MINISTRY OF EDUCATION AND TRAINING

VIETNAM ACADEMY OF SCIENCE AND TECHNOLOGY

GRADUATE UNIVERSITY OF SCIENCE AND TECHNOLOGY

DUONG XUAN HIEP

LEARNING MODELS FROM DATA WITH APPLICATIONS TO RIVER WATER QUALITY MODELS

MASTER THESIS IN APPLIED MATHEMATICS

Hanoi - 2023

MINISTRY OF EDUCATION
AND TRAINING

VIETNAM ACADEMY
OF SCIENCE AND TECHNOLOGY

GRADUATE UNIVERSITY OF SCIENCE AND TECHNOLOGY

DUONG XUAN HIEP

LEARNING MODELS FROM DATA WITH APPLICATIONS TO RIVER WATER QUALITY MODEL

MASTER THESIS IN APPLIED MATHEMATICS
Code: 8460112

ADVISOR:
PROF. DR. HABIL. DINH NHO HAO

Hanoi - 2023

Commitment

This work has been completed at Graduate University of Science and Technology, Vietnam Academy of Science and Technology under the supervision of Prof. Dr. Habil. Dinh Nho Hao. I commit hereby that the results in this thesis are of our own research and new.

Author: Duong Xuan Hiep

Acknowledgements

First of all, I want to express my deep gratitude to my supervior Prof. Dr. Habil. Dinh Nho Hao for his fullfil support and valuable advice from the beginning when I take my first steps to the Master program in Hanoi to complete this thesis. This thesis has never been accomplished without his tremendous guidence and assistance. Even though I have only studied under his supervision for two years, I have learned a lot from his introduction about exciting mathematical problems and had opportunities to expose vibrantly mathematical environment. It strengthens my determination in pursuing the mathematics career dramatically.

Secondly, I would like to give my special thanks to Assoc. Prof. Pham Quy Muoi, who is my undergraduate supervisor and who always encourages me to further my academic path in mathematics for his endless patience and support. He truly inspires and motivates me to apply Master program in Hanoi. In addition, this work does not complete without a number of his valuable suggestions and discussions in both theoretically and practically.

I am deeply grateful to the financial supports from the International Center for Research and Training in Mathematics under the auspices of UNESCO grant ICRTM03 2021.03 and the Master, PhD Scholarship Programme of Vingroup Innovation Foundation (VINIF), code VINIF.2022.ThS.VTH.05 which are really valuable aid for a period of two learning years.

I would like to thank the Institute of Mathematics, Vietnam Academy of Science and Technology as well as Graduate University of Science and Technology, Vietnam Academy of Science and Technology where I receive excellent conditions in learning and researching during my Master's course.

Last but the most important, I want to thank my family, relatives and close friends who have supported me to overcome touch times. This thesis is dedicated to them.

Hanoi, The second of September, 2023

Duong Xuan Hiep.

CONTENTS

List of Abbreviations and Symbols

$\mathbb{R}^n$	The n-dimensional Euclidean space
$L^2(0,T)$	The space of measurable, squared integrable function in $(0,T)$
$\|x\|_X$	Norm of x in the normed space X
$\|x\|$	Norm of x in the Euclidean space
$\langle x,y\rangle_X$	Inner product of x, y in Hilbert space X
$\langle x,y\rangle$, $x\cdot y$	Inner product of x, y in Euclidean space
$\|M\|$	Norm of matrix M
$\partial f(x)$	Subgradient of f at x
$S_{\omega,p}(\cdot)$	The shrinkage function
$\mathbb{S}_{\omega,p}(\cdot)$	The soft shrinkage operators
$\gamma(\delta)\sim\delta$	$\lim_{\delta\to 0}\gamma(\delta)/\delta = c > 0$
$[a]_+$	Maximum of real number a and 0
$[a]$	Integer part of real number a
BOD	Biochemical oxygen demand
DO	Disolved oxygen
PDEs	Partial differential equations

List of Figures

List of Tables

INTRODUCTION

Motivation

The light of proliferation of machine learning and data science have provided a revolution to analyse and understand the physical phenomena in nature underline complex data and extract patterns in vast multimodal data. However, although there is a rapid development in statisical tools to understand the data that derives from probability and statistics, distilling physical relationships of dynamic processes from big data is still standing in fundamental progresses. Moreover, the abundance of data together with the elusiveness of physical laws and governing equations are truly problematic for climate science, finance, epidemiology, and neuroscience. This leads to an limitation on the ability of data science models to deduce the dynamics and makes this data over the scenario of sample and construction. Therefore, discovering governing equations from data is a central task of scientists and engineers in various fields.

In recent years, environmental pollution has received the great attention of authorities, residents and scientists. The population explosion together with global warming significantly accelerate serious depression in living environment such as the decline of water quality especially in lakes and rivers. The human overactive in agricultural, industrial and daily activities results to contaminate the river water partially and completely. This leads to the study of water quality, mathematical modelling of water quality is the one among such a research.

Since the 19th century, people used aerobic biology to treat the waste water from daily activities. In 1930s, American scientists and engineers [1] have mainly employed it as a crucial technique to investigate and eliminate ordinary water contamination. To estimate water quality, the Biochemical Oxygen Demand (BOD) and Dissolved Oxygen (DO) index are usually used. BOD plays a significant role to measure the degradation of biological organics and the consumed oxygen in the degradable process of microbes. Furthermore, DO indicator is also a important parameter in chemical aspect to consolidate the healthy of water organisms and underwater creatures. In normal condition or unpolluted water, DO rates above the discharge will be near saturation [2, Lecture 19]. The higher DO indicator is, the higher biolocial diversity and the lower number of dead fish. As a result, the authorities and environmental scientists recommend the DO concentration in water is not smaller than 5.0 mg/L most of the time and 4.0 mg/L at all times, see [3, p.221].

$$\underset{\text{BOD}}{CH_2O} + \underset{\text{DO}}{O_2} \xrightarrow{\text{bacteria}} CO_2 + H_2O.$$

The first paper that considers the dissolved oxygen was published by Streeter and Phelps in 1925 when they conducted their research in Ohio River, see [1]. In this paper, Streeter and Phelps demonstrated the decrease of DO indicator which relates to the downstream distance because the dissolved organic BOD degenerates in this area. They also developed a mathematical model to present the phenomenon which was widely known as the Streeter-Phelps equation. In this model, biodegradable organic matter is taken into consideration by the BOD parameter, which is defined as the quantity of oxygen consumed from a unit volume of water by microorganisms during a specified period of time. The other process is the re-aeration oxygen across the water surface due to the turbulent motion of water and to molecular diffusion. It reduces the "oxygen deficit" of water, which is defined as the difference between the saturation oxygen content and the actual dissolved oxygen level. This model consists a set of dynamic equations governing the evolution of the BOD and DO. The BOD-DO model without diffusion process is given by

$$\frac{\partial b}{\partial t} + v(x,t)\frac{\partial b}{\partial x} = -k_1(x,t)b + s_1(x,t) \quad \text{in } (0,X)\times(0,T], \tag{1}$$

$$\frac{\partial d}{\partial t} + v(x,t)\frac{\partial d}{\partial x} = k_1(x,t)b - k_2(x,t)d + s_2(x,t) \quad \text{in } (0,X)\times(0,T], \tag{2}$$

together with initial and boundary conditions

$$b(x,0) = b_0(x)\, d(x,0) = d_0(x) \quad \text{on } (0,X), \tag{3}$$

$$b(0,t) = b_1(t),\ d(0,t) = d_1(t) \quad \text{on } (0,T], \tag{4}$$

where $X > 0$ is the length of river, $T > 0$ is the observational time, $b(x,t)$ represents the BOD concentration and $d(x,t)$ is the DO deficit in the water at time t and position x, respectively. Meanwhile, v is the velocity in the river, k_1 is the de-oxygenation rate (BOD decay rate) and k_2 is the re-aeration coefficient (the re-aeration rate), s_1 is the source of the BOD concentration and $s_2 = k_1 d_s$, where d_s is the saturation value of d. However, when we use the model (1)−(4) to approximate other models, the meaning of the parameters is no longer available. To apply this model, it requires to estimate v, k_1, k_2, s_1 and s_2. There are two basic problems of water quality modelling

1. The first one is the direct problem. This means that given the parameters v, k_1, k_2, s_1, s_2 as well as intial and boundary conditions, we solve numerically the model (1)−(4). There are various numerical algorithms to solve the direct problem such as the upwind scheme, forward time centered space scheme, Lax-Friedrichs' scheme, Leapfrog scheme, Lax-Wendroff scheme,…[4], two-step Lax-Friedrichs'scheme [5].

2. The second one is the inverse problem. It means that given b and d in the whole domain or in some part of the domain, we have to recover the parameters v, k_1, k_2, s_1 and s_2 in the model (1)−(4). These types of coefficents/source identification problems have recently been renamed as "learning models" or "learning partial differential equations/systems", see [6, 7, 8, 9, 10, 11, 12, 13, 14, 15].

In this thesis, we investigate the inverse problem of identifying the parameters of the BOD-DO model (1)−(4). By assuming that we known the data which approximates the solution (b, d) of the model in $[0, X] \times [0, T]$, we want to recover the constant parameters of the BOD-DO model. To this task, the model is rewritten in a general form in which the unknown coefficients to satisfy linear systems. Unfortunately, in this circumstance, these systems are typically over-determined ones. To deal with this situation, we apply l^1-weighted regularization to achieve the target parameters. Furthermore, to solve the optimization problems arising from l^1-weighted regularization we apply some efficient algorithms such as Nesterov's accelerated algorithm [16, 17].

Thesis's structure

Apart from the introduction and conclusion chapter, the thesis is organized as follow:

- Chapter 1 presents some basic definitions on inverse problems and ill-posed problems. Then, we recall some fundamental results in function spaces and optimization which are necessary to prove some results in the thesis. The soft shrinkage operators are also reconsidered in Euclidean space. Some properties about soft shrinkage operators are proven in the detail.

- Chapter 2 introduces a background of learning the partial differential equations from data. In this chapter, we represent an example of nonlinear dynamical system in constructing the underlying system from data by using sparse representation and some challenges in learning models via data and sparse optimization.

- Chapter 3 is the mainly theoretical results in the thesis. We firstly establish the exact solution of BOD-DO model in constant parameters case. Then, the BOD-DO model (1)−(4) is readjusted into a general form in which the unknown coefficients to satisfy linear systems and l^1-weighted regularization is added in the optimization problem for learning unknown parameters. In this chapter, the well-posedness and convergence of learning BOD-DO model by l^1-weighted regularization is also proven under some assumptions.

- Chapter 4 is devoted for the simulation and numerical algorithms. In this chapter, we apply the two-step Lax-Friedrichs method to solve the BOD-DO model in which the solution is used as the data for the inverse problem. After adding noise into the data, we analyse the convergence of Nesterov's accelerated method to find the unknown parameters.

- Chapter 5 presents some numerical experiments in learning BOD-DO model by l^1-weighted regularization and l^1-regularization. The numerical examples confirm the accuracy of our algorithm to solve the forward problem and restore the parameters of a given BOD-DO model.

Chapter 1

PRELIMINARIES

1.1 Ill-posed problems

In his lecture [18] published in the year 1923, Hadamard assumed that a mathematical model for a physical problem has to be well-posed in the sense that

1. There exists a solution of the problem (existence).
2. There is at most one solution of the problem (uniqueness).
3. The solution depends on continuously on the data (stability).

Definition 1.1. *(Well-posedness, see [19]) Let X and Y be normed spaces, $K : X \to Y$ a (linear or nonlinear) mapping. The equation $Kx = y$ is called well-posed if the following holds*

1. *Existence: For every $y \in Y$ there is (at least one) $x \in X$ such that $Kx = y$.*
2. *Uniqueness: For every $y \in Y$ there is at most one $x \in X$ with $Kx = y$.*
3. *Stability: The solution x depends continuously on y, that is, for every sequence $(x^n) \subset X$ with $Kx^n \to Kx\,(n \to \infty)$, it follows that $x^n \to x\,(n \to \infty)$.*

Definition 1.2. *(Ill-posedness, see [19]) Equations for which (at least) one of these properties does not hold are called ill-posed.*

1.2 Some results in function spaces

Definition 1.3. *(See [20])*

1. *The Euclidean norm (l_2-norm) of vector $x = (x_1, x_2, \dots, x_n) \in \mathbb{R}^n$ defined by*
$$\|x\| = \Big(\sum_{i=1}^{n} x_i^2\Big)^{1/2}.$$
2. *The sum-norm (l_1-norm) of vector $x = (x_1, x_2, \dots, x_n) \in \mathbb{R}^n$ defined by*
$$\|x\|_1 = \sum_{i=1}^{n} |x_i|.$$

3. The norm of $m \times n$ *matrix* $A = (a_{ij})$ $(a_{ij} \in \mathbb{R},\ 1 \le i \le m,\ 1 \le j \le n)$ *defined by*

$$\|A\| = \Big(\sum_{i=1}^{m}\sum_{j=1}^{n} a_{ij}^2\Big)^{1/2}.$$

Theorem 1.4. *(Weierstrass theorem, see [21, p.37]) Let* $f : M \to \mathbb{R}$ *be a functional on the compact nonempty subset* M *of a normed space. Then,* f *has a minimum and a maximum on* M*.*

Definition 1.5. *(See [22, Definition 32.11]) Let* $f : X \to \overline{\mathbb{R}} = \mathbb{R} \cup \{\pm\infty\}$ *be a functional in real Banach space* X*. The functional* f^* *in dual space* X^* *of* X *is called a subgradient of* f *at point* x *if and only if*

$$\begin{cases} f(x) \neq \pm\infty, \\ f(y) \ge f(x) + \langle f^*, y - x\rangle_X,\ \forall x \in X, \end{cases}$$

The set of all subgradients of f *at* x *is called the subdifferential* $\partial f(x)$ *at* x*.*

Theorem 1.6. *(See [22, Proposition 32.13]) Let* $f : X \to \mathbb{R}$ *be a convex functional on the real Banach space* X *and* f *is Gâteaux-differentiable at* $x \in X$ *with derivative* $f'(x)$*. Then,* $\partial f(x) = \{f'(x)\}$*.*

Theorem 1.7. *(Minimum principle, see [22, Proposition 32.14]) Let* $f : X \to (-\infty, +\infty]$ *be a functional on the real Banach space* X *with* $f \not\equiv +\infty$*. Then* x *is a solution of the minimum problem*

$$\min_{x \in X} f(x),$$

if and only if

$$0 \in \partial f(x).$$

Theorem 1.8. *(See [23, Theorem 4] and [22, Proposition 32.17]) Let* $f : X \to (-\infty, +\infty]$ *be convex and lower semicontinuous on the real Banach space* X *and* $f \not\equiv +\infty$*. Then, the subgradient* $\partial f : X \to 2^{X^*}$ *is maximal monotone.*

Theorem 1.9. *(Regularization, see [22, Corollary 32.30]) Let* X *be a real reflexive Banach space with dual space* X^**,* C *be a nonempty closed convex subset of* X *and let* $J : X \to X^*$ *be a duality map of space* X*. Assume that* X *and* X^* *are strictly convex and the mapping* $A : C \to 2^{X^*}$ *is maximal monotone. Then, for each* $\alpha > 0$*, the inverse operator*

$$(A + \alpha J)^{-1} : X^* \to X$$

is single-valued, demicontinuous and maximal monotone.

In the following, we recal some fundamental notations and inequalities which is used in the next chapter. The more details could be found in [24].

Definition 1.10. *(See [24, Definition, p.301]) Let X be a real Banach space with norm $\|\cdot\|$ and let $1 \le p \le \infty$. Then, the $L^p(0,T;X)$ consists of all strongly measurable function $u : [0,T] \to X$, i.e.,*

$$L^p(0,T;X) := \{u : [0,T] \to X : u \text{ strongly measurable}, \|u\|_{L^p(0,T;X)} < \infty\},$$

where

$$\|u\|_{L^p(0,T;X)} := \left(\int_0^T \|u(t)\|^p \, dt\right)^{1/p},$$

for $1 \le p < \infty$ and

$$\|u\|_{L^\infty(0,T;X)} := \operatorname*{ess\,sup}_{0 \le t \le T} \|u(t)\| < \infty.$$

Remark 1.11. *In this thesis, we briefly denote $L^p(0,T;\mathbb{R}^n)$ by $L^p(0,T)$.*

Theorem 1.12. *(Hölder's inequality, see [24, p.706]) Let $U \in \mathbb{R}^n$ be an open subset of $\mathbb{R}^n$. Assume that $1 \le p, q \le \infty$ and $\frac{1}{p} + \frac{1}{q} = 1$. Then if $u \in L^p(U)$, $v \in L^q(U)$, we have*

$$\int_U |uv| dx \le \|u\|_{L^p(U)} \|v\|_{L^q(U)}.$$

1.3 Soft shrinkage operators

To deal with sparsity constraints, soft shrinkage operators are usually used in pratical experiments. At first, they were introduced and investigated by [25] for linear inverse problems. After that, these operators are manipulated for gradient methods in nonlinear inverse problems, see [26, 27, 28, 16]. In this thesis, we only consider the soft shrinkage operators in real spaces for the application in Chapter 4.

Definition 1.13. *(Shrinkage function, see [26, 25]) Let $S_{\omega,p} : \mathbb{R} \to \mathbb{R}$ be a function defined by*

$$S_{\omega,p}(x) = \begin{cases} \operatorname{sgn}(x)\big[|x| - \omega\big]_+, & p = 1, \\ G^{-1}_{\omega,p}(x), & 1 < p \le 2, \end{cases} \tag{1.1}$$

where

$$G_{\omega,p}(x) = x + \omega p \operatorname{sgn}(x)|x|^{p-1} \quad \text{and} \quad \big[|x| - \omega\big]_+ = \max(|x| - \omega, 0).$$

Then, $S_{\omega,p}$ is called "shrinkage function".

Definition 1.14. *(Soft shrinkage operators, see [25, 29]) Let $\omega = \{\omega_i\}_{i=1}^N$ with $\omega_i > 0$ for all i. The soft shrinkage operators $\mathbb{S}_{\omega,p} : \mathbb{R}^N \to \mathbb{R}^N$ are defined by*

$$\mathbb{S}_{\omega,p}(x) = \big(S_{\omega_1,p}(x_1), S_{\omega_2,p}(x_2), \ldots, S_{\omega_N,p}(x_N)\big), \tag{1.2}$$

where $x = (x_1, x_2, \ldots, x_N) \in \mathbb{R}^N$ and $S_{\omega_k,p}$ is a shrinkage function which is defined by (1.1).

Next, we have some properties of the soft shrinkage operators which are briefly indicated in [25, Lemma 2.2].

Lemma 1.15. *(Nonexpansion, see [25]) The soft shrinkage operators* $\mathbb{S}_{\omega,p}$ *which are defined by* (1.2) *are non-expansive, i.e.*

$$\|\mathbb{S}_{\omega,p}(x) - \mathbb{S}_{\omega,p}(y)\| \leq \|x - y\|, \quad \forall x, y \in \mathbb{R}^N.$$

Proof. We will prove this result by dividing into two cases:

1. The case $p > 1$: Consider the functional $G_{\tau,p} : \mathbb{R} \to \mathbb{R}\,(\tau, p > 0)$ is given by
$$G_{\tau,p}(t) = t + \tau p \operatorname{sgn}(t)|t|^{p-1}.$$
For all $t_0 > 0$, we have
$$\begin{aligned}\lim_{t\to t_0} \frac{G_{\tau,p}(t) - G_{\tau,p}(t_0)}{t - t_0} &= \lim_{t\to t_0} \frac{(t - t_0) + \tau p(t^{p-1} - t_0^{p-1})}{t - t_0}\\ &= \lim_{t\to t_0} \left(1 + \tau p(t^{p-2} + t^{p-1}t_0 + \ldots + t_0^{p-2})\right)\\ &= 1 + \tau p(p-1)t_0^{p-2}.\end{aligned}$$
This means that $G_{\tau,p}$ is differentiable in $(0, +\infty)$.
Moreover, $G_{\tau,p}$ is an odd function. Then, it is also differentiable in $(-\infty, 0)$.
At the point 0, this follows that
$$\begin{aligned}\lim_{t\to 0^+} \frac{G_{\tau,p}(t) - G_{\tau,p}(0)}{t - 0} &= \lim_{t\to 0^+} \frac{t + \tau p t^{p-1}}{t} = 1,\\ \lim_{t\to 0^-} \frac{G_{\tau,p}(t) - G_{\tau,p}(0)}{t - 0} &= \lim_{t\to 0^-} \frac{t - \tau p(-t)^{p-1}}{t} = 1.\end{aligned}$$
Hence, $G_{\tau,p}$ is a smooth function in $\mathbb{R}$ with its derivative defined by
$$G'_{\tau,p}(t) = \begin{cases} 1 + \tau p(p-1)t^{p-2}, & \text{if } t > 0,\\ 1, & \text{if } t = 0,\\ 1 + \tau p(p-1)(-t)^{p-2}, & \text{if } t < 0.\end{cases}$$
From the derivative of $G_{\tau,p}$, it follows that for each $\tau, p > 0$,
$$G'_{\tau,p}(t) \geq 1,\ \forall t \in \mathbb{R}.$$
Return to the lemma 1.15. Let $u = \mathbb{S}_{\omega,p}(x)$ and $v = \mathbb{S}_{\omega,p}(y)$.
By (1.2), for all $i \in \{1, 2, \ldots, N\}$ we have
$$u_i = S_{\omega_i,p}(x_i) \quad \text{and} \quad v_i = S_{\omega_i,p}(y_i).$$
By (1.1), they are equivalent to
$$x_i = G_{\omega_i,p}(u_i) \quad \text{and} \quad y_i = G_{\omega_i,p}(v_i).$$
By the mean value theorem, for all $u_i, v_i \in \mathbb{R}$, there exist η_i between u_i and v_i such that
$$G_{\omega_i,p}(u_i) - G_{\omega_i,p}(v_i) = G'_{\omega_i,p}(\eta_i)(u_i - v_i).$$

According to above evaluation, $G'_{\omega_i,p}$ is uniformly bounded from below by 1, we get

$$|G_{\omega_i,p}(u_i) - G_{\omega_i,p}(v_i)| = |G'_{\omega_i,p}(\eta_i)||u_i - v_i| \geq |u_i - v_i|, \, \forall i \in \{1, 2, \ldots, N\}.$$

This means that for all $i \in \{1, 2, \ldots, N\}$, we have the following inequality

$$|S_{\omega_i,p}(x_i) - S_{\omega_i,p}(y_i)| = |u_i - v_i| \leq |G_{\omega_i,p}(u_i) - G_{\omega_i,p}(v_i)| = |x_i - y_i|.$$

As a result,

$$\|\mathbb{S}_{\omega,p}(x) - \mathbb{S}_{\omega,p}(y)\|^2 = \sum_{i=1}^{N} |S_{\omega_i,p}(x_i) - S_{\omega_i,p}(y_i)|^2 \leq \sum_{i=1}^{N} |x_i - y_i|^2 = \|x - y\|^2.$$

Then, for $p > 1$ we have

$$\|\mathbb{S}_{\omega,p}(x) - \mathbb{S}_{\omega,p}(y)\| \leq \|x - y\|, \quad \forall x, y \in \mathbb{R}^N.$$

2. The case $p = 1$: For $\tau > 0$ by (1.1), we have

$$S_{\tau,1}(t) = \operatorname{sgn}(t)\big[|x| - \omega\big]_+ = \begin{cases} t - \tau, & \text{if } t \geq \tau, \\ 0, & \text{if } -\tau < t < \tau, \\ t + \tau, & \text{if } t \leq -\tau. \end{cases}$$

Without loss of generality, we may assume that $t_1 < t_2$ for all $t_1, t_2 \in \mathbb{R}$. We will prove that

$$|S_{\tau,1}(t_1) - S_{\tau,1}(t_2)| \leq |t_1 - t_2|.$$

If $t_1 < t_2 \leq -\tau$ or $\tau \leq t_1 < t_2$, we have

$$|S_{\tau,1}(t_1) - S_{\tau,1}(t_2)| = |t_1 - t_2|.$$

If $t_1 \leq -\tau < t_2 < \tau$, it yields

$$t_1 - t_2 < t_1 + \tau \quad \text{and} \quad t_1 + \tau < t_2 - t_1.$$

Then,

$$|S_{\tau,1}(t_1) - S_{\tau,1}(t_2)| = |t_1 + \tau| < |t_1 - t_2|.$$

If $t_1 \leq -\tau < \tau \leq t_2$, we have

$$t_1 + \tau \leq 0 \quad \text{and} \quad -(t_2 - \tau) \leq 0.$$

It follows that

$$t_1 - t_2 + 2\tau \leq 0.$$

Then,

$$|S_{\tau,1}(t_1) - S_{\tau,1}(t_2)| = |t_1 - t_2 + 2\tau| = t_2 - t_1 - 2\tau < t_2 - t_2 = |t_1 - t_2|.$$

If $-\tau < t_1 < t_2 < \tau$, we have

$$0 = |S_{\tau,1}(t_1) - S_{\tau,1}(t_2)| < |t_1 - t_2|.$$

If $-\tau < t_1 < \tau \le t_2$, it gives

$$|S_{\tau,1}(t_1) - S_{\tau,1}(t_2)| = |-(t_2 - \tau)| = t_2 - \tau < t_2 - t_1 = |t_1 - t_2|.$$

Thus, we always have the following inequality

$$|S_{\tau,1}(t_1) - S_{\tau,1}(t_2)| \le |t_1 - t_2|, \quad \forall t_1, t_2 \in \mathbb{R}.$$

By (1.2), it follows that

$$\|\mathbb{S}_{\omega,1}(x) - \mathbb{S}_{\omega,1}(y)\| \le \|x - y\|, \quad \forall x, y \in \mathbb{R}^N.$$

□

The next lemma is crucial to prove the convergence of the Algorithm 4.2 in Chapter 4.

Lemma 1.16. *Let $\{x^n\}$, $\{y^n\}$ and $\{z^n\}$ be sequences in $\mathbb{R}^N$ and $\{\alpha_n\}$ be a positive sequence such that*

$$x^n = \mathbb{S}_{\alpha_n\omega,p}(y^n - \alpha_n z^n).$$

Assume that

$$\lim_{n\to\infty} x^n = \lim_{n\to\infty} y^n = x^*, \quad \lim_{n\to\infty} z^n = z^* \quad \text{and} \quad \lim_{n\to\infty} \alpha_n = \alpha^* > 0.$$

Then,

$$x^* = \mathbb{S}_{\alpha^*\omega,p}(x^* - \alpha^* z^*).$$

Proof. We seperate the proof into two cases.

1. For $p > 1$, by the assumption, we have

$$\lim_{n\to\infty} x_i^n = \lim_{n\to\infty} y_i^n = x_i^*, \quad \lim_{n\to\infty} z_i^n = z_i^*, \quad \forall i \in \{1, 2, \ldots, N\}.$$

Following (1.1) and (1.2), it yields

$$x_i^n + \alpha_n\omega_i p \operatorname{sgn}(x_i^n)|x_i^n|^{p-1} = y_i^n - \alpha_n z_i^n, \quad \forall i \in \{1, 2, \ldots, N\}.$$

Letting $n \to \infty$, we get

$$x_i^* + \alpha^*\omega_i p \operatorname{sgn}(x_i^*)|x_i^*|^{p-1} = x_i^* - \alpha^* z_i^*, \quad \forall i \in \{1, 2, \ldots, N\}.$$

Hence,

$$x^* = \mathbb{S}_{\alpha^*\omega,p}(x^* - \alpha^* z^*).$$

2. For $p = 1$, the assumption leads

$$x_i^n = \operatorname{sgn}(y_i^n - \alpha_n z_i^n)\max(|y_i^n - \alpha_n z_i^n| - \alpha_n\omega_i, 0), \forall i \in \{1, 2, \ldots, N\}. \tag{1.3}$$

Define

$$\Gamma_1 = \{i : |x_i^* - \alpha^* z_i^*| - \alpha^*\omega_i > 0\},$$
$$\Gamma_2 = \{i : |x_i^* - \alpha^* z_i^*| - \alpha^*\omega_i < 0\},$$
$$\Gamma_3 = \{i : |x_i^* - \alpha^* z_i^*| - \alpha^*\omega_i = 0\}.$$

Recall that

$$\lim_{n\to\infty}(y_i^n - \alpha_n z_i^n) = x_i^* - \alpha^* z_i^*\,,\ \lim_{n\to\infty}|y_i^n - \alpha_n z_i^n| - \alpha_n \omega_i = |x_i^* - \alpha^* z_i^*| - \alpha^* \omega_i.$$

If $i \in \Gamma_1$,for n large enough we have

$$\mathrm{sgn}(y_i^n - \alpha_n z_i^n) = \mathrm{sgn}(x_i^* - \alpha^* z_i^*) \quad \text{ and } \quad |y_i^n - \alpha_n z_i^n| - \alpha_n \omega_i > 0.$$

It deduces from (1.3) that

$$x_i^* = \mathrm{sgn}(x_i^* - \alpha^* z_i^*)\max(|x_i^* - \alpha^* z_i^*| - \alpha^* \omega_i, 0) = S_{\alpha^*\omega_i,1}(x_i^* - \alpha_i^* z_i^*),\ \forall i \in \Gamma_1.$$

If $i \in \Gamma_2$, we have $|y_i^n - \alpha_n z_i^n| - \alpha_n \omega_i < 0$ when n is large enough. From (1.3), it yields $x_i^n = 0$ when n is large enough. This gives $x_i^* = 0$ and implies

$$x_i^* = S_{\alpha^*\omega_i,1}(x_i^* - \alpha^* z_i^*),\ \forall i \in \Gamma_2.$$

If $i \in \Gamma_3$, $\mathrm{sgn}(y_i^n - \alpha_n z_i^n) = \mathrm{sgn}(x_i^* - \alpha^* z_i^*) = 1$ when n is large enough. Moreover,

$$\lim_{n\to\infty}(|y_i^n - \alpha_n z_i^n| - \alpha_n \omega_i) = |x_i^* - \alpha^* z_i^*| - \alpha^* \omega_i = 0.$$

It yields

$$x_i^* = \lim_{n\to\infty} x_i^n = \lim_{n\to\infty}[\mathrm{sgn}(y_i^n - \alpha_n z_i^n)\max(|y_i^n - \alpha_n z_i^n| - \alpha_n \omega_i, 0)] = 0.$$

Hence,

$$x_i^* = S_{\alpha^*\omega_i,1}(x_i^* - \alpha^* z_i^*),\ \forall i \in \Gamma_3.$$

In summary, we have

$$x_i^* = S_{\alpha^*\omega_i,1}(x_i^* - \alpha^* z_i^*),\ \forall i \in \Gamma_1 \cup \Gamma_2 \cup \Gamma_3,$$

which is equivalent to

$$x^* = \mathbb{S}_{\alpha^*\omega,1}(x^* - \alpha^* z^*).$$

□

Remark 1.17. *In the rest of this thesis, we succinctly denote* $\mathbb{S}_{\omega,1}$ *by* $\mathbb{S}_\omega$*.*

Chapter 2

LEARNING MODELS FROM DATA

In this chapter, we discuss the general background of determining the governing equations from the data which is dramatically developed and researched in recent years based on the proliferation of machine learning and data science.

2.1 Learning models from experimental data

The history of modeling dynamical systems from observational data experienced a prolonged and fruitful history and resulted in powerful techniques for system identification, with a rich history going back at least as far as the time of Kepler and Newton and the discovery of the laws of planetary motion [6]. A number of methods have beyonded the original purposes for understanding the complexity of flexible structures, such as the Hubble space telescope or the international space station. The resulting models have been widely applied in nearly every branch of engineering and applied mathematics, most notably for model-based feedback control. Nevertheless, to identify the model system, we need more assumptions on the form of the model and most of them often result in linear dynamics with their restricted effectiveness to small amplitude transient perturbations around a fixed point of the dynamics.

Learning partial differential equations/systems is a newly developed research direction in mathematics and applications since it can provide fundamental models in the physical and life sciences. Partial differential equations/systems also model complex behaviors in the social sciences, for example conservation laws for traffic flow, systems of equations for population dynamics, epidemic models and financial markets. The original discoveries of these equations/systems typically required a grasp of mathematics, an understanding of theory, and supportive evidence from experimental data. Any model is performed based on the data and the functions of the process under consideration if they are known. Data are abundant in many branches of science. Therefore, extracting governing equations from data is a central challenge in many diverse areas of science and engineering. The new approach to learn the underline behaviour of physical laws from experimental data is resulted from [30, 31]. According to [32, 33], system identification can be separated into two different categories:

1. Methods that are accurately reflected observed dynamics using black box functions (neural networks).
2. Methods that recover closed forms and expression of the dynamics by ordinary differential equations (ODEs) and partial differential equations (PDEs).

The first one is aimed at algorithmic models which need not reflecting the true mechanisms but are accurate in prediction. The second type of methods may assume a specific model for the data with known mechanisms. The advantages of the second approach include:

1. PDE-based models rely on well-established physical principles.
2. The number of parameters to be estimated in PDE-based models is usually much smaller than in neural network-based models. This reduces the need for a huge amount of sampling data.

For these reasons, researchers are currently considering the second approach in learning some models in which they study some theoretical aspects of learning PDE-based models using sparse optimization techniques. Learning PDE-based models has recently been attracted the attention of researchers (see [6, 7, 8, 9, 10, 11, 12, 13, 14]), this is an approach to modelling real life processes based on physical rules (PDEs) and data collected during the cause. To our knowledge, the first papers in this direction were published in about 2015 and mostly devoted to processes in fluid dynamics. There no research has been done in Vietnam.

2.2 Learning models via sparse optimization

A typical method in data-driven modeling based on PDE-models is the use of the observation of the state along the time and its derivative with respect to time to form a regression problem. For many systems of interest, the PDEs consist of only a few derivatives, making it sparse in the space of possible functions. The resulting sparse model identification inherently balances model complexity with accuracy, avoiding overfitting the model to data. Recently, some authors have proposed to use sparse optimization method to identify the parameters of the underlying PDEs in [6, 7, 11, 12, 14, 15].

Let considering an example, see [7], about the nonlinear dynamical system

$$\frac{d}{dt}\mathbf{x}(t) = \dot{\mathbf{x}}(t) = \mathbf{f}(\mathbf{x}(t)). \tag{2.1}$$

The vector $\mathbf{x}(t) = [x_1(t) \quad x_2(t) \quad \cdots \quad x_n(t)]^T \in \mathbb{R}^n$ represents the state of the system at time t, and the nonlinear function $\mathbf{f}(\mathbf{x}(t))$ represents the dynamic constraints that define the equations of motion of the system. To determine the function $\mathbf{f}$ from data, we collect a time-history of the state $\mathbf{x}(t)$ and either measure the derivative $\dot{\mathbf{x}}(t)$ or approximate it numerically from $\mathbf{x}$. The data is sampled at several times $t_1, t_2, \ldots, t_m$ and arranged into two large matrices:

$$\mathbf{X} = \begin{bmatrix} \mathbf{x}^T(t_1) \\ \mathbf{x}^T(t_2) \\ \vdots \\ \mathbf{x}^T(t_m) \end{bmatrix} = \begin{bmatrix} x_1(t_1) & x_2(t_1) & \cdots & x_n(t_1) \\ x_1(t_2) & x_2(t_2) & \cdots & x_n(t_2) \\ \vdots & \vdots & \ddots & \vdots \\ x_1(t_m) & x_2(t_m) & \cdots & x_n(t_m) \end{bmatrix},$$

$$\dot{\mathbf{X}} = \begin{bmatrix} \dot{\mathbf{x}}^T(t_1) \\ \dot{\mathbf{x}}^T(t_2) \\ \vdots \\ \dot{\mathbf{x}}^T(t_m) \end{bmatrix} = \begin{bmatrix} \dot{x}_1(t_1) & \dot{x}_2(t_1) & \cdots & \dot{x}_n(t_1) \\ \dot{x}_1(t_2) & \dot{x}_2(t_2) & \cdots & \dot{x}_n(t_2) \\ \vdots & \vdots & \ddots & \vdots \\ \dot{x}_1(t_m) & \dot{x}_2(t_m) & \cdots & \dot{x}_n(t_m) \end{bmatrix}.$$

Next, an augmented library $\boldsymbol{\Theta}(\mathbf{X})$ is constructed from possible nonlinear functions of the columns of $\mathbf{X}$. For example, $\boldsymbol{\Theta}(\mathbf{X})$ may consist of constant, polynomial and trigonometric terms, i.e.,

$$\boldsymbol{\Theta}(\mathbf{X}) = \begin{bmatrix} | & | & | & | & & | & | & | & & \\ \mathbf{1} & \mathbf{X} & \mathbf{X}^{P_2} & \mathbf{X}^{P_3} & \cdots & \sin(\mathbf{X}) & \cos(\mathbf{X}) & \sin(2\mathbf{X}) & \cos(2\mathbf{X}) & \cdots \\ | & | & | & | & & | & | & | & & \end{bmatrix}.$$

Here, higher polynomials are denoted as $\mathbf{X}^{P_2}, \mathbf{X}^{P_3}$, etc. For example, $\mathbf{X}^{P_2}$ denotes the quadratic nonlinearities in the state variable $\mathbf{x}$, given by:

$$\mathbf{X}^{P_2} = \begin{bmatrix} x_1^2(t_1) & x_1(t_1)x_2(t_1) & \cdots & x_2^2(t_1) & x_2(t_1)x_3(t_1) & \cdots & x_n^2(t_1) \\ x_1^2(t_2) & x_1(t_2)x_2(t_2) & \cdots & x_2^2(t_2) & x_2(t_2)x_3(t_2) & \cdots & x_n^2(t_2) \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ x_1^2(t_m) & x_1(t_m)x_2(t_m) & \cdots & x_2^2(t_m) & x_2(t_m)x_3(t_m) & \cdots & x_n^2(t_m) \end{bmatrix}.$$

Each column of $\boldsymbol{\Theta}(\mathbf{X})$ represents a possible function for the right hand side of (2.1). We may construct a sparse regression problem to determine the sparse vectors of coefficients $\boldsymbol{\Xi} = [\boldsymbol{\xi}_1 \quad \boldsymbol{\xi}_2 \quad \cdots \quad \boldsymbol{\xi}_n]$.

$$\dot{\mathbf{X}} = \boldsymbol{\Theta}(\mathbf{x})\boldsymbol{\Xi}.$$

Each column $\boldsymbol{\xi}_k$ of $\boldsymbol{\Xi}$ represents a sparse vector of coefficients determining which terms areactive in the right hand side for one of the row equations $\dot{\mathbf{x}}_k = \mathbf{f}_k(\mathbf{x})$ in (2.1), i.e.,

$$\dot{\mathbf{x}}_k = \mathbf{f}_k(\mathbf{x}) = \boldsymbol{\Theta}(\mathbf{x}^T)\boldsymbol{\xi}_k. \tag{2.2}$$

A model will provide an accurate model fit in (2.2) with as few terms as possible in $\boldsymbol{\Xi}$. To overcome the over-determined system, we use the l^1-regularized sparse regression:

$$\boldsymbol{\xi}_k = \text{argmin}_{\boldsymbol{\xi}'_k} \left\| \dot{\mathbf{x}}_k - \boldsymbol{\Theta}(\mathbf{x}^T)\boldsymbol{\xi}'_k \right\| + \lambda \left\| \boldsymbol{\xi}'_k \right\|_1.$$

Here, $\|\cdot\|$ is the Euclidean norm and $\|\cdot\|_1$ is the l_1-norm.

2.3 The challenges of learning models from data and sparse optimization

As noted above, the PDEs describing the processes consist of only a few derivatives, making it sparse in the space of possible functions. However, if data are collected from a numerical discretization or from experimental measurements on a spatial grid, the state dimension may be extremely large. Hence, the sparse regression problems are computationally challenging. One of the ways to overcome this difficulty is to approximate the large problems by low-rank approximation using dimensionality reduction techniques, such as the proper orthogonal decomposition (see [34, 35, 36]). Further, the following questions are the main challenges in the sparse regression framework for identifying the PDE-based models:

1. How much and which kinds of data do we need for the model identification problem to uniquely solvable?

2. How do we efficiently solve the discretized model?

3. How do we solve the nonlinear optimization problem resulting from the regression problem?

4. How do we choose regularization parameters in the regression problem to obtain accurate models or to overcome the overfitting problem?

Therefore, research in these problems is desired.

Chapter 3

LEARNING CONSTANT PARAMETERS IN THE BOD-DO MODEL WITH l^1-WEIGHTED REGULARIZATION

In this chapter, we establish the exact solution for BOD-DO system (1)−(4) for the case of constant parameters. We then apply l^1-weighted regularization technique for learning BOD-DO systems.

3.1 The solution of the BOD-DO model with constant parameters

In this paragraph, we establish the exact solution of the BOD-DO system with constant parameters. The BOD-DO model with initial and boundary conditions is given by (1)−(4) with the constant parameters v, k_1, k_2, s_1, s_2 has the form

$$\frac{\partial b}{\partial t} + v\frac{\partial b}{\partial x} = -k_1 b + s_1 \qquad \text{in } (0,X)\times(0,T], \tag{3.1}$$

$$\frac{\partial d}{\partial t} + v\frac{\partial d}{\partial x} = k_1 b - k_2 d + s_2 \qquad \text{in } (0,X)\times(0,T], \tag{3.2}$$

$$b(x,0) = b_0(x),\ d(x,0) = d_0(x) \qquad \text{on } (0,X), \tag{3.3}$$

$$b(0,t) = b_1(t),\ d(0,t) = d_1(t) \qquad \text{on } (0,T]. \tag{3.4}$$

Theorem 3.1. *Let v, k_1, k_2, s_1, s_2 be real constants and $v > 0$. Assume that b_0, d_0, b_1, d_1 are continuous functions, $b_0(0) = b_1(0)$ and $d_0(0) = d_1(0)$. Then, system (3.1)–(3.4) has a unique solution that is given by*

(i) For $x - vt \in [0, X]$,

$$b(x,t) = \frac{s_1}{k_1} + e^{-k_1 t}\left[-\frac{s_1}{k_1} + b_0(x-vt)\right]$$

$$d(x,t) = \begin{cases} \dfrac{s_1+s_2}{k_1} + e^{-k_1 t}\left[d_0(x-vt) - \dfrac{s_1+s_2}{k_1} + \big(-s_1 + k_1 b_0(x-vt)\big)t\right], \\ \hfill \text{if } k_1 = k_2, \\ \dfrac{s_1+s_2}{k_2} + \dfrac{e^{-k_1 t}}{k_2-k_1}\big(-s_1 + k_1 b_0(x-vt)\big) + e^{-k_2 t}\left[d_0(x-vt) \right. \\ \qquad \left. -\dfrac{s_1+s_2}{k_2} + \dfrac{1}{k_1-k_2}\big(-s_1+k_1 b_0(x-vt)\big)\right], \ \text{if } k_1 \neq k_2. \end{cases}$$

(ii) For $t - \dfrac{x}{v} \in [0,T]$,

$$b(x,t) = \frac{s_1}{k_1} + e^{\frac{-k_1 x}{v}}\left[-\frac{s_1}{k_1} + b_1\left(t - \frac{x}{v}\right)\right]$$

$$d(x,t) = \begin{cases} \dfrac{s_1+s_2}{k_1} + e^{\frac{-k_1 x}{v}}\Bigg[d_1\left(t-\dfrac{x}{v}\right) - \dfrac{s_1+s_2}{k_1} \\ \qquad\qquad + \left(-s_1 + k_1 b_1\left(t - \dfrac{x}{v}\right)\right)\dfrac{x}{v}\Bigg], \ \textit{if } k_1 = k_2, \\ \dfrac{s_1+s_2}{k_2} + \dfrac{e^{\frac{-k_1 x}{v}}}{k_2 - k_1}\left(-s_1 + k_1 b_1\left(t-\dfrac{x}{v}\right)\right) + e^{\frac{-k_2 x}{v}}\Bigg[d_1\left(t - \dfrac{x}{v}\right) \\ \qquad -\dfrac{s_1+s_2}{k_1} + \dfrac{1}{k_1-k_2}\left(-s_1 + k_1 b_1\left(t-\dfrac{x}{v}\right)\right)\Bigg], \ \textit{if } k_1 \neq k_2. \end{cases}$$

Proof. - First, we consider the initial-boundary problem

$$\begin{cases} \dfrac{\partial b}{\partial t} + v\dfrac{\partial b}{\partial x} = -k_1 b + s_1 & \text{in } (0,X)\times(0,T), \\ b(x,0) = b_0(x) & \text{for } x \in [0,X], \\ b(0,t) = b_1(t) & \text{for } t \in [0,T]. \end{cases} \tag{3.5}$$

Let $z(s) = b(x+vs, t+s)$ $(s \in \mathbb{R})$. Then from the first equation of (3.5) it follows that

$$\frac{\partial b(x+vs,t+s)}{\partial(t+s)} + v\frac{\partial b(x+vs,t+s)}{\partial(x+vs)} = -k_1 b(x+vs,t+s) + s_1,$$

and so

$$\frac{\partial z(s)}{\partial s} + k_1 z(s) = s_1.$$

Multiplying above equation by $e^{k_1 s}$ we have that

$$\frac{\partial(e^{k_1 s}z(s))}{\partial s} = e^{k_1 s}\frac{\partial z(s)}{\partial s} + k_1 e^{k_1 s} z(s) = e^{k_1 s} s_1.$$

It yields

$$e^{k_1 s} z(s) = \int e^{k_1 s} s_1 ds + c_1 = \frac{1}{k_1} e^{k_1 s} s_1 + c,$$

then

$$z(s) = \frac{s_1}{k_1} + ce^{-k_1 s},$$

or

$$b(x+vs, t+s) = \frac{s_1}{k_1} + ce^{-k_1 s}, \quad s \in \mathbb{R}. \tag{3.6}$$

From the second condition of (3.5) and (3.6), letting $s = -t$ we get

$$b_0(x - vt) = b(x-vt, 0) = \frac{s_1}{k_1} + ce^{k_1 t}, \quad x - vt \in [0,X].$$

Then

$$c = e^{-k_1 t}\left[-\frac{s_1}{k_1} + b_0(x - vt)\right].$$

From (3.6) letting $s = 0$, we conclude that

$$b(x,t) = \frac{s_1}{k_1} + e^{-k_1 t}\left[-\frac{s_1}{k_1} + b_0(x-vt)\right], \quad x - vt \in [0,X].$$

From the third condition of (3.5) and (3.6), letting $s = \dfrac{-x}{v}$ we get

$$b_1\left(t - \frac{x}{v}\right) = b\left(0, t - \frac{x}{v}\right) = \frac{s_1}{k_1} + ce^{\frac{-k_1 x}{v}}, \quad t - \frac{x}{v} \in [0, T].$$

Then

$$c = e^{\frac{-k_1 x}{v}}\left[-\frac{s_1}{k_1} + b_1\left(t - \frac{x}{v}\right)\right].$$

From (3.6) letting $s = 0$, we conclude that

$$b(x,t) = \frac{s_1}{k_1} + e^{\frac{-k_1 x}{v}}\left[-\frac{s_1}{k_1} + b_1\left(t - \frac{x}{v}\right)\right], \quad t - \frac{x}{v} \in [0, T].$$

Hence,

$$b(x,t) = \begin{cases} \dfrac{s_1}{k_1} + e^{-k_1 t}\left[-\dfrac{s_1}{k_1} + b_0(x - vt)\right], \text{ if } x - vt \in [0, X], \\ \dfrac{s_1}{k_1} + e^{\frac{-k_1 x}{v}}\left[-\dfrac{s_1}{k_1} + b_1\left(t - \dfrac{x}{v}\right)\right], \text{ if } t - \dfrac{x}{v} \in [0, T]. \end{cases}$$

Now, we consider the initial-boundary problem in DO model

$$\begin{cases} \dfrac{\partial d}{\partial t} + v\dfrac{\partial d}{\partial x} = k_1 b - k_2 d + s_2 & \text{in } (0, X) \times (0, T), \\ d(x, 0) = d_0(x) & \text{for } x \in [0, X], \\ d(0, t) = d_1(t) & \text{for } t \in [0, T]. \end{cases} \tag{3.7}$$

Let $u(s) = d(x + vs, t + s)\,(s \in \mathbb{R})$. Then it follows from the first equation of (3.7) that

$$\frac{\partial d(x + vs, t + s)}{\partial(t + s)} + v\frac{\partial d(x + vs, t + s)}{\partial(x + vs)} = k_1 b(x + vs, t + s) - k_2 d(x + vs, t + s) + s_2,$$

and so

$$\frac{\partial u(s)}{\partial s} + k_2 u(s) = k_1 z(s) + s_2.$$

Multiplying the above equation by $e^{k_2 s}$ we have that

$$\frac{\partial(e^{k_2 s}u(s))}{\partial s} = e^{k_2 s}\frac{\partial u(s)}{\partial s} + k_2 e^{k_2 s}u(s) = e^{k_2 s}k_1 z(s) + e^{k_2 s}s_2$$

It yields

$$e^{k_2 s}u(s) = \int [e^{k_2 s}k_1 z(s) + e^{k_2 s}s_2]ds + m_1 = \int e^{k_2 s}k_1 z(s)ds + \frac{s_2}{k_2}e^{k_2 s} + m_2.$$

Then

$$u(s) = e^{-k_2 s}\int e^{k_2 s}k_1 z(s)ds + \frac{s_2}{k_2} + m_2 e^{-k_2 s}. \tag{3.8}$$

For this equation we consider two cases:

+ *Case 1:* For $x - vt \in [0, X]$, according to (3.6) and taking into account that

$$c = e^{-k_1 t}\left[-\frac{s_1}{k_1} + b_0(x - vt)\right],$$

we have

$$k_1 z(s) = s_1 + e^{-k_1(t+s)}\big[-s_1 + k_1 b_0(x - vt)\big].$$

Substituting the last to (3.8), we get

$$
\begin{aligned}
u(s) &= e^{-k_2 s}\int e^{k_2 s}\Big[s_1 + e^{-k_1(t+s)}\big(-s_1 + k_1 b_0(x-vt)\big)\Big]ds + \frac{s_2}{k_2} + m_2 e^{-k_2 s}\\
&= e^{-k_2 s}\Big[s_1\int e^{k_2 s}ds + e^{-k_1 t}\big(-s_1 + k_1 b_0(x-vt)\big)\int e^{(k_2-k_1)s}ds\Big]\\
&\qquad\qquad + \frac{s_2}{k_2} + m_2 e^{-k_2 s}\\
&= \begin{cases} e^{-k_2 s}\Big[\dfrac{s_1 e^{k_2 s}}{k_2} + e^{-k_1 t}\big(-s_1 + k_1 b_0(x-vt)\big)s + m_3\Big] + \dfrac{s_2}{k_2} + m_2 e^{-k_2 s}, \\ \qquad\qquad \text{if } k_1 = k_2,\\ e^{-k_2 s}\Big[\dfrac{s_1 e^{k_2 s}}{k_2} + e^{-k_1 t}\big(-s_1 + k_1 b_0(x-vt)\big)\dfrac{e^{(k_2-k_1)s}}{k_2-k_1} + m_3\Big] \\ \qquad\qquad + \dfrac{s_2}{k_2} + m_2 e^{-k_2 s}, \text{ if } k_1 \neq k_2. \end{cases}\\
&= \begin{cases} \dfrac{s_1+s_2}{k_1} + e^{-k_1(t+s)}\big(-s_1 + k_1 b_0(x-vt)\big)s + m e^{-k_1 s}, \text{ if } k_1 = k_2,\\ \dfrac{s_1+s_2}{k_2} + \dfrac{e^{-k_1(t+s)}}{k_2-k_1}\big(-s_1 + k_1 b_0(x-vt)\big) + m' e^{-k_2 s}, \text{ if } k_1 \neq k_2. \end{cases}
\end{aligned}
$$

or

$$
d(x+vs,t+s) = \begin{cases} \dfrac{s_1+s_2}{k_1} + e^{-k_1(t+s)}\big(-s_1 + k_1 b_0(x-vt)\big)s + m e^{-k_1 s}, \text{ if } k_1 = k_2,\\ \dfrac{s_1+s_2}{k_2} + \dfrac{e^{-k_1(t+s)}}{k_2-k_1}\big(-s_1 + k_1 b_0(x-vt)\big) + m' e^{-k_2 s}, \text{ if } k_1 \neq k_2. \end{cases} \tag{3.9}
$$

- From (3.9) letting $s = -t$ we get

$$
d_0(x-vt) = d(x-vt,0) = \begin{cases} \dfrac{s_1+s_2}{k_1} + \big(-s_1 + k_1 b_0(x-vt)\big)(-t) + m e^{k_1 t}, \\ \qquad\qquad \text{if } k_1 = k_2,\\ \dfrac{s_1+s_2}{k_2} + \dfrac{1}{k_2-k_1}\big(-s_1 + k_1 b_0(x-vt)\big) + m' e^{k_2 t}, \\ \qquad\qquad \text{if } k_1 \neq k_2. \end{cases}
$$

  It follows that

$$
m = e^{-k_1 t}\Big[d_0(x-vt) - \frac{s_1+s_2}{k_1} + \big(-s_1 + k_1 b_0(x-vt)\big)t\Big],
$$

  and

$$
m' = e^{-k_2 t}\Big[d_0(x-vt) - \frac{s_1+s_2}{k_2} + \frac{1}{k_1-k_2}\big(-s_1 + k_1 b_0(x-vt)\big)\Big].
$$

- Letting $s = 0$ in (3.9), we get

$$d(x,t) = \begin{cases} \dfrac{s_1+s_2}{k_1} + e^{-k_1 t}\Big[d_0(x-vt) - \dfrac{s_1+s_2}{k_1} + \big(-s_1 + k_1 b_0(x-vt)\big)t\Big], \\ \hfill \text{if } k_1 = k_2, \\ \dfrac{s_1+s_2}{k_2} + \dfrac{e^{-k_1 t}}{k_2-k_1}\big(-s_1 + k_1 b_0(x-vt)\big) + e^{-k_2 t}\Big[d_0(x-vt) - \dfrac{s_1+s_2}{k_2} \\ \qquad + \dfrac{1}{k_1-k_2}\big(-s_1 + k_1 b_0(x-vt)\big)\Big], \text{ if } k_1 \neq k_2. \end{cases}$$

+*Case 2:* For $t - \dfrac{x}{v} \in [0,T]$ according to (3.6) and

$$c = e^{\frac{-k_1 x}{v}}\left[-\frac{s_1}{k_1} + b_1\left(t - \frac{x}{v}\right)\right],$$

we have

$$k_1 z(s) = s_1 + e^{-k_1(\frac{x}{v}+s)}\left[-s_1 + k_1 b_1\left(t - \frac{x}{v}\right)\right].$$

Substituting the last to (3.8), we get

$$\begin{aligned} u(s) &= e^{-k_2 s}\int e^{k_2 s}\left[s_1 + e^{-k_1(\frac{x}{v}+s)}\left(-s_1 + k_1 b_1\left(t-\frac{x}{v}\right)\right)\right]ds + \frac{s_2}{k_2} + n_2 e^{-k_2 s} \\ &= e^{-k_2 s}\left[s_1\int e^{k_2 s}ds + e^{-k_1\frac{x}{v}}\left(-s_1 + k_1 b_1\left(t-\frac{x}{v}\right)\right)\int e^{(k_2-k_1)s}ds\right] \\ &\qquad + \frac{s_2}{k_2} + n_2 e^{-k_2 s} \\ &= \begin{cases} e^{-k_2 s}\left[\dfrac{s_1 e^{k_2 s}}{k_2} + e^{-k_1\frac{x}{v}}\left(-s_1 + k_1 b_1\left(t-\dfrac{x}{v}\right)\right)s + n_3\right] + \dfrac{s_2}{k_2} + n_2 e^{-k_2 s}, \\ \hfill \text{if } k_1 = k_2, \\ e^{-k_2 s}\left[\dfrac{s_1 e^{k_2 s}}{k_2} + e^{-k_1\frac{x}{v}}\left(-s_1 + k_1 b_1\left(t-\dfrac{x}{v}\right)\right)\dfrac{e^{(k_2-k_1)s}}{k_2-k_1} + n_3\right] \\ \qquad + \dfrac{s_2}{k_2} + n_2 e^{-k_2 s}, \text{ if } k_1 \neq k_2. \end{cases} \\ &= \begin{cases} \dfrac{s_1+s_2}{k_1} + e^{-k_1(\frac{x}{v}+s)}\left(-s_1 + k_1 b_1\left(t-\dfrac{x}{v}\right)\right)s + ne^{-k_1 s}, \text{ if } k_1 = k_2, \\ \dfrac{s_1+s_2}{k_2} + \dfrac{e^{-k_1(\frac{x}{v}+s)}}{k_2-k_1}\left(-s_1 + k_1 b_1\left(t-\dfrac{x}{v}\right)\right) + n'e^{-k_2 s}, \text{ if } k_1 \neq k_2. \end{cases} \end{aligned}$$

or

$$d(x+vs, t+s) = \begin{cases} \dfrac{s_1+s_2}{k_1} + e^{-k_1(\frac{x}{v}+s)}\left(-s_1 + k_1 b_1\left(t-\dfrac{x}{v}\right)\right)s + ne^{-k_1 s}, \text{ if } k_1 = k_2, \\ \dfrac{s_1+s_2}{k_2} + \dfrac{e^{-k_1(\frac{x}{v}+s)}}{k_2-k_1}\left(-s_1 + k_1 b_1\left(t-\dfrac{x}{v}\right)\right) + n'e^{-k_2 s}, \text{ if } k_1 \neq k_2. \end{cases} \tag{3.10}$$

- Letting $s = \dfrac{-x}{v}$ in (3.10), we get

$$d_1\Big(t-\frac{x}{v}\Big) = d\Big(0, t-\frac{x}{v}\Big) = \begin{cases} \dfrac{s_1+s_2}{k_1} + \Big(-s_1 + k_1 b_1\Big(t-\dfrac{x}{v}\Big)\Big)\dfrac{-x}{v} + n e^{\frac{k_1 x}{v}}, \\ \hfill \text{if } k_1 = k_2, \\ \dfrac{s_1+s_2}{k_2} + \dfrac{1}{k_2-k_1}\Big(-s_1 + k_1 b_1\Big(t-\dfrac{x}{v}\Big)\Big) \\ \hfill + n' e^{\frac{k_2 x}{v}}, \text{ if } k_1 \neq k_2. \end{cases}$$

It follows that

$$n = e^{\frac{-k_1 x}{v}}\left[d_1\Big(t-\frac{x}{v}\Big) - \frac{s_1+s_2}{k_1} + \Big(-s_1 + k_1 b_1\Big(t-\frac{x}{v}\Big)\Big)\frac{x}{v}\right],$$

and

$$n' = e^{\frac{-k_2 x}{v}}\left[d_1\Big(t-\frac{x}{v}\Big) - \frac{s_1+s_2}{k_1} + \frac{1}{k_1-k_2}\Big(-s_1 + k_1 b_1\Big(t-\frac{x}{v}\Big)\Big)\right].$$

- From (3.10) letting $s = 0$, we get

$$d(x,t) = \begin{cases} \dfrac{s_1+s_2}{k_1} + e^{\frac{-k_1 x}{v}}\Big[d_1\Big(t-\dfrac{x}{v}\Big) - \dfrac{s_1+s_2}{k_1} \\ \hfill + \Big(-s_1 + k_1 b_1\Big(t-\dfrac{x}{v}\Big)\Big)\dfrac{x}{v}\Big], \text{ if } k_1 = k_2, \\ \dfrac{s_1+s_2}{k_2} + \dfrac{e^{\frac{-k_1 x}{v}}}{k_2-k_1}\Big(-s_1 + k_1 b_1\Big(t-\dfrac{x}{v}\Big)\Big) + e^{\frac{-k_2 x}{v}}\Big[d_1\Big(t-\dfrac{x}{v}\Big) \\ \qquad - \dfrac{s_1+s_2}{k_1} + \dfrac{1}{k_1-k_2}\Big(-s_1 + k_1 b_1\Big(t-\dfrac{x}{v}\Big)\Big)\Big], \text{ if } k_1 \neq k_2. \end{cases}$$

□

3.2 Learning the BOD-DO model by l^1-weighted regularization

We suppose that the general model for a BOD-DO water quality model has the form

$$\begin{aligned} \frac{\partial b}{\partial t} &= \alpha_1 + \alpha_2 b + \alpha_3 d + \alpha_4 \frac{\partial b}{\partial x} + \alpha_5 \frac{\partial d}{\partial x}, \\ \frac{\partial d}{\partial t} &= \beta_1 + \beta_2 b + \beta_3 d + \beta_4 \frac{\partial b}{\partial x} + \beta_5 \frac{\partial d}{\partial x}. \end{aligned}$$

This system can be rewritten by the product of (b, d, b_x, d_x)-dependent terms with fixed coefficients

$$b_t = [1 \quad b \quad d \quad b_x \quad d_x]\alpha, \tag{3.11}$$

$$d_t = [1 \quad b \quad d \quad b_x \quad d_x]\beta, \tag{3.12}$$

where $\alpha = [\alpha_1 \quad \alpha_2 \quad \alpha_3 \quad \alpha_4 \quad \alpha_5]^T$ and $\beta = [\beta_1 \quad \beta_2 \quad \beta_3 \quad \beta_4 \quad \beta_5]^T$.
These equations hold at all points (x, t) in the domain $(0, X) \times (0, T)$. Each of

(b, d, b_x, d_x)-dependent terms in system (3.11)−(3.12) represents a potential feature that decribes the intrinsic dynamics of data. The feature vectors $f_1(t)$, $f_2(t)$, $f_3(t)$, $f_4(t)$, $f_5(t)$ are defined by

$$f_1(t) = \begin{bmatrix} | \\ 1 \\ | \end{bmatrix}, \; f_2(t) = \begin{bmatrix} | \\ b(t) \\ | \end{bmatrix}, \; f_3(t) = \begin{bmatrix} | \\ d(t) \\ | \end{bmatrix}, \; f_4(t) = \begin{bmatrix} | \\ b_x(t) \\ | \end{bmatrix}, \; f_5(t) = \begin{bmatrix} | \\ d_x(t) \\ | \end{bmatrix},$$

where each time-stepping feature vector is the vectorization of the term in system (3.11)−(3.12). Each component of the feature vector $f_i(t)$ is the evaluation of the corresponding term from system (3.11)−(3.12) at a pre-determined point in space, i.e.

$$f_1(t) = \begin{bmatrix} 1 \\ | \\ 1 \\ | \\ 1 \end{bmatrix}, \quad f_2(t) = \begin{bmatrix} b(x_1,t) \\ | \\ b(x_j,t) \\ | \\ b(x_n,t) \end{bmatrix}, \quad f_3(t) = \begin{bmatrix} d(x_1,t) \\ | \\ d(x_j,t) \\ | \\ d(x_n,t) \end{bmatrix},$$

$$f_4(t) = \begin{bmatrix} b_x(x_1,t) \\ | \\ b_x(x_j,t) \\ | \\ b_x(x_n,t) \end{bmatrix}, \quad f_5(t) = \begin{bmatrix} d_x(x_1,t) \\ | \\ d_x(x_j,t) \\ | \\ d_x(x_n,t) \end{bmatrix}.$$

The collection of feature vectors defines the feature matrix $F(t) = [f_i(t)]$,

$$F(t) = [f_1(t) \quad f_2(t) \quad f_3(t) \quad f_4(t) \quad f_5(t)].$$

Set

$$V_1(t) = \begin{bmatrix} | \\ b_t(t) \\ | \end{bmatrix}, \quad V_2(t) = \begin{bmatrix} | \\ d_t(t) \\ | \end{bmatrix}.$$

Then for $t \geq 0$, the following system of equations holds

$$\left[V_1(t) | V_2(t) \right] = F(t) \left[\alpha | \beta \right]. \tag{3.13}$$

Here, $V_i(t)$ and $F(t)$ are known while the coefficient vectors α, β are unknown.

In partical scenario, the exact data of solution (b, d) is unknown. The system (3.13) may not produce a unique solution (α, β) beacause we only measure the noisy data (b^δ, d^δ). Then, $V_1(t), V_2(t)$ and $F(t)$ are only approximated by $V_1^\delta(t), V_2^\delta(t)$ and $F^\delta(t)$. As a result, the problem (3.13) is ill-posed. To carry out this situation, we use the l^1-weighted regularization and solve the following minimization problem

$$\min_{\alpha,\beta \in \mathbb{R}^5} \left\{ \int_0^T \left\| V_1^\delta(t) - F^\delta(t)\alpha \right\|^2 dt + \int_0^T \left\| V_2^\delta(t) - F^\delta(t)\beta \right\|^2 dt + \gamma_1 \Phi_1(\alpha) + \gamma_2 \Phi_2(\beta) \right\}. \tag{3.14}$$

where the norm $\|\cdot\|$ here is the Euclidean norm space in $\mathbb{R}^n$ and $\gamma_1, \gamma_2 > 0$ are regularization parameters and

$$\Phi_1(\alpha) = \sum_{j=1}^{5} \omega_j^1 |\alpha_j|, \quad \Phi_2(\beta) = \sum_{j=1}^{5} \omega_j^2 |\beta_j|, \tag{3.15}$$

with $\omega_j^1, \omega_j^2 \geq \omega_0 > 0$ being weighted numbers.
Note that if $\omega_j^1, \omega_j^2 = 1$ for all j then it is called l^1-regularization, otherwise it is called l^1-weighted regularization. The choice of weighted numbers is very important. They affect the quality of recovered parameters and depend on a priori information on the solution of the model. We remark that with only one data set we could successfully recover the parameters in the BOD-DO model if the weights are set suitably. We will illustrate and analysis the effect of choosing the weights in Chapter 5.

3.3 The well-posedness and convergence of learning the BOD-DO model by l^1-weighted regularization

In the following, we will investigate the well-posedness problem (3.14) and the convergence of its solution to system (3.13). To this end, we need the following assumption:

Assumption 3.2. *Suppose that V_1, V_2 and F in problem (3.13) belong to $L^2(0,T)$. Let V_1^δ, V_2^δ and F^δ in $L^2(0,T)$ be, respectively, the noisy data of V_1, V_2 and F satisfying*

$$\int_0^T \left\|V_1^\delta(t) - V_1(t)\right\|^2 dt < \delta, \int_0^T \left\|V_2^\delta(t) - V_2(t)\right\|^2 dt < \delta, \int_0^T \left\|F^\delta(t) - F(t)\right\|^2 dt < \delta, \tag{3.16}$$

where $\delta > 0$ is a known noise level.

To overcome the ill-posedness, stable numerical techniques are required, among them Tikhonov and sparse regularization are the most well-known. There are a number of works to deal with problem (3.14) such as [37, 38, 39].

Theorem 3.3. *Let $J_1^\delta : \mathbb{R}^5 \to \mathbb{R}$ be defined by*

$$J_1^\delta(\alpha) = \int_0^T \left\|V_1^\delta(t) - F^\delta(t)\alpha\right\|^2 dt. \tag{3.17}$$

Then the functional J_1^δ is continuous and convex. Moreover, the derivative of J_1^δ is continuous.

Proof. (i) First, we prove that the functional J_1^δ is continuous. Indeed, suppose that the sequence $\{\alpha^n\}_{n\in\mathbb{N}} \subset \mathbb{R}^5$ converges to α in $\mathbb{R}^5$. Using Hölder's inequality, we have

$$
\begin{aligned}
|J_1^\delta(\alpha^n) - J_1^\delta(\alpha)| &= \int_0^T \left\|V_1^\delta(t) - F^\delta(t)\alpha^n\right\|^2 dt - \int_0^T \left\|V_1^\delta(t) - F^\delta(t)\alpha\right\|^2 dt \\
&= \int_0^T F^\delta(t)(\alpha^n - \alpha)\cdot\Big(F^\delta(t)\alpha^n + F^\delta(t)\alpha - 2V^\delta(t)\Big)dt \\
&\le \Big(\int_0^T \left\|F^\delta(t)(\alpha^n-\alpha)\right\|^2 dt\Big)^{1/2}\cdot \\
&\qquad \Big(\int_0^T \left\|F^\delta(t)\alpha^n + F^\delta(t)\alpha - 2V^\delta(t)\right\|^2 dt\Big)^{1/2}. \qquad (3.18)
\end{aligned}
$$

By the definition of matrix norm, it implies that

$$
\begin{aligned}
\int_0^T \left\|F^\delta(t)(\alpha^n-\alpha)\right\|^2 dt &\le \int_0^T \left\|F^\delta(t)\right\|^2 \|\alpha^n-\alpha\|^2 dt \\
&\le \|\alpha^n-\alpha\|^2 \int_0^T \left\|F^\delta(t)\right\|^2 dt \to 0 \text{ as } n\to\infty. \qquad (3.19)
\end{aligned}
$$

By using the Cauchy-Schwartz inequality, the second term of the right-side in inequality (3.18) is estimated as follows

$$
\begin{aligned}
&\int_0^T \left\|F^\delta(t)\alpha^n + F^\delta(t)\alpha - 2V^\delta(t)\right\|^2 dt \\
&\le 9\bigg(\int_0^T \left\|F^\delta(t)\alpha^n\right\|^2 dt + \int_0^T \left\|F^\delta(t)\alpha\right\|^2 dt + 4\int_0^T \left\|V^\delta(t)\right\|^2 dt\bigg) \\
&\le 9\bigg(\|\alpha^n\|^2\int_0^T \left\|F^\delta(t)\right\|^2 dt + \|\alpha\|^2\int_0^T \left\|F^\delta(t)\right\|^2 dt + 4\int_0^T \left\|V^\delta(t)\right\|^2 dt\bigg) \\
&\le M. \qquad (3.20)
\end{aligned}
$$

From (3.18)−(3.20), it follows that

$$|J_1^\delta(\alpha^n) - J_1^\delta(\alpha)| \to 0 \ \text{ as } n\to\infty.$$

Thus the functional J_1^δ is continuous.

(ii) We can see that $J_1^\delta(\cdot)$ is differentiable and its derivative is defined by

$$(J_1^\delta)'(\alpha) = 2\int_0^T \Big(F^\delta(t)\alpha - V_1^\delta(t)\Big)\cdot F^\delta(t)dt.$$

Then for all $\vartheta\in\mathbb{R}^5$, it yields

$$
\begin{aligned}
(J_1^\delta)'(\alpha)\vartheta &= 2\int_0^T \Big(F^\delta(t)\alpha - V_1^\delta(t)\Big)\cdot F^\delta(t)\vartheta dt \\
&= 2\bigg(\alpha\int_0^T F^\delta(t)\cdot F^\delta(t)\vartheta dt - \int_0^T V_1^\delta(t)\cdot F^\delta(t)\vartheta dt\bigg).
\end{aligned}
$$

This implies that the second derivative of $J_1^\delta(\cdot)$ is given by

$$(J_1^\delta)''(\alpha)(\vartheta,\eta) = 2\int_0^T F^\delta(t)\vartheta\cdot F^\delta(t)\eta dt.$$

Therefore,

$$(J_1^\delta)''(\alpha)(\vartheta,\vartheta) = 2\int_0^T F^\delta(t)\vartheta \cdot F^\delta(t)\vartheta dt \geq 0, \forall \vartheta \in \mathbb{R}^5.$$

This means that the function $J_1^\delta(\cdot)$ is convex on $\mathbb{R}^5$.

(iii) Suppose that there exists a sequence $\{\alpha^n\}_{n\in\mathbb{N}} \subset \mathbb{R}^5$ converges to $\alpha \in \mathbb{R}^5$. Then

$$\begin{aligned}\left\|(J_1^\delta)'(\alpha^n) - (J_1^\delta)'(\alpha)\right\| &= 2\left\|\int_0^T F^\delta(t)(\alpha^n - \alpha)\cdot F^\delta(t)dt\right\| \\ &\leq 2\left\|\alpha^n - \alpha\right\| \int_0^T \left\|F^\delta(t)\right\|^2 dt \to 0 \text{ as } n \to \infty.\end{aligned}$$

Hence, $(J_1^\delta)'$ is continuous on $\mathbb{R}^5$.

□

Remark 3.4. *From Theorem 3.3, we also have the continuity and convexity of the functional* $J_2^\delta : \mathbb{R}^5 \to \mathbb{R}$ *be defined by*

$$J_2^\delta(\beta) = \int_0^T \left\|V_2^\delta(t) - F^\delta(t)\beta\right\|^2 dt.$$

Futhermore, $(J_2^\delta)'$ *is continuous on* $\mathbb{R}^5$.

In the following, we consider some properties of the function Φ_i in (3.15). The first property is obvious and the second one is a special case as $p = 1$ in [40, Remark 3]. For convenience, we recall it with the detailed proof.

Theorem 3.5. *The functions* $\Phi_i : \mathbb{R}^5 \to \mathbb{R}\,(i = 1,2)$ *defined by (3.15) have the following properties*

(1) Φ_i *is non-negative, convex and continuous* $(i = 1,2)$.

(2) There exists a positive constant C *such that*

$$\Phi_i(\vartheta) \geq C\left\|\vartheta\right\|, \forall \vartheta \in \mathbb{R}^5 \text{ and } i = 1,2,$$

where $\left\|\cdot\right\|$ *is the Euclidean norm in* $\mathbb{R}^5$. *This implies that* Φ_i *is coercive, i.e.,* $\Phi_i(\vartheta) \to \infty$ *as* $\left\|\vartheta\right\| \to \infty$.

Proof. It is trivial that $\Phi_i (i = 1,2)$ are non-negative by their definition. Furthermore, since $|x|$ is convex and continuous, so is Φ_i.
Since

$$\Phi_i(\vartheta) \geq \omega_0 \sum_{j=1}^5 |\vartheta_j| \geq \omega_0 \sqrt{\sum_{j=1}^5 \vartheta_j^2} = \omega_0\|\vartheta\|,$$

it is coercive. □

Theorem 3.6. *(Existence) Problem (3.14) has at least one solution.*

Proof. Since the function $J_1^\delta, J_2^\delta, \Phi_1$ and Φ_2 are convex and continuous on $\mathbb{R}^5$, they are lower semi-continuous in $\mathbb{R}^5$. It implies the objective funtion of Problem (3.14) is convex and lower semi-continuous on $\mathbb{R}^{10}$. Futhermore, the objective function is also coercive by Theorem 3.5. As a results, Problem (3.14) has at least one solution [41]. $\square$

In Theorem 2.1, we have proven that problem (3.14) has at least one solution, but the solution may not unique as the functions J_1^δ and J_2^δ are convex but may not strongly convex

Theorem 3.7. *(Stability) For a fixed regularization $\gamma_1, \gamma_2 > 0$. If the sequence $\{V_1^n\}, \{V_2^n\}$ and $\{F^n\}$ converge to V_1^δ, V_2^δ and F^δ in $L^2(0,T)$, respectively and*

$$(\alpha^n, \beta^n) \in \arg\min_{\alpha,\beta\in\mathbb{R}^5}\{J_1^n(\alpha) + J_2^n(\beta) + \gamma_1\Phi_1(\alpha) + \gamma_2\Phi_2(\beta)\},$$

then there exist a subsequence $\{(\alpha^{n_k}, \beta^{n_k})\}$ of $\{(\alpha^n, \beta^n)\}$ and a minimizer $(\alpha^\delta_{\gamma_1,\gamma_2}, \beta^\delta_{\gamma_1,\gamma_2})$ of (3.14) such that

$$\left\|(\alpha^{n_k}, \beta^{n_k}) - (\alpha^\delta_{\gamma_1,\gamma_2}, \beta^\delta_{\gamma_1,\gamma_2})\right\| \to 0.$$

In addition, if the minimizer $(\alpha^\delta_{\gamma_1,\gamma_2}, \beta^\delta_{\gamma_1,\gamma_2})$ is unique, then the sequence $\{(\alpha_n, \beta_n)\}$ converges to $(\alpha^\delta_{\gamma_1,\gamma_2}, \beta^\delta_{\gamma_1,\gamma_2})$.

Proof. By the definition of α^n and β^n

$$\begin{aligned}
&J_1^n(\alpha^n) + J_2^n(\beta^n) + \gamma_1\Phi_1(\alpha^n) + \gamma_2\Phi_2(\beta^n)\\
=&\int_0^T \|V_1^n(t) - F^n(t)\alpha^n\|^2\,dt + \int_0^T \|V_2^n(t) - F^n(t)\beta^n\|^2\,dt + \gamma_1\Phi_1(\alpha^n) + \gamma_2\Phi_2(\beta^n)\\
\le&\int_0^T \|V_1^n(t) - F^n(t)\alpha\|^2\,dt + \int_0^T \|V_2^n(t) - F^n(t)\beta\|^2\,dt + \gamma_1\Phi_1(\alpha) + \gamma_2\Phi_2(\beta)\\
\le&C\Bigg(\int_0^T \left\|V_1^n(t) - V_1^\delta(t)\right\|^2 dt + \left\|V_1^\delta\right\|^2_{L^2(0,T)} + \|\alpha\|^2\int_0^T \left\|F^\delta(t)\right\|^2 dt\\
&+ \|\alpha\|^2\int_0^T \left\|F^n(t) - F^\delta(t)\right\|^2 dt + \int_0^T \left\|V_2^n(t) - V_2^\delta(t)\right\|^2 dt + \left\|V_2^\delta\right\|^2_{L^2(0,T)}\\
&+ \|\beta\|^2\int_0^T \left\|F^\delta(t)\right\|^2 dt + \|\beta\|^2\int_0^T \left\|F^n(t) - F^\delta(t)\right\|^2 dt\Bigg) + \gamma_1\Phi_1(\alpha) + \gamma_2\Phi_2(\beta).
\end{aligned} \tag{3.21}$$

Since the sequence $\{V_1^n\}, \{V_2^n\}$ and $\{F^n\}$ converge to V_1^δ, V_2^δ and F^δ in $L^2(0,T)$, the sequences $\{\Phi(\alpha^n)\}$ and $\{\Phi(\beta^n)\}$ are bounded. According to Theorem 3.5, the sequence $\{\alpha^n\}$ and $\{\beta^n\}$ are also bounded in $\mathbb{R}^5$. Hence, there exits subsequences $\{\alpha^{n_k}\}$ of $\{\alpha^n\}$ and $\{\beta^{n_k}\}$ of $\{\beta^n\}$ such that $\{\alpha^{n_k}\}$ and $\{\beta^{n_k}\}$ converge to $\alpha^\delta_{\gamma_1,\gamma_2}$ and $\beta^\delta_{\gamma_1,\gamma_2}$, respectively. In addition, since $J_1^\delta(\cdot)$, $J_2^\delta(\cdot)$, $\Phi_1(\cdot)$ and $\Phi_2(\cdot)$ are continuous on $\mathbb{R}^5$, we have

$$J_1^\delta(\alpha^\delta_{\gamma_1,\gamma_2}) = \liminf_k J_1^\delta(\alpha^{n_k}),\ J_2^\delta(\beta^\delta_{\gamma_1,\gamma_2}) = \liminf_k J_2^\delta(\beta^{n_k}) \tag{3.22}$$

and

$$\Phi_1(\alpha^\delta_{\gamma_1,\gamma_2}) = \liminf_k \Phi_1(\alpha^{n_k}),\ \Phi_2(\beta^\delta_{\gamma_1,\gamma_2}) = \liminf_k \Phi_2(\beta^{n_k}). \tag{3.23}$$

Furthermore,

$$\begin{aligned} J_1^\delta(\alpha^{n_k}) = J_1^{n_k}(\alpha^{n_k}) + \Bigg[\int_0^T & \Big(V_1^\delta(t) + V_1^{n_k}(t)\Big) \cdot \Big(V_1^\delta(t) - V_1^{n_k}(t)\Big) dt \\ - \int_0^T & \alpha^{n_k}\Big(F^\delta(t) + F^{n_k}(t)\Big) \cdot \Big(V_1^\delta(t) - V_1^{n_k}(t)\Big) dt \\ + \int_0^T & \alpha^{n_k}\Big(V_1^\delta(t) + V_1^{n_k}(t)\Big) \cdot \Big(F^{n_k}(t) - F^\delta(t)\Big) dt \\ - \int_0^T & \|\alpha^{n_k}\|^2 \Big(F^\delta(t) - F^{n_k}(t)\Big) \cdot \Big(F^\delta(t) + F^{n_k}(t)\Big) dt \Bigg]. \end{aligned} \tag{3.24}$$

Since $V_1^{n_k} \to V_1^\delta$ and $F^{n_k} \to F^\delta$ in $L^2(0,T)$, the term in brackets on the right-hand side of (3.24) goes to zero when k tends to ∞. Therefore, we get

$$\liminf_k J_1^{n_k}(\alpha^{n_k}) = \liminf_k J_1^\delta(\alpha^{n_k}),\ \limsup_k J_1^{n_k}(\alpha^{n_k}) = \limsup_k J_1^\delta(\alpha^{n_k}). \tag{3.25}$$

Similarly, we have

$$\liminf_k J_2^{n_k}(\beta^{n_k}) = \liminf_k J_2^\delta(\beta^{n_k}),\ \limsup_k J_2^{n_k}(\beta^{n_k}) = \limsup_k J_2^\delta(\beta^{n_k}). \tag{3.26}$$

From (3.22), (3.23), (3.25) and (3.26), we obtain

$$\begin{aligned} & J_1^\delta(\alpha^\delta_{\gamma_1,\gamma_2}) + J_2^\delta(\beta^\delta_{\gamma_1,\gamma_2}) + \gamma_1\Phi_1(\alpha^\delta_{\gamma_1,\gamma_2}) + \gamma_2\Phi_2(\beta^\delta_{\gamma_1,\gamma_2}) \\ =& \liminf_k J_1^\delta(\alpha^{n_k}) + \liminf_k J_2^\delta(\beta^{n_k}) + \gamma_1 \liminf_k \Phi_1(\alpha^{n_k}) + \gamma_2 \liminf_k \Phi_2(\beta^{n_k}) \\ =& \liminf_k J_1^{n_k}(\alpha^{n_k}) + \liminf_k J_2^{n_k}(\beta^{n_k}) + \gamma_1 \liminf_k \Phi_1(\alpha^{n_k}) + \gamma_2 \liminf_k \Phi_2(\beta^{n_k}) \\ \le& \liminf_k \Bigg(J_1^{n_k}(\alpha^{n_k}) + J_2^{n_k}(\beta^{n_k}) + \gamma_1\Phi_1(\alpha^{n_k}) + \gamma_2\Phi_2(\beta^{n_k}) \Bigg) \\ \le& \limsup_k \Bigg(J_1^{n_k}(\alpha^{n_k}) + J_2^{n_k}(\beta^{n_k}) + \gamma_1\Phi_1(\alpha^{n_k}) + \gamma_2\Phi_2(\beta^{n_k}) \Bigg) \\ \le& \limsup_k J_1^{n_k}(\alpha^{n_k}) + \limsup_k J_2^{n_k}(\beta^{n_k}) + \gamma_1 \limsup_k \Phi_1(\alpha^{n_k}) + \gamma_2 \limsup_k \Phi_2(\beta^{n_k}) \\ =& \limsup_k J_1^\delta(\alpha^{n_k}) + \limsup_k J_2^\delta(\beta^{n_k}) + \gamma_1 \limsup_k \Phi_1(\alpha^{n_k}) + \gamma_2 \limsup_k \Phi_2(\beta^{n_k}) \\ =& J_1^\delta(\alpha) + J_2^\delta(\beta) + \gamma_1\Phi_1(\alpha) + \gamma_2\Phi_2(\beta), \quad \forall \alpha, \beta \in \mathbb{R}^5. \end{aligned} \tag{3.27}$$

This means that $(\alpha^\delta_{\gamma_1,\gamma_2}, \beta^\delta_{\gamma_1,\gamma_2})$ is a minimizer of (3.14).

In addition we have

$$\left\|\alpha^{n_k} - \alpha^\delta_{\gamma_1,\gamma_2}\right\| \to 0 \quad \text{and} \quad \left\|\beta^{n_k} - \beta^\delta_{\gamma_1,\gamma_2}\right\| \to 0, \qquad \text{as } k \to \infty.$$

Thus,

$$\left\|(\alpha^{n_k}, \beta^{n_k}) - (\alpha^\delta_{\gamma_1,\gamma_2}, \beta^\delta_{\gamma_1,\gamma_2})\right\| \to 0, \qquad \text{as } k \to \infty.$$

In the cases the minimizer $(\alpha^\delta_{\gamma_1,\gamma_2}, \beta^\delta_{\gamma_1,\gamma_2})$ is unique, the convergence of the original sequence $\{(\alpha^{n_k}, \beta^{n_k})\}$ to $(\alpha^\delta_{\gamma_1,\gamma_2}, \beta^\delta_{\gamma_1,\gamma_2})$ follows by a subsequence argument. □

Lemma 3.8. *The set*

$$\Pi(V) = \{(\alpha,\beta) \in \mathbb{R}^5 \times \mathbb{R}^5 : F(t)\left[\alpha|\beta\right] = \left[V_1(t)|V_2(t)\right] = V(t)\}$$

is non-empty, closed and convex. Then, there exists a solution (α^+,β^+) *of the problem*

$$\min_{(\alpha,\beta)\in\Pi(V)} \{\Phi_1(\alpha) + \Phi_2(\beta)\},$$

which is called a Φ_1, Φ_2*-minimizing solution of problem (3.13).*

Proof. It is trivial that $\Pi(V)$ is nonempty and convex by the definition of set $\Pi(V)$ and the assumption of problem. Suppose that the sequence $\{(\alpha^n,\beta^n)\} \subset \Pi(V)$ converges to (α,β) in $\mathbb{R}^5 \times \mathbb{R}^5$. We prove that $(\alpha,\beta) \in \Pi(V)$. Note that $F \in L^2(0,T)$, we have

$$0 \le \int_0^T \left\|V(t) - F(t)\left[\alpha|\beta\right]\right\|^2 dt = \int_0^T \left\|F(t)\left[\alpha^n - \alpha|\beta^n - \beta\right]\right\|^2 dt$$
$$\le C\Big(\|\alpha^n - \alpha\|^2 + \|\beta^n - \beta\|^2\Big) \int_0^T \|F(t)\|^2\, dt \to 0.$$

Then,

$$V(t) = F(t)\left[\alpha|\beta\right].$$

Hence, $\Pi(V)$ is closed in $\mathbb{R}^5 \times \mathbb{R}^5$.

According to Theorem 3.5, $\Phi_1(\cdot)$ and $\Phi_2(\cdot)$ are continuous and weakly coercive over nonempty, closed convex set $\Pi(V)$ then it yields that $\Phi_1(\cdot) + \Phi_2(\cdot)$ has a global minimum point over $\Pi(V)$, see [42, Theorem 2.32], [43, Proposition 1.2]. □

In the following theorem, the notation $\gamma(\delta) \sim \delta$ means that there exists a constant $c > 0$ such that $\lim_{\delta\to 0} \gamma(\delta)/\delta = c$.

Theorem 3.9. *(Convergence) Suppose that the operator equation*

$$F(t)\left[\alpha|\beta\right] = \left[V_1(t)|V_2(t)\right]$$

has a solution in $\mathbb{R}^5 \times \mathbb{R}^5$ *and* $\gamma_1, \gamma_2 : \mathbb{R}^+ \to \mathbb{R}^+$ *satisfy*

$$\gamma_1(\delta) \sim \delta \text{ and } \gamma_2(\delta) \sim \delta.$$

Let $\delta_n \to 0$, $\|V_1^n(t) - V_1(t)\|^2_{L^2(0,T)} \le \delta_n$, $\|V_2^n(t) - V_2(t)\|^2_{L^2(0,T)} \le \delta_n$ *and* $\|F^n(t) - F(t)\|^2_{L^2(0,T)} \le \delta_n$. *Moreover, let* $\gamma_{1,n} = \gamma_1(\delta_n)$, $\gamma_{2,n} = \gamma_2(\delta_n)$ *and*

$$(\alpha^n,\beta^n) \in \arg\min_{\alpha,\beta\in\mathbb{R}^5}\{J_1^n(\alpha) + J_2^n(\beta) + \gamma_{1,n}\Phi_1(\alpha) + \gamma_{2,n}\Phi_2(\beta)\}.$$

Then, there exist a Φ_1, Φ_2*-minimizing solution* (α^+,β^+) *of* $F(t)\left[\alpha|\beta\right] = \left[V_1(t)|V_2(t)\right]$ *and a subsequence of* $\{(\alpha^n,\beta^n)\}$ *which converges to* (α^+,β^+) *on* $\mathbb{R}^5 \times \mathbb{R}^5$.

Proof. We note that

$$\min_{\alpha\in\mathbb{R}^5}\left\{J_1^n(\alpha)+\gamma_{1,n}\Phi_1(\alpha)\right\}+\min_{\beta\in\mathbb{R}^5}\left\{J_2^n(\beta)+\gamma_{2,n}\Phi_2(\beta)\right\}$$
$$\leq \min_{(\alpha,\beta)\in\mathbb{R}^5\times\mathbb{R}^5}\left\{J_1^n(\alpha)+J_2^n(\beta)+\gamma_{1,n}\Phi_1(\alpha)+\gamma_{2,n}\Phi_2(\beta)\right\}.$$

Following Theorem 3.6, each problem in the left-hand side has its minimal solution. Thus, from the definition of (α^n,β^n) it deduces that

$$\alpha^n\in\arg\min_{\alpha\in\mathbb{R}^5}\{J_1^n(\alpha)+\gamma_{1,n}\Phi_1(\alpha)\}\quad\text{and}\quad\beta^n\in\arg\min_{\beta\in\mathbb{R}^5}\{J_2^n(\beta)+\gamma_{2,n}\Phi_2(\beta)\}.\quad(3.28)$$

Let $(\tilde{\alpha},\tilde{\beta})\in\mathbb{R}^5\times\mathbb{R}^5$ be a solution of $F(t)\left[\alpha|\beta\right]=\left[V_1(t)|V_2(t)\right]$.
By the definition of (α^n,β^n) and above statement, it implies that

$$J_1^n(\alpha^n)+\gamma_{1,n}\Phi_1(\alpha^n)\leq J_1^n(\tilde{\alpha})+\gamma_{1,n}\Phi_1(\tilde{\alpha}),\quad(3.29)$$

and

$$J_2^n(\beta^n)+\gamma_{2,n}\Phi_2(\beta^n)\leq J_2^n(\tilde{\beta})+\gamma_{2,n}\Phi_2(\tilde{\beta}).\quad(3.30)$$

Here,

$$\begin{aligned}J_1^n(\tilde{\alpha})&=\int_0^T\|V_1^n(t)-F^n(t)\tilde{\alpha}\|^2\,dt\\&\leq C_1\bigg(\int_0^T\|V_1^n(t)-V_1(t)\|^2\,dt+\int_0^T\|V_1(t)-F(t)\tilde{\alpha}\|^2\,dt\\&\qquad\qquad+\int_0^T\|F(t)\tilde{\alpha}-F^n(t)\tilde{\alpha}\|^2\,dt\bigg)\\&\leq C_1\delta_n.\end{aligned}$$

Similarly, we have $J_2^n(\tilde{\beta})\leq C_2\delta_n$.
From above estimate, (3.29) and (3.30), it follows that

$$J_1^n(\alpha^n)+\gamma_{1,n}\Phi_1(\alpha^n)\leq C_1\delta_n+\gamma_{1,n}\Phi_1(\tilde{\alpha}),\quad(3.31)$$

and

$$J_2^n(\beta^n)+\gamma_{2,n}\Phi_2(\beta^n)\leq C_2\delta_n+\gamma_{2,n}\Phi_2(\tilde{\beta}).\quad(3.32)$$

In particular, when $\delta_n\to 0$, $\gamma_{1,n}\sim\delta_n$ and $\gamma_{2,n}\sim\delta_n$, we deduce that

$$J_1^n(\alpha^n)\to 0\quad,\quad\lim_n\Phi_1(\alpha^n)=\limsup_n\Phi_1(\alpha^n)\leq\Phi_1(\tilde{\alpha}),\quad(3.33)$$

and

$$J_2^n(\beta^n)\to 0\quad,\quad\lim_n\Phi_2(\beta^n)=\limsup_n\Phi_2(\beta^n)\leq\Phi_2(\tilde{\beta}).\quad(3.34)$$

It implies that $\{\Phi_1(\alpha^n)\}$ and $\{\Phi_2(\beta^n)\}$ are bounded. Thus, $\{\gamma_1\Phi_1(\alpha^n)+\gamma_2\Phi_2(\beta^n)\}$ is bounded. Futhermore, $\gamma_1\Phi_1(\cdot)+\gamma_2\Phi_2(\cdot)$ is coercive, $\{(\alpha^n,\beta^n)\}$ is bounded, too. This leads to the existence of a subsequence $\{(\alpha^{n_k},\beta^{n_k})\}$ of $\{(\alpha^n,\beta^n)\}$ such that $(\alpha^{n_k},\beta^{n_k})$ converges to (α^+,β^+).

From (3.33), we deduce that

$$
\begin{aligned}
J_1(\alpha^{n_k}) &= \int_0^T \|V_1(t) - F(t)\alpha^{n_k}\|^2\,dt \\
&\le C_3\bigg(\int_0^T \left\|V_1(t) - V_1^{n_k}(t)\right\|^2 dt + \int_0^T \left\|V_1^{n_k}(t) - F^{n_k}(t)\alpha^{n_k}\right\|^2 dt \\
&\qquad + \int_0^T \|F(t)\alpha^{n_k} - F^{n_k}(t)\alpha^{n_k}\|^2\,dt\bigg) \\
&\le C_3\left(\delta_n^2 + J_1^{n_k}(\alpha^{n_k}) + C_3'\delta_n^2\right) \to 0 \quad (k \to \infty).
\end{aligned}
$$

It is similar to see from (3.34) that $J_2(\alpha^{n_k}) \to 0$ as $k \to \infty$.
Since $J_1(\cdot)$ and $J_2(\cdot)$ are continuous, we have

$$J_1(\alpha^+) = \lim_k J_1(\alpha^{n_k}) = 0 \quad \text{and} \quad J_2(\beta^+) = \lim_k J_2(\beta^{n_k}) = 0.$$

It implies that $F(t)(\alpha^+) = V_1(t)$ and $F(t)(\beta^+) = V_2(t)$ or (α^+, β^+) is a solution of $F(t)\left[\alpha|\beta\right] = \left[V_1(t)|V_2(t)\right]$. Since $\Phi_1(\cdot), \Phi_2(\cdot)$ are convex and continuous on $\mathbb{R}^5$ (Theorem 3.5), (3.33), (3.34) we get

$$\Phi_1(\alpha^+) = \lim_k \Phi_1(\alpha^{n_k}) \le \Phi_1(\tilde{\alpha}) \quad \text{and} \quad \Phi_2(\beta^+) = \lim_k \Phi_2(\beta^{n_k}) \le \Phi_2(\tilde{\beta}).$$

Hence,

$$\gamma_1\Phi_1(\alpha^+) + \gamma_2\Phi_2(\beta^+) \le \gamma_1\Phi_1(\tilde{\alpha}) + \gamma_2\Phi_2(\tilde{\beta}), \quad \forall(\tilde{\alpha}, \tilde{\beta}) \in \Pi(V).$$

It implies that (α^+, β^+) is a Φ_1, Φ_2-minimizing solution.
Moreover, we have that $\{\alpha^{n_k}\}$ and $\{\beta^{n_k}\}$ converge to α^+, β^+. Then, we obtain

$$\left\|(\alpha^{n_k}, \beta^{n_k}) - (\alpha^+, \beta^+)\right\| \to 0.$$

In the case the minimizer (α^+, β^+) is unique, the original sequence $\{(\alpha^n, \beta^n)\}$ converges to (α^+, β^+) as a subsequence argument. □

THIS CHAPTER WAS WRITTEN BASED ON THE PAPER

[15] Hao D.N., Hiep D.X., Muoi P.Q., 2023, Learning river water quality models by l^1-weighted regularization. Published in *IMA Journal of Applied Mathematics.*

Chapter 4

SIMULATION AND NUMERICAL ALGORITHMS

To generate data, we have to solve system (3.1)−(3.4). In doing so we shall use the two-step Lax-Friedrichs method. After having generated data, we shall apply the method in Chapter 3 to verify the theory.

4.1 Two-step Lax–Friedrichs method

There are many available algorithms for solving the direct problem (3.1)−(3.4), such as the upwind scheme, forward time centered space scheme, Lax-Friedrichs algorithm, Leapfrog scheme, Lax-Wendroff scheme,... [4] or two-step Lax-Friedrichs scheme [5]. Among them, the two-step Lax-Friedrichs algorithm is a simple one and has the convergence rate of the second order. Thus, we use the two-step Lax-Friedrichs algorithm mentioned in [5]. Let us consider the BOD-DO system (3.1)−(3.4):

$$\begin{aligned}
&\frac{\partial b}{\partial t} + v\frac{\partial b}{\partial x} = -k_1 b + s_1 && \text{in } (0,X)\times(0,T],\\
&\frac{\partial d}{\partial t} + v\frac{\partial d}{\partial x} = k_1 b - k_2 d + s_2 && \text{in } (0,X)\times(0,T],\\
&b(x,0) = b_0(x),\ d(x,0) = d_0(x) && \text{on } (0,X),\\
&b(0,t) = b_1(t),\ d(0,t) = d_1(t) && \text{on } (0,T].
\end{aligned}$$

Setting

$$u(x,t) = \begin{bmatrix} b(x,t) \\ d(x,t) \end{bmatrix} \quad \text{and} \quad F(x,t,u,u_x) = \begin{bmatrix} -k_1 b(x,t) - v b_x(x,t) + s_1(x,t) \\ k_1 b(x,t) - k_2 d(x,t) - v d_x(x,t) + s_2(x,t) \end{bmatrix},$$

the BOD-DO system could be rewritten by

$$\begin{aligned}
u_t(x,t) &= F(x,t,u,u_x), \qquad (4.1)\\
u(0,t) &= [b(0,t), d(0,t)],\\
u(x,0) &= [b(x,0), d(x,0)].
\end{aligned}$$

The data are simulated from a grid mesh with $n+1$ timesteps $t_0, t_1, t_2, \ldots, t_n$ and $m+1$ grid points $x_0, x_1, x_2, \ldots, x_m$ and constants increments Δt, Δx, respectively. Here, we set $t_0 = 0$, $t_n = T$, $x_0 = 0$, $x_m = X$

For $i \in \{1, 2, \ldots, m-1\}$ and $j \in \{1, 2, \ldots, n\}$, by using a second-order Taylor expansion around $(x_i + \Delta x, t_j)$ and (x_i, t_j) by a half in space, we get

$$u(x_i + \Delta x, t_j) = u\left(x_i + \frac{\Delta x}{2}, t_j\right) + u_x\left(x_i + \frac{\Delta x}{2}, t_j\right)\frac{\Delta x}{2} + \frac{1}{2}u_{xx}\left(x_i + \frac{\Delta x}{2}, t_j\right)\frac{\Delta x^2}{4} + O(\Delta x^3),$$

$$u(x_i, t_j) = u\left(x_i + \frac{\Delta x}{2}, t_j\right) - u_x\left(x_i + \frac{\Delta x}{2}, t_j\right)\frac{\Delta x}{2} + \frac{1}{2}u_{xx}\left(x_i + \frac{\Delta x}{2}, t_j\right)\frac{\Delta x^2}{4} + O(\Delta x^3).$$

Subtracting the two expressions with note that $x_i \pm \Delta x = x_{i\pm 1}$ and $x_i \pm \dfrac{\Delta x}{2} = x_{i\pm\frac{1}{2}}$ we get

$$u(x_{i+1}, t_j) - u(x_i, t_j) = u_x(x_{i+\frac{1}{2}}, t_j)\Delta x + O(\Delta x^3),$$

Thus

$$u_x(x_{i+\frac{1}{2}}, t_j) = \frac{u(x_{i+1}, t_j) - u(x_i, t_j)}{\Delta x} + O(\Delta x^3). \tag{4.2}$$

Similarly, using a first-order Taylor expansion around (x_i, t_j) we have

$$u(x_i + \Delta x, t_j) = u\left(x_i + \frac{\Delta x}{2}, t_j\right) + u_x\left(x_i + \frac{\Delta x}{2}, t_j\right)\frac{\Delta x}{2} + O(\Delta x^2),$$

$$u(x_i, t_j) = u\left(x_i + \frac{\Delta x}{2}, t_j\right) - u_x\left(x_i + \frac{\Delta x}{2}, t_j\right)\frac{\Delta x}{2} + O(\Delta x^2).$$

Adding side by side two expressions, we get

$$u(x_{i+1}, t_j) + u(x_i, t_j) = 2u(x_{i+\frac{1}{2}}, t_j) + O(\Delta x^2).$$

Thus

$$u(x_{i+\frac{1}{2}}, t_j) = \frac{1}{2}[u(x_{i+1}, t_j) + u(x_i, t_j)] + O(\Delta x^2). \tag{4.3}$$

From (4.1), (4.2) and (4.3), we have

$$\frac{u\big(x_{i+\frac{1}{2}}, t_{j+\frac{1}{2}}\big) - u\big(x_{i+\frac{1}{2}}, t_j\big)}{\frac{\Delta t}{2}} = F\left(x_{i+\frac{1}{2}}, t_j, u(x_{i+\frac{1}{2}}, t_j), u_x(x_{i+\frac{1}{2}}, t_j)\right).$$

Thus

$$u\big(x_{i+\frac{1}{2}}, t_{j+\frac{1}{2}}\big) = \frac{u(x_{i+1}, t_j) + u(x_i, t_j)}{2} + F\left(x_{i+\frac{1}{2}}, t_j, \frac{u(x_{i+1}, t_j) + u(x_i, t_j)}{2}, \frac{u(x_{i+1}, t_j) - u(x_i, t_j)}{\Delta x}\right)\frac{\Delta t}{2}. \tag{4.4}$$

Analogously, we have

$$\frac{u(x_i, t_{j+1}) - u(x_i, t_{j+\frac{1}{2}})}{\frac{\Delta t}{2}} = F\left(x_i, t_{j+\frac{1}{2}}, u(x_i, t_{j+\frac{1}{2}}), u_x(x_i, t_{j+\frac{1}{2}})\right).$$

Thus

$$
\begin{aligned}
u(x_i, t_{j+1}) &= u(x_i, t_{j+\frac{1}{2}}) + F\left(x_i, t_{j+\frac{1}{2}}, u(x_i, t_{j+\frac{1}{2}}), u_x(x_i, t_{j+\frac{1}{2}})\right)\frac{\Delta t}{2} \\
&= \frac{u(x_{i+\frac{1}{2}}, t_{j+\frac{1}{2}}) + u(x_{i-\frac{1}{2}}, t_{j+\frac{1}{2}})}{2} + \\
&\quad F\left(x_i, t_{j+\frac{1}{2}}, \frac{u(x_{i+\frac{1}{2}}, t_{j+\frac{1}{2}}) + u(x_{i-\frac{1}{2}}, t_{j+\frac{1}{2}})}{2}, \frac{u(x_{i+\frac{1}{2}}, t_{j+\frac{1}{2}}) - u(x_{i-\frac{1}{2}}, t_{j+\frac{1}{2}})}{\Delta x}\right)\frac{\Delta t}{2}
\end{aligned}
\tag{4.5}
$$

Hence, we arrive at the two-step Lax-Friedrich method.

Algorithm 4.1 Two-step Lax-Friedrich method

Input: Endpoint X, maximum time T, $m+1$ spatial points, $n+1$ temporal points, initial condition $u_i^0 = \begin{bmatrix} b_i^0 & d_i^0 \end{bmatrix}^T$, boundary condition $u_0^j = \begin{bmatrix} b_0^j & d_0^j \end{bmatrix}^T$.

Set $\Delta x = \dfrac{X}{m}, \Delta t = \dfrac{T}{n}$.
for $j = 0, 1, 2, \ldots, n-1$ **do**
  for $i = 0, 1, 2, \ldots, m-1$ **do**
    $u_{i+\frac{1}{2}}^j \leftarrow \dfrac{u_{i+1}^j + u_i^j}{2}$
    $(u_x)_{i+\frac{1}{2}}^j \leftarrow \dfrac{u_{i+1}^j - u_i^j}{\Delta x}$
    $u_{i+\frac{1}{2}}^{j+\frac{1}{2}} \leftarrow u_{i+\frac{1}{2}}^j + \dfrac{\Delta t}{2} F_{i+\frac{1}{2}}^j$
  end for
  for $i = 1, 2, \ldots, m-1$ **do**
    $u_i^{j+\frac{1}{2}} \leftarrow \dfrac{u_{i+\frac{1}{2}}^{j+\frac{1}{2}} + u_{i-\frac{1}{2}}^{j+\frac{1}{2}}}{2}$
    $(u_x)_i^{j+\frac{1}{2}} \leftarrow \dfrac{u_{i+\frac{1}{2}}^{j+\frac{1}{2}} - u_{i-\frac{1}{2}}^{j+\frac{1}{2}}}{\Delta x}$
    $u_i^{j+1} = u_i^{j+\frac{1}{2}} + \dfrac{\Delta t}{2} F_i^{j+\frac{1}{2}}$
  end for
  $u_m^j = 2u_{m-1}^j - u_{m-2}^j$
  $b_i^j = (u_i^j)_1$, $d_i^j = (u_i^j)_2$
end for

Output: $b_i^j = b(x_i, t_j)$, $d_i^j = d(x_i, t_j)$, $\forall (i,j) \in \{0, 1, 2, \ldots, m\} \times \{0, 1, 2, \ldots, n\}$.

Remark 4.1. *According to [5], Algorithm 4.1 converges if*

$$\left| v \frac{\Delta t}{\Delta x} \right| \leq 1,$$

with the truncation error $O(\Delta t^2, \Delta x^2)$.

4.2 Data generation

After obtaining approximate BOD-DO solution b, d by the two-step Lax-Friedrich method (Algorithm 4.1), we create the new feature matrix b_{feature} and d_{feature} from b and d by eliminating some nodes in spatial and temporal domain.

For instance, we will get data at $x_1, x_{1+k}, x_{1+2k}, \ldots, x_{1+l_1 k}$ of the spatial set $\{x_i\}_{i=0}^m$ which correspond to $t_1, t_{1+k}, t_{1+2k}, \ldots, t_{1+l_2 k}$ of the temporal set $\{t_j\}_{j=0}^n$ with $k > 0$ and $l_1 = \left[\frac{m}{k}\right], l_2 = \left[\frac{n}{k}\right]$ ($[x]$ denotes the integer part of real number x). From the new feature data matrix b_{feature} and d_{feature}, we will add $\varepsilon\%$ observation noise into these matrices such that

$$b'_{\text{feature}} = b_{\text{feature}} + \frac{R_1}{\|R_1\|}\varepsilon\% \quad , \quad d'_{\text{feature}} = d_{\text{feature}} + \frac{R_2}{\|R_2\|}\varepsilon\%, \tag{4.6}$$

where R_1, R_2 are random matrices which have same size with b_{feature} and d_{feature}, respectively. Note that (4.6) is necessary for the convergence of problem (3.14) by the condition in Assumption 3.2.

Next, the feature matrices b_x, d_x, b_t and d_t are directly computed by b'_{feature} and d'_{feature} based on the finite difference method and the linear extrapolation to construct the feature matrix $F(t)$ and the vector $V_1(t), V_2(t)$ in (3.13).

4.3 Nesterov's accelerated method

Consider the uniform partition of $[0, T]$ with the mesh point

$$t_i = ih,$$

where $h = \frac{T}{M}$ is the small enough stepsize, Problem (3.14) is approximated by

$$\min_{\alpha,\beta\in\mathbb{R}^5} \left\{ \frac{T}{M} \sum_{i=1}^{M} \left(\left\|V_1^\delta(t_i) - F^\delta(t_i)\alpha\right\|^2 + \left\|V_2^\delta(t_i) - F^\delta(t_i)\beta\right\|^2 \right) + \gamma_1\Phi_1(\alpha) + \gamma_2\Phi_2(\beta) \right\}. \tag{4.7}$$

The solution (α^*, β^*) of Problem (4.7) is equivalent to

$$\alpha^* \in \operatorname*{argmin}_{\alpha\in\mathbb{R}^5} \left\{ \frac{T}{M} \sum_{i=1}^{M} \left\|V_1^\delta(t_i) - F^\delta(t_i)\alpha\right\|^2 + \gamma_1\Phi_1(\alpha) \right\},$$

and

$$\beta^* \in \operatorname*{argmin}_{\beta\in\mathbb{R}^5} \left\{ \frac{T}{M} \sum_{i=1}^{M} \left\|V_1^\delta(t_i) - F^\delta(t_i)\beta\right\|^2 + \gamma_2\Phi_2(\beta) \right\}.$$

In the remain of this section, we only concern about solving Problem (4.7).
This means that our task is to solve minimization problems which have the form

$$\min_{u\in\mathbb{R}^5} \Theta(u) := G(u) + \Psi(u), \tag{4.8}$$

with

$$G(u)=\frac{T}{M}\sum_{i=1}^{M}\left\|V_l^\delta(t_i)-F^\delta(t_i)u\right\|^2 \quad \text{and} \quad \Psi(u)=\gamma\Phi_l(u)=\gamma\sum_{k=1}^{5}\omega_k|u_k| \quad (l=1,2).$$

In particular, G is convex, differentiable (Theorem 3.3) and Ψ is convex and continuous (Theorem 3.5) on $\mathbb{R}^5$. Therefore, the problem (4.8) has at least one solution.

The problem (4.8) is investigated in numerous previous articles based on the gradient method in signal and image processing, see [44, 26, 27, 16]. One of the motivated ideas in this research is originated from [17] in which Problem (4.8) was replaced by the quadratic approximate functional of $\Theta(v)$ at a given point u

$$\min_{v\in\mathbb{R}^5}\Theta_s(v,u):=G(u)+\big\langle G'(u),v-u\big\rangle+\frac{s}{2}\|v-u\|^2+\Psi(v). \tag{4.9}$$

Compared to problem (4.8), the functional $\Theta_{s_n}(\cdot,u^n)$ is strictly convex and it has a unique solution that makes us to find its minimizers easily. In addition, the minimizers $u^{n+1}=\underset{v\in\mathbb{R}^5}{\text{argmin}}\Theta_{s_n}(v,u^n)$ converges to a minimizer of Problem (4.8).

Lemma 4.2. *Given a fixed $u\in\mathbb{R}^5$ and $s,\gamma>0$, the functional $\Theta_s(v,u)$ has a unique minimizer. Moreover, the unique solution of Problem (4.9) is $\mathbb{S}_{\frac{\gamma\omega}{s}}\Big(u-\frac{1}{s}G'(u)\Big)$, where $\mathbb{S}_{\frac{\gamma\omega}{s}}$ is defined by (1.2) and $\omega=\{\omega_k\}_{k=1}^5$ with $\omega_k>0$ for all $k\in\{1,2,\ldots,5\}$.*

Proof. The proof of this lemma is similar to that of Lemma 2.1 in [16].

The funtional Θ_s in (4.9) can be rewritten by

$$\begin{aligned}\Theta_s(v,u)&=G(u)-\frac{1}{2s}\|G'(u)\|^2+s\Big[\frac{1}{2}\Big(\|v-u\|^2+2\frac{1}{s}\langle G'(u),v-u\rangle+\frac{1}{s^2}\|G'(u)\|^2\Big)\\&\qquad+\frac{1}{s}\Psi(v)\Big]\\&=G(u)-\frac{1}{2s}\|G'(u)\|^2+s\Big(\frac{1}{2}\Big\|v-u+\frac{1}{s}G'(u)\Big\|^2+\frac{1}{s}\Psi(v)\Big).\end{aligned}$$

For each fixed $u\in\mathbb{R}^5$ and $s>0$, Problem (4.9) is equivalent to

$$\min_{v\in\mathbb{R}^5}\Big(\frac{1}{2}\Big\|v-\Big(u-\frac{1}{s}G'(u)\Big)\Big\|^2+\frac{1}{s}\Psi(v)\Big). \tag{4.10}$$

From Theorem 3.5, it follows that for each $s>0$ and $w\in\mathbb{R}^5$ the functional

$$v\mapsto\frac{1}{2}\|v-w\|^2+\frac{1}{s}\Psi(v)$$

is strictly convex, bounded below and continuous on $\mathbb{R}^5$. Hence, it has a unique solution. Therefore, Problem (4.9) has a unique minimizer.

Since G is convex and differentiable then according to Theorem 1.6 and Theorem 1.7, the necessary and sufficient condition for the solution $\overline{u}$ of Problem (4.9) is

$$0\in\partial\Theta_s(\overline{u},u)=G'(u)+s(\overline{u}-u)+\partial\Psi(\overline{u}).$$

It implies that

$$u-\frac{1}{s}G'(u)\in\overline{u}+\frac{1}{s}\partial\Psi(\overline{u})=\Big(I+\frac{1}{s}\partial\Psi\Big)(\overline{u}).$$

From Theorem 1.8, $\partial\Psi$ is maximal monotone. Then, it follows from Theorem 1.9 that the inverse operator $\left(I + \frac{1}{s}\partial\Psi\right)^{-1}$ is single-valued. Hence, we have

$$\overline{u} = \left(I + \frac{1}{s}\partial\Psi\right)^{-1}\left(u - \frac{1}{s}G'(u)\right). \tag{4.11}$$

Consider the functional $\psi : \mathbb{R} \to \mathbb{R}$ is defined by

$$\psi(x) = |x|, \quad x \in \mathbb{R}.$$

It is well-known (see [45, Example 16.15]) that

$$\partial\psi(x) = \begin{cases} \{1\}, & \text{if } x > 0, \\ [-1, 1], & \text{if } x = 0, \\ \{-1\}, & \text{if } x < 0. \end{cases}$$

For all $\alpha > 0$, it follows

$$\left(I + \frac{1}{s}\alpha\partial\psi\right)(x) = \begin{cases} \{x + s^{-1}\alpha\}, & \text{if } x > 0, \\ [-s^{-1}\alpha, s^{-1}\alpha], & \text{if } x = 0, \\ \{x - s^{-1}\alpha\}, & \text{if } x < 0. \end{cases}$$

In addition, from (1.1) as $p = 1$, we have

$$S_{\frac{\alpha}{s}}(x) = \begin{cases} x - s^{-1}\alpha, & \text{if } x - s^{-1}\alpha > 0, \\ 0, & \text{if } -s^{-1}\alpha \le x \le s^{-1}\alpha, \\ x + s^{-1}\alpha, & \text{if } x + s^{-1}\alpha < 0. \end{cases}$$

It yields

$$\left(I + \frac{1}{s}\alpha\partial\psi\right)\left(S_{\frac{\alpha}{s}}(x)\right) = \begin{cases} x, & \text{if } x - s^{-1}\alpha > 0 \\ [-s^{-1}\alpha, s^{-1}\alpha], & \text{if } -s^{-1}\alpha \le x \le s^{-1}\alpha, \\ x, & \text{if } x + s^{-1}\alpha < 0. \end{cases}$$

It implies that

$$\left(I + \frac{1}{s}\alpha\partial\psi\right)\left(S_{\frac{\alpha}{s}}(x)\right) = x \quad \text{or} \quad \left(I + \frac{1}{s}\alpha\partial\psi\right)^{-1}(x) = S_{\frac{\alpha}{s}}(x).$$

Thus, (4.11) has the form

$$\overline{u} = \mathbb{S}_{\frac{\gamma\omega}{s}}\left(u - \frac{1}{s}G'(u)\right).$$

□

Lemma 4.3. *Given a positive arbitrary number β and $\omega = \{\omega_k\}_{k=1}^5$. Then, u^* is a minimizer of Θ defined in (4.8) if and only if $u^* = \mathbb{S}_{\beta\omega}(u^* - \beta G'(u^*))$.*

Proof. By Theorem 1.7 and Theorem 1.6, u^* is a solution of (4.8) if and only if

$$0 \in \partial\Theta(u^*) = G'(u^*) + \partial\Psi(u^*).$$

Then,

$$-G'(u^*) \in \partial\Psi(u^*).$$

Multiplying with $\beta > 0$ and adding u^* into both sides, we get

$$u^* - \beta G'(u^*) \in u^* + \beta\partial\Psi(u^*) = (I + \partial\Psi)(u^*).$$

Following Theorem 1.9 and the proof of Lemma 4.2, the inverse operator $\left(I + \frac{1}{s}\partial\Psi\right)^{-1}$ is single-valued and

$$u^* = (I + \partial\Psi)^{-1}(u^* - \beta G'(u^*)) = \mathbb{S}_{\beta\omega}(u^* - \beta G'(u^*)).$$

□

Denote the minimizer of (4.10) in the proof of Lemma 4.2 by

$$H_s(u) = \underset{v\in\mathbb{R}^5}{\operatorname{argmin}}\Theta_s(v,u) = \underset{v\in\mathbb{R}^5}{\operatorname{argmin}}\left(\frac{1}{2}\left\|v - \left(u - \frac{1}{s}G'(u)\right)\right\|^2 + \frac{1}{s}\Psi(v)\right). \tag{4.12}$$

Lemma 4.4. *Let $u \in \mathbb{R}^5$ and $s > 0$ be such that*

$$\Theta(H_s(u)) \leq \Theta_s(H_s(u), u), \tag{4.13}$$

Then, for all $v \in \mathbb{R}^5$ we have

$$\Theta(v) - \Theta(H_s(u)) \geq \frac{s}{2}\left\|H_s(u) - u\right\|^2 + s\langle H_s(u) - u, u - v\rangle.$$

Proof. By the definition of $H_s(u)$ in (4.12), there exists $\psi' \in \partial\Psi(H_s(u))$ such that

$$G'(u) + s(H_s(u) - u) + \psi' = 0,$$

It follows that

$$\psi' = -G'(u) - s(H_s(u) - u). \tag{4.14}$$

Futhermore, since G is convex, by the definition of subgradient we get

$$G(v) \geq G(u) + \langle G'(u), v - u\rangle,$$
$$\Psi(v) \geq \Psi(H_s(u)) + \langle \psi', v - H_s(u)\rangle.$$

Summing two above inequalities and replacing (4.14) into this, it yields

$$G(v) + \Psi(v) \geq G(u) + \Psi(H_s(u)) + \langle G'(u), H_s(u) - u\rangle + s\langle u - H_s(u), v - H_s(u)\rangle$$

Thus,

$$\Theta(v) - \Theta_s(H_s(u), u) \geq -\frac{s}{2}\left\|H_s(u) - u\right\|^2 + s\langle H_s(u) - u, H_s(u) - v\rangle.$$

Hence

$$\Theta(v) - \Theta_s(H_s(u), u) \geq \frac{s}{2}\left\|H_s(u) - u\right\|^2 + s\langle H_s(u) - u, u - v\rangle. \tag{4.15}$$

From (4.13), we get

$$\Theta(v) - \Theta(H_s(u)) \geq \Theta(v) - \Theta_s(H_s(u), u). \tag{4.16}$$

From (4.15) and (4.16), we arrive at the assertion of the lemma. □

Lemma 4.5. *Let $\{u^n\}$ be a sequence given by the gradient-type iteration*

$$u^{n+1} = H_{s_n}(u^n) = \mathbb{S}_{\frac{\gamma\omega}{s_n}}\left(u^n - \frac{1}{s_n}G'(u^n)\right), \tag{4.17}$$

where $\{s_n\}$ *be a positive sequence that satisfies* $s_n \in [\underline{s}, \overline{s}]$ $(\overline{s} > \underline{s} > 0)$ *and*

$$\Theta(u^{n+1}) \leq \Theta_{s_n}(u^{n+1}, u^n).$$

Then, the sequence $\{\Theta(u^n)\}$ *decreases monotonically and* $\lim_{n\to\infty} \left\|u^{n+1} - u^n\right\| = 0$. *In addition, the sequence* $\{u^n\}$ *is bounded.*

Proof. The proof of this lemma is adopted from [44, Lemma 2.3].

By the assumption and the definition of u^{n+1}, it follows that

$$\Theta(u^{n+1}) \leq \Theta_{s_n}(u^{n+1}, u^n) \leq \Theta_{s_n}(u^n, u^n) = \Theta(u^n).$$

This implies that $\{\Theta(u^n)\}$ is a motonomically decreasing sequence.

Applying Lemma 4.4 with $v = u = u^k$ and $s = s_k$ for each $k \in \{0, 1, \dots, n\}$, we get

$$\Theta(u^k) - \Theta(u^{k+1}) \geq \frac{s_k}{2} \left\|u^{k+1} - u^k\right\|^2.$$

It follows that

$$\frac{2}{\underline{s}}(\Theta(u^k) - \Theta(u^{k+1})) \geq \frac{2}{s_k}(\Theta(u^k) - \Theta(u^{k+1})) \geq \left\|u^{k+1} - u^k\right\|^2.$$

Summing these above inequalities over $k = 0, 1, \cdots, n$, we obtain

$$\frac{2}{\underline{s}}(\Theta(u^0) - \Theta(u^{n+1})) \geq \sum_{k=0}^{n} \left\|u^{k+1} - u^k\right\|^2, \quad \forall n.$$

One infers that the series $\sum_{k=0}^{\infty} \left\|u^{k+1} - u^k\right\|^2$ converges. Thus, $\lim_{n\to\infty} \left\|u^{n+1} - u^n\right\| = 0$. Since the functional Θ is coercive by Theorem 3.5 and the sequence $\{\Theta(u^n)\}$ is monotonically decreasing, it leads to the boundedness of the sequence $\{u^n\}$. □

Theorem 4.6. *Let* $\{u^n\}$ *be a sequence that given in Lemma 4.5. Then, there exists a subsequence* $\{u^{n_k}\}$ *of* $\{u^n\}$ *such that* $\{u^{n_k}\}$ *converges to a stationary point* u^* *of* Θ.

Proof. By Lemma 4.5, the sequence $\{u^n\}$ is bounded, then there exists a subsequence $\{u^{n_k}\}$ of $\{u^n\}$ such that $\{u^{n_k}\}$ converges to u^* in $\mathbb{R}^5$. Moreover, it follows from Lemma 4.5 that

$$\lim_{k\to\infty} u^{n_k+1} = u^*.$$

From (4.17), we have

$$u^{n_k+1} = \mathbb{S}_{\frac{\gamma\omega}{s_{n_k}}}\Big(u^{n_k} - \frac{1}{s_{n_k}} G'(u^{n_k})\Big).$$

According to Theorem 3.3, G' is continuous. Choosing a positive sequence $\{s_n\}$ such that there exists a subsequence $\{s_{n_k}\}$ of $\{s_n\}$ converges to $s^* \in [\underline{s}, \overline{s}]$, we get

$$\lim_{k\to\infty} u^{n_k} = u^* \quad , \quad \lim_{k\to\infty} G'(u^{n_k}) = G'(u^*) \quad , \quad \lim_{k\to\infty} s_{n_k} = s^*.$$

It follows from Lemma 1.16 that

$$u^* = \mathbb{S}_{\frac{\gamma\omega}{s^*}}\Big(u^* - \frac{1}{s^*} G'(u^*)\Big).$$

By Lemma 4.3, u^* is a stationary point of Θ. □

Theorem 4.7. *Let $\{u^n\}$ and $\{s_n\}$ be sequences defined in Lemma 4.5 and u^* is a minimizer of Θ. Then, we get*

$$\Theta(u^n) - \Theta(u^*) \leq \frac{\overline{s}\left\|u^0 - u^*\right\|^2}{2n}, \forall n \geq 1.$$

Proof. The proof of this theorem is adopted from [44, Lemma 3.1] and [28, Theorem 3.12]. By the definition of u^{k+1} and Lemma 4.4, for all $k \geq 0$,

$$\begin{aligned}\frac{2}{s_k}(\Theta(u^*) - \Theta(u^{k+1})) &\geq \left\|u^{k+1} - u^k\right\|^2 + 2\langle u^{k+1} - u^k, u^k - u^*\rangle \\ &= \left\|u^{k+1} - u^*\right\|^2 - \left\|u^k - u^*\right\|^2.\end{aligned}$$

In fact, we have $\Theta(u^*) - \Theta(u^{k+1}) \leq 0$ (by Lemma 4.5) and $\underline{s} \leq s_k \leq \overline{s}$. Then, it yields

$$\frac{2}{\overline{s}}(\Theta(u^*) - \Theta(u^{k+1})) \geq \frac{2}{s_k}(\Theta(u^*) - \Theta(u^{k+1})) \geq \left\|u^{k+1} - u^*\right\|^2 - \left\|u^k - u^*\right\|^2.$$

Summing these above inequatilies over $k = 0, 1, \ldots, n-1$, we obtain

$$\frac{2}{\overline{s}}\left(n\Theta(u^*) - \sum_{k=0}^{n-1}\Theta(u^{k+1})\right) \geq \|u^n - u^*\|^2 - \left\|u^0 - u^*\right\|^2. \tag{4.18}$$

By Lemma 4.4 and the definition of u^{k+1}, we also have

$$\frac{2}{s_k}(\Theta(u^k) - \Theta(u^{k+1})) \geq \left\|u^{k+1} - u^k\right\|^2, \forall k \geq 0.$$

Similarly, the fact that $\{\Theta(u^k)\}$ is a decreasing sequence (by Lemma 4.5) and $s^k \in [\underline{s}, \overline{s}]$ yields

$$\frac{2}{\underline{s}}(\Theta(u^k) - \Theta(u^{k+1})) \geq \frac{2}{s_k}(\Theta(u^k) - \Theta(u^{k+1})) \geq \left\|u^{k+1} - u^k\right\|^2, \forall k \geq 0.$$

Then,

$$\frac{2}{\underline{s}}(k\Theta(u^k) - k\Theta(u^{k+1})) \geq k\left\|u^{k+1} - u^k\right\|^2, \forall k \geq 0.$$

Summing these above inequalities over $k = 0, 1, \cdots, n-1$ we obtain

$$\frac{2}{\underline{s}}\sum_{k=0}^{n-1}\left(k\Theta(u^k) - k\Theta(u^{k+1})\right) \geq \sum_{k=0}^{n-1}k\left\|u^{k+1} - u^k\right\|^2.$$

Note that

$$\begin{aligned}\sum_{k=0}^{n-1}\left(k\Theta(u^k) - k\Theta(u^{k+1})\right) &= \sum_{k=0}^{n-1}\left(k\Theta(u^k) - (k+1)\Theta(u^{k+1}) + \Theta(u^{k+1})\right) \\ &= -n\Theta(u^n) + \sum_{k=0}^{n-1}\Theta(u^{k+1}).\end{aligned}$$

It follows that

$$\frac{2}{\underline{s}}\left(-n\Theta(u^n) + \sum_{k=0}^{n-1}\Theta(u^{k+1})\right) \geq \sum_{k=0}^{n-1}k\left\|u^{k+1} - u^k\right\|^2.$$

Then,

$$\frac{2}{\overline{s}}\left(-n\Theta(u^n) + \sum_{k=0}^{n-1}\Theta(u^{k+1})\right) \geq \frac{\underline{s}}{\overline{s}}\sum_{k=0}^{n-1}k\left\|u^{k+1} - u^k\right\|^2. \tag{4.19}$$

Summing (4.18) and (4.19), we get

$$\frac{2n}{\overline{s}}\big(\Theta(u^*)-\Theta(u^n)\big) \geq \|u^n - u^*\|^2 - \left\|u^0 - u^*\right\|^2 + \frac{\underline{s}}{\overline{s}}\sum_{k=0}^{n-1} k\left\|u^{k+1} - u^k\right\|^2 \geq -\left\|u^0 - u^*\right\|^2.$$

Therefore,

$$\Theta(u^n) - \Theta(u^*) \leq \frac{\overline{s}\left\|u^0 - u^*\right\|^2}{2n}, \ \forall n \geq 1.$$

□

Next, we present Nesterov's accelerated algorithm that was firstly introduced in [17]. This algorithm used gradient-type methods achieves the great convergent rate $O\left(\frac{1}{n^2}\right)$ in convex cases.

Algorithm 4.2 Nesterov's accelerated method

Input: Initial value $u^0 \in \mathbb{R}^5$, $A_0 = 0$, $v^0 = u^0, \eta \in (1, \infty)$, $s_0 \in [\underline{s}, \overline{s}]$
  and $\varphi_0(u) = \frac{1}{2}\left\|u - u^0\right\|^2$.

for $n = 0, 1, 2, \ldots$ **do**
  repeat
    $a_{n+1} \leftarrow \dfrac{1 + \sqrt{1 + 2A_n s_n}}{s_n}$.
    $y^n \leftarrow \dfrac{A_n u^n + a_{n+1} v^n}{A_n + a_{n+1}}$.
    $u^{n+1} \leftarrow \mathbb{S}_{\frac{\gamma\omega}{s_n}}\left(y^n - \dfrac{1}{s_n}G'(y^n)\right)$.
    if $\left\|G'(u^{n+1}) - G'(y^n)\right\|^2 > s_n\langle G'(u^{n+1}) - G'(y^n), u^{n+1} - y^n\rangle$ **then**
      $s_n \leftarrow s_n\eta$
    end if
  until $\left\|G'(u^{n+1}) - G'(y^n)\right\|^2 \leq s_n\langle G'(u^{n+1}) - G'(y^n), u^{n+1} - y^n\rangle$ or $s_n \notin [\underline{s}, \overline{s}]$.
  $A_{n+1} \leftarrow A_n + a_{n+1}$.
  $v^{n+1} \leftarrow \operatorname*{argmin}_{u\in\mathbb{R}^5} \varphi_{n+1}(u)$ with

$$\varphi_{n+1}(u) = \varphi_n(u) + a_{n+1}\big(G(u^{n+1}) + \langle G'(u^{n+1}), u - u^{n+1}\rangle + \gamma\Phi(u)\big).$$

  $s_{n+1} \leftarrow P_{[\underline{s},\overline{s}]}\dfrac{\left\|G'(u^{n+1}) - G'(y^n)\right\|^2}{\langle G'(u^{n+1}) - G'(y^n), u^{n+1} - y^n\rangle}$.
end for
Output: $u = \lim u^n$.

Remark 4.8. *In Algorithm 4.2, we use the Barzilai-Borwein rule to choose the step size s_n that is indicated in [46], i.e., s_n satisfies*

$$s_n = \max\left(\overline{s}, \min\left(\underline{s}, \frac{\left\|G'(u^{n+1}) - G'(y^n)\right\|^2}{\langle G'(u^{n+1}) - G'(y^n), u^{n+1} - y^n\rangle}\right)\right).$$

Remark 4.9. *The stepsize s_n in Algorithm 4.2 is to belong to $[\underline{s}, \overline{s}]$ in the cases the Lipschitz constant L_G of the derivative of the functional G is unknown. However, if L_G is known, $\underline{s}$ and $\overline{s}$ would be satisfied $\underline{s} \leq \eta L_G \leq \overline{s}$, $\eta > 0$, see [28].*

Lemma 4.10. *The solution v^n of the functional φ_n in Step 11 of Algorithm 4.2 is defined by*

$$v^n = \mathbb{S}_{\gamma A_n \omega}\left(u^0 - \sum_{k=1}^{n} a_k G'(u^k)\right), \ (n > 0).$$

Proof. From Algorithm 4.2, we have

$$\varphi_n(u) = \frac{1}{2}\left\|u - u^0\right\|^2 + \sum_{k=1}^{n} a_k\Big(G(u^k) + \langle G'(u^k), u - u^k\rangle\Big) + \gamma \sum_{k=1}^{n} a_k \Phi(u).$$

Similar to Lemma 4.2, the functional φ_n is convex because it is the sum of convex functionals and by Theorem 1.7 ([22, Proposition 31.14]), the necessary and sufficient condition for v^n be a minimizer of φ_n is

$$0 \in \partial\varphi_n(v^n) = (v^n - u^0) + \sum_{k=1}^{n} a_k G'(u^k) + \gamma \sum_{k=1}^{n} a_k \partial\Phi(v^n).$$

This follows that

$$u^0 - \sum_{k=1}^{n} a_k G'(u^k) \in v^n + \gamma \sum_{k=1}^{n} a_k \partial\Phi(v^n) = (I + \gamma A_n \partial\Phi)(v^n).$$

Similarly, $\gamma A_n \partial\Phi$ is maximal monotone by Theorem 1.8 and by Theorem 1.9 the operator $(I + \gamma A_n \partial\Phi)$ is invertible. Furthermore, its inverse is single-valued. Thus, we deduce that

$$v^n = (I + \gamma A_n \partial\Phi)^{-1}\left(u^0 - \sum_{k=1}^{n} a_k G'(u^k)\right).$$

Following the proof in Lemma 4.2, we also have

$$(I + \gamma A_n \partial\Phi)^{-1}\left(u^0 - \sum_{k=1}^{n} a_k G'(u^k)\right) = \mathbb{S}_{\gamma A_n \omega}\left(u^0 - \sum_{k=1}^{n} a_k G'(u^k)\right).$$

Hence, the proof is completed. □

Lemma 4.11. *The functional φ_n $(n \geq 0)$ defined in Step 11 of Algorithm 4.2 is a 1-strongly convex functional.*

Proof. By [47, Theorem 5.24], we need only to prove that

$$\varphi_n(v) \geq \varphi_n(u) + \langle \xi, v - u\rangle + \frac{1}{2}\|v - u\|^2, \ \forall u, v \in \mathbb{R}^5, \xi \in \partial\varphi_n(u).$$

This is equivalent to

$$\frac{1}{2}\left\|v-u^0\right\|^2+\sum_{k=1}^{n}a_k\Big(G(u^k)+\langle G'(u^k),v-u^k\rangle\Big)+\gamma A_n\Phi(v)$$

$$\geq\frac{1}{2}\left\|u-u^0\right\|^2+\sum_{k=1}^{n}a_k\Big(G(u^k)+\langle G'(u^k),u-u^k\rangle\Big)+\gamma A_n\Phi(u)$$

$$+\langle v-u^0+\sum_{k=1}^{n}a_kG'(u^k)+\gamma A_n\nu,v-u\rangle+\frac{1}{2}\left\|v-u\right\|^2,\ \text{where } \nu\in\partial\Phi(u).$$

By simple calculation, it is reduced to

$$\frac{1}{2}\left\|v-u^0\right\|^2+\gamma A_n\Phi(v)\geq\frac{1}{2}\left\|u-u^0\right\|^2+\gamma A_n\Phi(u)+\langle v-u^0,v-u\rangle$$

$$+\gamma A_n\langle\nu,v-u\rangle+\frac{1}{2}\left\|v-u\right\|^2,\ \text{where } \nu\in\partial\Phi(u).$$

The above inequality is true by the definition of $\partial\Phi(u)$ and the fact that

$$\frac{1}{2}\left\|v-u^0\right\|^2=\frac{1}{2}\left\|u-u^0\right\|^2+\langle v-u^0,v-u\rangle+\frac{1}{2}\left\|v-u\right\|^2.$$

□

Next, we will prove that the Algorithm 4.2 satisfies the following relations:

$$\mathcal{R}_n^1: \qquad A_n\Theta(u^n)\leq\varphi_n^*\equiv\min_{u\in\mathbb{R}^5}\varphi_n(u), \tag{4.20}$$

$$\mathcal{R}_n^2: \qquad \varphi_n(u)\leq A_n\Theta(u)+\frac{1}{2}\left\|u-u^0\right\|^2, \forall u\in\mathbb{R}^5. \tag{4.21}$$

Lemma 4.12. *The sequence $\{u^n\}$, $\{A_n\}$ and $\{\varphi_n\}$ generated by Algorithm 4.2 satisfy the relations* (4.20) *and* (4.21).

Proof. We will prove this Lemma by induction. First, the relations (4.20) and (4.21) are true by the initial setting of Algorithm 4.2. Assume that the relations (4.20) and (4.21) are valid for some $n\geq 0$.

1. By the hypothesis of the induction and the definition of φ_n in Step 11 of Algorithm 4.2, it follows

   $$\varphi_{n+1}(u)\leq A_n\Theta(u)+\frac{1}{2}\left\|u-u^0\right\|^2+a_{n+1}\big(G(u^{n+1})+\langle G'(u^{n+1}),u-u^{n+1}\rangle+\gamma\Phi(u)\big).$$

   Since G is convex and differentiable, we have

   $$G(u)\geq G(u^{n+1})+\langle G'(u^{n+1}),u-u^{n+1}\rangle,\ \forall u\in\mathbb{R}^5.$$

   It yields

   $$\begin{aligned}\varphi_{n+1}(u)&\leq A_n\Theta(u)+\frac{1}{2}\left\|u-u^0\right\|^2+a_{n+1}(G(u)+\gamma\Phi(u))\\&=(A_n+a_{n+1})\Theta(u)+\frac{1}{2}\left\|u-u^0\right\|^2\\&=A_{n+1}\Theta(u)+\frac{1}{2}\left\|u-u^0\right\|^2.\end{aligned}$$

   Thus, the relation (4.21) is true for $n+1$.

2. According to Lemma 4.10, φ_n is a 1−strongly convex function and v^n is a unique minimizer of φ_n. Then by the property of strongly convex functions, see [47, Theorem 5.25], and the inducted hypothesis, we have
$$\varphi_n(u) \geq \varphi_n^* + \frac{1}{2}\|u - v^n\|^2 \geq A_n\Theta(u^n) + \frac{1}{2}\|u - v^n\|^2, \forall u \in \mathbb{R}^5. \tag{4.22}$$
By the definition of φ_{n+1}^* in (4.20) and (4.22), we have
$$\begin{aligned}\varphi_{n+1}^* &= \min_{u\in\mathbb{R}^5}\left\{\varphi_n(u) + a_{n+1}\big(G(u^{n+1}) + \langle G'(u^{n+1}), u - u^{n+1}\rangle + \gamma\Phi(u)\big)\right\}\\ &\geq \min_{u\in\mathbb{R}^5}\Big\{A_n\Theta(u^n) + \frac{1}{2}\|u - v^n\|^2 + a_{n+1}\big(G(u^{n+1})\\ &\qquad + \langle G'(u^{n+1}), u - u^{n+1}\rangle + \gamma\Phi(u)\big)\Big\}.\end{aligned}$$
For all $u \in \mathbb{R}^5$, there exists $\nu^{n+1} \in \partial\Phi(u^{n+1})$ such that
$$\Phi(u) \geq \Phi(u^{n+1}) + \langle \nu^{n+1}, u - u^{n+1}\rangle.$$
Define the notation $\Theta'(\cdot)$ by
$$\Theta'(u) = G'(u) + \gamma\nu,\ \nu \in \partial\Phi(u).$$
By the convexity of the functional Θ, we have
$$\Theta(u^n) \geq \Theta(u^{n+1}) + \langle \Theta'(u^{n+1}), u^n - u^{n+1}\rangle.$$
Using these estimate, we get
$$\begin{aligned}\varphi_{n+1}^* &\geq \min_{u\in\mathbb{R}^5}\Big\{A_n(\Theta(u^{n+1}) + \langle \Theta'(u^{n+1}), u^n - u^{n+1}\rangle) + \frac{1}{2}\|u - v^n\|^2\\ &\quad + a_{n+1}\big(G(u^{n+1}) + \langle G'(u^{n+1}), u - u^{n+1}\rangle + \gamma\Phi(u^{n+1}) + \langle \nu^{n+1}, u - u^{n+1}\rangle\big)\Big\}\\ &= \min_{u\in\mathbb{R}^5}\Big\{(A_n + a_{n+1})\Theta(u^{n+1}) + \frac{1}{2}\|u - v^n\|^2 + A_n\langle \Theta'(u_{n+1}), u^n - u^{n+1}\rangle\\ &\qquad + a_{n+1}\langle \Theta'(u^{n+1}), u - u^{n+1}\rangle\Big\}.\end{aligned}$$
Due to the Step 4 in Algorithm 4.2, this follows that
$$\begin{aligned}\varphi_{n+1}^* &\geq \min_{u\in\mathbb{R}^5}\Big\{A_{n+1}\Theta(u^{n+1}) + \frac{1}{2}\|u - v^n\|^2 + a_{n+1}\langle \Theta'(u^{n+1}), u - u^{n+1}\rangle\\ &\qquad + \langle \Theta'(u^{n+1}), A_{n+1}y^n - a_{n+1}v^n - A_n u^{n+1}\rangle\Big\}\\ &= \min_{u\in\mathbb{R}^5}\Big\{A_{n+1}\Theta(u^{n+1}) + \frac{1}{2}\|u - v^n\|^2 + a_{n+1}\langle \Theta'(u^{n+1}), u - v^n\rangle\\ &\qquad + A_{n+1}\langle \Theta'(u^{n+1}), y^n - u^{n+1}\rangle\Big\}.\end{aligned} \tag{4.23}$$

The minimizer of the right hand side (4.23) attains at
$$u = v^n - a_{n+1}\Theta'(u^{n+1}).$$

Then,

$$\varphi_{n+1}^* \geq A_{n+1}\Theta(u^{n+1}) + \frac{a_{n+1}^2}{2}\left\|\Theta'(u^{n+1})\right\|^2 - a_{n+1}^2\left\|\Theta'(u^{n+1})\right\|^2 + A_{n+1}\langle\Theta'(u^{n+1}), y^n - u^{n+1}\rangle$$
$$= A_{n+1}\Theta(u^{n+1}) - \frac{a_{n+1}^2}{2}\left\|\Theta'(u^{n+1})\right\|^2 + A_{n+1}\langle\Theta'(u^{n+1}), y^n - u^{n+1}\rangle. \tag{4.24}$$

Since u^{n+1} is the minimizer of $\Theta_{s_n}(\cdot, y^n)$, we have

$$0 \in \partial\Theta_{s_n}(u^{n+1}, y^n) = G'(y^n) + s_n(u^{n+1} - y^n) + \partial\Psi(u^{n+1}).$$

It implies that there exists $\xi^{n+1} \in \partial\Phi(u^{n+1})$ satisfies

$$G'(y^n) + s_n(u^{n+1} - y^n) + \gamma\xi^{n+1} = 0.$$

Then,

$$\Theta'(u^{n+1}) = G'(u^{n+1}) + \gamma\xi^{n+1} = s_n(y^n - u^{n+1}) + G'(u^{n+1}) - G'(y^n).$$

This implies that

$$\langle\Theta'(u^{n+1}), y^n - u^{n+1}\rangle = s_n\left\|y^n - u^{n+1}\right\|^2 + \langle G'(u^{n+1}) - G'(y^n), y^n - u^{n+1}\rangle. \tag{4.25}$$

Moreover, we have

$$\left\|y^n - u^{n+1}\right\|^2 = \frac{1}{s_n^2}\Big(\left\|s_n(y^n - u^{n+1}) + G'(u^{n+1}) - G'(y^n)\right\|^2 - 2s_n\left\langle G'(u^{n+1}) - G'(y^n), y^n - u^{n+1}\right\rangle - \left\|G'(u^{n+1}) - G'(y^n)\right\|^2\Big)$$
$$= \frac{1}{s_n^2}\Big(\left\|\Theta'(u^{n+1})\right\|^2 - 2s_n\left\langle G'(u^{n+1}) - G'(y^n), y^n - u^{n+1}\right\rangle - \left\|G'(u^{n+1}) - G'(y^n)\right\|^2\Big).$$

Substituting it into (4.25), we get

$$\left\langle\Theta'(u^{n+1}), y^n - u^{n+1}\right\rangle = \frac{1}{s_n}\left\|\Theta'(u^{n+1})\right\|^2 - \left\langle G'(u^{n+1}) - G'(y^n), y^n - u^{n+1}\right\rangle - \frac{1}{s_n}\left\|G'(u^{n+1}) - G'(y^n)\right\|^2.$$

Due to Step 9 in Algorithm 4.2

$$\frac{1}{s_n}\left\|G'(u^{n+1}) - G'(y^n)\right\|^2 \leq \left\langle G'(y^n) - G'(u^{n+1}), y^n - u^{n+1}\right\rangle.$$

It implies that

$$\left\langle\Theta'(u^{n+1}), y^n - u^{n+1}\right\rangle \geq \frac{1}{s_n}\left\|\Theta'(u^{n+1})\right\|^2. \tag{4.26}$$

From (4.24) and (4.26), we get

$$\varphi_{n+1}^* \geq A_{n+1}\Theta(u^{n+1}) + \left(\frac{A_{n+1}}{s_n} - \frac{a_{n+1}^2}{2}\right)\left\|\Theta'(u^{n+1})\right\|^2.$$

Following Step 3 in Algorithm 4.2, a_{n+1} is the positive solution of the quadratic equation

$$a_{n+1}^2 - \frac{2}{s_n}a_{n+1} - \frac{2A_n}{s_n} = 0.$$

This leads to

$$\frac{a_{n+1}^2}{2} = \frac{A_n + a_{n+1}}{s_n} = \frac{A_{n+1}}{s_n}. \tag{4.27}$$

We deduce that

$$\varphi_{n+1}^* \geq A_{n+1}\Theta(u^{n+1}).$$

Therefore, the relation (4.20) is also valid for $n+1$.

□

Lemma 4.13. *The positive sequence $\{A_n\}$ generated by Algorithm 4.2 satisfies*

$$A_n \geq \frac{n^2}{2\overline{s}}, \quad \forall n \geq 0.$$

In addition, if G has a Lipschitz continuous derivative with Lipschitz constant L_G, we have

$$A_n \geq \frac{n^2}{2\eta L_G}, \quad \forall n \geq 0,$$

where $\eta > 0$ such that $\eta L_G \geq s_n$.

Proof. Following (4.27) in the proof of Lemma 4.11, we have

$$\begin{aligned} A_{n+1} &= \frac{s_n}{2}a_{n+1}^2 = \frac{s_n}{2}(A_{n+1} - A_n)^2 = \frac{s_n}{2}\left(A_{n+1}^{1/2} - A_n^{1/2}\right)^2\left(A_{n+1}^{1/2} + A_n^{1/2}\right)^2 \\ &\leq 2\overline{s}A_{n+1}\left(A_{n+1}^{1/2} - A_n^{1/2}\right)^2. \end{aligned}$$

Thus

$$A_{n+1}^{1/2} - A_n^{1/2} \geq \frac{1}{\sqrt{2\overline{s}}}, \quad \forall n \geq 0.$$

Summing these inequalities, we get

$$A_{n+1}^{1/2} - A_0^{1/2} \geq \frac{n+1}{\sqrt{2\overline{s}}}, \quad \forall n \geq 0.$$

It yields

$$A_{n+1}^{1/2} \geq \frac{n+1}{\sqrt{2\overline{s}}}, \quad \forall n \geq 0.$$

Thus, for all $n \geq 0$, we have

$$A_n \geq \frac{n^2}{2\overline{s}}.$$

The rest of Lemma is directly resulted from the above inequality. □

The next theorem is a result of the one in [17, Theorem 6].

Theorem 4.14. *Assume that the sequence $\{u^n\}$ generated by Algorithm 4.2 converges to the solution u^* of problem* (4.8)*. Then the following inequality is satisfied*

$$\Theta(u^n) - \Theta(u^*) \leq \frac{\overline{s}\left\|u^0 - u^*\right\|^2}{n^2}, \quad \forall n \geq 0.$$

If G has a Lipschitz continuous derivative with Lipschitz constant L_G, we also have

$$\Theta(u^n) - \Theta(u^*) \leq \frac{\eta L_G\left\|u^0 - u^*\right\|^2}{n^2}, \quad \forall n \geq 0.$$

Proof. By Lemma 4.12, we have the relations (4.20) and (4.21):

$$A_n\Theta(u^n) \le \varphi_n^* \equiv \min_{u\in\mathbb{R}^5} \varphi_n(u),$$

$$\varphi_n(u) \le A_n\Theta(u) + \frac{1}{2}\left\|u - u^0\right\|^2, \forall u \in \mathbb{R}^5.$$

In the second relation, choosing $u = u^*$ and combining with the first relation, we get

$$A_n\Theta(u^n) \le A_n\Theta(u^*) + \frac{1}{2}\left\|u^* - u^0\right\|^2, \forall n \ge 0.$$

Thus,

$$\Theta(u^n) - \Theta(u^*) \le \frac{\left\|u^* - u^0\right\|^2}{2A_n}, \forall n \ge 0.$$

Following Lemma 4.12, we obtain

$$\Theta(u^n) - \Theta(u^*) \le \frac{\overline{s}\left\|u^0 - u^*\right\|^2}{n^2}, \quad \forall n \ge 0.$$

The remained inequality is a result from above estimate and the second inequality of Lemma 4.12. □

THIS CHAPTER WAS WRITTEN BASED ON THE PAPER

[15] Hao D.N., Hiep D.X., Muoi P.Q., 2023, Learning river water quality models by l^1-weighted regularization. Published in *IMA Journal of Applied Mathematics*.

Chapter 5

NUMERICAL EXAMPLES

In this chapter, we will present some numerical examples illustrating the theoretical results of Chapter 3 and Chapter 4. All examples were written in Python 3.10.1 software.

Example 5.1. *Consider the BOD-DO model*

$$\begin{cases} \dfrac{\partial b(x,t)}{\partial t} + 3\dfrac{\partial b(x,t)}{\partial x} = -2b(x,t), & \textit{in } (0,1)\times(0,2), \\ \dfrac{\partial d(x,t)}{\partial t} + 3\dfrac{\partial d(x,t)}{\partial x} = 2b(x,t) - d(x,t), & \textit{in } (0,1)\times(0,2), \\ b(x,0) = d(x,0) = 0, & \textit{on } (0,1), \\ b(0,t) = d(0,t) = \sin(2\pi t), & \textit{on } (0,2). \end{cases}$$

In this example, parameter vectors α^T, β^T are $\alpha^T = [0, -2, 0, -3, 0]$ and $\beta^T = [0, 2, -1, 0, -3]$, respectively. Moreover, the parameters are $v = 3, k_1 = 2$, $k_2 = 1$, $s_1 = s_2 = 0$, the initial conditions are $b_0 = d_0 = 0$ and the boundary conditions are $b_1 = d_1 = \sin(2\pi t)$.

According to Theorem 3.1, the exact solution to this problem is

$$b(x,t) = \begin{cases} 0 & , \text{ if } x - 3t \in [0,1], \\ e^{\frac{-2x}{3}} \sin\left[2\pi\left(t - \dfrac{x}{3}\right)\right] & , \text{ if } t - \dfrac{x}{3} \in [0,2], \end{cases}$$

and,

$$d(x,t) = \begin{cases} 0 & , \text{ if } x - 3t \in [0,1], \\ 2e^{\frac{-2x}{3}} \sin\left[2\pi\left(t - \dfrac{x}{3}\right)\right] + 2e^{\frac{-x}{3}}\left[\sin\left(2\pi\left(t - \dfrac{x}{3}\right)\right)\right] & , \text{ if } t - \dfrac{x}{3} \in [0,2]. \end{cases}$$

Figure 5.2 (Top) represents the solution to this example.

Then, we use the two-step Lax-Friedrichs method (Algorithm 4.1) to solve this problem. First, we divide the spatial domain $[0, 1]$ into 801 uniformly distributed points $\{x_i\}_{i=0}^{800}$ with $\Delta x = \dfrac{1}{800}$ and equally distribute temporal domain $[0, 2]$ into 4801 points $\{t_j\}_{j=0}^{4800}$ with $\Delta t = \dfrac{2}{4800}$, see Figure 5.1. In this case, $\left|v\dfrac{\Delta t}{\Delta x}\right| = 1$. By Remark 4.1 in Chapter 4, the two-step Lax-Friedrichs method (Algorithm 4.1) converges. The approximate solution for the BOD-DO model in Example 5.1 is illustrated in Figure 5.2 (Bottom).

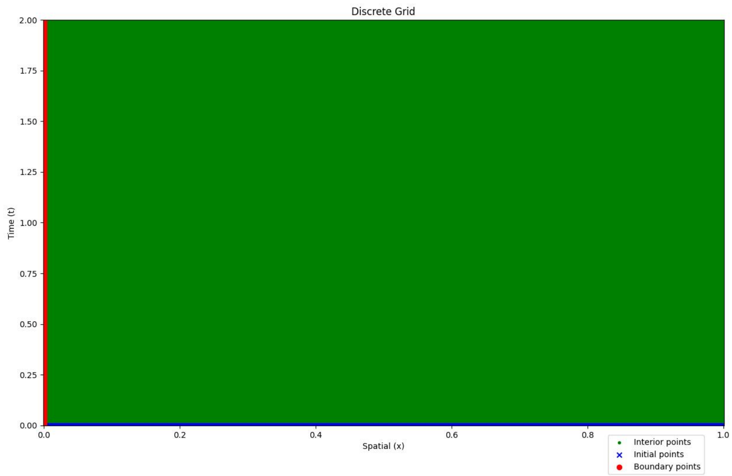

Figure 5.1: Grid mesh of Example 5.1.

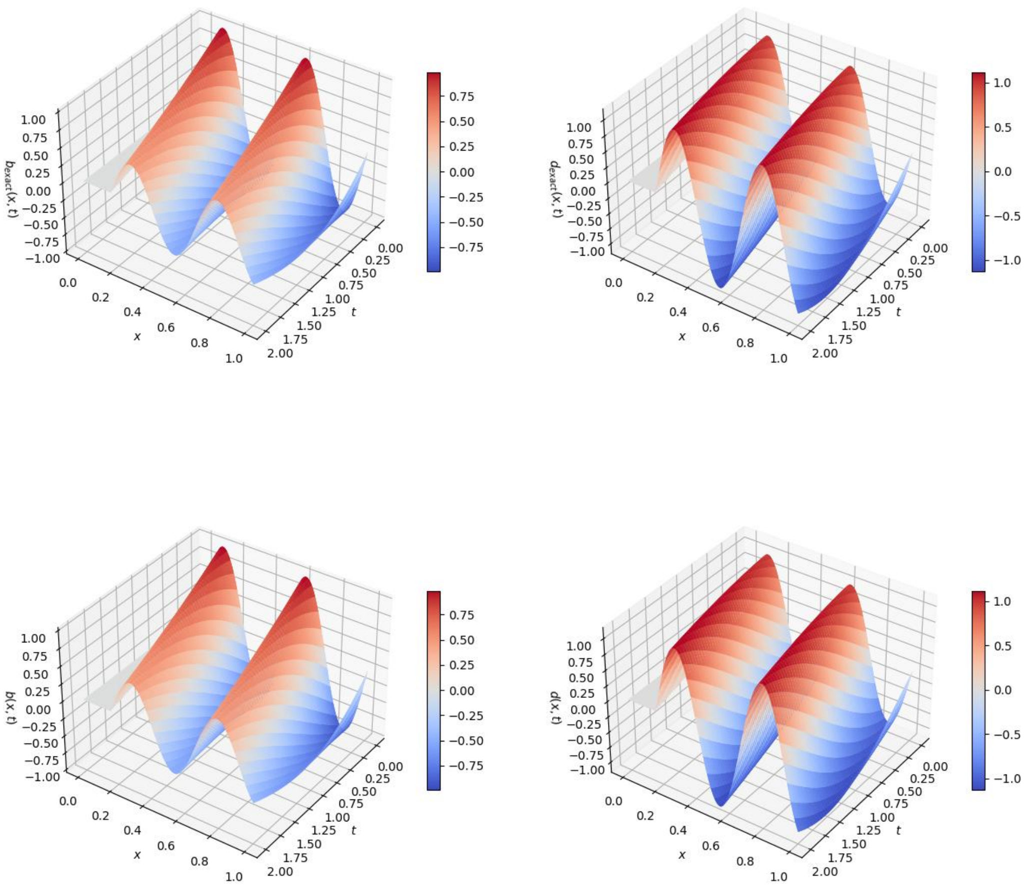

Figure 5.2: Example 5.1: Exact solution (top) and approximate solution (bottom) of the BOD-DO model.

It can be seen from Figure 5.3 that the approximate solution well approximate the exact solution. Indeed, the errors between exact and approximate solutions are not greater than 0.06% and 0.02%, respectively.

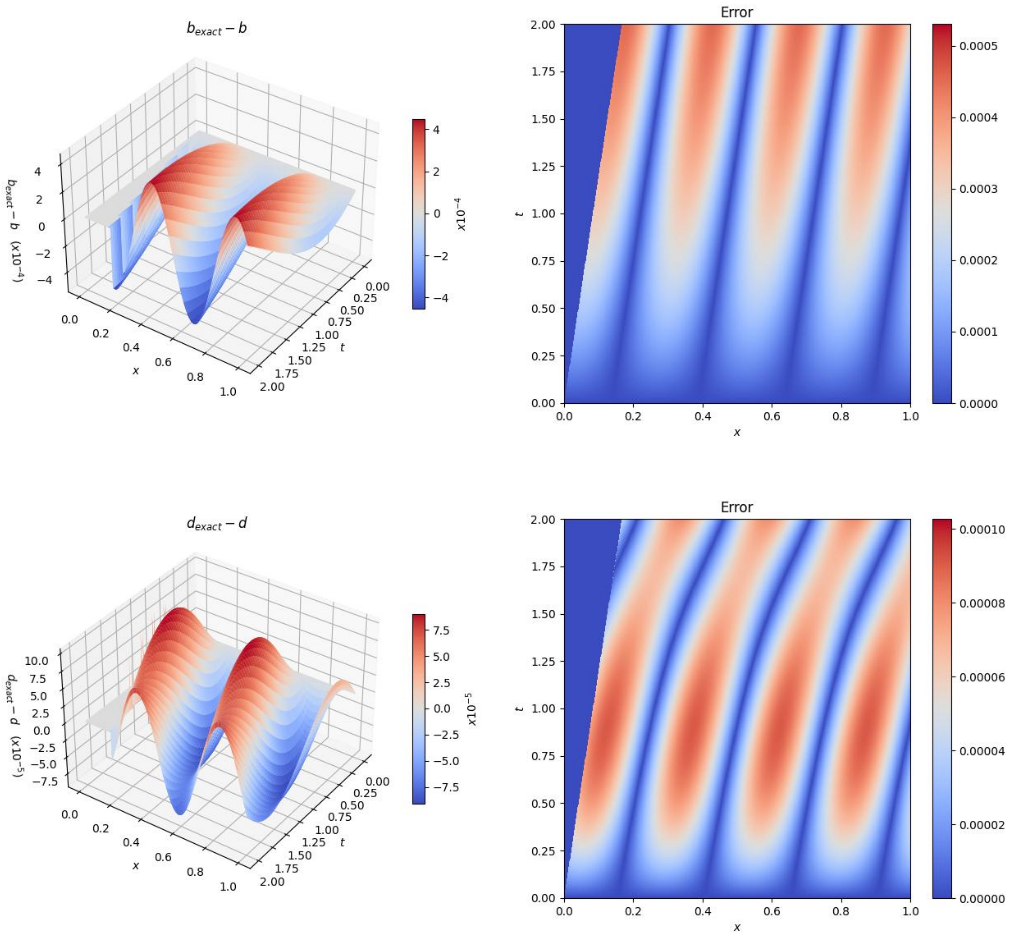

Figure 5.3: Example 5.1: The misfit and error between the exact and numerical solutions.

Next, we collect data of the solution $b_{\text{feature}}(x,t)$ and $d_{\text{feature}}(x,t)$ from approximate solution b, d with $k = 2$ (see Section 4.2) and add 1% noise in the data. Following Section 4.2, the derivatives of b_x, d_x, b_t and d_t are also directly calculated by the finite difference method and the linear extrapolation to construct the feature matrix F as well as the feature matrix V.

First, we use l^1-weighted regularization for these problems with $\gamma_1 = \gamma_2 = 10^{-1}$. As a priori information, we know the form of BOD-DO model, i.e., only parameters s_1, s_2, k_1, k_2, v are possibly nonzero while the other must be zero. Based on this information, we set $\omega_j^1 = 1$ for $j = 2, 4$, $\omega_j^2 = 1$ for $j = 2, 3, 5$ (corresponding the parameters which are possible nonzeros) and set high values for weighted $\omega_j^1 = \omega_j^2 = 10$ for the others. We also apply Nesterov's accelerated method (Algorithm 4.2) for with the initial point

$[0, 0, 0, 0, 0]$ for two problems (5.1) and (5.2) in 800 iterations.

$$\min_{\alpha\in\mathbb{R}^5}\{G_1(\alpha)+\gamma_1\Phi_1(\alpha)\} = \min_{\alpha\in\mathbb{R}^5}\left\{\frac{2}{2400}\sum_{i=1}^{2400}\left(\left\|V_1^\delta(t_i)-F^\delta(t_i)\alpha\right\|^2+\gamma_1\sum_{j=1}^{5}w_j^1|\alpha_j|\right)\right\}, \tag{5.1}$$

and

$$\min_{\beta\in\mathbb{R}^5}\{G_2(\beta)+\gamma_2\Phi_2(\beta)\} = \min_{\beta\in\mathbb{R}^5}\left\{\frac{2}{2400}\sum_{i=1}^{2400}\left(\left\|V_2^\delta(t_i)-F^\delta(t_i)\beta\right\|^2+\gamma_2\sum_{j=1}^{5}w_j^2|\beta_j|\right)\right\}. \tag{5.2}$$

Next, we use l^1-regularization to learn the problems (5.1) and (5.2) with $\omega_j^1 = \omega_j^2 = 1$ for all j. We also set $\gamma_1 = \gamma_2 = 10^{-1}$. Then, we exploit Nesterov's accelerated method (Algorithm 4.2) with the initial point $[0, 0, 0, 0, 0]$ in 800 iterations.

In both cases, the value of $G_1(\alpha)$ and $G_2(\beta)$ dramatically decrease to 0 after fewer than 50 iterations, see Figure 5.4. Hence, the problems (5.1) and (5.2) get their minimizer in l^1-weighted regularization and l^1-regularization.

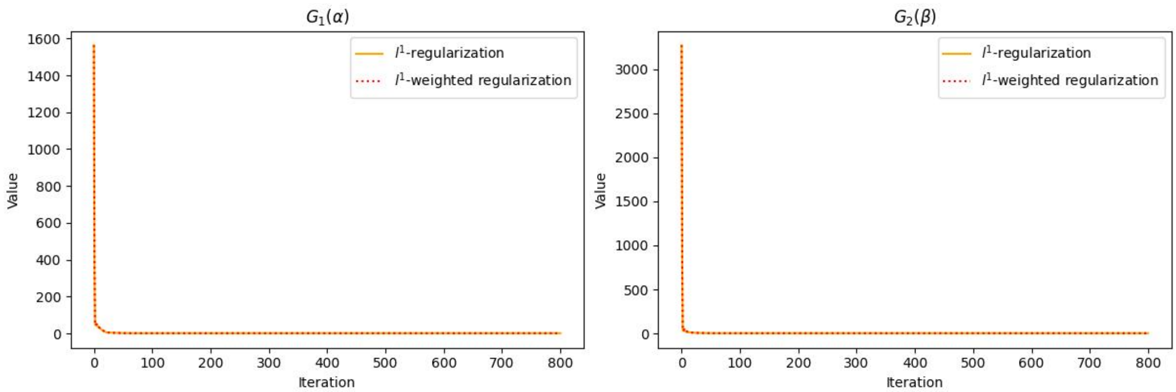

Figure 5.4: Example 5.1: Objective functions with $k = 2$ and 1% noise.

In addition, the recovered processes of parameters by using l^1-regularization and l^1-weighted regularization are successful and the convergence of α and β are illustrated in Figure 5.5. The parameter vector α converges to the solution $(\alpha^*)^T = [0, -2, 0, -3, 0]$ after no more than 200 iterations. Meanwhile, the parameter vector β need more than 300 iterations to obtain its convergence, $(\beta^*)^T = [0, 2, -1, 0, -3]$.

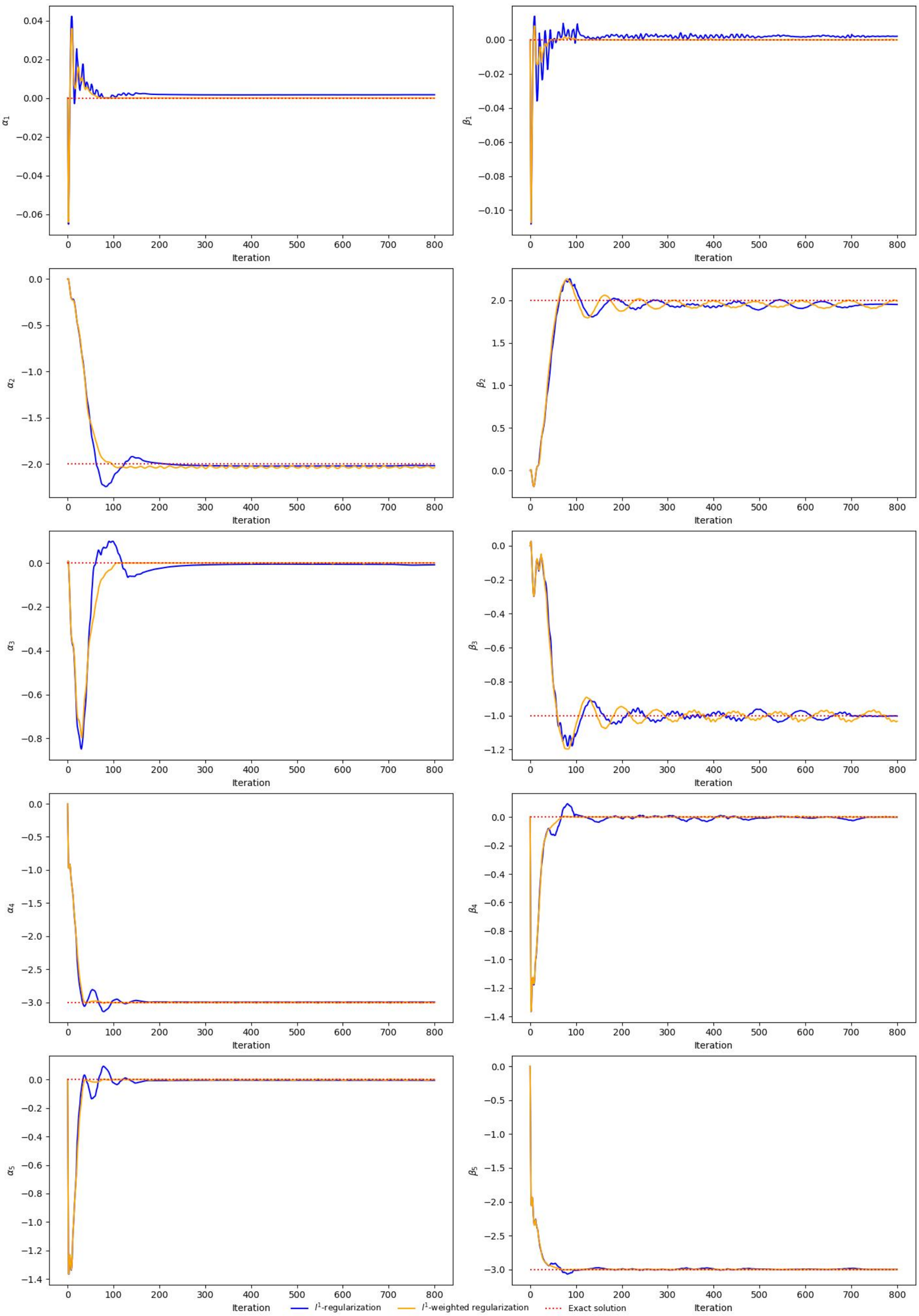

Figure 5.5: Example 5.1: The convergence of parameters α and β by using l^1-weighted regularization and l^1-regularization with $k = 2$ and 1% noise.

Tables 5.1 and 5.2 represent the error between the recovered parameters α, β and their exact solutions by using l^1-weighted regularization with $k = 2$ and 1% noise.

α^{800}	0	−2.03642653	0	−3.00516974	0
α^*	0	−2	0	3	0
Error	0	0.03642653	0	0.00516974	0

Table 5.1: Example 5.1: The recovered parameters α using l^1-weighted regularization with $k = 2$ and 1% noise.

β^{800}	0	1.99049525	−1.03311082	−0.00082761	−2.99778389
β^*	0	2	−1	0	−3
Error	0	0.00950475	0.03311082	0.00082762	0.00221611

Table 5.2: Example 5.1: The recovered parameters β using l^1-weighted regularization with $k = 2$ and 1% noise.

Tables 5.3 and 5.4 represent the error between the recovered parameters α, β and their exact solutions by using l^1-regularization with $k = 2$ and 1% noise.

α^{800}	0.00169953	−2.01628592	−0.00819491	−2.99847880	−0.00513288
α^*	0	−2	0	3	0
Error	0.00169953	0.01628592	0.00819492	0.0015212	0.00513288

Table 5.3: Example 5.1: The recovered parameters α using l^1-regularization with $k = 2$ and 1% noise.

β^{800}	0.00211043	1.95032496	−1.00248516	−0.00148607	−2.99763948
β^*	0	2	−1	0	−3
Error	0.00211043	0.04967504	0.00248516	0.00148608	0.00236052

Table 5.4: Example 5.1: The recovered parameters β using l^1-regularization with $k = 2$ and 1% noise.

From above estimate, we release that l^1-weighted regularization recover the parameters better than l^1-regularization, especially the sparse vector parameters.

The errors between the recovered parameter vectors α, β and the exact parameter vectors α^*, β^* by Algorithm 4.2 with l^1-weighted regularization and l^1-regularization after 800 iterations are illustrated in Figure 5.6.

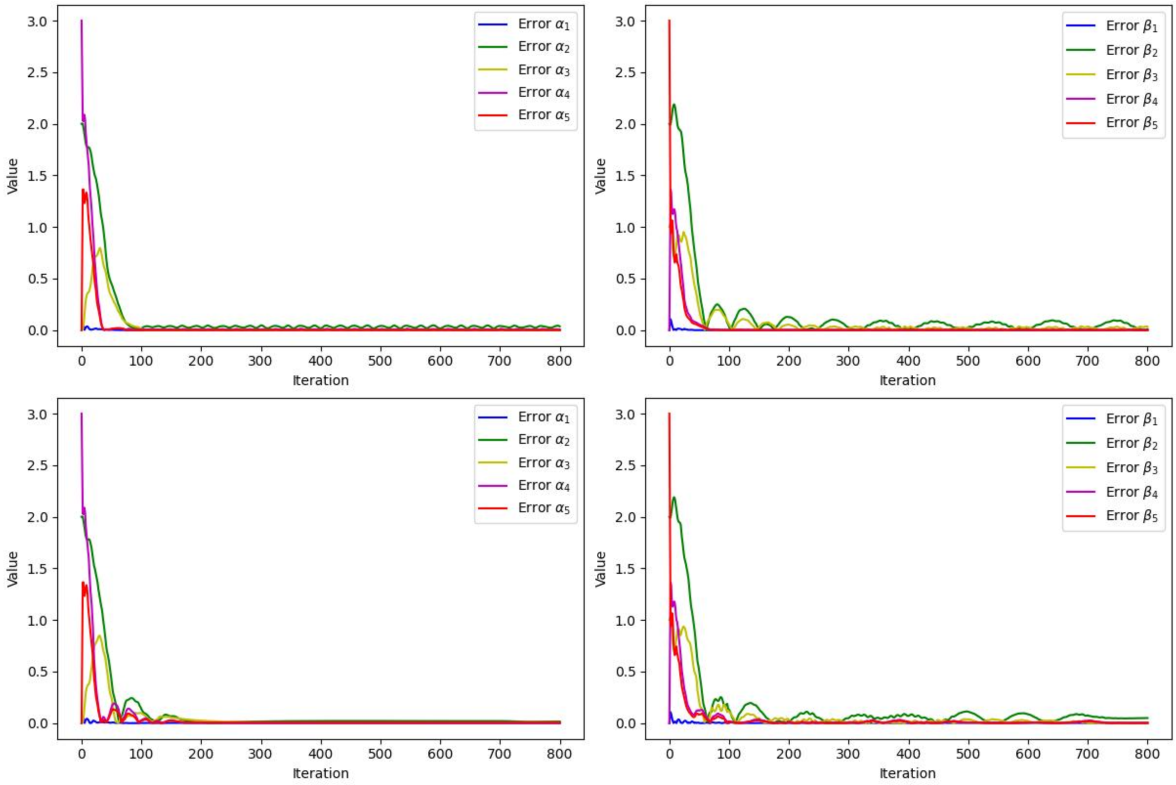

Figure 5.6: Example 5.1: The error between the recovered parameters α, β and the exact ones by using l^1-weighted regularization (top) and l^1-regularization (bottom) with $k = 2$ and 1% noise.

Table 5.5 and Table 5.6 show the recovered parameters of BOD-DO model in Example 5.1 when choosing collected data with $k = 2$, $k = 5$ and $k = 10$ to create the feature matrix F, V_1 and V_2 and exploiting Nesterov's accelerated method with l^1-weighted regularization, l^1-regularization after 800 iterations.

Exact parameters		$k = 2$		$k = 5$		$k = 10$	
α^*	β^*	α	β	α	β	α	β
0	0	0	0	0	0	0	0
-2	2	-2.03642	1.99049	-2.06182	1.88249	-2.16183	1.81195
0	-1	0	-1.03311	-0.00038	-1.00524	0	-1.04580
-3	0	-3.00516	-0.00082	-3.00928	0	-3.02205	0
0	-3	0	-2.99778	0	-2.99677	0	-2.99322

Table 5.5: Example 5.1: The recovered parameters with different input data using l^1-weighted regularization with 1% noise.

With 1% observation noise, Tables 5.5 and 5.6 indicate that although with different input data, the algorithm converges to the exact solution and learning process provides the reliable information in practical prediction. Moreover, it also points out that the

Exact parameters		$k = 2$		$k = 5$		$k = 10$	
α^*	β^*	α	β	α	β	α	β
0	0	0.00169	0.00211	0.00430	0.00525	0.00903	0.01146
-2	2	-2.01628	1.95032	-2.04122	1.87515	-2.07794	1.76940
0	-1	-0.00819	-1.00248	-0.01995	-1.00038	-0.04272	-1.02475
-3	0	-2.99847	-0.00148	-2.99482	-0.00128	-2.98836	-0.01516
0	-3	-0.00513	-2.99763	-0.01265	2.99574	-0.02545	-2.98363

Table 5.6: Example 5.1: The recovered parameters with different input data using l^1-regularization with $\varepsilon = 1\%$ noise.

more input data are collected, the more accurate solutions are received.

The same result is also obtained if we increase the observation noise at 5% (see Tables 5.7 and 5.8) and 10% (see Tables 5.9 and 5.10). Furthermore, from above observation, we conclude that l^1-weighted regularization recovers the form of the model better than l^1-regularization.

Exact parameters		$k = 2$		$k = 5$		$k = 10$	
α^*	β^*	α	β	α	β	α	β
0	0	0	0	0	0	0	0
-2	2	-2.04242	1.89316	-2.07196	1.82582	-2.12490	1.71230
0	-1	-0.00023	-0.96709	0	-0.96632	-0.00243	-0.97412
-3	0	-3.00625	0	-3.00112	0	-3.02122	0
0	-3	0	-2.99766	-0.00303	-2.99599	0	-2.99251

Table 5.7: Example 5.1: The recovered parameters with different input data using l^1-weighted regularization with 5% noise.

Exact parameters		$k = 2$		$k = 5$		$k = 10$	
α^*	β^*	α	β	α	β	α	β
0	0	0.00174	0.00161	0.00438	0.00499	0.00901	0.01177
-2	2	-2.01224	1.92222	-2.03795	1.85515	-2.08129	1.78092
0	-1	-0.00891	-0.98355	-0.02160	-0.99608	-0.04089	-1.02989
-3	0	-2.99459	-0.00612	-2.99383	-0.01049	-2.98886	-0.00141
0	-3	-0.00773	-2.99411	-0.01334	-2.98962	-0.02507	-2.99329

Table 5.8: Example 5.1: The recovered parameters with different input data using l^1-regularization with 5% noise.

Exact parameters		$k=2$		$k=5$		$k=10$	
α^*	β^*	α	β	α	β	α	β
0	0	0	0	0	0	0	0
-2	2	-2.03195	1.90489	-2.07915	1.85972	-2.15614	1.83558
0	-1	0	-0.96888	0	-0.98980	0	-1.06142
-3	0	-3.00504	0	-2.99280	0	-3.01381	0
0	-3	0	-2.99778	-0.00320	-2.99619	-0.02554	-2.99340

Table 5.9: Example 5.1: The recovered parameters with different input data using l^1-weighted regularization with 10% noise.

Exact parameters		$k=2$		$k=5$		$k=10$	
α^*	β^*	α	β	α	β	α	β
0	0	0.00194	0.00199	0.00440	0.00525	0.00906	0.01112
-2	2	-1.99544	1.94566	-2.03926	1.87793	-2.07961	1.77782
0	-1	-0.01337	-1.00203	-0.02033	-1.00560	-0.04193	-1.02350
-3	0	-2.98095	-0.00503	-2.99280	-0.00492	-2.98822	-0.01183
0	-3	-0.01691	-2.99494	-0.01400	-2.99334	-0.02554	-2.98603

Table 5.10: Example 5.1: The recovered parameters with different input data using l^1-regularization with 10% noise.

Example 5.2. *Consider the BOD-DO model*

$$\begin{cases} \dfrac{\partial b(x,t)}{\partial t} + 2\dfrac{\partial b(x,t)}{\partial x} = -2b(x,t), & \textit{in } (0,1)\times(0,3), \\ \dfrac{\partial d(x,t)}{\partial t} + 2\dfrac{\partial d(x,t)}{\partial x} = 2b(x,t) - d(x,t), & \textit{in } (0,1)\times(0,3), \\ b(x,0) = d(x,0) = 0, & \textit{on } (0,1), \\ b(0,t) = d(0,t) = 5(1-e^{-t}), & \textit{on } (0,3). \end{cases}$$

In this example, parameter vectors α^T, β^T are $\alpha^T = [0,-2,0,-2,0]$ and $\beta^T = [0,2,-1,0,-2]$, respectively. Moreover, the parameters are $v = 2, k_1 = 2$, $k_2 = 1$, $s_1 = s_2 = 0$, the initial conditions are $b_0 = d_0 = 0$ and the boundary conditions are $b_1 = d_1 = 5(1-e^{-t})$.

According to Theorem 3.1, the exact solution to this problem is

$$b(x,t) = \begin{cases} 0 & , \text{ if } x - 2t \in [0,1], \\ 5e^{-x}\left[1 - e^{-(t-\frac{x}{2})}\right] & , \text{ if } t - \dfrac{x}{2} \in [0,3], \end{cases}$$

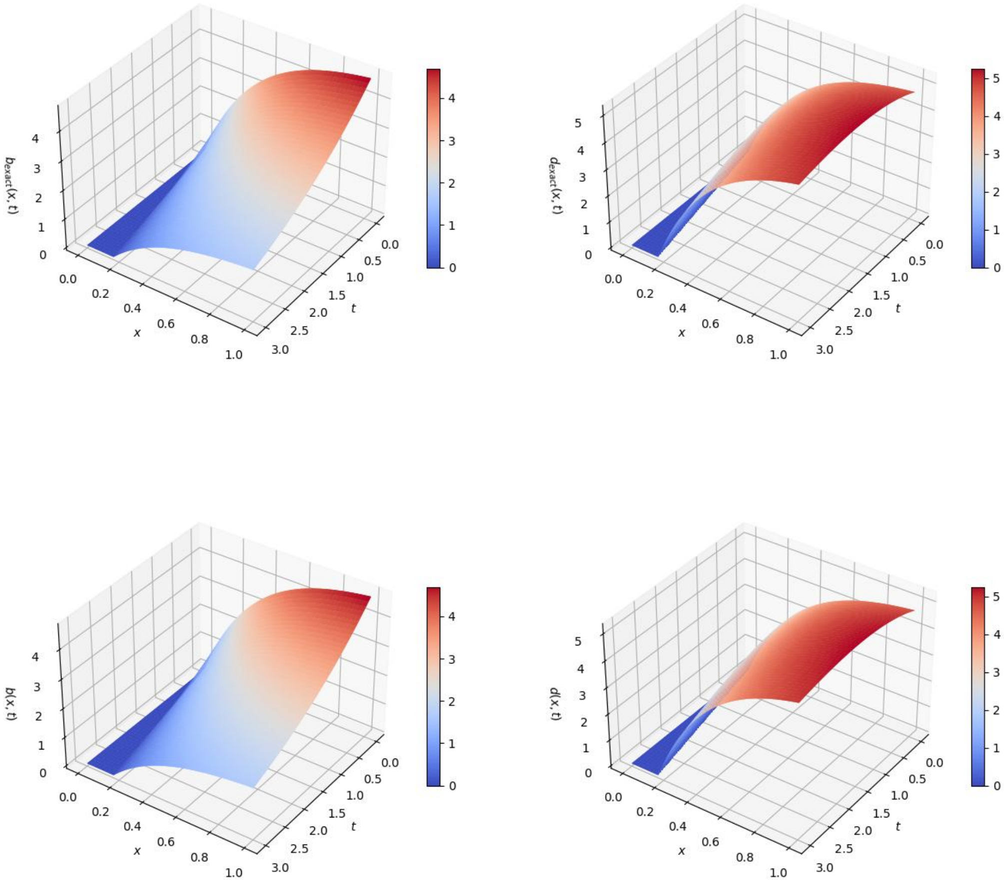

Figure 5.7: Example 5.2: Exact solution (top) and approximate solution (bottom) of the BOD-DO model.

and,

$$d(x,t) = \begin{cases} 0 & \text{, if } x - 2t \in [0,1], \\ -10e^{-x}\left[1 - e^{-(t-\frac{x}{2})}\right] + 5e^{-\frac{x}{2}}\left[e^{-(t-\frac{x}{2})} - 1\right] & \text{, if } t - \dfrac{x}{3} \in [0,3]. \end{cases}$$

Next, we make the same mesh points as the ones in Example 5.1 and apply the two-step Lax-Friedrichs method (Algorithm 4.1) to obtain the numerical solution to the BOD-DO model in Example 5.2. Figures 5.7 illustrates the numerical solution by Algorithm 4.1 for the exact and approximate solutions of the model. From Figure 5.8, it can be seen that the error between the exact and approximate solutions is not greater than 0.07%.

Similarity with Example 5.1, we collect data of the solutions $b_{\text{feature}}(x,t)$ and the $d_{\text{feature}}(x,t)$ from approximate solutions b, d with $k = 2$ (see Section 4.2), add 1% noise in the data and compute the derivatives of b_x, d_x, b_t and d_t to make the feature matrix

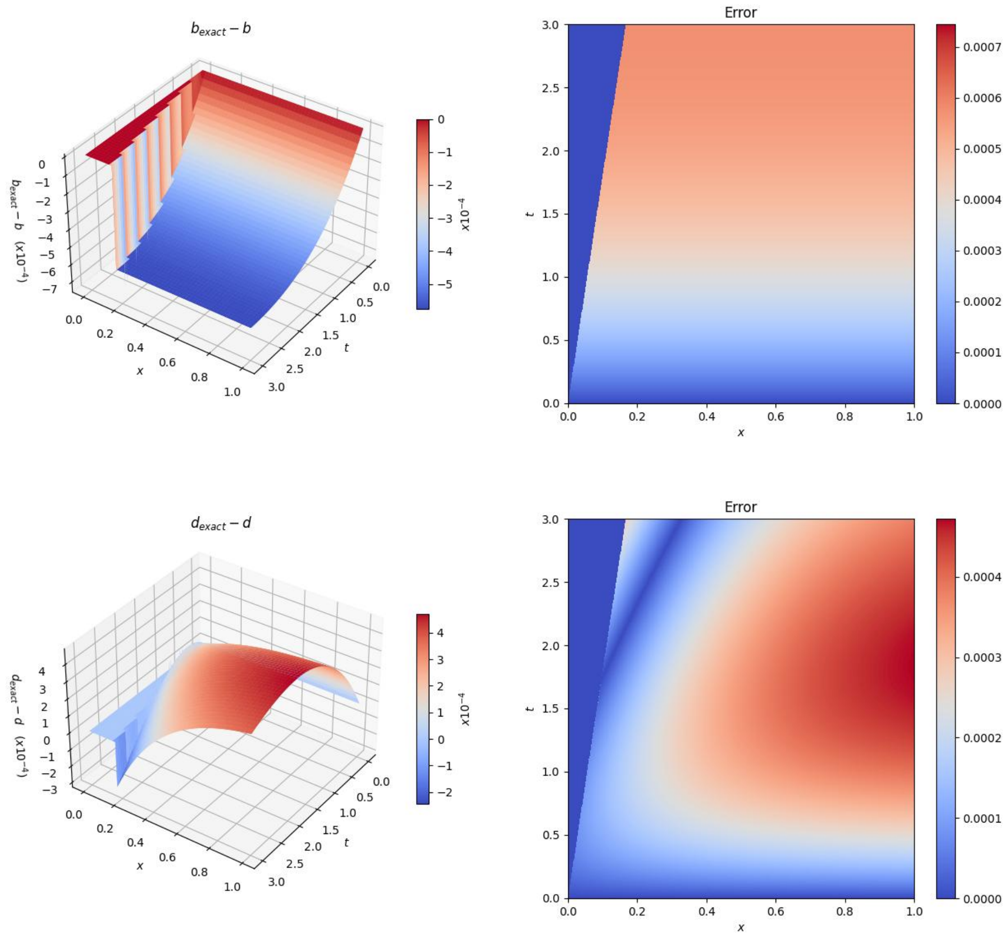

Figure 5.8: Example 5.2: The misfit and error between the exact and numerical solutions.

F as well as the matrix V. This leads to solve two minimization problems

$$\min_{\alpha\in\mathbb{R}^5}\{G_1(\alpha)+\gamma_1\Phi_1(\alpha)\} = \min_{\alpha\in\mathbb{R}^5}\left\{\frac{3}{1600}\sum_{i=1}^{1600}\left(\left\|V_1^\delta(t_i)-F^\delta(t_i)\alpha\right\|^2+\gamma_1\sum_{j=1}^{5}w_j^1|\alpha_j|\right)\right\}, \tag{5.3}$$

and

$$\min_{\beta\in\mathbb{R}^5}\{G_2(\beta)+\gamma_2\Phi_2(\beta)\} = \min_{\beta\in\mathbb{R}^5}\left\{\frac{3}{1600}\sum_{i=1}^{1600}\left(\left\|V_2^\delta(t_i)-F^\delta(t_i)\beta\right\|^2+\gamma_2\sum_{j=1}^{5}w_j^2|\beta_j|\right)\right\}. \tag{5.4}$$

We set $\gamma_1 = \gamma_2 = 10^{-2}$. Unlike Example 5.1, l^1-regularization does not achieve the convergence to the minimizer of problems (5.2) and (5.4). Thus, we have to use l^1-weighted regularization for these problems. Because of a priori information, we know that only parameters s_1, s_2, k_1, k_2, v are possibly nonzero while the other must be zero. From this information, we set

$$\omega_j^1 = \begin{cases} 1, & \text{for } j = 2, 4, \\ 100, & \text{otherwise,} \end{cases} \quad \text{and} \quad \omega_j^2 = \begin{cases} 1000, & \text{for } j = 1, 4, \\ 10, & \text{for } j = 2, 3, \\ 200, & \text{for } j = 5. \end{cases}$$

Then, we exploit Nesterov's accelerated method (Algorithm 4.2) for with the initial point $[0, 0, 0, 0, 0]$ for two problems (5.3) and (5.4) in 500 iterations.

Figure 5.9 shows that it takes more than 200 iterations to achieve the optimal value.

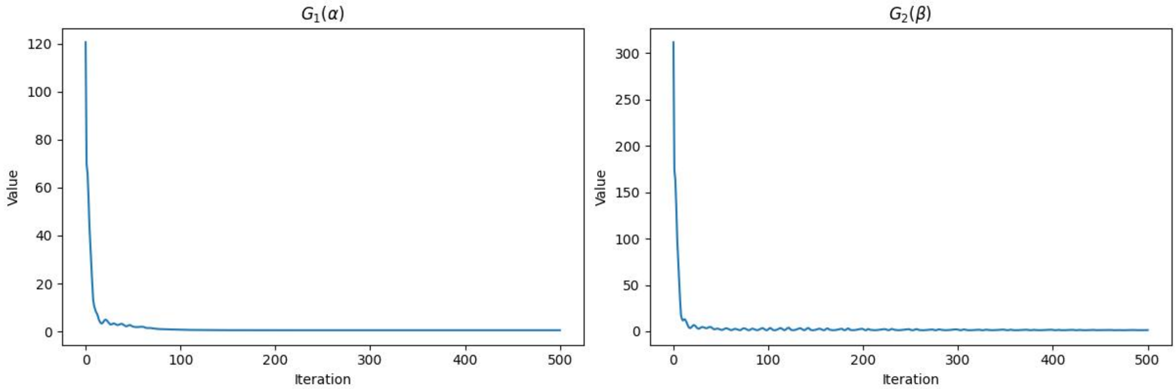

Figure 5.9: Example 5.2: Objective functions with $k = 2$ and 1% noise.

From Figure 5.10, Tables 5.11 and 5.12, we can see that the parameters are recovered successfully, i.e., the model (5.3)−(5.4) has been learned successfully.

α^{800}	0	−2.00409935	0.00077286	−2.00479683	0
α^*	0	−2	0	−2	0
Error	0	0.004099357	0.00077287	0.00479683	0

Table 5.11: Example 5.2: The recovered parameters α using l^1-weighted regularization with $k = 2$ and 1% noise.

β^{800}	0	1.93016615	−0.95407856	0	−1.94127817
β^*	0	2	−1	0	−2
Error	0	0.06983385	0.04592144	0	0.05872183

Table 5.12: Example 5.2: The recovered parameters β using l^1-weighted regularization with $k = 2$ and 1% noise.

The errors between the recovered parameter vectors α, β and the exact parameter vectors α^*, β^* by Algorithm 4.2 with l^1-weighted regularization after 500 iterations are illustrated in Figure 5.11.

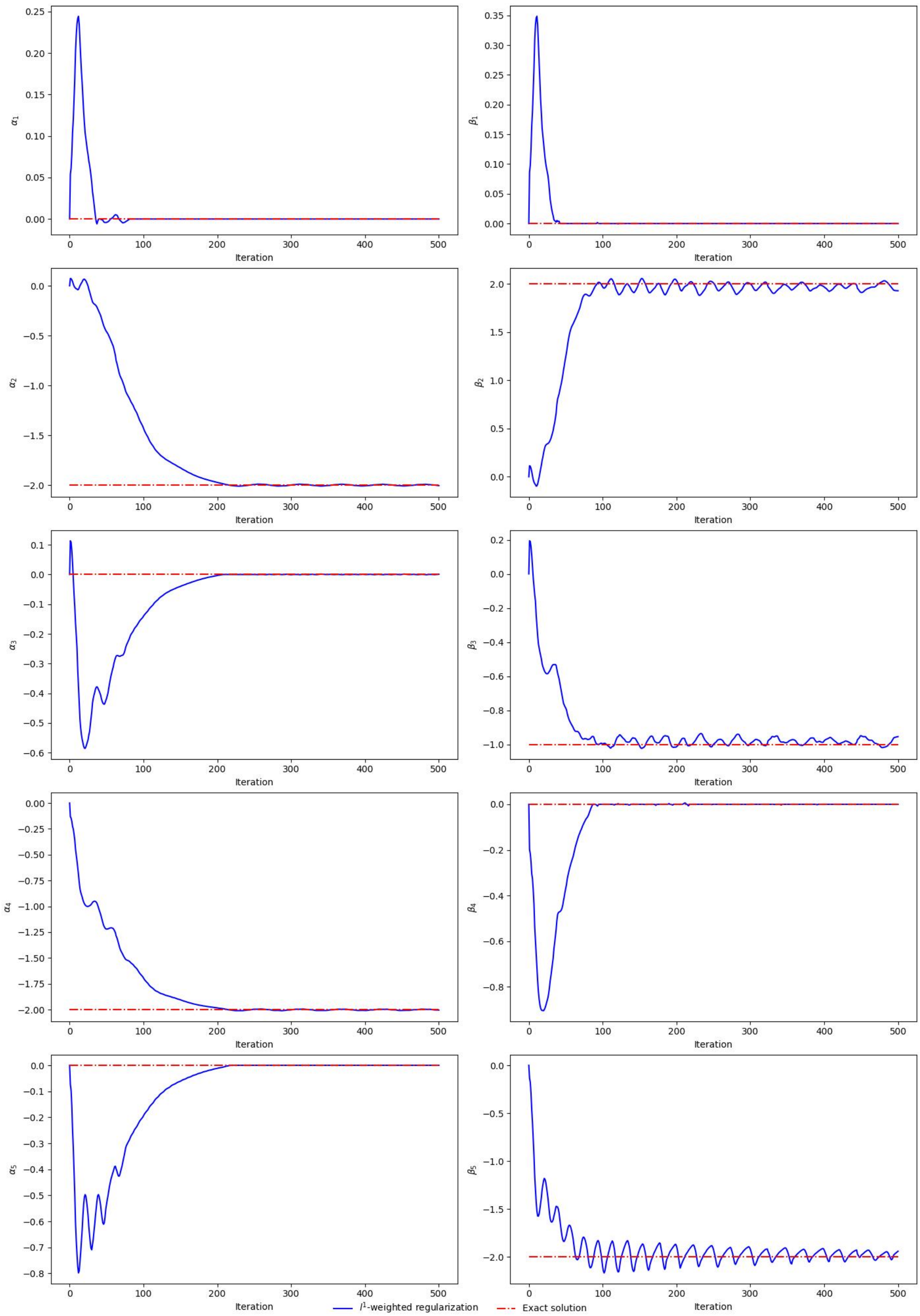

Figure 5.10: Example 5.2: The convergence of parameters α and β by using l^1-weighted regularization with $k = 2$ and 1% noise.

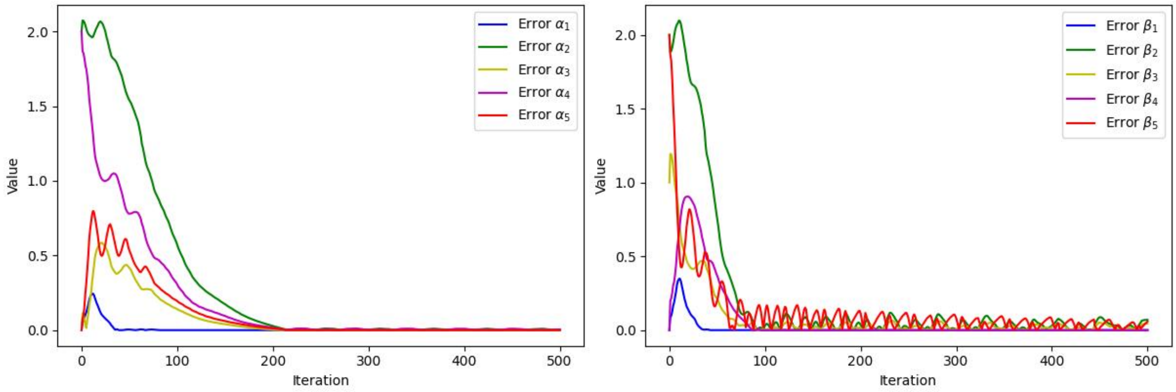

Figure 5.11: Example 5.2: The error between the recovered parameters α, β and the and exact ones by using l^1-weighted regularization with $k = 2$ and 1% noise.

Table 5.13 shows the recovered parameters of BOD-DO model when we choose collected data with $k = 2$, $k = 5$ and $k = 10$ to create the feature matrix F, V_1 and V_2 by Nesterov's accelerated method (Algorithm 4.2) with l^1-weighted regularization after 800 iterations. It is similar to Example 5.1, if we have more data, the accuracy of approximate solution is more improvable.

Exact parameters		$k = 2$		$k = 5$		$k = 10$	
α^*	β^*	α	β	α	β	α	β
0	0	0	0	0	0	0	0
-2	2	-2.00409	1.93016	-2.00443	1.87582	-1.99515	1.83560
0	-1	0.00077	-0.95407	0.00207	-0.92844	0.00135	-0.90391
-2	0	-2.00479	0	-2.00597	0	-2.00378	0
0	-2	0	1.94127	0	-1.97346	0	-1.89327

Table 5.13: Example 5.2: The recovered parameters with different input data using l^1-weighted regularization with 1% noise.

Exact parameters		$k = 2$		$k = 5$		$k = 10$	
α^*	β^*	α	β	α	β	α	β
0	0	0	0	0	0	0	0
-2	2	-1.99398	1.99662	-2.01226	1.83645	-1.98861	1.86889
0	-1	0.00031	-0.99676	0.00125	-0.89668	-0.00188	-0.91926
-2	0	-1.99718	0	-2.01368	0	-2.00715	0
0	-2	0	-1.99150	0	-1.93755	0	-1.86488

Table 5.14: Example 5.2: The recovered parameters with different input data using l^1-weighted regularization with 5% noise.

The same result is also obtained if we increase the noise at 5% (see Table 5.14) and

Exact parameters		$k = 2$		$k = 5$		$k = 10$	
α^*	β^*	α	β	α	β	α	β
0	0	0	0	0	0	0	0
-2	2	-1.99563	1.94133	-2.00571	1.83026	-1.98604	1.80944
0	-1	0.00030	-0.96131	0.00161	-0.89476	-0.00184	-0.88721
-2	0	-1.99779	0	-2.00768	0	-2.01261	0
0	-2	0	-1.93156	0	-1.91683	0	-1.91842

Table 5.15: Example 5.2: The recovered parameters with different input data using l^1-weighted regularization with 10% noise.

10% (see Table 5.15). From these tables, we get the higher error if the noise increases but the model has been learned successfully.

THIS CHAPTER WAS WRITTEN BASED ON THE PAPER

[15] Hao D.N., Hiep D.X., Muoi P.Q., 2023, Learning river water quality models by l^1-weighted regularization. Published in *IMA Journal of Applied Mathematics*.

CONCLUSIONS

In this thesis, we have introduced the learning models from data and investigated its applications in inverse problem: learning constants parameters in BOD-DO model

$$
\begin{aligned}
&\frac{\partial b}{\partial t} + v\frac{\partial b}{\partial x} = -k_1 b + s_1 && \text{in } (0, X) \times (0, T],\\
&\frac{\partial d}{\partial t} + v\frac{\partial d}{\partial x} = k_1 b - k_2 d + s_2 && \text{in } (0, X) \times (0, T],\\
&b(x, 0) = b_0(x),\ d(x, 0) = d_0(x) && \text{on } (0, X),\\
&b(0, t) = b_1(t),\ d(0, t) = d_1(t) && \text{on } (0, T].
\end{aligned}
$$

First, we establish the exact solution of above model and prove the well-posedness and convergence in learning the BOD-DO model based on l^1-weighted regularization.

Second, we propose the numerical algorithms for the BOD-DO model in the direct and inverse problem with constant parameters. In the direct problem, we apply the two-step Lax-Friedrichs method. Then, Nesterov's accelerated method is applied to solve the minimization problem to get the unknown parameters in learning BOD-DO model with l^1-weighted regularization.

Finally, we test our algorithms in some examples in Python software with different noise. These numerical examples result the efficiency of our approach in approximating the unknown parameters of BOD-DO model. Moreover, the numerical examples also illustrate that l^1-weighted regularization overwhelms l^1-regularization in learning and maintaining the form of BOD-DO model.

There is potential development in future work derive from this thesis. The theorerical analysis in general models, e.g., the BOD-DO model with variable parameters is still open. The study that involves to determine a source term in this model is also one of promising work. Moreover, the problem of choosing the weighted parameters in l^1-weighted regularization have not been solved yet.

THE AUTHOR'S PUBLICATION RELATED TO THE THESIS

[1] Hao D.N., Hiep D.X., Muoi P.Q., 2023, Learning river water quality models by l^1-weighted regularization. Published in *IMA Journal of Applied Mathematics*.

BIBLIOGRAPHY

[1] Streeter H.W., Phelps E.B., 1925, A study of the Pollution and Natural Purification of the Ohio Rivers, *US Public Health Service Bulletin*, 146.

[2] Chapra S.C., 1997, *Surface Water-Quality Modeling*, Waveland Press, Inc.

[3] Schnoor J.L., 1996, *Environmental Modeling: Fate and Transport of Pollutants in Water, Air and Soil*, John Wiley & Sons, Inc.

[4] Rezzolla L., 2011, Numerical methods for the solution of partial differential equations, *Lecture Notes for the COMPSTAR School on Computational Astrophysics*, pp. 21–22.

[5] Shamphine L.F., 2005, Two-step Lax-Friedrichs method, *Applied Mathematics Letters*, 18(10), pp. 1134–1136.

[6] Brunton S.L., Kutz J.N., 2022, *Data-Driven Science and Engineering: Machine Learning, Dynamical Systems and Control*, volume 2nd, Cambridge University Press.

[7] Brunton S.L., Proctor J.L., Kutz J.N., 2016, Discovering governing equations from data by sparse identification of nonlinear dynamical systems, *Proceedings of the National Academy of Sciences*, 113(15), pp. 3932–3937.

[8] Champion K., Lusch B., Kutz J.N., Brunton S.L., 2019, Data-driven discovery of coordinates and governing equations, *Proceedings of the National Academy of Sciences*, 116(45), pp. 22445–22451.

[9] Kutz J.N., 2013, *Data-Driven Modeling and Scientific Computation: Methods for Complex Systems and Big Data*, volume 1st, Oxford University Press.

[10] Mendez M.A., Ianiro A., Noack B.R., Brunton S.L., 2023, *Data-Driven Fluid Mechanics Combining First Principles and Machine Learning*, volume 1st, Cambridge University Press.

[11] Rudy S.H., Alla A., Brunton S.L., Kutz J.N., 2019, Data-driven identification of parametric partial differential equations, *SIAM Journal on Applied Dynamical Systems*, 18(2), pp. 643–660.

[12] Rudy S.H., Brunton S.L., Proctor J.L., Kutz J.N., 2017, Data-driven discovery of partial differential equations, *Science Advances*, 3(4), p. e1602614.

[13] Rudy S.H., Kutz J.N., Alla A., Brunton S.L., 2019, Deep learning of dynamics and signal-noise decomposition with time-stepping constraints, *Journal of Computational Physics*, 396(1), pp. 483–506.

[14] Schaeffer H., 2017, Learning partial differential equations via data discovery and sparse optimization, *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 473(2197), p. 20160446.

[15] Hao D.N., Hiep D.X., Muoi P.Q., 2023, Learning river water quality models by l^1-weighted regularization, *Published in IMA Journal of Applied Mathematics.*

[16] Muoi P.Q., Mass P., Hao D.N., Pidcock M., 2016, Descent gradient methods for nonsmooth minimization problems in ill-posed problems, *Journal of Computational and Applied Mathematics*, 298(15), pp. 105–122.

[17] Nesterov Y., 2013, Gradient methods for minimizing composite functions, *Mathematical Programming*, 140, pp. 125–161.

[18] Hadamard J., 1923, *Lectures on the Cauchy problem in Linear Partial Differential Equations*, Yale University Press.

[19] Kirsch A., 2021, *An Introduction to the Mathematical Theory of Inverse Problems*, volume 3rd, Springer Nature Switzerland.

[20] Trefethen L.N., Bau III D., 1997, *Numerical Linear Algebra*, Society for Industrial and Applied Mathematics.

[21] Zeidler E., 1995, *Applied Functional Analysis: Applications to Mathematical Physics*, Springer-Verlag.

[22] Zeidler E., 1990, *Nonlinear Functional Analysis and its Applications II/B: Nonlinear Monotone Operators*, Springer New York, NY.

[23] Rockafellar R.T., 1966, Characterization of the subdifferentials of convex functions, *Pacific journal of mathematics*, 17(3), pp. 497–510.

[24] Evans L.C., 2010, *Partial Differential Equations*, volume 2nd, American Mathematical Society.

[25] Daubechies I., Defrise M., De Mol C., 2004, An iterative thresholding algorithm for linear inverse problems with a sparsity constraint, *Communications on Pure and Applied Mathematics*, LVII, pp. 1413–1457.

[26] Bonesky T., Bredies K., Lorenz D.A., Maass P., 2007, A generalized conditional gradient method fornonlinear operator equations with sparsity constraints, *Inverse problems*, 23(5), pp. 2041–2058.

[27] Bredies K., Lorenz D.A., Maass P., 2009, A generalized conditional gradient method and its connection to an iterative shrinkage method, *Computational Optimization and Applications*, 42, pp. 173–193.

[28] Lorenz D.A., Mass P., Muoi P.Q., 2012, Gradient descent for Tikhonov functionals with sparsity constrains: Theory and numerical comparison step size rules, *Electronic Transactions on Numerical Analysis*, 39, pp. 437–463.

[29] Muoi P.Q., 2012, *Sparsity Constraints and Regularization for Nonlinear Inverse Problems*, PhD thesis, University of Bremen.

[30] Bongard J., Lipson H., 2007, Automated reverse engineering of nonlinear dynamical systems, *Proceedings of the National Academy of Sciences*, 104(24), pp. 9943–9948.

[31] Schmidt M., Lipson H., 2009, Distilling free-form natural laws from experimental data, *Science*, 324(5923), pp. 81–85.

[32] Breiman L., 2001, Statistical modeling: The two cultures, *Statistical science*, 16(3), pp. 199–231.

[33] Crutchfield J.P., McNamara B.S., 1987, Equations of motion from a data series, *Complex Systems*, 1, pp. 417–452.

[34] Benner P., Gugercin S., Willcox K., 2015, A survey of projection-based model reduction methods for parametric dynamical systems, *SIAM Review*, 57(4), p. 483–531.

[35] Kepler G.M., Tran T.H., Banks H.T., 2000, Reduced order model compensator control of species transport in a CVD reactor, *Optimal Control Applications and Methods*, 21, pp. 143–160.

[36] Kepler G.M., Tran T.H., Banks H.T., 2001, Compensator control for chemical vapor deposition film growth using reduced-order design models, *IEEE Transactions on Semiconductor Manufacturing*, 14(3), pp. 231–241.

[37] Hao D.N., Quyen T.N.T., 2010, Convergence rates for Tikhonov regularization of coefficient identification problems in Laplace-type equation, *Inverse problems*, 26(12), p. 125014.

[38] Hao D.N., Quyen T.N.T., 2011, Convergence rates for total variation regularization of coefficient identification problems in elliptic equations I, *Inverse problems*, 27(7), p. 075008.

[39] Muoi P.Q., 2015, Reconstructing conductivity coefficients based on sparsity regularization and measured data in electrical impedance tomography, *Inverse Problems in Science and Engineering*, 23(8), pp. 1366–1387.

[40] Grasmair M., Haltmier M., Scherzer O., 2008, Sparsity regularization with l^p penalty term, *Inverse problems*, 24(5), p. 055020.

[41] Ito K., Jin B., Takeuchi T., 2011, Multi-parameter Tikhonov regularization, *Methods and Applications of Analysis*, 18, pp. 31–46.

[42] Beck A., 2014, *Introduction to Nonlinear Optimization: Theory, Algorithms, and Applications with MATLAB*, Society for Industrial and Applied Mathematics.

[43] Ekaland I., Témam R., 1999, *Convex Analysis and Variational Problems*, Society for Industrial and Applied Mathematics.

[44] Beck A., Teboulle M., 2009, A Fast Iterative Shrinkage-Thresholding Algorithm for Linear Inverse Problems, *SIAM Journal on Imaging Science*, 2(1), pp. 183–202.

[45] Bauschke H.H., Combettes P.L., 2011, *Convex Analysis and Monotone Operator Theory in Hilbert*, Springer New York, NY.

[46] Barzilai J., Borwein J.M., 1988, Two-point step size gradient methods, *IMA Journal of Numerical Analysis*, 8(1), pp. 141–148.

[47] Beck A., 2017, *First-order Methods in Optimization*, Society for Industrial and Applied Mathematics.