–95, 2017. 43

[55] G. Pang, L. Lu, and G. E. Karniadakis. fpinns: Fractional physics-informed neural networks. *SIAM Journal on Scientific Computing*, 41(4):A2603–A2626, 2019. 16

[56] I. Podlubny. *Fractional Differential Equations*, volume 198. Academic Press, 1999. 31

[57] I. Podlubny. Fractional differential equations, mathematics in science and engineering, 1999. 12, 29

[58] A. G. C. Ralf Metzler, Jae Hyung Jeon and E. Barkai. Anomalous diffusion models and their properties: non-stationarity, non-ergodicity, and ageing at the centenary of single particle tracking. *Phys. Chem. Chem. Phys*, 16(44):24128–24164, 2014. 13

[59] B. Ross. A brief history and exposition of the fundamental theory of fractional calculus. *Fractional calculus and its applications*, pages 1–36, 1975. 12

[60] K. Sakamoto and M. Yamamoto. Initial value/boundary value problems for fractional diffusion-wave equations and applications to some inverse problems. *J. Math. Anal. Appl*, 382(1):426–447, 2011. 13, 31

[61] S. G. Samko, A. A. Kilbas, O. I. Marichev, et al. *Fractional integrals and derivatives*, volume 1. Gordon and breach science publishers, Yverdon Yverdon-les-Bains, Switzerland, 1993. 33

[62] J. Simon. Compact sets in the space $l^p(0,t;b)$. *Annali di Matematica pura ed applicata*, 146(1):65–96, 1986. 28

[63] I. M. Sokolov and J. Klafter. From diffusion to anomalous diffusion: a century after einstein's brownian motion. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 15(2):026103, 2005. 13

[64] Y. Song, X. Yuan, and H. Yue. Implementation of the admm to parabolic optimal control problems with control constraints and beyond. *arXiv preprint arXiv:2005.01582*, 2020. 16, 17, 88

[65] T. Wei and J. Wang. Determination of robin coefficient in a fractional diffusion problem. *Appl. Math. Model*, 40(17–18):7948–7961, 2016. 13

[66] C. X. Xianjuan Li. Existence and uniqueness of the weak solution of the space-time fractional diffusion equation and a spectral method approximation. *Commun. Comput. Phys.*, 8(5):1016–1051, 2010. 32, 55

[67] M. Yamamoto and Y. Zhang. Conditional stability in determining a zeroth-order coefficient in a half-order fractional diffusion equation by a carleman estimate. *Inverse Problems*, 28(10):1–10, 2012. 13

[68] M. Yamamoto and Y. Zhang. Conditional stability in determining a zeroth-order coefficient in a half-order fractional diffusion equation by a carleman estimate. *Inverse problems*, 28(10):105010, 2012. 13

[69] X. Yan and T. Wei. Inverse space-dependent source problem for a time-fractional diffusion equation by an adjoint problem approach. *J. Inverse Ill-Posed Probl*, 65(20-21):1–16, 2019. 13

[70] X. Yan and T. Wei. Determine a space-dependent source term in a time fractional diffusion-wave equation. *Acta Applicandae Mathematicae*, 165, 02 2020. 13

[71] P. P. Yong-Ki Ma and A. Deiveegan. Optimization method for determining the source term in fractional diffusion equation. *Mathematics and Computers in Simulation*, 155:168–176, 2019. 32

[72] A. Zeghal. Existence results for inverse problems associated with a nonlinear parabolic equation. *J. Math. Anal. Appl*, 272(1):240–248, 2002. 14

[73] E. Zeidler. *Nonlinear functional analysis and its applications: III: variational methods and optimization*. Springer Science & Business Media, 2013. 26, 28

[74] D. Zhang, L. Lu, L. Guo, and G. E. Karniadakis. Quantifying total uncertainty in physics-informed neural networks for solving forward and inverse stochastic problems. *Journal of Computational Physics*, 397:108850, 2019. 17

[75] Y. Zhang and X. Xu. Inverse source problem for a fractional diffusion equation. *Inverse Problems*, 27(3):1–13, 2011. 13

[76] X. Zheng, X. Cheng, and R. Gong. A coupled complex boundary method for parameter identification in elliptic problems. *International Journal of Computer Mathematics*, 97(5):998–1015, 2020. 78

[77] P. Zhuang and F. Liu. Implicit difference approximation for the time fractional diffusion equation. *Journal of Applied Mathematics and Computing*, 22(3):87–99, 2006. 44, 98

[78] J.-N. Y. Zui-Cha Deng, Liu Yang and G.-W. Luo. An inverse problem of identifying the coefficient in a nonlinear parabolic equation. *Nonlinear Analysis*, 71:6212–6221, 2009. 14