TURUN YLIOPISTO
UNIVERSITY OF TURKU

GENERALIZED ORLICZ SPACES AND PARTIAL DIFFERENTIAL EQUATIONS WITH APPLICATION TO IMAGE RESTORATION

Debangana Baruah

TURUN YLIOPISTON JULKAISUJA – ANNALES UNIVERSITATIS TURKUENSIS
SARJA – SER. AI OSA – TOM. 677 | ASTRONOMICA – CHEMICA – PHYSICA – MATHEMATICA | TURKU 2022

TURUN
YLIOPISTO
UNIVERSITY
OF TURKU

GENERALIZED ORLICZ SPACES AND PARTIAL DIFFERENTIAL EQUATIONS WITH APPLICATION TO IMAGE RESTORATION

Debangana Baruah

TURUN YLIOPISTON JULKAISUJA – ANNALES UNIVERSITATIS TURKUENSIS
SARJA – SER. AI OSA – TOM. 677 | ASTRONOMICA – CHEMICA – PHYSICA – MATHEMATICA | TURKU 2022

University of Turku

Faculty of Science
Department of Mathematics and Statistics
Mathematics
Doctoral Programme in Exact Sciences (EXACTUS)

Supervised by

Professor Peter Hästö
University of Turku

Docent Petteri Harjulehto
University of Turku

Docent Napsu Karmitsa
University of Turku

Professor Marko Mäkelä
University of Turku

Reviewed by

Professor Petri Juutinen
University of Jyväskylä

Associate Professor Anna Zatorska-Goldstein
University of Warsaw

Opponent

Professor Juha Kinnunen
Aalto University

ISBN 978-951-29-9035-1 (PRINT)
ISBN 978-951-29-9036-8 (PDF)
ISSN 0082-7002 (PRINT)
ISSN 2343-3175 (ONLINE)
Painosalama, Turku, Finland, 2022

Dedicated to my Math teachers
(St. Mary's school, Guwahati)

UNIVERSITY OF TURKU
Faculty of Science
Department of Mathematics and Statistics
Mathematics
BARUAH, DEBANGANA: Generalized Orlicz Spaces and Partial Differential Equations with Application to Image Restoration
Doctoral dissertation, 116 pp.
Doctoral Programme in Exact Sciences (EXACTUS)
October 2022

ABSTRACT

The research area of this thesis is nonlinear functional analysis, a branch of Mathematics which examines questions related to qualitative aspects of solution of a differential equation, such as existence, uniqueness, stability, solvability conditions. Owning to the rapid progress in image processing research involving variational problems, this research work deals with the study of existence and properties of minimizer for image restoration model under Sobolev-Orlicz function space setting.

The nature of image restoration that we deal here is noise reduction, where the observed image is assumed to be degraded by a random noise. The noise reduction problem is formulated as a minimization problem consisting of a least squares fit and a regularization term. In the proposed image denoising model, the regularization term represent a double-phase functional that serve the purpose of anisotropic diffusion along with isotropic smoothing, for piecewise smoothing and edge preservation. The mathematical modeling of image restoration problem requires the setting of the domain function space to permit discontinuities of the solution. In this respect, Sobolev-Orlicz function space, which consists of functions having weak derivatives and satisfy certain integrability conditions, provide a favorable framework. For solving such minimization problems, the so-called *direct method in the Calculus of Variations* is widely used, whose basic topological ingredients are the lower semicontinuity of the functional and the compactness of the lower level sets of the regularization functional. The natural question which then arises here is to study the regularity of such solutions and to establish under which conditions on the data and domain, we have a solution in the sense of distributions. This forms the main objective of my research from the theoretical perspective. Although considerable contributions have been devoted to this challenging question, investigating new approaches under the Sobolev-Orlicz space setting provide new insight into the matter.

In this thesis, the study of image restoration problem is carried out in two approaches: variational and PDE-based. The variational approach presents restoration through minimization, where the existence and uniqueness of minimizer is established using the direct Method of the Calculus of Variations. This approach gives information about the qualitative aspects of the model in the Sobolev-Orlicz space setting. In the PDE-based approach, we consider models in the form of heat flow differential equation, where the image is embedded in an evolution process in both space and time dimensions. This yields a quasilinear parabolic boundary value problem. However, due to the degenerate behavior of the PDE, it is not possible to ap-

ply general results from classical parabolic equations theory. Thus, to formulate a well-adapted framework, we regularize the PDE using approximations to obtain appropriate solvability conditions. The idea is to construct an approximated boundary problem whose solution converges to the solution of the heat flow problem, under the suitable conditions. Further, to prove the existence and uniqueness of the solution, we derive a few a priori estimates, which gives information about the qualitative behavior of the boundary function. This approach is particularly useful in determining the nature of the domain, where the image corresponds to a feasible solution, that is usually required for numerical purposes.

Finally, after proving the existence and uniqueness of the solution, we discretize the problem in order to find a numerical solution. The behaviour and efficiency of the model is then tested and illustrated through numerical experiments.

KEYWORDS: image restoration, double-phase, Sobolev-Orlicz space, minimizer, heat flow, PDE

TURUN YLIOPISTO
Tieteen Tiedekunta
Matematiikan ja Tilastotieteen Laitos

BARUAH, DEBANGANA: Generalized Orlicz Spaces and Partial Differential Equations with Application to Image Restoration
Väitöskirja, 116 s.

Lokakuu 2022

TIIVISTELMÄ

Tämän opinnäytetyön tutkimusalueena on epälineaarinen funktionaalinen analyysi, matematiikan haara, joka tutkii differentiaaliyhtälön ratkaisun laadullisiin näkökohtiin liittyviä kysymyksiä, kuten olemassaoloa, yksikäsitteisyyttä ja stabiilisuutta. Johtuen nopeasta edistymisestä kuvankäsittelytutkimuksessa, joka liittyy variaatio-ongelmiin, tämä tutkimustyö käsittelee kuvan restaurointimallin minimoijan olemassaoloa ja ominaisuuksia Sobolev-Orliczin funktioavaruuksissa.

Tässä käsiteltävä kuvan restauroinnin luonne on kohinanvaimennus, jossa oletetaan havaitun kuvan sisältävän satunnaista kohinaa. Kohinanvaimennusongelma on muotoiltu minimoimisongelmaksi, joka koostuu pienimmän neliösumman sovituksesta ja regularisointitermistä. Ehdotetussa kuvan kohinanpoistomallissa regularisointitermi edustaa kaksivaiheista (double phase) funktiota, jossa on sekä anisotrooppinen että isotrooppinen diffuusio paloittaista tasoittamista ja reunan säilyttämistä varten. Kuvan restaurointiongelman matemaattinen mallinnus edellyttää funktioavaruutta, joka mahdollistaa epäjatkuvuudet ratkaisussa. Tässä suhteessa Sobolev-Orlicz-funktioavaruus, joka koostuu funktioista, joilla on heikko derivaatta ja tietyt integroitavuusehdot, tarjoavat suotuisat puitteet. Tällaisten minimointiongelmien ratkaisemiseksi ns. variaatiolaskun suora menetelmä on laajalti käytössä, ja sen topologiset perusainekset ovat funktionaalin alhaalta puolijatkuvuus ja kompaktisuus. Luonnollinen kysymys on tutkia tällaisten ratkaisujen säännöllisyyttä ja pyrkiä määrittämään, millä ehdoilla meillä on ratkaisu distribuution mielessä. Tämä on tutkimukseni päätavoite teoreettisesta näkökulmasta. Vaikka tätä haastavaa kysymystä on tutkittu paljon, uudet lähestymistavat Sobolev-Orliczin avaruuksissa antavat uutta näkemystä asiaan.

Tässä opinnäytetyössä kuvan restaurointiongelmaa tutkitaan kahdella lähestymistavalla: variaatio- ja ODY-perustaisesti. Variaatiolaskennassa restaurointia lähestytään minimoinnin kautta, jossa olemassaolo ja minimoinnin yksikäsitteisyys sadaan käyttämällä suoraa variaatiolaskentamenetelmää. Tämä lähestymistapa antaa tietoa mallin laadullisista näkökohdista Sobolev-Orliczin avaruudessa. ODY-pohjaisessa lähestymistavassa tutkimme lämpöyhtälön muotoisia malleja, joissa on evoluutioprosessi sekä tila- että aikaulottuvuuksissa. Tämä tuottaa kvasilineaarisen parabolisen reuna-arvo-ongelman. ODY:n degeneroidun käyttäytymisen vuoksi siihen ei kuitenkaan voida soveltaa yleisiä tuloksia klassisesta parabolisten yhtälöiden teoriasta. Joten muotoillaksemme sopivan kehyksen, normalisoimme ODY:n käyttämällä approksimaatioita saadaksemme sopivat ratkeavuusehdot. Ideana on rakentaa likimääräiset reuna-arvo-

ongelmat, joiden ratkaisut suppenevat lämpöyhtälön ratkaisua kohti sopivissa olosuhteissa. Lisäksi todistaaksemme ratkaisun olemassaolon ja yksikäsitteisyyden johdamme a priori arvioita, jotka antavat tietoa rajafunktion laadullisesta käyttäytymisestä. Tämä lähestymistapa on erityisen hyödyllinen määritettäessä alueen luonnetta, jossa kuva vastaa mielekästä ratkaisua, jota yleensä tarvitaan numeerisissa sovelluksissa.

Lopuksi, kun olemme todistaneet ratkaisun olemassaolon ja yksikäsitteisyyden, diskretisoimme ongelman löytääksemme numeerisen ratkaisun. Sen jälkeen mallin käyttäytymistä ja tehokkuutta testataan ja havainnollistetaan numeerisilla kokeilla.

Acknowledgements

There may be just one word to describe my journey of PhD as "extraordinary" but many to be thankful of behind the milestone. First, I would like to express my profound gratitude to my PhD supervisors: Professor Peter Hästö and Docent Petteri Harjulehto, along with Docent Napsu Karmitsa and Professor Marko Mäkelä. I feel priviledged to have worked with Professor Peter Hästö whose guidance have been monumental in building my learning approach and critical thinking. I am thankful to Docent Petteri Harjulehto for his continuous guidance and comments that have shaped my PhD thesis. A special thank you to Docent Napsu Karmitsa and Professor Marko Mäkelä for giving me the opportunity to work with them and discover application areas for my work. I would like to thank the pre-examiners, Professor Petri Juutinen and Associate Professor Anna Zatorska-Goldstein, for reviewing my thesis and providing useful input, and also Professor Juha Kinnunen for being my opponent. I am thankful for the financial support received for my PhD research from the Doctoral Programme (previously MATTI and now EXACTUS) of University of Turku, Department of Mathematics and Statistics, the Vilho, Yrjö and Kalle Väisälä Foundation.

I would like to extend my sincere thanks to Professor Iiro Honkala for his continuous support in the department. I wish to express my gratitude to Docent Yury Nikulin for the opportunity of taking majority of the courses during my PhD studies, in the areas of optimization and partial differential equations. In the department, my time would not have been complete without the cheerful company of my fellow colleagues, including Shemunyenge Hamukwaya and especially Bishwesvar Pratap Singh. I would also like to thank Professor Heiko Gimperlein, my former supervisor for Masters thesis, for motivating me to pursue PhD in this field.

My time at University of Turku has been a pleasant experience, and moreover, the beautiful city of Turku is one of the best places for me to live and study. And last but not the least, I will always be grateful to my parents, Gayatri Baruah and Amlan Jyoti Baruah, and my sister Dyutismita Baruah, for supporting me all the way; thank you for being my strength even after being miles away in India.

My sincere thanks to all of my friends and family for the good wishes and optimism, without whom this journey would not be the same.

Turku, October 2022
Debangana Baruah

Table of Contents

1 Introduction

"In Mathematics, the art of proposing a question must be held of higher value than solving it."

— Georg Cantor (1845-1918)

We begin by presenting an intuitive introduction to the class of generalized Orlicz spaces, briefly sketching its background on variational problems, followed by an outline of the thesis.

1.1 A brief background

In the theory of existence of solutions to minimization problems, functionals with non-standard growth have been studied extensively due to their application in modelling physical phenomena. Such classes of functionals are mostly covered by generalized Orlicz spaces, as we look back how it was developed.

The study of generalized Orlicz spaces L^φ can be traced back to the 1940s. These spaces are similar to Orlicz spaces [50], but defined by a more general function $\varphi(x,t)$, that may vary with the location in space with the norm defined by means of the integral

$$\int_{\mathbb{R}^n} \varphi(x, |u(x)|)\, dx,$$

whereas in Orlicz space, the function φ would be independent of x, that is $\varphi(|u(x)|)$. When $\varphi(t) = t^p$, we obtain the Lebesgue spaces, L^p. Other principal classes of examples of generalized Orlicz spaces include variable exponent spaces, where $\varphi(x,t) := t^{p(x)}$ [14], and double phase spaces, where $\varphi(x,t) := t^p + a(x)t^q$ [3; 13].

Historically, variable exponent Lebesgue spaces $L^{p(\cdot)}(\Omega)$, where Ω is an open subset of $\mathbb{R}^n$, appeared in the literature for the first time in 1931 in an article written by Orlicz [49]. However, after this one paper, the study of variable exponent Lebesgue spaces was then abandoned by Orlicz to concentrate on the theory of function spaces consisting of those measurable functions $u : \Omega \to \mathbb{R}$ with a modular of the form

$$\varrho(\lambda u) := \int_\Omega \varphi(\lambda |u(x)|)\, dx < \infty,$$

for some $\lambda > 0$ (φ has to satisfy certain conditions), which are now called Orlicz spaces. Abstracting certain central properties of ϱ, one is lead to a more general class of so-called modular function spaces which were first systematically studied by Nakano [47]. Following the work of Nakano, modular spaces were investigated by several people, most importantly by groups at Sapporo (Japan), Voronezh (USSR) and Leiden (Netherlands). From 1970s, these function spaces were extensively studied by Polish mathematicians, for instance Hudzik, Kamińska [31; 32] and notably Musielak in his monograph [46] with a comprehensive presentation of modular function spaces.

In the mid-80s, Zhikov [64; 65] started studying variational integrals with non-standard growth conditions including the variable exponent case $\varphi(x,t) = t^{p(x)}$ and the double phase case $\varphi(x,t) = t^p + a(x)t^q$. This developed into a new line of investigation, intimately related to the study of variable exponent spaces. While during 80s and 90s, Kováĉik and Rákosník [36] also found a few results of variable exponent Lebesgue and Sobolev spaces, concerning completeness, density, reflexivity and separability.

Further, in the calculus of variations, minimization problems of the following form have been considered,

$$\min_{u \in W^{1,1}(\Omega)} \int_\Omega F(x, \nabla u)\, dx.$$

If $F(x,t) \approx |t|^p$ (in some suitable sense), Marcellini [43] called this the *standard growth* case. While, for more general (p,q)-growth case, $t^p \lesssim F(x,t) \lesssim t^q + 1, q > p > 1$ for all $t \geqslant 0$, Marcellini [44] found that results from standard growth case generalize when the ratio $\frac{q}{p}$ is sufficiently close to 1 with different bounds for different properties. For instance, in [44], the minimizers in $W^{1,q}$ are Lipschitz if $\frac{q}{p} \leqslant \frac{n}{n-2}$, where $n > 2$ is the dimension. Regularity properties for such minimization problems have been widely studied and proved in the recent decades, as found in the works of Baroni, Colombo and Mingione [3; 13], especially for the cases

$$\min_u \int |\nabla u|^{p(x)} \log(e + |\nabla u|)\, dx \quad \text{and} \quad \min_u \int |\nabla u|^p + a(x)|\nabla u|^q\, dx,$$

respectively. They showed that the regularity of the minimizer depends on the choice of the exponents p, q and the weight a. Both of these are special cases of generalized Orlicz growth. Notably, several of those results on regularity of minimizers have been extended to the generalized Orlicz space, as studied by Hästö and Harjulehto in [24], and also in [60].

Růžička and his collaborators [55; 54] studied equations with non-standard growth in the modelling of so-called electrorheological fluids, as such materials with inhomogeneities requires that the exponent should be able to vary, in case of which classical Lebesgue and Sobolev spaces does not prove to be useful. This leads to the study of those materials in Lebesgue and Sobolev spaces with variable exponent. As another application, Chen, Levine and Rao [6] studied a model involving (p, q)-growth functionals for image restoration, in variable exponent Lebesgue and Sobolev spaces.

When we try to integrate both the two functional structures of variable exponent Lebesgue spaces and Orlicz spaces, we are led to the so-called generalized Orlicz spaces, as it naturally generalizes the Lebesgue spaces with constant exponent. In addition to being a natural generalization which covers results from both variable exponent and Orlicz spaces, the study of generalized Orlicz spaces are motivated by applications to image processing [25; 27], fluid dynamics [61], differential equations [8; 21; 26].

A different approach to differential equations is based on (nonlinear) potential theory. The foundation of nonlinear potential theory includes general notions of a Radon measure, a capacity and generalized functions. The sets of capacity zero are the exceptional sets for representatives of the function.

The basic properties of both Sobolev capacity and relative capacity in the generalized Orlicz setting, are presented here, as stated in [4]. It should be noted that T. Ohno and T. Shimomura [48] have also produced similar results on the (Sobolev) capacity of generalized Orlicz space, however, they consider the capacity in a metric measure space setting with Hajłasz gradients. These results therefore work in the Euclidean setting only when the maximal operator is bounded, since the Hajłasz gradient corresponds to $M(|\nabla f|)$.

The goal of this thesis is to study the generalized Orlicz and Sobolev-Orlicz spaces, in the context of variational problems and the associated PDEs. We show that, for a given minimization problem, one can find for a unique solution in the domain of Sobolev-Orlicz space, satisfying certain regularity conditions. This finds application in the field of image restoration, where we conduct a few numerical experiments.

1.2 Outline of the thesis

The structure of the thesis is organized as follows.

Chapter 2 consists of the preliminary definitions and concepts relevant to the work. We first introduce generalized Orlicz and Sobolev-Orlicz spaces, and also the notion of parabolic Hölder and Sobolev spaces involving time. Further, we review the necessary background material on functions belonging to these spaces, along with some standard regularity theory for elliptic and parabolic PDEs on Euclidean domain. The material discussed in this first part is mostly standard and can be found in the literature. This lays the foundation for study of variational and boundary value problems for image restoration model, in the domain of Sobolev-Orlicz space.

Chapter 3 consists of a few original results on capacities published in [4], as part of the doctoral research work. Further, extended results on quasicontinuous representative are newly added.

In chapter 4, we study the minimization problem under the setting of Sobolev-Orlicz space, and the associated boundary problems. We begin by introducing the variational problem in line with image restoration model, which involves double-phase functional. Next, the existence and uniqueness of the minimizer is established, using the direct method of the Calculus of Variations. Thereafter, we formulate the associated parabolic boundary value problems, of which the existence and regularity results are proved using approximation theory. In addition, the behaviour of the minimizing sequence is also studied. Finally, we have the numerical section that presents experimental results obtained using optimization algorithms along with images, to understand the working and efficiency of the image restoration model.

The final chapter 5, appendix, consists of supplementary results and arguments for the proofs presented in the preceding chapters.

2 Preliminaries: Function spaces and PDE

"The organic unity of mathematics is inherent in the nature of this science, for mathematics is the foundation of all exact knowledge of natural phenomena."

— David Hilbert (1862-1943)

This chapter of preliminaries is organized as follows. In section 2.1, there is a collection of notations used throughout this thesis, along with some standard formulae and results. The definitions of Φ-functions and some of its generalizations are stated here, along with the main assumptions on (generalized weak) Φ-functions required in the thesis. In section 2.2, the generalized Orlicz and the associated Sobolev Orlicz function spaces are defined, along with discussing a few of their properties. Finally, to study boundary value problems involving PDEs, a few function spaces involving time are briefly introduced.

2.1 Notation and definitions

We denote by $\mathbb{R}^n$ the n-dimensional real Euclidean space, and $n \in \mathbb{N}$ stands for the dimension of the space. By $B(x, r)$, we denote an open ball in $\mathbb{R}^n$ centred at $x \in \mathbb{R}^n$ with radius $r > 0$, and $|x| := \sqrt{\sum_{i=1}^n x_i^2}$ for $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$. Let $\Omega \subset \mathbb{R}^n$ be a nonempty open bounded set, which is mainly used throughout, unless otherwise stated. We denote by $\overline{\Omega}$ its closure and refer to $\partial\Omega := \overline{\Omega} \setminus \Omega$ as its boundary. Further, $|\Omega|$ stands for its (Lebesgue) measure and χ_Ω denotes its characteristic function. We denote by $L^0(\Omega)$ the space of all Lebesgue measurable functions on Ω, and by $L^1_{\text{loc}}(\Omega)$ the space of all locally integrable functions on Ω.

For constants, we use the letters $c, c_1, c_2, C, C_1, C_2, \ldots$, or other letters specifically mentioned to be constants. The symbol C without index stands for a generic constant which may vary between appearances. We use $f \approx g$ and $f \lesssim g$ if there exists constants $c_1, c_2 > 0$ such that $c_1 f \leqslant g \leqslant c_2 f$ and $f \leqslant c_2 g$, respectively. In a given interval I, a function $f(x)$ is said to be *increasing* if $s < t$ in I implies $f(s) \leqslant f(t)$ and *strictly increasing* if $s < t$ in I implies $f(s) < f(t)$, for all $s, t \in I$. The terms *decreasing* and *strictly decreasing* are defined analogously.

A function $u : (0, \infty) \to \mathbb{R}$ is *almost increasing* if there exists a constant $C \geqslant 1$ such that $u(s) \leqslant Cu(t)$ for all $0 < s < t$ (abbreviated C-almost increasing). *Almost* *decreasing* is defined analogously. Increasing and decreasing functions are special cases when $C = 1$. For $p, q > 0$, we say that u satisfies (*Inc*)$_p$ if $\frac{u(t)}{t^p}$ is increasing and (*aInc*)$_p$ if $\frac{u(t)}{t^p}$ is almost increasing; similarly, u satisfies (*Dec*)$_q$ if $\frac{u(t)}{t^q}$ is decreasing and (*aDec*)$_q$ if $\frac{u(t)}{t^q}$ is almost decreasing. While for $p > 1, q < \infty$, we drop the subscripts p, q from the notations above, that is, we say, u satisfies (*Inc*), (*aInc*) for some $p > 1$ or satisfies (*Dec*), (*aDec*) for some $q < \infty$, which will be mainly used in this thesis.

Let $k = 1, 2, \ldots$, and $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_n) \in \mathbb{N}_0^n$ be a multi-index, that is, a vector of n non-negative integers. We set the order of the multi-index as $|\alpha| := \alpha_1 + \alpha_2 + \ldots + \alpha_n$. We call $\partial^\alpha := \frac{\partial^{\alpha_1}}{\partial x_1^{\alpha_1}} \frac{\partial^{\alpha_2}}{\partial x_2^{\alpha_2}} \cdots \frac{\partial^{\alpha_n}}{\partial x_n^{\alpha_n}}$ the partial derivative of order α. For sufficiently smooth functions, all partial derivatives commute. Provided that the derivatives exist, the *gradient* of a function $u : \Omega \to \mathbb{R}$ is the vector $\nabla u(x) := \big(\frac{\partial u}{\partial x_1}(x), \ldots, \frac{\partial u}{\partial x_n}(x)\big)$, for $x \in \Omega$. Let $F : \Omega \to \mathbb{R}^n$ be a vector field, defined as $F := (F_1, F_2, \ldots, F_n)$. The *divergence* of the vector field F, denoted as $\operatorname{div} F$, is the function $\nabla \cdot F(x) := \sum_{i=1}^{n} \frac{\partial F_i}{\partial x_i}(x)$, for $x \in \Omega$. The *Laplacian* operator is denoted by $\Delta := \nabla \cdot \nabla$.

The space of uniformly continuous functions on $\overline{\Omega}$, denoted by, $C(\overline{\Omega})$, is equipped with the supremum norm $\|u\|_\infty := \sup_{x \in \overline{\Omega}} |u(x)|$. While, $C^k(\overline{\Omega}), k \in \mathbb{N}$, denote the space of functions u, such that the partial derivatives $\frac{\partial^\alpha u}{\partial x^\alpha} \in C(\overline{\Omega})$ for all $|\alpha| \leqslant k$. This is a Banach space, equipped with the norm,

$$\|u\|_{C^k(\overline{\Omega})} := \sup_{|\alpha| \leqslant k, x \in \overline{\Omega}} \left| \frac{\partial^\alpha u}{\partial x^\alpha} \right|. \tag{2.1.1}$$

Further, we denote $C^\infty(\overline{\Omega})$ as the space of functions with continuous partial derivatives of all orders in $\overline{\Omega}$. While, $C_0^\infty(\Omega)$ denotes the set of $C^\infty(\Omega)$ functions with compact support in Ω.

We also recall the definition of weak derivative. Let $u \in L^1_{\text{loc}}(\Omega)$, then a function $w \in L^1_{\text{loc}}(\Omega)$ is called the α^{th}-*weak derivative* of u, if

$$\int_\Omega wh \, dx = (-1)^{|\alpha|} \int_\Omega u \, \partial^\alpha h \, dx, \quad \forall \, h \in C_0^\infty(\Omega).$$

For $|\alpha| = 1$ in the equation above, the function w is called the *weak gradient* of u, that is,

$$\int_\Omega wh \, dx = -\int_\Omega u \, \nabla h \, dx, \quad \forall \, h \in C_0^\infty(\Omega).$$

Further, $w \in L^1_{\text{loc}}(\Omega)$ is called the *weak divergence* of vector function $u \in L^2(\Omega; \mathbb{R}^n)$ if

$$\int_\Omega wh \, dx = -\int_\Omega u \cdot \nabla h \, dx, \quad \forall \, h \in C_0^\infty(\Omega)$$

Here, if each component u_i of u is weakly differentiable with weak derivatives $\frac{\partial u_i}{\partial x_i} \in L^1_{\text{loc}}(\Omega)$, then $w = \operatorname{div} u = \nabla \cdot u = \sum_{i=1}^{n} \frac{\partial u_i}{\partial x_i}(x)$.

Next we have that, $W^{1,p}(\Omega), p \geqslant 1$, denotes the Sobolev space, consisting of functions $u \in L^p(\Omega)$ for which the partial derivatives of u of order $\alpha \in \mathbb{N}_0^n$ where $|\alpha| = 1$ exists in the weak sense and belong to the Lebesgue space $L^p(\Omega)$. Its norm is defined as,

$$\|u\|_{W^{1,p}(\Omega)} := \begin{cases} \left(\int_\Omega |u|^p \, dx + \int_\Omega |\nabla u|^p \, dx \right)^{\frac{1}{p}} & (1 \leqslant p < \infty) \\ \operatorname*{ess\,sup}_{\Omega} \left(|u| + |\nabla u| \right) & (p = \infty). \end{cases}$$

The Sobolev space $W_0^{1,p}(\Omega)$ is the completion of $C_0^\infty(\Omega)$ in the norm of $W^{1,p}(\Omega)$.

For $f \in L^1_{\text{loc}}(\Omega)$, we define

$$Mf(x) := \sup_{x \in B} \frac{1}{|B|} \int_{B \cap \Omega} |f(y)| \, dy \tag{2.1.2}$$

and call M the *maximal operator*, where the supremum is taken over all open balls B containing x.

We also define *convolution* as follows,

$$f * g(x) := \int_\Omega f(y) \, g(x - y) \, dy, \quad x \in \Omega$$

where $f \in L^1_{\text{loc}}(\Omega)$ and $g \in L^1_{\text{loc}}(\mathbb{R}^n)$.

Let $N(x) := (N_1(x), \ldots, N_n(x))$ denote the outward pointing unit normal vector defined almost everywhere for $x \in \partial\Omega$ and dS denotes the surface measure on the C^1 boundary $\partial\Omega$. Let $u \in C^1(\overline{\Omega})$ and $\varphi \in C^2(\overline{\Omega})$. The *Green's first identity* reads as

$$\int_\Omega (\nabla u \cdot \nabla \varphi + u \Delta \varphi) \, dx = \int_{\partial\Omega} u \frac{\partial \varphi}{\partial N} \, dS \tag{2.1.3}$$

where $\frac{\partial \varphi}{\partial N} = \nabla \varphi \cdot N$ is the outward normal derivative of φ.

Next, we have the integration by parts formula stated as follows.

***Lemma* 2.1.1. (Integration by parts formula)** [15, Theorem 2, C.2., pp. 628] *Let $\Omega \subset \mathbb{R}^n$ be a bounded open set with C^1 boundary. For every $u, v \in C^1(\overline{\Omega})$, the following formula holds,*

$$\int_\Omega \frac{\partial u}{\partial x_i} v \, dx = -\int_\Omega \frac{\partial v}{\partial x_i} u \, dx + \int_{\partial\Omega} uv \, N_i \, dS, \quad i = 1, \ldots, n. \tag{2.1.4}$$

Moreover, for $v \in C^2(\overline{\Omega})$, replacing v in the above equality (2.1.4) with $\frac{\partial v}{\partial x_i}$ and summing over $i = 1, \ldots, n$ to obtain

$$\int_\Omega \nabla u \cdot \nabla v \, dx = -\int_\Omega u \Delta v \, dx + \int_{\partial\Omega} \frac{\partial v}{\partial N} u \, dS, \tag{2.1.5}$$

where $\frac{\partial}{\partial N} := \sum_{i=1}^n N_i \frac{\partial}{\partial x_i}$ is the operator of differentiation in the direction of the vector N. Note that, the above formula (2.1.5) is in line with the Green's first identity in (2.1.3). While using Dirichlet's boundary, $u = 0$ on $\partial\Omega$, the integration by parts formula (2.1.4) can be equivalently written as,

$$\int_\Omega \frac{\partial u}{\partial x_i} v \, dx = -\int_\Omega u \frac{\partial v}{\partial x_i} \, dx, \quad i = 1, 2, \ldots, n.$$

The integration by parts formula is satisfied for Sobolev functions $u \in W_0^{1,p}(\Omega), p \geqslant 1$, in bounded domain $\Omega \subset \mathbb{R}^n$ [38, Chapter 2, section 2], that is,

$$\int_\Omega \frac{\partial u}{\partial x_i} v \, dx = -\int_\Omega \frac{\partial v}{\partial x_i} u \, dx, \tag{2.1.6}$$

where $v \in W^{1,q}(\Omega), q \geqslant 1$, such that $\frac{1}{p} + \frac{1}{q} = 1$.

***Lemma* 2.1.2. (Fundamental Lemma of Calculus of Variations)** [18, Chapter 1, Lemma 1] *If $f \in L^1_{\text{loc}}(\Omega)$ satisfies,*

$$\int_\Omega f \varphi \, dx = 0,$$

for every $\varphi \in C_0^\infty(\Omega)$, then $f \equiv 0$ almost everywhere in Ω.

For use during calculations, we also state the *Young's inequality* for products [15, B.2 (c), pp. 622],

$$ab \leqslant \frac{a^p}{p} + \frac{b^q}{q}, \quad \forall \, a, b \geqslant 0, \tag{2.1.7}$$

where $p, q \in (1, \infty)$ satisfy $\frac{1}{p} + \frac{1}{q} = 1$.

In the case of convolution functions, for $f \in L^p(\mathbb{R}^n)$, $1 \leqslant p \leqslant \infty$ and $\zeta_\delta \in L^1(\mathbb{R}^n)$, the *Young's convolution inequality* result implies that $(f * \zeta_\delta)(x)$ exists for almost all $x \in \mathbb{R}^n$ and we have

$$\|f * \zeta_\delta\|_p \leq \|f\|_p \|\zeta_\delta\|_1. \tag{2.1.8}$$

2.1.1 Φ-functions Let $\varphi : [0, \infty) \to [0, \infty]$ be increasing with $\varphi(0) = 0$, $\lim\limits_{t \to 0^+} \varphi(t) = 0$ and $\lim\limits_{t \to \infty} \varphi(t) = \infty$. Such φ is called a Φ-*prefunction*. While by $\varphi^{-1} : [0, \infty] \to [0, \infty]$ we denote the *left-continuous inverse* of $\varphi : [0, \infty] \to [0, \infty]$, $\varphi^{-1}(\tau) := \inf\{t \geqslant 0 : \varphi(t) \geqslant \tau\}$, for all $\tau \in [0, \infty]$.

We say that a Φ-prefunction φ is a (*weak*) Φ*-function* if it satisfies $(\mathrm{aInc})_1$ on $(0,\infty)$; a *convex* Φ*-function* if it is left-continuous in the topology of $[0,\infty]$ and convex; a *strong* Φ*-function* if it is continuous in the topology of $[0,\infty]$ and convex. The sets of weak, convex and strong Φ-functions are denoted by Φ_w, Φ_c and Φ_s, respectively.

Remark 2.1.9. If φ is convex and $\varphi(0)=0$, then for $0<s<t$, we obtain $\varphi(s) = \varphi\left(\frac{s}{t}t+0\right) \leqslant \frac{s}{t}\varphi(t) + \left(1-\frac{s}{t}\right)\varphi(0) = \frac{s}{t}\varphi(t)$, which implies φ satisfies $(\mathrm{aInc})_1$. Hence, it follows from the above definition that $\Phi_s \subset \Phi_c \subset \Phi_w$.

In order that Φ-functions depend on the spatial variable in set $\Omega \subset \mathbb{R}^n$, we need some generalization of Φ-functions, as introduced in the following.

Let $u : \Omega \times [0,\infty) \to \mathbb{R}$ and $p, q > 0$. We say that u satisfies $(\mathrm{aInc})_p$ or $(\mathrm{aDec})_q$, if there exists a constant $C \geqslant 1$ such that the function $t \mapsto u(x,t)$ satisfies $(\mathrm{aInc})_p$ or $(\mathrm{aDec})_q$ with the same constant C for almost every $x \in \Omega$. When $C=1$, we use the notations $(\mathrm{Inc})_p$ and $(\mathrm{Dec})_q$, respectively.

A function $\varphi : \Omega \times [0,\infty) \to [0,\infty]$ is said to be a (*generalized*) Φ*-prefunction* on Ω if $x \mapsto \varphi(x,|u(x)|)$ is measurable for every $u \in L^0(\Omega)$ and $\varphi(x,\cdot)$ is a Φ-prefunction for almost every $x \in \Omega$. We say that the Φ-prefunction φ is a (*generalized*) *weak* Φ*-function* if $\varphi(x,\cdot)$ satisfies $(\mathrm{aInc})_1$ with the same constant for almost every $x \in \Omega$ with the same constant, denoted as $\varphi \in \Phi_w(\Omega)$. In a similar way, φ is called a (*generalized*) *convex* Φ*-function* if $\varphi(x,\cdot) \in \Phi_c$ for almost all $x \in \Omega$, denoted as $\varphi \in \Phi_c(\Omega)$; and a (*generalized*) *strong* Φ*-function* if $\varphi(x,\cdot) \in \Phi_s$ for almost all $x \in \Omega$, denoted as $\varphi \in \Phi_s(\Omega)$. And, by φ^{-1} we denote the inverse of $\varphi \in \Phi_w(\Omega)$ with respect to the second variable, that is, for $\tau \geqslant 0$, $\varphi^{-1}(x,\tau) := \inf\{t \geqslant 0 : \varphi(x,t) \geqslant \tau\}$.

By the above definition, it is clear that the properties of (weak) Φ-functions carry over to generalized (weak) Φ-functions point-wise uniformly. Also, $\Phi_s(\Omega) \subset \Phi_c(\Omega) \subset \Phi_w(\Omega)$.

We say that $\varphi \in \Phi_c(\Omega)$ is *uniformly convex*, by [24, Definition 3.6.1], if for every $\varepsilon > 0$ there exists $\delta \in (0,1)$ such that

$$\varphi\left(x, \frac{s+t}{2}\right) \leqslant (1-\delta)\frac{\varphi(x,s)+\varphi(x,t)}{2},$$

for almost every $x \in \Omega$ whenever $s,t \geqslant 0$, and $|s-t| \geqslant \varepsilon \max\{s,t\}$. By [24, Proposition 3.6.2], the function $\varphi \in \Phi_w(\Omega)$ is equivalent to a uniformly convex Φ-function if and only if it satisfies (aInc). On the other hand, a vector space X is *uniformly convex* if it has a norm $\|\cdot\|$ such that for every $\varepsilon > 0$ there exists $\delta > 0$ with $\|x-y\| \geqslant \varepsilon$ or $\|x+y\| \leqslant 2(1-\delta)$ for all $x, y \in X$ with $\|x\| = \|y\| = 1$.

Next, we have the following assumptions for Φ-function in set $\Omega \subset \mathbb{R}^n$, which enables the use of certain useful properties such as density of smooth functions, as discussed in [24].

Definition 2.1.10. (A0) condition [24, Definition 3.7.1]: We say that $\varphi \in \Phi_w(\Omega)$ satisfies (A0), if there exists a constant $\beta \in (0,1]$ such that $\beta \leqslant \varphi^{-1}(x,1) \leqslant \frac{1}{\beta}$ for almost every $x \in \Omega$.

Equivalently, this means that there exists $\beta \in (0,1]$ such that $\varphi(x,\beta) \leqslant 1 \leqslant \varphi(x,\frac{1}{\beta})$ for almost every $x \in \Omega$ (cf. [24, Corollary 3.7.4]).

Definition 2.1.11. (A1) condition [24, Definition 4.1.1]: We say that $\varphi \in \Phi_w(\Omega)$ satisfies (A1) if there exists $\beta \in (0,1)$ such that $\beta\varphi^{-1}(x,t) \leqslant \varphi^{-1}(y,t)$ for every $t \in [1, \frac{1}{|B|}]$, almost every $x, y \in B \cap \Omega$ and every ball B with $|B| \leqslant 1$.

Definition 2.1.12. (A2) condition [24, Definition 4.2.1]: We say that $\varphi \in \Phi_w(\Omega)$ satisfies (A2) if for every $s > 0$ there exist $\beta \in (0,1]$ and $h \in L^1(\Omega) \cap L^\infty(\Omega)$ such that $\beta\varphi^{-1}(x,t) \leqslant \varphi^{-1}(y,t)$ for almost every $x, y \in \Omega$ and every $t \in [h(x)+h(y), s]$.

***Lemma* 2.1.3.** [24, Lemma 4.2.3] *If $\Omega \subset \mathbb{R}^n$ is bounded, then every $\varphi \in \Phi_w(\Omega)$ satisfies* (A2).

2.2 Generalized Orlicz and Sobolev Orlicz spaces

In an open bounded set $\Omega \subset \mathbb{R}^n$, consider $\varphi \in \Phi_w(\Omega)$. We then define the modular ϱ_φ for $u \in L^0(\Omega)$ as

$$\varrho_\varphi(u) := \int_\Omega \varphi(x, |u(x)|)\, dx.$$

The set

$$L^\varphi(\Omega) := \{u \in L^0(\Omega) : \varrho_\varphi(\lambda u) < \infty \text{ for some } \lambda > 0\}$$

is called *generalized Orlicz space*, also known as *Musielak-Orlicz space*, and it is equipped with the (quasi) norm

$$\|u\|_\varphi := \inf_{\lambda > 0} \left\{ \lambda : \varrho_\varphi\left(\frac{u}{\lambda}\right) \leqslant 1 \right\},$$

while this yields a norm if φ is convex [24, Lemma 3.2.2(b)].

We say that $\varphi \in \Phi_w(\Omega)$ satisfies the Δ_2-*condition* if there exists a constant $L \geqslant 2$ such that $\varphi(x,2t) \leqslant L\,\varphi(x,t)$ for all $x \in \Omega$ and all $t \geqslant 0$. In such case, the semimodular ϱ_φ also satisfies the Δ_2-condition with the same constant. Here, Δ_2 is equivalent to (aDec) (cf. [24, Lemma 2.2.6(a)]).

***Lemma* 2.2.1.** [24, Theorem 3.3.7, Corollary 3.6.7] *Let* $\varphi \in \Phi_w(\Omega)$.

(i) Then $L^\varphi(\Omega)$ is a quasi-Banach space.
(ii) If φ is convex, then $L^\varphi(\Omega)$ is a Banach space.
(iii) If φ satisfies (aDec)*, then $L^\varphi(\Omega)$ is separable.*
(iv) If φ satisfies (aInc) *and* (aDec)*, then $L^\varphi(\Omega)$ is uniformly convex and reflexive.*

The following results examine the relationship between norm and modular, to be used later.

Lemma 2.2.2. [24, Lemma 3.1.3, 3.2.3] *Let $\varphi \in \Phi_w(\Omega)$, then the following properties hold.*

(i) If φ satisfies the Δ_2-condition, then $L^\varphi(\Omega) = \{u \in L^0(\Omega) \mid \varrho_\varphi(u) < \infty\}$.

(ii) $\|u\|_\varphi < 1 \Rightarrow \varrho_\varphi(u) \leqslant 1 \Rightarrow \|u\|_\varphi \leqslant 1$.
If φ is left-continuous, then, $\varrho_\varphi(u) \leqslant 1 \iff \|u\|_\varphi \leqslant 1$.

Lemma 2.2.3. [24, Lemma 3.1.6] *Let $\varphi \in \Phi_w(\Omega)$ satisfy (Dec). Let $f_j, g_j \in L^\varphi(\mathbb{R}^n)$ for $j = 1, 2, \ldots$ with $(\varrho_\varphi(f_j))_{j=1}^\infty$ bounded. If $\varrho_\varphi(f_j - g_j) \to 0$ as $j \to \infty$, then*

$$|\varrho_\varphi(f_j) - \varrho_\varphi(g_j)| \to 0 \quad \text{as } j \to \infty.$$

Lemma 2.2.4. [24, Lemma 3.2.9] *Let $\varphi \in \Phi_w(\Omega)$ satisfy (aInc)$_p$ and (aDec)$_q$, $1 \leqslant p \leqslant q < \infty$. Then*

$$\min\{\big(\tfrac{1}{c}\varrho_\varphi(u)\big)^{\frac{1}{p}}, \big(\tfrac{1}{c}\varrho_\varphi(u)\big)^{\frac{1}{q}}\} \leqslant \|u\|_\varphi \leqslant \max\{\big(c\varrho_\varphi(u)\big)^{\frac{1}{p}}, \big(c\varrho_\varphi(u)\big)^{\frac{1}{q}}\},$$

for $u \in L^0(\Omega)$, where c is the maximum of the constants from (aInc)$_p$ and (aDec)$_q$.

For the next result, we first note that, for two normed spaces X and Y, the intersection $X \cap Y$ and the sum $X + Y := \{g + h : g \in X, h \in Y\}$ are equipped with the norms

$$\|f\|_{X\cap Y} := \max\{\|f\|_X, \|f\|_Y\} \quad \text{and} \quad \|f\|_{X+Y} := \inf_{f=g+h, g\in X, h\in Y}(\|g\|_X + \|h\|_Y).$$

Lemma 2.2.5. [24, Lemma 3.7.7] *Let $\varphi \in \Phi_w(\Omega)$ satisfy* (A0), (aInc)$_p$ *and* (aDec)$_q$, *with $p \in [1, \infty)$ and $q \in [1, \infty]$. Then*

$$L^p(\Omega) \cap L^q(\Omega) \hookrightarrow L^\varphi(\Omega) \hookrightarrow L^p(\Omega) + L^q(\Omega)$$

and the embedding constants depend only on (A0), (aInc)$_p$ *and* (aDec)$_q$.

For $\varphi \in \Phi_w(\Omega)$, we now define the related *generalized Sobolev-Orlicz space* $W^{1,\varphi}(\Omega)$ as the set of functions $u \in L^\varphi(\Omega)$ with weak partial derivatives $\frac{\partial u}{\partial x_i}, i = 1, 2, \ldots,$ belonging to the same $L^\varphi(\Omega)$, that is,

$$W^{1,\varphi}(\Omega) := \left\{u \in W^{1,1}_{\mathrm{loc}}(\Omega) : u, |\nabla u| \in L^\varphi(\Omega)\right\}.$$

The semimodular on $W^{1,\varphi}(\Omega)$ is defined by

$$\varrho_{1,\varphi}(u) := \varrho_\varphi(u) + \varrho_\varphi(\nabla u) = \int_\Omega \varphi(x, |u|)\, dx + \int_\Omega \varphi(x, |\nabla u|)\, dx$$

which induces a (quasi)norm on $W^{1,\varphi}(\Omega)$ given by

$$\|u\|_{1,\varphi} := \inf_{\lambda>0} \left\{\lambda : \varrho_{1,\varphi}\left(\frac{u}{\lambda}\right) \leqslant 1\right\}.$$

By [24, Lemma 6.1.5], the above norm is estimated as,

$$\|u\|_{1,\varphi} \approx \|u\|_{\varphi} + \||\nabla u|\|_{\varphi}.$$

For later use, the norm of the gradient is abbreviated as $\|\nabla u\|_{\varphi}$.

Here, we also define the zero boundary valued Sobolev-Orlicz space $W_0^{1,\varphi}(\Omega)$ as the closure of $C_0^{\infty}(\Omega) \cap W^{1,\varphi}(\Omega)$ in $W^{1,\varphi}(\Omega)$. This is often required for studying boundary value problems in situations requiring density of smooth functions in Sobolev-Orlicz space.

Lemma 2.2.6. [24, Theorem 6.1.4, Theorem 6.1.9] *Let $\Omega \subset \mathbb{R}^n$ and $\varphi \in \Phi_w(\Omega)$.*

(i) Then $W^{1,\varphi}(\Omega)$ is a quasi-Banach space.
(ii) If φ is convex, then $W^{1,\varphi}(\Omega)$ is a Banach space.
(iii) If φ satisfies (aDec)*, then $W^{1,\varphi}(\Omega)$ is separable.*
(iv) If φ satisfies (aInc) *and* (aDec)*, then $W^{1,\varphi}(\Omega)$ is uniformly convex and reflexive.*

The above conditions hold true also in the case of $W_0^{1,\varphi}(\Omega)$.

The following lemma implies that $W^{1,\varphi}(\Omega)$ is a lattice, that is, the pointwise minimum and maximum of its elements belong to $W^{1,\varphi}(\Omega)$, provided $L^{\varphi} \subset L^1_{\mathrm{loc}}(\Omega)$.

Lemma 2.2.7. [29, Theorem 1.20] *If $u, v \in W^{1,1}_{\mathrm{loc}}(\Omega)$, then $\max\{u,v\}$ and $\min\{u,v\}$ are in $W^{1,1}_{\mathrm{loc}}(\Omega)$ with*

$$\nabla \max\{u,v\}(x) = \begin{cases} \nabla u(x), & \text{for almost every } x \in \{u \geqslant v\}; \\ \nabla v(x), & \text{for almost every } x \in \{v \geqslant u\}; \end{cases}$$

and

$$\nabla \min\{u,v\}(x) = \begin{cases} \nabla u(x), & \text{for almost every } x \in \{u \leqslant v\}; \\ \nabla v(x), & \text{for almost every } x \in \{v \leqslant u\}; \end{cases}$$

In particular, $|u|$ belongs to $W^{1,1}_{\mathrm{loc}}(\Omega)$ and $|\nabla|u|| = |\nabla u|$ almost everywhere in Ω.

Lemma 2.2.8. [24, Lemma 6.1.6] *For bounded set $\Omega \subset \mathbb{R}^n$, assume that $\varphi \in \Phi_w(\Omega)$ satisfy* (A0) *and* $(\text{aInc})_p$*, $p \in [1,\infty)$. Then $W^{1,\varphi}(\Omega) \hookrightarrow W^{1,p}(\Omega)$.*

The following results provide density properties of Sobolev-Orlicz spaces, to be used later.

Lemma 2.2.9. [24, Theorem 6.4.2] *Let $\Omega \subset \mathbb{R}^n$ be a bounded domain and let $\varphi \in \Phi_w(\Omega)$ satisfy* (aDec) *and $L^{\varphi} \subset L^1_{\mathrm{loc}}$. Then bounded Sobolev functions with compact support in $\mathbb{R}^n$ are dense in $W^{1,\varphi}(\Omega)$.*

***Lemma* 2.2.10.** [24, Theorem 6.4.7] *Let $\Omega \subset \mathbb{R}^n$ be a bounded domain and let $\varphi \in \Phi_w(\Omega)$ satisfy* (A0), (A1), (A2) *and* (aDec). *Then $C^\infty(\Omega) \cap W^{1,\varphi}(\Omega)$ is dense in $W^{1,\varphi}(\Omega)$.*

A few convolution properties associated with generalized Orlicz functions from [24] are useful in this thesis. To state so, we use the notion of *bell shaped function*, defined as a non-negative function $\sigma \in L^1(\mathbb{R}^n)$ if it is radially decreasing and radially symmetric. The function

$$\Sigma(x) := \sup_{y \notin B(0,|x|)} |\sigma(y)|$$

is called the *least bell shaped majorant of* σ. Also consider L^1-scaling $\sigma_\varepsilon(x) := \frac{1}{\varepsilon^n}\sigma\left(\frac{x}{\varepsilon}\right)$, for $\varepsilon > 0$. By change of variables, we have $\|\sigma_\varepsilon\|_1 = \|\sigma\|_1$.

***Lemma* 2.2.11.** [24, Lemma 4.4.6, Theorem 6.4.5] *Let $\Omega \subset \mathbb{R}^n$ be a bounded domain and let $\varphi \in \Phi_w(\Omega)$ satisfy* (A0), (A1), (A2) *and* (aDec). *Let $\sigma \in L^1(\mathbb{R}^n), \sigma \geqslant 0$, have integrable least bell shaped majorant $\sum$. Then*

$$\|f * \sigma_\varepsilon\|_{1,\varphi} \lesssim \|\Sigma\|_1 \|f\|_{1,\varphi}$$

for all $f \in W^{1,\varphi}(\Omega)$. Further, for $D \subset\subset \Omega$, we have $\sigma_\epsilon * f \to f$ in $W^{1,\varphi}(D)$ as $\epsilon \to 0^+$.

2.3 Function spaces involving time

In this thesis, we study time-dependent partial differential equations of second order, as well. In order to find solutions to parabolic problems, we need to define special class of functions depending on the spatial variable x and the time variable t. In the time variable t, these functions have values in a Banach space of functions in x, namely in Sobolev spaces for the applications we have in view.

We only consider here spaces of vector-valued functions defined on a bounded interval I of $\mathbb{R}$ of the form $I = (0,T), 0 < T < \infty$, the details of which can be found in [15; 62]. In order to define the function spaces involving time, we first recall the definitions of simple functions and strongly measurable functions.

Suppose that X is a real Banach space with norm $\|\cdot\|$. A function $f : [0,T] \to X$ is called a *simple function* if it is of the form

$$f(t) := \sum_{i=1}^{n} \chi_{E_i}(t)\, u_i, \quad 0 \leqslant t \leqslant T,$$

where every E_i is a Lebesgue measurable subset of $[0,T]$ and $u_i \in X$ for all $i = 1, \ldots, n$. Further, a function $f : [0,T] \to X$ is called *strongly measurable* if there exists a sequence of simple functions $f_i : [0,T] \to X$, $i \in \mathbb{N}$, such that $f_i(t) \to f(t)$ in X, for almost every $0 \leqslant t \leqslant T$.

Recall that, we will be using $\nabla := (\frac{\partial}{\partial x_1} \cdots \frac{\partial}{\partial x_n})$ for weak gradient with respect to the space variables, and $\frac{\partial}{\partial t}$ for weak derivatives with respect to the time variable t. We will also sometimes use the notation $u_t := \frac{\partial u}{\partial t}$.

Let $\Omega \subset \mathbb{R}^n$ be a bounded, connected open set. We denote $\Omega_T := \Omega \times (0,T)$ and $\overline{\Omega}_T := \overline{\Omega} \times [0,T]$.

The space $C(\overline{\Omega}_T)$ consists of all continuous functions on Ω_T, which can be extended continuously on $\overline{\Omega}_T$. Higher order spaces, $C^k(\overline{\Omega}_T)$, $k \in \mathbb{N}$, are defined similarly, as follows,

$$C^k(\overline{\Omega}_T) := \{u \in C(\overline{\Omega}_T) : \frac{\partial^{|\alpha|} u}{\partial x^\alpha} \in C(\overline{\Omega}_T) \text{ for all } |\alpha| \leqslant k\}$$

We define $C^{2,1}(\overline{\Omega}_T)$ to be the class of functions $u : \overline{\Omega} \times [0,T] \to \mathbb{R}$ with continuous spatial partial derivatives up to order 2, and which are once continuously differentiable in time, that is,

$$C^{2,1}(\overline{\Omega}_T) = \{u \in C(\overline{\Omega}_T) : u_t \in C(\overline{\Omega}_T) \text{ and } \frac{\partial^{|\alpha|} u}{\partial x^\alpha} \in C(\overline{\Omega}_T) \text{ for all } |\alpha| \leqslant 2\}.$$

The space $C_0^\infty(\Omega_T)$ is defined as the set of all maps $u : \Omega_T \to \mathbb{R}^n$ such that u and its derivatives of all orders are continuous in Ω_T with compact support in Ω_T.

We now introduce the L^p-spaces for vector-valued functions.

If X is a Banach space, then $L^p(0,T;X)$ is defined as the space consisting of all strongly measurable functions $u : (0,T) \to X$ for which $t \mapsto \|u(t)\|_X \in L^p(0,T)$, endowed with the finite norm

$$\|u\|_{L^p(0,T;X)} := \begin{cases} \left(\int_0^T \|u(t)\|_X^p \, dt\right)^{\frac{1}{p}}, & (1 \leqslant p < \infty) \\ \operatorname*{ess\,sup}_{0<t<T} \|u(t)\|_X, & (p = \infty). \end{cases}$$

For $1 \leqslant p \leqslant \infty$, $L^p_{\text{loc}}(0,T;X)$ denotes the space of strongly measurable functions $u : [0,T] \to X$ so that $u \in L^p(a,b;X)$, for all $[a,b] \subset (0,T), a, b \in \mathbb{R}^+$.

The space $L^p(0,T;X)$ has properties similar to the L^p-spaces for real-valued functions, as follows.

***Proposition* 2.3.1.** *[63, Proposition 23.2, Proposition 23.7] Let X and Y be Banach spaces. Assume that $0 < T < \infty$. Then:*

(i) $L^p(0,T;X)$ *is a Banach space, for* $1 \leqslant p < \infty$.

(ii) $L^p(0,T;X)$ *is separable and reflexive in the case where X is separable and reflexive, for* $1 < p < \infty$.

(iii) $L^p(0,T;X)$ *is uniformly convex in the case where X is uniformly convex and* $1 < p < \infty$.

(iv) *If the embedding $X \subset Y$ is continuous, then the embedding*

$$L^r(0,T;X) \subset L^q(0,T;Y), \quad 1 \leqslant q \leqslant r \leqslant \infty$$

is also continuous.

Another example of $L^p(0,T;X)$ Banach space is $L^2(0,T;W^{1,2}(\Omega))$, that is used in this thesis. For bounded set $\Omega \subset \mathbb{R}^n$, the function space $L^2(0,T;W^{1,2}(\Omega))$, thus, consists of strongly measurable functions $u : (0,T) \to W^{1,2}(\Omega)$ such that $t \mapsto \|u(t)\|_{W^{1,2}(\Omega)} \in L^2(0,T)$, for which the norm,

$$\|u\|_{L^2(0,T;W^{1,2}(\Omega))} = \left(\int_0^T \int_\Omega (|u(t)|^2 + |\nabla u(t)|^2)\,dx\,dt\right)^{\frac{1}{2}}$$

is finite. The space $L^2(0,T;W_0^{1,2}(\Omega))$, equipped with zero boundary value, is defined analogously.

Suppose X and Y are two Banach spaces such that $X \subset Y$ with continuous embedding, then by [63, Proposition 23.2 (h)], $L^p(0,T;X) \subset L^p(0,T;Y)$ is also a continuous embedding. However, this property does not extend to compact embeddings.

For any Lebesgue or Sobolev space X, an element v of $L^p(0,T;X)$ can be regarded as a function of the ordered pair $(x,t) \in \Omega \times (0,T)$ and identified with the functions $v_1(x,t)$ defined by

$$v_1(x,t) = v(t)(x), \quad \text{for a.e. } (x,t) \in \Omega \times (0,T). \tag{2.3.1}$$

This point of view is especially well-adapted to evolution problems, where the space and time variables play a different role and the involved functions have different properties with respect to x and t. This property will be systematically used throughout this thesis.

In particular, for $p = 2, X = L^2(\Omega)$ and v given by (2.3.1), we have that the space $L^2(0,T;L^2(\Omega))$ can be identified with the space $L^2(\Omega \times (0,T))$, also denoted as $L^2(\Omega_T)$, via the map

$$v \in L^2(0,T;L^2(\Omega)) \mapsto v_1 \in L^2(\Omega_T),$$

with v_1 given by (2.3.1), which is an isometry.

To have a weaker notion of the differentiability of vector-valued function, we use the notion of a distributional or weak derivative, which is a natural generalization of the definition for real-valued functions. A function $u \in L^1_{\text{loc}}(0,T;X)$ is weakly differentiable with weak derivative $u_t \in L^1_{\text{loc}}(0,T;X)$ if

$$\int_0^T uh'(t)\,dt = -\int_0^T u_t h(t)\,dt, \quad \text{for every } h \in C_0^\infty(0,T;\mathbb{R}).$$

Thus, $u_t(x,t)$ is the partial derivative of $u(x,t)$ with respect to time t that exists in a weak sense, which will be referred to as the time derivative of u throughout the thesis.

2.3.1 Hölder spaces. In this subsection, we introduce Hölder spaces on Euclidean domain as well as parabolic Hölder spaces, which are comprehensively found in [37; 38; 62]. The Hölder spaces that we mainly deal with here are $C^{2+\gamma,\gamma}(\Omega)$ and $C^{2+\gamma,1+\frac{\gamma}{2}}(\Omega_T)$, where $\gamma \in (0,1)$.

Let $\Omega_T := \Omega \times (0,T)$ where $\Omega \subset \mathbb{R}^n$ is an open bounded set and $T > 0$. First, we define Hölder continuous function as follows: a function $u : \Omega \to \mathbb{R}$ is *uniformly Hölder continuous* with exponent $\gamma \in (0,1)$ in Ω, if the quantity (semi-norm)

$$[u]_{\gamma;\Omega} := \sup_{x,y\in\Omega} \frac{|u(x)-u(y)|}{|x-y|^\gamma}, x \neq y \tag{2.3.2}$$

is finite, and the space of such functions u is denoted by $C^\gamma(\Omega)$ for which the norm

$$\|u\|_{C^\gamma(\Omega)} := \sup_\Omega |u| + [u]_{\gamma;\Omega}$$

is finite. If the right-hand side of (2.3.2) is finite for $\gamma = 1$, the function u is called *Lipschitz continuous* in Ω. While, u is said to be *locally uniformly Hölder continuous* with exponent γ in Ω if the quantity $[u]_{\gamma;\Omega'}$ is finite for every $\Omega' \Subset \Omega$.

For $0 < \gamma < 1$ and $k = 0,1,2,\ldots$, the *Hölder space* $C^{k+\gamma}(\Omega)$ is the space of all functions $u \in C^k(\Omega)$ for which the norm

$$\|u\|_{C^{k+\gamma}(\Omega)} := \|u\|_{C^k(\Omega)} + \max_{|\alpha|=k} \left[\frac{\partial^{|\alpha|}u}{\partial x^\alpha}\right]_{\gamma;\Omega} \tag{2.3.3}$$

is finite. The function space $C^{k+\gamma}(\Omega)$ is a Banach space [15, Theorem 1, Section 5.1]. Note that in [15, Section 5.1], the notation $C^{k+\gamma}(\overline{\Omega})$ is same as the notation $C^{k+\gamma}(\Omega)$ used here.

Taking $k = 2$, the Hölder space $C^{2+\gamma}(\Omega)$ is defined as

$$C^{2+\gamma}(\Omega) = \{u \mid \frac{\partial^{|\alpha|}u}{\partial x^\alpha} \in C^\gamma(\Omega) \text{ for any } \alpha \in \mathbb{N}_0^n \text{ such that } |\alpha| \leqslant 2\}.$$

This is a Banach space equipped for which the norm

$$\|u\|_{C^{2+\gamma}(\Omega)} = \|u\|_{C^2(\Omega)} + \max_{|\alpha|=2} \sup_{x,y\in\Omega,\, x\neq y} \frac{\left|\frac{\partial^{|\alpha|}u(x)}{\partial x^\alpha} - \frac{\partial^{|\alpha|}u(y)}{\partial y^\alpha}\right|}{|x-y|^\gamma},$$

is finite, where $\|u\|_{C^2(\Omega)} = \max_{0\leqslant|\alpha|\leqslant 2} \sup_{x\in\Omega} \left|\frac{\partial^{|\alpha|}u}{\partial x^\alpha}\right|$, as defined in (2.1.1).

Now to define the parabolic Hölder space $C^{2+\gamma,1+\frac{\gamma}{2}}(\Omega_T)$, we first consider any two points $P(x,t), Q(y,s) \in \Omega_T$ and define a metric as,

$$d(P,Q) := \max\{|x-y|, |t-s|^{\frac{1}{2}}\}.$$

Let $u(x,t)$ be a function in Ω_T. For $0 < \gamma < 1$, we define the parabolic Hölder semi-norm as

$$[u]_{\gamma,\frac{\gamma}{2};\Omega_T} := \sup_{P,Q\in\Omega_T;P\neq Q} \frac{|u(x,t)-u(y,s)|}{(d(P,Q))^{\gamma}},$$

and denote by $C^{\gamma,\frac{\gamma}{2}}(\Omega_T)$ the set of all functions $u(x,t)$ satisfying $[u]_{\gamma,\frac{\gamma}{2};\Omega_T} < +\infty$, endowed with the norm

$$\|u\|_{\gamma,\frac{\gamma}{2};\Omega_T} := \sup_{(x,t)\in\Omega_T} |u(x,t)| + [u]_{\gamma,\frac{\gamma}{2};\Omega_T}. \tag{2.3.4}$$

For $\gamma \in (0,1)$, the *parabolic Hölder space* $C^{2+\gamma,1+\frac{\gamma}{2}}(\Omega_T)$ is defined as,

$$C^{2+\gamma,1+\frac{\gamma}{2}}(\Omega_T) := \{u \mid \frac{\partial^{\rho}}{\partial t^{\rho}}\frac{\partial^{|\beta|}}{\partial x^{\beta}}u \in C^{\gamma,\frac{\gamma}{2}}(\Omega_T) \text{ for any } \rho\in\mathbb{N}_0 \text{ and } \beta\in\mathbb{N}_0^n \text{ such that } 2\rho+|\beta|\leqslant 2\}.$$

Thus, the space $C^{2+\gamma,1+\frac{\gamma}{2}}(\Omega_T)$ consists of partial derivatives of u with respect to t upto order 1 and with respect to x upto order 2, respectively, which are Hölder continuous, such that the norm,

$$\|u\|_{C^{2+\gamma,1+\frac{\gamma}{2}}(\Omega_T)} := \|u\|_{\gamma,\frac{\gamma}{2};\Omega_T} + \max_{|\alpha|=1}\left\|\frac{\partial^{|\alpha|}u}{\partial x^{\alpha}}\right\|_{\gamma,\frac{\gamma}{2};\Omega_T} + \max_{|\alpha|=2}\left\|\frac{\partial^{|\alpha|}u}{\partial x^{\alpha}}\right\|_{\gamma,\frac{\gamma}{2};\Omega_T} + \|u_t\|_{\gamma,\frac{\gamma}{2};\Omega_T},$$

is finite, where $\|\cdot\|_{\gamma,\frac{\gamma}{2};\Omega_T}$ follows from (2.3.4).

3 Capacities and quasicontinuity

"The shortest route between two truths in the real domain passes through the complex domain."

— Jacques Salomon Hadamard (1865-1963)

In Mathematics, the notion of capacity of a set in Euclidean space is related to the measure of size of the set, introduced by Gustave Choquet [10] in 1950. It was inspired by classical (electrostatic) capacity but is now a general basis tool in analysis. In Calculus of Variations, the characterization of capacity of a set can be expressed as the minimization of an energy functional satisfying certain boundary values. This plays a significant role in extending its properties to other (generalized) functionals as well, which lays the foundation of (nonlinear) potential theory. The sets of capacity zero are the exceptional sets for representatives of the function space, especially with zero boundary values.

In this chapter, the properties of capacities associated with the generalized Orlicz and Sobolev-Orlicz function spaces, along with results related to their relationship with Hausdorff measure and quasicontinuity are stated from [4]. These results are applicable in the study of boundary behavior of solutions to PDEs, but not used in this thesis though. Additionally, the existence and uniqueness of the quasicontinuous represenatative are proved.

3.1 Capacity: definition and properties

In generalized Orlicz setting, the concept of capacity is introduced as follows.
First, we define a set of test-functions for the capacity of a set $E \subset \mathbb{R}^n$ as,

$$S_{1,\varphi}(E) := \{u \in W^{1,\varphi}(\mathbb{R}^n) \mid u \geqslant 1 \text{ in an open set containing } E \text{ and } u \geqslant 0\}.$$

Here $u \in S_{1,\varphi}(E)$ are said to be *φ-admissible* for the capacity of the set E. The *generalized Orlicz* φ-capacity of E is defined by

$$C_\varphi(E) := \inf_{u \in S_{1,\varphi}(E)} \int_{\mathbb{R}^n} \varphi(x,u) + \varphi(x, |\nabla u|)\, dx.$$

We now have the following properties for the set function $E \mapsto C_\varphi(E)$.

***Proposition* 3.1.1.** [4, Propositions 7, 8, 9] *The following properties of Sobolev capacity hold.*

(S1) *If* $\varphi \in \Phi_w(\mathbb{R}^n)$*, then* $C_\varphi(\emptyset) = 0$.

(S2) *If* $\varphi \in \Phi_w(\mathbb{R}^n)$ *and* $E_1 \subset E_2 \subset \mathbb{R}^n$*, then* $C_\varphi(E_1) \leqslant C_\varphi(E_2)$.

(S3) *If* $\varphi \in \Phi_w(\mathbb{R}^n)$ *and* $E \subset \mathbb{R}^n$*, then* $C_\varphi(E) = \inf_{U \supset E \text{ open}} C_\varphi(U)$.

(S4) *If* $\varphi \in \Phi_w(\mathbb{R}^n)$ *and* $E_1, E_2 \subset \mathbb{R}^n$*, then* $C_\varphi(E_1 \cup E_2) + C_\varphi(E_1 \cap E_2) \leqslant C_\varphi(E_1) + C_\varphi(E_2)$.

(S5) *If* $\varphi \in \Phi_w(\mathbb{R}^n)$ *and* $K_1 \supset K_2 \supset \ldots$ *are compact, then*

$$\lim_{i\to\infty} C_\varphi(K_i) = C_\varphi\Big(\bigcap_{i=1}^{\infty} K_i\Big).$$

(S6) *If* $\varphi \in \Phi_c(\mathbb{R}^n)$ *satisfies* (aInc) *and* (aDec) *and* $E_1 \subset E_2 \subset \cdots \subset \mathbb{R}^n$*, then*

$$\lim_{i\to\infty} C_\varphi(E_i) = C_\varphi\bigg(\bigcup_{i=1}^{\infty} E_i\bigg).$$

(S7) *If* $\varphi \in \Phi_w(\mathbb{R}^n)$ *satisfies* (aInc) *and* (aDec) *and* $E_1, E_2, \ldots \subset \mathbb{R}^n$*, then*

$$C_\varphi\bigg(\bigcup_{i=1}^{\infty} E_i\bigg) \leqslant \sum_{i=1}^{\infty} C_\varphi(E_i).$$

Remark 3.1.1. A set function satisfying the properties (S1), (S2) and (S7) is called an *outer measure*. This holds if φ is satisfies (aInc) and (aDec). If φ is convex and satisfies (aInc) and (aDec), then it is a *Choquet capacity*, [9], i.e. it satisfies (S1), (S2), (S5) and (S6). Then for every Borel $E \subset \Omega$,

$$C_\varphi(E) = \sup\big\{C_\varphi(K) : K \text{ is compact and } K \subset E\big\}.$$

Further, the following results discuss the connection between the generalized Orlicz capacity and the Lebesgue and Hausdorff measures. We also use the Sobolev p-capacity, denoted as $C_p(E)$, in which case $\varphi(x,t) := t^p$, for any fixed exponent $p \in [1,\infty)$, that is, we have

$$C_p(E) := \inf_{u \in \mathcal{S}(E)} \int_{\mathbb{R}^n} |u|^p + |\nabla u|^p \, dx,$$

where $\mathcal{S}(E) := \{u \in W^{1,p}(\mathbb{R}^n) : u \geqslant 1 \text{ on a neighbourhood of } E \text{ and } u \geqslant 0\}$.

***Proposition* 3.1.2.** [4, Proposition 12] *Let* $\varphi \in \Phi_w(\mathbb{R}^n)$ *satisfies* (A0)*,* (aInc)$_p$*,* $p > 1$*, and* (aDec)*. If* $C_\varphi(E) = 0$*, then* $C_p(E) = 0$.

Moreover, we have $|E| \leqslant C\, C_\varphi(E)$ for every $E \subset \mathbb{R}^n$ with constant $C > 0$, provided φ satisfies (aDec) and (A0) conditions. This implies that the sets of capacity zero are of measure zero.

The *s-dimensional Hausdorff measure* of a set $E \subset \mathbb{R}^n$, denoted by $\mathcal{H}^s(E)$, is defined as

$$\mathcal{H}^s(E) := \lim_{\delta\to 0} \mathcal{H}^s_\delta(E) = \sup_{\delta>0} \mathcal{H}^s_\delta(E),$$

where $\mathcal{H}^s_\delta(E) := \inf\{\sum_{i=1}^\infty r_i^s : E \subset \cup_{i=1}^\infty B(x_i, r_i), r_i \leqslant \delta\}$.

***Lemma* 3.1.3.** [4, Corollary 14] *Let $\varphi \in \Phi_w(\mathbb{R}^n)$ satisfies* (A0)*,* (aInc)$_p$ *with $p > 1$ and* (aDec)*.*

(i) If $p \leqslant n$ and $E \subset \mathbb{R}^n$ with $C_\varphi(E) = 0$, then $\mathcal{H}^s(E) = 0$ for all $s > n - p$.
(ii) If $p > n$, then $C_\varphi(E) = 0$ if and only if $E = \emptyset$, where $E \subset \mathbb{R}^n$.

The above result implies that capacity is a useful tool only when $p \leqslant n$.

***Lemma* 3.1.4.** [4, Corollary 15] *Let $\varphi \in \Phi_w(\mathbb{R}^n)$ satisfies* (A0)*,* (aInc) *and* (aDec)$_q$ *with $q > 1$. Let $E \subset \mathbb{R}^n$ be bounded. If $C_q(E) = 0$ or $\mathcal{H}^{n-q}(E) < \infty$, then $C_\varphi(E) = 0$.*

Next, we have another notion of capacity, known as relative capacity, in which the capacity of a set is taken relative to a surrounding open subset of $\mathbb{R}^n$, defined as follows.

Definition 3.1.2. Given an open set, $\Omega \subset \mathbb{R}^n$, and $\varphi \in \Phi_w(\Omega)$, suppose that K is a compact subset of Ω. We denote,

$$R_\varphi(K, \Omega) := \{u \in W^{1,\varphi}(\Omega) \cap C_0(\Omega) : u > 1 \text{ in } K \text{ and } u \geqslant 0\}$$

and, define,

$$\operatorname{cap}^*_\varphi(K, \Omega) := \inf_{u\in R_\varphi(K,\Omega)} \varrho_\varphi(|\nabla u|).$$

Further, if $U \subset \Omega$ is open, then we set

$$\operatorname{cap}_\varphi(U, \Omega) := \sup_{K\subset U \text{ compact}} \operatorname{cap}^*_\varphi(K, \Omega),$$

and, for an arbitrary set $E \subset \Omega$, we define

$$\operatorname{cap}_\varphi(E, \Omega) := \inf_{U\supset E \text{ open}} \operatorname{cap}_\varphi(U, \Omega).$$

The number $\operatorname{cap}_\varphi(\mathrm{E}, \Omega)$ is called the relative φ-capacity of E with respect to Ω.

For a compact set $K \subset \Omega$, $\operatorname{cap}^*_\varphi(K, \Omega)$ and $\operatorname{cap}_\varphi(K, \Omega)$ are the same, that is, the relative capacity is well defined on compact sets [4, Proposition 21]. Moreover, the relative φ-capacity has the same basic properties as the Sobolev capacity, as stated below.

***Proposition* 3.1.5.** [4, Proposition 22] *For an open set $\Omega \subset \mathbb{R}^n$ and $\varphi \in \Phi_w(\mathbb{R}^n)$, the set function $E \mapsto \operatorname{cap}_\varphi(E,\Omega)$ satisfies the following properties:*

(R1) $\operatorname{cap}_\varphi(\emptyset,\Omega) = 0$.

(R2) If $E_1 \subset E_2 \subset \Omega_2 \subset \Omega_1$, then $\operatorname{cap}_\varphi(E_1,\Omega_1) \leqslant \operatorname{cap}_\varphi(E_2,\Omega_2)$.

(R3) For an arbitrary set $E \subset \Omega$,

$$\operatorname{cap}_\varphi(E,\Omega) = \inf_{U \supset E \text{ open}} \operatorname{cap}_\varphi(U,\Omega).$$

(R4) If $E_1, E_2 \subset \mathbb{R}^n$, then $\operatorname{cap}_\varphi(E_1 \cup E_2,\Omega) + \operatorname{cap}_\varphi(E_1 \cap E_2,\Omega) \leqslant \operatorname{cap}_\varphi(E_1,\Omega) + \operatorname{cap}_\varphi(E_2,\Omega)$.

(R5) If $K_1 \supset K_2 \supset \dots$ are compact, then $\lim_{i\to\infty} \operatorname{cap}_\varphi(K_i,\Omega) = \operatorname{cap}_\varphi\Big(\bigcap_{i=1}^\infty K_i,\Omega\Big)$.

(R6) If $E_1 \subset E_2 \subset \dots \subset \Omega$, then $\displaystyle\lim_{i\to\infty} \operatorname{cap}_\varphi(E_i,\Omega) = \operatorname{cap}_\varphi\Big(\bigcup_{i=1}^\infty E_i,\Omega\Big)$.

(R7) If $E_i \subset \Omega, i = 1,2,\dots$, then $\displaystyle\operatorname{cap}_\varphi\Big(\bigcup_{i=1}^\infty E_i,\Omega\Big) \leqslant \sum_{i=1}^\infty \operatorname{cap}_\varphi(E_i,\Omega)$.

Remark 3.1.3. A set function satisfying the properties (R1), (R2), (R5) and (R6) is a *Choquet capacity*. Then for every Borel $E \subset \Omega$,

$$\operatorname{cap}_\varphi(E,\Omega) = \sup\big\{ \operatorname{cap}_\varphi(K,\Omega) : K \text{ is compact and } K \subset E \big\}.$$

Finally, we have a few results on relationships between the Sobolev capacity and relative capacity.

***Lemma* 3.1.6.** [4, Lemma 26] *Assume that $\varphi \in \Phi_w(\mathbb{R}^n)$ satisfies* (A0) *and* (aDec)$_q$. *If Ω is bounded and $K \subset \Omega$ is compact, then,*

$$C_\varphi(K) \leqslant C \max\{\operatorname{cap}_\varphi(K,\Omega)^{\frac{1}{q}}, \operatorname{cap}_\varphi(K,\Omega)\},$$

where the constant C depends on the dimension n, $|\Omega|$ and the constants in (A0) *and* (aDec).

***Theorem* 3.1.7.** [4, Theorem 27] *Assume that $\varphi \in \Phi(\mathbb{R}^n)$ satisfies* (A0) *and* (aDec)$_q$. *If Ω is bounded and $E \subset \Omega$, then*

$$C_\varphi(E) \leqslant C \max\{\operatorname{cap}_\varphi(E,\Omega)^{\frac{1}{q}}, \operatorname{cap}_\varphi(E,\Omega)\},$$

where the constant C depends on the dimension n, $|\Omega|$, and the constants in (A0) *and* (aDec).

From the above result, it can be concluded that $C_\varphi(E) = 0$ if $\operatorname{cap}_\varphi(E,\Omega) = 0$. The converse implication is established in the following result.

***Theorem* 3.1.8.** [4, Proposition 29] *Let $\Omega \subset \mathbb{R}^n$ be bounded and $\varphi \in \Phi_w(\mathbb{R}^n)$ satisfies* (aDec). *Assume that $C(\mathbb{R}^n)$ is dense in $W^{1,\varphi}(\mathbb{R}^n)$. If $E \subset \Omega$ with $C_\varphi(E) = 0$, then* $\operatorname{cap}_\varphi(E,\Omega) = 0$.

3.2 Quasicontinuity

Here, we have the relation between the generalized Orlicz capacity and quasicontinuity as studied in [4]. In general, a function $u : \mathbb{R}^n \to [-\infty, \infty]$ is *φ-quasicontinuous* in $\mathbb{R}^n$ if for every $\epsilon > 0$ there exists an open set E such that $C_\varphi(E) < \epsilon$ and the restriction of u to $\mathbb{R}^n \setminus E$ is continuous. We say that a claim holds φ-quasieverywhere if it holds everywhere except in a set of Sobolev φ-capacity zero. The following result establishes a quasieverywhere converging subsequence.

***Theorem* 3.2.1.** [4, Theorem 16] *Let $\varphi \in \Phi_w(\mathbb{R}^n)$ satisfies (aInc) and (aDec). Then for each Cauchy sequence in $C(\mathbb{R}^n) \cap W^{1,\varphi}(\mathbb{R}^n)$, there is a subsequence which converges pointwise φ-quasieverywhere in $\mathbb{R}^n$. Moreover, the convergence is uniform outside a set of arbitrarily small Sobolev φ-capacity.*

The next result shows that under suitable conditions, every Sobolev-Orlicz function has a φ-quasicontinuous representative.

***Theorem* 3.2.2.** [4, Theorem 17] *Let $\varphi \in \Phi_w(\mathbb{R}^n)$ satisfies* (aInc) *and* (aDec). *Assume that $C(\mathbb{R}^n) \cap W^{1,\varphi}(\mathbb{R}^n)$ is dense in $W^{1,\varphi}(\mathbb{R}^n)$. Then, for each $u \in W^{1,\varphi}(\mathbb{R}^n)$, there exists a φ-quasicontinuous function $g \in W^{1,\varphi}(\mathbb{R}^n)$ such that $u = g$ almost everywhere in $\mathbb{R}^n$.*

Analogously, we also have that φ-quasicontinuous functions in Sobolev-Orlicz spaces with zero boundary values are zero φ-quasieverywhere in the complement, from the following result.

***Theorem* 3.2.3.** *Let $\varphi \in \Phi_w(\mathbb{R}^n)$ satisfies* (A0), (aInc) *and* (aDec). *Assume that $C^\infty(\mathbb{R}^n) \cap W^{1,\varphi}(\mathbb{R}^n)$ is dense in $W^{1,\varphi}(\mathbb{R}^n)$. Then $u \in W_0^{1,\varphi}(\Omega)$ if and only if there exists a φ-quasicontinuous function $v \in W^{1,\varphi}(\mathbb{R}^n)$ such that $v = u$ almost everywhere in Ω and $v = 0$ φ-quasieverywhere in $\mathbb{R}^n \setminus \Omega$.*

Proof. Let $u \in W_0^{1,\varphi}(\Omega)$. Then there exists a sequence $\{u_i\}, u_i \in C_0^\infty(\Omega), i = 1, 2, \ldots$, such that $u_i \to u$ in $W^{1,\varphi}(\Omega)$ as $i \to \infty$. Since $\{u_i\}$ is a Cauchy sequence in $W^{1,\varphi}(\mathbb{R}^n)$, then by Theorem 3.2.1, there exists a subsequence of $\{u_i\}$ that converges pointwise φ-quasieverywhere in $\mathbb{R}^n$ to a function $v \in W^{1,\varphi}(\mathbb{R}^n)$. Moreover, the convergence is uniform outside a set of arbitrarily small capacity. Uniform convergence implies continuity of the limit function and thus the function v is continuous outside a set of arbitrarily small φ-capacity. Hence, by definition, $v \in W^{1,\varphi}(\mathbb{R}^n)$ is a φ-quasicontinuous function such that $v = u$ almost everywhere in Ω and $v = 0$ φ-quasieverywhere in $\mathbb{R}^n \setminus \Omega$.

Next, we prove the converse, that if $u \in W^{1,\varphi}(\mathbb{R}^n)$ is φ-quasicontinuous and $u = 0$ φ-quasieverywhere in $\mathbb{R}^n \setminus \Omega$, then $u \in W_0^{1,\varphi}(\Omega)$. In order for that, we show that $u \in W^{1,\varphi}(\mathbb{R}^n)$ can be approximated by $W^{1,\varphi}(\mathbb{R}^n)$ functions with compact support in Ω, and then use the assumption of $C^\infty(\mathbb{R}^n) \cap W^{1,\varphi}(\mathbb{R}^n)$ being dense in $W^{1,\varphi}(\mathbb{R}^n)$. If such sequence of functions u exist for $u_+ := \max\{u, 0\}$, we can similarly claim so for $u_- := \min\{u, 0\}$. Thus, we may assume that $u \geqslant 0$. Since φ satisfies (aDec) property, we may further assume that u is bounded and has compact support in $\mathbb{R}^n$, by [24, Lemma 6.4.2].

Let $\delta > 0$ and $U \subset \mathbb{R}^n$ be an open set such that $C_\varphi(U) < \delta$ and the restriction of u to $\mathbb{R}^n \setminus U$ is continuous. Denote

$$E := \{x \in \mathbb{R}^n \setminus \Omega : u(x) \neq 0\}.$$

By assumption, we have $C_\varphi(E) = 0$. This yields $C_\varphi(U \cup E) \leqslant C_\varphi(U) + C_\varphi(E) < \delta$. Then we choose $w_\delta \in S_{1,\varphi}(U \cup E)$ such that $0 \leqslant w_\delta \leqslant 1$ and $\varrho_{1,\varphi}(w_\delta) < \delta$. Then $w_\delta = 1$ in an open set V containing $U \cup E$. For $0 < \epsilon < 1$, define

$$u_\epsilon(x) := \max\{u(x) - \epsilon, 0\}. \tag{3.2.1}$$

Since $u(x) = 0$ for $x \in \partial\Omega \setminus V$ and the restriction of u to $\mathbb{R}^n \setminus V$ is continuous, there exists $r_x > 0$ such that $u_\epsilon = 0$ in $B(x, r_x) \setminus V$. Thus $(1 - w_\delta)u_\epsilon = 0$ in $B(x, r_x) \cup V$ for each $x \in \partial\Omega \setminus V$. This shows that $(1 - w_\delta)u_\epsilon$ is zero in a neighbourhood of $\mathbb{R}^n \setminus \Omega$, which implies that $(1 - w_\delta)u_\epsilon$ is compactly supported in Ω. Hence, by [24, Lemma 6.1.10], we have

$$(1 - w_\delta)u_\epsilon \in W_0^{1,\varphi}(\Omega).$$

We next show that this kind of functions converge to u in $W^{1,\varphi}(\mathbb{R}^n)$.

From (3.2.1), we conclude using Lemma 2.2.7 that

$$\nabla u_\epsilon(x) = \begin{cases} \nabla u(x), & \text{almost everywhere in } \{x \in \mathbb{R}^n : u(x) > \epsilon\}, \\ 0, & \text{almost everywhere in } \{x \in \mathbb{R}^n : u(x) \leqslant \epsilon\}. \end{cases}$$

We have,

$$\|u - (1 - w_\delta)u_\epsilon\|_{W^{1,\varphi}(\mathbb{R}^n)} \leqslant \|u - u_\epsilon\|_{W^{1,\varphi}(\mathbb{R}^n)} + \|w_\delta u_\epsilon\|_{W^{1,\varphi}(\mathbb{R}^n)}. \tag{3.2.2}$$

Using the facts that $|u - u_\epsilon| \leqslant \epsilon$ and $\operatorname{supp}(u - u_\epsilon) \subset \operatorname{supp} u$, the first expression on the right-hand side of (3.2.2) implies

$$\begin{aligned} \|u - u_\epsilon\|_{W^{1,\varphi}(\mathbb{R}^n)} &= \|u - u_\epsilon\|_{L^\varphi(\mathbb{R}^n)} + \|\nabla(u - u_\epsilon)\|_{L^\varphi(\mathbb{R}^n)} \\ &\leqslant \epsilon\|\chi_{\operatorname{supp} u}\|_{L^\varphi(\mathbb{R}^n)} + \|\chi_{\{0<u\leqslant\epsilon\}}\nabla u\|_{L^\varphi(\mathbb{R}^n)}. \end{aligned} \tag{3.2.3}$$

Since φ satisfy (A0) property, hence the quantity $\|\chi_{\operatorname{supp} u}\|_{L^\varphi(\mathbb{R}^n)}$ on the right above, is finite. On the other hand, by the dominated convergence property for the modular [24, Lemma 3.1.4(c)], we have,

$$\lim_{\epsilon\to 0} \varrho_\varphi(\chi_{\{0<u\leqslant\epsilon\}}\nabla u) = 0$$

where $|\chi_{\{0<u\leqslant\epsilon\}}\nabla u| \leqslant |\nabla u| \in L^1(\mathbb{R}^n)$ since ∇u belongs to $L^\varphi(\mathbb{R}^n)$ with compact support in $\mathbb{R}^n$, which may be used as an integrable majorant. Further, since norm convergence and modular convergence are equivalent by [24, Corollary 3.3.4], we thus obtain $\lim_{\epsilon\to 0} \|\chi_{\{0<u\leqslant\epsilon\}}\nabla u\|_{L^\varphi(\mathbb{R}^n)} = 0$. Hence, from (3.2.3), we have

$$\|u - u_\epsilon\|_{W^{1,\varphi}(\mathbb{R}^n)} \to 0, \quad \text{as } \epsilon \to 0.$$

On the other hand, for the second expression in (3.2.2), we have

$$\begin{aligned} \|w_\delta u_\epsilon\|_{W^{1,\varphi}(\mathbb{R}^n)} &\leqslant \|w_\delta u_\epsilon\|_{L^\varphi(\mathbb{R}^n)} + \|\nabla(w_\delta u_\epsilon)\|_{L^\varphi(\mathbb{R}^n)} \\ &\leqslant \|w_\delta u_\epsilon\|_{L^\varphi(\mathbb{R}^n)} + \|u_\epsilon \nabla w\|_{L^\varphi(\mathbb{R}^n)} + \|w_\delta \nabla u_\epsilon\|_{L^\varphi(\mathbb{R}^n)}. \end{aligned} \tag{3.2.4}$$

Since $|u_\epsilon| \leqslant |u| \leqslant \|u\|_{L^\infty}$, by solid property of L^φ, we have $\|w_\delta u_\epsilon\|_{L^\varphi(\mathbb{R}^n)} \leqslant \big\|w_\delta \|u\|_{L^\infty}\big\|_{L^\varphi(\mathbb{R}^n)}$ and $\|u_\epsilon \nabla w_\delta\|_{L^\varphi(\mathbb{R}^n)} \leqslant \big\|\|u\|_{L^\infty}\nabla w_\delta\big\|_{L^\varphi(\mathbb{R}^n)}$. With $|\nabla u_\epsilon| \leqslant |\nabla u|$, we similarly obtain the estimate $\|w_\delta \nabla u_\epsilon\|_{L^\varphi(\mathbb{R}^n)} \leqslant \|w_\delta \nabla u\|_{L^\varphi(\mathbb{R}^n)}$. Then (3.2.4) implies

$$\begin{aligned} \|w_\delta u_\epsilon\|_{W^{1,\varphi}(\mathbb{R}^n)} &\leqslant \big\|w_\delta \|u\|_{L^\infty}\big\|_{L^\varphi(\mathbb{R}^n)} + \big\|\nabla w_\delta \|u\|_{L^\infty}\big\|_{L^\varphi(\mathbb{R}^n)} + \|w_\delta \nabla u\|_{L^\varphi(\mathbb{R}^n)} \\ &\leqslant \|u\|_{L^\infty}\|w_\delta\|_{L^\varphi(\mathbb{R}^n)} + \|u\|_{L^\infty}\|\nabla w_\delta\|_{L^\varphi(\mathbb{R}^n)} + \|w_\delta \nabla u\|_{L^\varphi(\mathbb{R}^n)} \end{aligned}$$

which gives

$$\begin{aligned} \|w_\delta u_\epsilon\|_{W^{1,\varphi}(\mathbb{R}^n)} &\leqslant \|u\|_{L^\infty}\|w_\delta\|_{W^{1,\varphi}(\mathbb{R}^n)} + \|w_\delta \nabla u\|_{L^\varphi(\mathbb{R}^n)} \\ &\leqslant C\delta^{\frac{1}{p}}\|u\|_{L^\infty} + \|w_\delta \nabla u\|_{L^\varphi(\mathbb{R}^n)}, \end{aligned}$$

since $\varrho_{1,\varphi}(w_\delta) < \delta$, which implies $\|w_\delta\|_{W^{1,\varphi}(\mathbb{R}^n)} \leqslant C\delta^{\frac{1}{q}}$ following the (aInc) and (aDec) properties, where $C > 0$ is the maximum of the constants from (aInc) and (aDec). As $\delta \to 0$, we have $w_\delta \to 0$ in $L^\varphi(\mathbb{R}^n)$, hence there exists a subsequence $\{w_{\delta_i}\}_{i=1}^\infty \subset w_\delta$ which tends to 0 pointwise almost everywhere, as $\delta_i \to 0$. Again using the dominated convergence property of the modular, we have

$$\lim_{\delta_i\to 0} \varrho_\varphi(w_{\delta_i}\nabla u) = 0,$$

where $|w_{\delta_i}\nabla u| \leqslant |\nabla u|$, so that $|\nabla u| \in L^1(\mathbb{R}^n)$ may be used as an integrable majorant. Since modular and norm convergence are equivalent, the above limit implies $\lim_{\delta_i\to 0}\|w_{\delta_i}\nabla u\|_{L^\varphi(\mathbb{R}^n)} = 0$. Thus, we conclude that,

$$\lim_{\delta_i\to 0}\|w_{\delta_i}\nabla u\|_{W^{1,\varphi}(\mathbb{R}^n)} \leqslant \lim_{\delta_i\to 0}(C\delta^p\|u\|_{L^\infty} + \|w_{\delta_i}\nabla u\|_{L^\varphi(\mathbb{R}^n)}) = 0.$$

Hence,

$$\|u - (1 - w_{\delta_i})u_\epsilon\|_{W^{1,\varphi}(\mathbb{R}^n)} \to 0$$

as $\epsilon, \delta_i \to 0$. Since $(1 - w_{\delta_i})u_\epsilon \in W_0^{1,\varphi}(\Omega)$ and $(1 - w_{\delta_i})u_\epsilon \to u$ in $W^{1,\varphi}(\mathbb{R}^n)$ as $\epsilon, \delta_i \to 0$, we conclude that $u \in W_0^{1,\varphi}(\Omega)$. □

The above result can be used to prove that a given function belongs to the Sobolev-Orlicz space with zero boundary values without constructing an approximating sequence of compactly supported smooth functions.

Further, to prove the uniqueness of the quasicontinuous representative of φ-capacity, we first have the following lemma.

***Lemma* 3.2.4.** *Let $\varphi \in \Phi_w(\mathbb{R}^n)$. Assume that $G \subset \mathbb{R}^n$ is open and $E \subset \mathbb{R}^n$ with $|E| = 0$. Then $C_\varphi(G) = C_\varphi(G \setminus E)$.*

Proof. By monotonocity property of C_φ from Proposition 3.1.1(S2), we have,

$$C_\varphi(G \setminus E) \leqslant C_\varphi(G). \tag{3.2.5}$$

Let $\epsilon > 0$ and let $u \in S_{1,\varphi}(G \setminus E)$ be such that,

$$\varrho_{1,\varphi}(u) \leqslant C_\varphi(G \setminus E) + \epsilon.$$

Then there exists an open set $O \subset \mathbb{R}^n$, with $(G \setminus E) \subset O$ and $u \geqslant 1$ in O. Since $O \cup G$ is open and taking $u \geqslant 1$ in $O \cup (G \setminus E)$, we have that $u \geqslant 1$ almost everywhere in $O \cup G$ since $|E| = 0$. Hence, we conclude that $u \in S_{1,\varphi}(G)$, which implies

$$C_\varphi(G) \leqslant \varrho_{1,\varphi}(u) \leqslant C_\varphi(G \setminus E) + \epsilon.$$

Then letting $\epsilon \to 0$ to obtain,

$$C_\varphi(G) \leqslant C_\varphi(G \setminus E). \tag{3.2.6}$$

Hence, from (3.2.5) and (3.2.6), we conclude that

$$C_\varphi(G) = C_\varphi(G \setminus E).$$

□

Finally, we state the uniqueness property of φ-quasicontinuous representative, which follows as a special case from [35, Theorem] that shows two quasicontinuous functions that agree almost everywhere coincide quasieverywhere.

***Theorem* 3.2.5.** *Let $\varphi \in \Phi_w(\mathbb{R}^n)$ satisfies* (aInc) *and* (aDec). *Assume that u and v are φ- quasicontinuous functions on $\mathbb{R}^n$. If $u = v$ almost everywhere in $\mathbb{R}^n$, then $u = v$ φ-quasieverywhere in $\mathbb{R}^n$.*

4 Double phase growth functionals in image restoration

"As far as the laws of mathematics refer to reality, they are not certain, and as far as they are certain, they do not refer to reality."

— Albert Einstein (1879-1955)

4.1 Introduction

One of the most active research areas in mathematical image processing and computer vision is image restoration, which deals with the recovery of images corrupted due to noise, data errors and geometric distortions. A major concern in designing image denoising models for feature extraction and target detection is to preserve significant image features, such as edges, while removing noise. Thus, the important task is how to preserve edges in restored images. Variational approach have been shown to be extremely successful for a wide variety of image restoration problems (see [19; 30]), not only limited to the fundamental problem of image denoising, but also other restoration tasks such as deblurring, blind deconvolution and inpainting. Variational models exhibit the solution of these problems as minimizers of appropriately chosen functionals, which involve the solution of nonlinear partial differential equations (PDEs) derived as necessary optimality conditions.

The pioneering works in this field mainly consist of Wiener filter [53], Tikhonov regularization and its extensions [59]. However, experiments have shown that these traditional methods commonly suffer from noise amplification or ringing-like artifacts, which degrade the image quality significantly. In order to overcome these limitations, *total variation* (TV) based image restoration models were first introduced by Rudin, Osher and Fatemi (ROF) in their work [56] on PDE based edge preserving denoising. It was designed with the explicit goal of preserving sharp discontinuities (edges) in images while removing noise and other unwanted fine scale detail, along with smoothing flat regions. However, these restored images consist of undesirable staircase-like features as the TV regularizer favors solutions which are piecewise constant.

Over the years, the ROF model has been extended to many other image restoration tasks, and has been modified in a variety of ways to improve its performance. Another growing interest in the literature is replacing the TV-norm by higher order norms in order to maintain sharp edges and avoid staircase effects on the smooth part. This procedure is introduced, e.g. in [5], where properties of higher order TV have been established. The higher order norms involving second-order differential operators usually lead to piecewise-linear solutions as discussed in [40], and this is useful for image restoration problems and in applications such as biomedical imaging.

4.2 Problem formulation

In this section, we deal with variational integral functionals having non-standard growth conditions [1; 43; 45], specifically of the type $\mathcal{F}(u) := \int_\Omega \mathcal{H}(x, |\nabla u(x)|)\, dx$, where $\mathcal{H} : \Omega \times [0,\infty) \to [0,\infty)$, with $r \in \mathbb{R}^+$, defined as,

$$\mathcal{H}(x,r) := r^p + a(x)r^q, \quad \text{in } \Omega \times [0,\infty) \tag{4.2.1}$$

with $1 < p \leqslant q \leqslant 2$ and $\Omega \subset \mathbb{R}^n$ is a nonempty, bounded, open set with $n \geqslant 2$. Here, $0 \leqslant a(\cdot) \in L^\infty(\Omega)$ is Lipschitz continuous. Such class of double-phase functionals, introduced by Zhikov [64; 65], are basically characterised by the fact of having the energy density switching between two different types of elliptic behaviours, in accordance to the size of the "modulating coefficient $a(\cdot)$" that determines the phase. Further study in the framework of regularity theory for minimizers of this class of double-phase integrals, can be found e.g. in [3; 12; 13].

The derivative of the functional $\mathcal{H}(x,r)$ with respect to r is given by

$$\mathcal{H}_r(x,r) := pr^{p-1} + a(x)qr^{q-1}.$$

Note that, we denote the derivative of any $\mathcal{H}$ functional as $\mathcal{H}_r$, hence the derivative of $\mathcal{H}(x, |\nabla u|)$ with respect to $|\nabla u|$ is $\mathcal{H}_r(x, |\nabla u|) = p|\nabla u|^{p-1} + a(x)q|\nabla u|^{q-1}$.

We also have $\lim_{r\to 0^+} \mathcal{H}(x,r) = 0$ and $\lim_{r\to\infty} \mathcal{H}(x,r) = \infty$. The convexity of $r \mapsto r^p$ and $r \mapsto r^q$, where $1 < p \leqslant q \leqslant 2$, implies that $\mathcal{H}(x,r)$ is convex. As $\mathcal{H}(x,0) = 0$, then for $0 < s < r$, $s \in \mathbb{R}$, we have

$$\mathcal{H}(x,s) = \mathcal{H}(x, \tfrac{s}{r}r + 0) \leqslant \tfrac{s}{r}\mathcal{H}(x,r) + \left(1 - \tfrac{s}{r}\right)\mathcal{H}(x,0) = \tfrac{s}{r}\mathcal{H}(x,r),$$

which implies (Inc)$_1$ holds on $(0,\infty)$. Thus, with $x \mapsto \mathcal{H}(x, |\nabla u(x)|)$ measurable for every $|\nabla u(x)| \in L^0(\Omega)$, $\mathcal{H} : \Omega \times [0,\infty) \to [0,\infty)$ is a generalized weak Φ-function on Ω, that is $\mathcal{H} \in \Phi_w(\Omega)$.

The functional $\mathcal{H}(x,r)$ satisfies the Δ_2-condition, that is,

$$\mathcal{H}(x,2r) = 2^p r^p + a(x)2^q r^q \leqslant 2^q \mathcal{H}(x,r),$$

holds, for every $x \in \Omega$ and $r \in \mathbb{R}^+$. Moreover, $\mathcal{H}$ satisfies (aDec) property, which implies that $\mathcal{H}$ is finite. Hence, $\mathcal{H}(x,\cdot)$ being continuous and convex for almost all $x \in \Omega$, we have $\mathcal{H} \in \Phi_s(\Omega)$.

Additionally, $\mathcal{H}$ satisfies (A0) condition, by [24, Proposition 7.2.1]. On the other hand, the function $a(x)$ being Lipschitz continuous, we have that $a \in C^{\frac{n}{p}(q-p)}(\Omega)$ for $\frac{n}{p}(q-p) \leqslant 1$. Then by [24, Proposition 7.2.2], $\mathcal{H}$ satisfies the (A1) condition, provided $q < \frac{3}{2}p$ for $n = 2$.

Switching the symbolic representations of Φ-functions, from φ to $\mathcal{H}$, for convenient understanding, we have the generalized Orlicz space denoted as $L^{\mathcal{H}}(\Omega), \|\cdot\|_{\mathcal{H}}$ and the associated Sobolev-Orlicz spaces as $W^{1,\mathcal{H}}(\Omega), W_0^{1,\mathcal{H}}(\Omega)$, respectively. We now have the following useful preliminary results in regard to these spaces.

***Lemma* 4.2.1.** *For $\Omega \subset \mathbb{R}^n$ and $\mathcal{H} : \Omega \times [0,\infty) \to [0,\infty)$ as defined in* (4.2.1)*, the spaces $L^{\mathcal{H}}(\Omega)$, $W^{1,\mathcal{H}}(\Omega)$ and $W_0^{1,\mathcal{H}}(\Omega)$ are separable, uniformly convex and reflexive Banach spaces.*

Proof. Since $\mathcal{H} \in \Phi_w(\Omega)$ from (4.2.1) is convex and satisfies (aInc) and (aDec) conditions, we can conclude that $L^{\mathcal{H}}(\Omega)$ and $W^{1,\mathcal{H}}(\Omega)$ are separable, uniformly convex and reflexive Banach spaces by Lemma 2.2.1 and Lemma 2.2.6, respectively. Similarly, by [24, Theorem 6.1.9], $W_0^{1,\mathcal{H}}(\Omega)$ is also separable, uniformly convex and a reflexive Banach space. □

***Lemma* 4.2.2.** *Let $\Omega \subset \mathbb{R}^n$ be a bounded open set, and function $\mathcal{H} : \Omega \times [0,\infty) \to [0,\infty)$ as defined in* (4.2.1)*. Assume that $\{u_i\}_{i\in\mathbb{N}}$ is a bounded sequence in $W^{1,\mathcal{H}}(\Omega)$ such that its subsequence $u_i \rightharpoonup u$ weakly in $L^{\mathcal{H}}(\Omega)$ as $i \to \infty$. Then $u \in W^{1,\mathcal{H}}(\Omega)$ such that the partial derivatives $\partial_j u_i \rightharpoonup \partial_j u$ weakly in $L^{\mathcal{H}}(\Omega)$ as $i \to \infty$. Moreover, if $u_i \in W_0^{1,\mathcal{H}}(\Omega)$, then we have $u \in W_0^{1,\mathcal{H}}(\Omega)$.*

Proof. Since $W^{1,\mathcal{H}}(\Omega)$ is a reflexive Banach space, by Lemma 4.2.1, then for every bounded sequence $\{u_i\}_{i=1}^{\infty}$ in $W^{1,\mathcal{H}}(\Omega)$, there exists a subsequence of $\{u_i\}$, denoted by same $\{u_i\}$, and $u \in W^{1,\mathcal{H}}(\Omega)$, such that $u_i \to u$ weakly in $L^{\mathcal{H}}(\Omega)$ and the partial derivatives $\partial_j u_i \to \partial_j u$ weakly in $L^{\mathcal{H}}(\Omega)$, as $i \to \infty$.
Analogously, the claim holds for $u_i \in W_0^{1,\mathcal{H}}(\Omega)$. □

Next we consider the property of lower semicontinuity with respect to weak convergence, as in [14, Theorem 2.2.8.] stated for a normed space X with semimodular ϱ [14, Definition 2.1.1.]. Since the properties of ϱ are satisfied by the modular $\varrho_{\mathcal{H}}$, the lower semicontinuity property holds for generalized Orlicz space $L^{\mathcal{H}}(\Omega)$ as follows.

***Lemma* 4.2.3.** *The modular $\varrho_{\mathcal{H}}$ on $L^{\mathcal{H}}(\Omega)$ is weakly (sequentially) lower semicontinuous, that is, if $u_k \rightharpoonup u$ weakly in $L^{\mathcal{H}}(\Omega)$, then $\varrho_{\mathcal{H}}(u) \leqslant \liminf_{k\to\infty} \varrho_{\mathcal{H}}(u_k)$.*

***Lemma* 4.2.4.** *Let $\Omega \subset \mathbb{R}^2$ be a bounded open set, and function $\mathcal{H} : \Omega \times [0,\infty) \to [0,\infty)$ as defined in* (4.2.1)*. Then $W^{1,\mathcal{H}}(\Omega) \hookrightarrow L^2(\Omega)$ holds.*

Proof. By [24, Lemma 6.1.6], we have

$$W^{1,\mathcal{H}}(\Omega) \hookrightarrow W^{1,p}(\Omega). \tag{4.2.2}$$

While, Sobolev embedding implies,

$$W^{1,p}(\Omega) \hookrightarrow L^{p^*}(\Omega), \tag{4.2.3}$$

where $p^* = \frac{2p}{2-p} \geqslant 2$ is the Sobolev conjugate exponent. Due to the boundedness of Ω, we have from Hölder's inequality for Lebesgue spaces,

$$L^{p^*}(\Omega) \hookrightarrow L^2(\Omega). \tag{4.2.4}$$

Thus, from (4.2.2),(4.2.3) and (4.2.4), we obtain,

$$W^{1,\mathcal{H}}(\Omega) \hookrightarrow L^2(\Omega).$$

□

The following result on density of smooth functions in parabolic spaces is proved using a similar approach from [51, Theorem 4.3].

***Lemma* 4.2.5.** *Let $\Omega \subset \mathbb{R}^n$ and $\mathcal{H} : \Omega \times [0,\infty) \to [0,\infty)$ as defined in* (4.2.1). *Then, for $T > 0$, $C_0^\infty(\Omega_T)$ is dense in $L^2(0,T;W_0^{1,\mathcal{H}}(\Omega))$, where $\Omega_T := \Omega \times (0,T)$.*

Proof. By the definition of the space $L^2(0,T;W_0^{1,\mathcal{H}}(\Omega))$, we have strongly measurable functions u such that $u(t) : (0,T) \mapsto W_0^{1,\mathcal{H}}(\Omega)$. First we prove that any function $u \in L^2(0,T;W_0^{1,\mathcal{H}}(\Omega))$, denoted by $u(t) = u(x,t)$, can be approximated with simple functions.

Now, by separability property of $W_0^{1,\mathcal{H}}(\Omega)$ (Lemma 4.2.1), we choose a countable, dense set $\{a_k\}_{k=1}^\infty \subset u(0,T)$. Fixing $n \in \mathbb{N}$, we partition $W_0^{1,\mathcal{H}}(\Omega)$ into,

$$\begin{aligned}
F_1^n :=& \{f \in W_0^{1,\mathcal{H}}(\Omega) : \|f - a_1\|_{W^{1,\mathcal{H}}(\Omega)} \leqslant \min_{1\leqslant i\leqslant n} \|f - a_i\|_{W^{1,\mathcal{H}}(\Omega)}\} \\
F_2^n :=& \{f \in W_0^{1,\mathcal{H}}(\Omega) : \|f - a_2\|_{W^{1,\mathcal{H}}(\Omega)} \leqslant \min_{1\leqslant i\leqslant n} \|f - a_i\|_{W^{1,\mathcal{H}}(\Omega)}\} \setminus F_1^n \\
&\vdots \\
F_k^n :=& \{f \in W_0^{1,\mathcal{H}}(\Omega) : \|f - a_k\|_{W^{1,\mathcal{H}}(\Omega)} \leqslant \min_{1\leqslant i\leqslant n} \|f - a_i\|_{W^{1,\mathcal{H}}(\Omega)}\} \setminus \big(\cup_{i=1}^{k-1} F_i^n\big).
\end{aligned}$$

Here, F_1^n is closed, hence a Borel set. Since the countable union of Borel sets is Borel and the difference of two Borel sets is a Borel, then each $F_j^n, j = 2,\ldots,k$ is a Borel set. Hence, $(F_j^n)_{1\leqslant j\leqslant k}$ is a family of mutually disjoint, Borel sets such that $\cup_{j=1}^k F_j^n = W_0^{1,\mathcal{H}}(\Omega)$. While, denote

$$B_k^n := u^{-1}(F_k^n), \quad D_1^n := B_1^n, \quad D_k^n := B_k^n \setminus (\cup_{i=1}^{k-1} B_i^n) \quad \text{for } k = 2,3,\ldots,n.$$

The sets B_k^n are measurable, since the pre-image of Borel sets are measurable due to measurability of function u. Further, the measurability of the sets D_k^n follows from the measurability of B_k^n. We then define the approximating simple function $u_n : (0,T) \to W_0^{1,\mathcal{H}}(\Omega)$ as

$$u_n(t) := \sum_{k=1}^{n} a_k \chi_{D_k^n}(t).$$

Here, $\{a_k\}$ being a dense set in $u(0,T)$, we have $\min_{1\leqslant k\leqslant n} \|a_k - u(t)\|_{W^{1,\mathcal{H}}(\Omega)} \to 0$, as $n \to \infty$. Since B_k^n is pre-image of F_k^n under u, we get those values of t for which the distance between a_k and $u(t)$ under the $W^{1,\mathcal{H}}$ norm is minimum. We then have

$$\begin{aligned}\|u_n(t) - u(t)\|_{W^{1,\mathcal{H}}(\Omega)} &= \|\sum_{k=1}^{n} a_k \chi_{D_k^n}(t) - u(t)\|_{W^{1,\mathcal{H}}(\Omega)} \\ &= \min_{1\leqslant k\leqslant n} \|a_k - u(t)\|_{W^{1,\mathcal{H}}(\Omega)} \to 0\end{aligned}$$

almost everywhere, as $n \to \infty$.

Now, in order to use Lebesgue's dominated convergence theorem, we modify u_n whenever $\|u_n\|_{W^{1,\mathcal{H}}(\Omega)}$ is large compared to $\|u\|_{W^{1,\mathcal{H}}(\Omega)}$. So, to have an upper bound, we define, for fixed t,

$$v_n(t) := \begin{cases} u_n(t), & \text{if } \|u_n(t)\|_{W^{1,\mathcal{H}}(\Omega)} \leqslant 2\|u(t)\|_{W^{1,\mathcal{H}}(\Omega)} \\ 0, & \text{if } \|u_n(t)\|_{W^{1,\mathcal{H}}(\Omega)} > 2\|u(t)\|_{W^{1,\mathcal{H}}(\Omega)}. \end{cases} \tag{4.2.5}$$

If, for some t, $\|u(t)\|_{W^{1,\mathcal{H}}(\Omega)} = 0$, then $v_n(t) = 0$ a.e., and, while for some t, if $\|u(t)\|_{W^{1,\mathcal{H}}(\Omega)} > 0$, then $v_n(t) = u_n(t)$ a.e., for n large enough, thus $|v_n(t)| \leqslant |u_n(t)|$. Moreover, we can deduce that $v_n(t) \to u(t)$ in $W^{1,\mathcal{H}}(\Omega)$, as $n \to \infty$. Due to the uniform bound $\|v_n(t)\|_{W^{1,\mathcal{H}}(\Omega)} \leqslant 2\|u(t)\|_{W^{1,\mathcal{H}}(\Omega)}$, we can then apply the dominated convergence theorem to obtain

$$\int_0^T \|v_n(t) - u(t)\|^2_{W^{1,\mathcal{H}}(\Omega)}\, dt \to 0, \quad \text{in } L^2(0,T; W_0^{1,\mathcal{H}}(\Omega)),$$

as $n \to \infty$.

Next, consider

$$\check{D}_k^n := D_k^n \setminus \{t \in (0,T) : \|v_n(t)\|_{W^{1,\mathcal{H}}(\Omega)} > 2\|u(t)\|_{W^{1,\mathcal{H}}(\Omega)}\}, \quad k = 1, 2, \ldots, n,$$

where $\check{D}_k^n$ are mutually disjoint. By (4.2.5), we get $v_n(t) = \sum_{k=1}^{n} a_k \chi_{\check{D}_k^n}(t)$.

Since $C_0^\infty(\Omega)$ is dense in $W_0^{1,\mathcal{H}}(\Omega)$, we can choose $d_k \in C_0^\infty(\Omega)$ such that,

$$\|d_k - a_k\|^2_{W^{1,\mathcal{H}}(\Omega)} < \frac{\epsilon}{2T}.$$

Further, since $\check{D}_k^n$ are disjoint sets, we have $\sum_{k=1}^n \chi_{\check{D}_k^n}(t) \leqslant 1$. Also, d_k being constant with respect to t variable, then

$$\int_0^T \| \sum_{k=1}^n d_k \chi_{\check{D}_k^n}(t) - \sum_{k=1}^n a_k \chi_{\check{D}_k^n}(t) \|^2_{W^{1,\mathcal{H}}(\Omega)} \, dt$$
$$= \|d_k - a_k\|^2_{W^{1,\mathcal{H}}(\Omega)} \int_0^T \big| \sum_{k=1}^n \chi_{\check{D}_k^n}(t) \big|^2 \, dt < \frac{\epsilon}{2},$$

that is,

$$\int_0^T \| \sum_{k=1}^n d_k \chi_{\check{D}_n^k}(t) - \sum_{k=1}^n a_k \chi_{\check{D}_k^n}(t) \|^2_{W^{1,\mathcal{H}}(\Omega)} \, dt < \frac{\epsilon}{2}. \tag{4.2.6}$$

Next, to mollify the above in t, consider a mollification function $\delta_n \in C_0^\infty(\mathbb{R})$ such that for each $k = 1, 2, \ldots, n$, we have $\delta_n * (d_k \chi_{\check{D}_k^n}) = d_k(\delta_n * \chi_{\check{D}_k^n}) \in C_0^\infty(\mathbb{R} \times \Omega)$ and choosing

$$\int_0^T |\delta_n * (\chi_{\check{D}_k^n}(t)) - \chi_{\check{D}_k^n}(t)|^2 \, dt < \frac{\epsilon}{2n\|d_k\|^2_{W^{1,\mathcal{H}}(\Omega)}}.$$

Then the following expression implies

$$\int_0^T \|\delta_n * (d_k \chi_{\check{D}_k^n}) - d_k \chi_{\check{D}_k^n}\|^2_{W^{1,\mathcal{H}}(\Omega)} \, dt$$
$$= \int_0^T \|d_k\big(\delta_n * \chi_{\check{D}_k^n}(t) - \chi_{\check{D}_k^n}(t)\big)\|^2_{W^{1,\mathcal{H}}(\Omega)} \, dt$$
$$= \|d_k\|^2_{W^{1,\mathcal{H}}(\Omega)} \int_0^T |\delta_n * (\chi_{\check{D}_k^n}(t)) - \chi_{\check{D}_k^n}(t)|^2 \, dt < \frac{\epsilon}{2n}$$

that is,

$$\int_0^T \|\delta_n * (d_k \chi_{\check{D}_k^n}) - d_k \chi_{\check{D}_k^n}\|^2_{W^{1,\mathcal{H}}(\Omega)} \, dt < \frac{\epsilon}{2n}. \tag{4.2.7}$$

Denote $g := \sum_{k=1}^n \delta_n * (d_k \chi_{\check{D}_k^n})$. Then, we have,

$$\int_0^T \|g - v_n\|^2_{W^{1,\mathcal{H}}(\Omega)} \, dt = \int_0^T \| \sum_{k=1}^n \delta_n * (d_k \chi_{\check{D}_k^n}) - \sum_{k=1}^n a_k \chi_{\check{D}_k^n} \|^2_{W^{1,\mathcal{H}}(\Omega)} \, dt$$
$$\leqslant \int_0^T \| \sum_{k=1}^n \delta_n * (d_k \chi_{\check{D}_k^n}) - \sum_{k=1}^n d_k \chi_{\check{D}_k^n} \|^2_{W^{1,\mathcal{H}}(\Omega)} \, dt$$
$$+ \int_0^T \| \sum_{k=1}^n d_k \chi_{\check{D}_k^n} - \sum_{k=1}^n a_k \chi_{\check{D}_k^n} \|^2_{W^{1,\mathcal{H}}(\Omega)} \, dt,$$

which follows from triangle inequality and further applying (4.2.7) and (4.2.6), to obtain

$$\begin{aligned}\int_0^T \|g - v_n\|^2_{W^{1,\mathcal{H}}(\Omega)}\, dt &\leqslant \sum_{k=1}^n \int_0^T \|\delta_n * (d_k\chi_{\check{D}_k^n}) - d_k\chi_{\check{D}_k^n}\|^2_{W^{1,\mathcal{H}}(\Omega)}\, dt \\ &\quad + \int_0^T \|\sum_{k=1}^n d_k\chi_{\check{D}_k^n} - \sum_{k=1}^n a_k\chi_{\check{D}_k^n}\|^2_{W^{1,\mathcal{H}}(\Omega)}\, dt \\ &< \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon.\end{aligned}$$

Thus, we can conclude that $C_0^\infty(\Omega_T)$ is dense in $L^2(0,T;W_0^{1,\mathcal{H}}(\Omega))$. □

***Lemma* 4.2.6.** *Let $\Omega \subset \mathbb{R}^n$ be a bounded open set and $\mathcal{H} \in \Phi_w(\Omega)$ as defined in* (4.2.1). *For each $t > 0$, assume that $f_j(\cdot,t), g_j(\cdot,t) \in L^{\mathcal{H}}(\mathbb{R}^n)$ for $j = 1,2,\ldots$, with $\int_0^s \varrho_{\mathcal{H}}(f_j)\,dt$ bounded, where $s > 0$. If $\int_0^s \varrho_{\mathcal{H}}(f_j - g_j)\,dt \to 0$ as $j \to \infty$, then*

$$\Big|\int_0^s \varrho_{\mathcal{H}}(f_j)\,dt - \int_0^s \varrho_{\mathcal{H}}(g_j)\,dt\Big| \to 0 \quad \text{as } j \to \infty.$$

Proof. Since $\mathcal{H}$ is increasing and satisfies (Dec) property, we obtain,

$$\begin{aligned}\mathcal{H}(x,|g_j|) &\leqslant \mathcal{H}(x,|g_j - f_j| + |f_j|) \leqslant \mathcal{H}(x,2|g_j - f_j|) + \mathcal{H}(x,2|f_j|) \\ &\leqslant 2^q\mathcal{H}(x,|g_j - f_j|) + 2^q\mathcal{H}(x,|f_j|).\end{aligned}$$

Integrating both sides with respect to x over Ω and with respect to t over $[0,s]$, $s > 0$, to have

$$\int_0^s \varrho_{\mathcal{H}}(g_j)\,dt \leqslant 2^q \int_0^s \varrho_{\mathcal{H}}(g_j - f_j)\,dt + 2^q \int_0^s \varrho_{\mathcal{H}}(f_j)\,dt.$$

This implies that $\int_0^s \varrho_{\mathcal{H}}(g_j)\,dt$ is bounded. Then choosing $c > 0$ such that

$$\int_0^s \varrho_{\mathcal{H}}(f_j)\,dt \leqslant c \quad \text{and} \quad \int_0^s \varrho_{\mathcal{H}}(g_j)\,dt \leqslant c.$$

Let $\lambda > 0$ and note that $|f_j| \leqslant |f_j - g_j| + |g_j|$.

If $|f_j - g_j| \leqslant \lambda|g_j|$, then by (Dec), we have

$$\mathcal{H}(x,|f_j|) \leqslant \mathcal{H}(x,(1+\lambda)|g_j|) \leqslant (1+\lambda)^q\mathcal{H}(x,|g_j|). \tag{4.2.8}$$

If $|f_j - g_j| > \lambda|g_j|$, then we estimate by (Dec),

$$\mathcal{H}(x,|f_j|) \leqslant \mathcal{H}(x,(1+\tfrac{1}{\lambda})|f_j - g_j|) \leqslant \big(1+\tfrac{1}{\lambda}\big)^q\mathcal{H}(x,|f_j - g_j|). \tag{4.2.9}$$

Combining both the cases above and integrating over $x \in \Omega$, we find that,

$$\varrho_{\mathcal{H}}(f_j) - \varrho_{\mathcal{H}}(g_j) \leqslant \big(1+\tfrac{1}{\lambda}\big)^q\varrho_{\mathcal{H}}(f_j - g_j) + (1+\lambda)^q\varrho_{\mathcal{H}}(g_j) - \varrho_{\mathcal{H}}(g_j).$$

Further, integrating with respect to t over $[0, s], s > 0$, gives

$$\int_0^s (\varrho_{\mathcal{H}}(f_j) - \varrho_{\mathcal{H}}(g_j))\,dt \\ \leqslant \big(1 + \tfrac{1}{\lambda}\big)^q \int_0^s \varrho_{\mathcal{H}}(f_j - g_j)\,dt + ((1+\lambda)^q - 1) \int_0^s \varrho_{\mathcal{H}}(g_j)\,dt.$$

Swapping f_j and g_j in both (4.2.8) and (4.2.9) gives a similar inequality, and combining the inequalities, we find that,

$$\Big|\int_0^s (\varrho_{\mathcal{H}}(f_j) - \varrho_{\mathcal{H}}(g_j))\,dt\Big| \leqslant \big(1 + \tfrac{1}{\lambda}\big)^q \int_0^s \varrho_{\mathcal{H}}(f_j - g_j)\,dt \\ + ((1+\lambda)^q - 1)\Big(\int_0^s \varrho_{\mathcal{H}}(f_j)\,dt + \int_0^s \varrho_{\mathcal{H}}(g_j)\,dt\Big).$$

Let $\epsilon > 0$. Since $\int_0^s \varrho_{\mathcal{H}}(f_j)\,dt + \int_0^s \varrho_{\mathcal{H}}(g_j)\,dt \leqslant 2c$, we can choose λ so small that,

$$((1+\lambda)^q - 1)\Big(\int_0^s \varrho_{\mathcal{H}}(f_j)\,dt + \int_0^s \varrho_{\mathcal{H}}(g_j)\,dt\Big) \leqslant \tfrac{\epsilon}{2},$$

when $j \geqslant j_0$, and it follows that,

$$\Big|\int_0^s \varrho_{\mathcal{H}}(f_j)\,dt - \int_0^s \varrho_{\mathcal{H}}(g_j)\,dt\Big| \leqslant \epsilon.$$

□

4.3 Minimization of image restoration model

Let us consider any image f as a scalar function defined on a bounded and piecewise smooth open set Ω of $\mathbb{R}^N$ -typically a rectangle in $\mathbb{R}^2$. Most commonly, image restoration models (refer [2; 33]) consist of an original image u in the domain Ω, describing a real scene, and the observed image f of the same scene, which is a degradation of u resulting due to the presence of additive Gaussian noise, denoted as η, with mean zero and variance σ^2. Thus the process of image restoration is modeled as

$$f = Au + \eta \quad \text{in } \Omega, \tag{4.3.1}$$

where A is a linear operator representing the blur (usually a convolution). The objective is to recover u, with known f and some statistics of η. Taking A as an identity operator, we assume that the model of degradation (4.3.1) is valid.

Mostly, restoring image u from (4.3.1) is an ill-condition problem, and prior information on the underlying image is required to find an acceptable solution. For this purpose, many variational methods are developed, the most popular one being the TV-based (ROF) model proposed by Rudin, Osher and Fatemi [56] in the BV space as follows:

$$\min_{u \in BV(\Omega)} \left\{ \int_\Omega |Du|^p + \frac{\lambda}{2} \int_\Omega (u-f)^2 \, dx \right\}, \tag{4.3.2}$$

where $p = 1$ and Du is the BV-gradient. Based on the above TV-based diffusion, an adaptive total variation model:

$$\min_{u \in BV(\Omega)} \int_\Omega a(x)|Du| \, dx,$$

was proposed by Chan and Strong [58]. Here, a control factor $a(x)$ slows the diffusion at likely edges, thus controlling the speed of the diffusion and reconstructing edges.

Now, we consider the types of diffusion arising from the minimization problem (4.3.2) with different values of p. At $p = 2$, it results in isotropic diffusion which smooths the 'staircasing effect' caused by the TV-based diffusion at $p = 1$, but not able to preserve edges. While different values of $1 < p < 2$ result in anisotropic diffusion, originally presented by Perona and Malik [52], which lies between TV-based and isotropic smoothing. Such type of anisotropic diffusion proved to be more effective in reconstructing piecewise smooth regions with an edge preserving inhomogenous technique.

Motivated by the above models, using (4.2.1), we define the following functional

$$E(u) := \mathcal{F}(u) + \frac{\lambda}{2}\|u-f\|^2 = \int_\Omega \mathcal{H}(x, |\nabla u|) + \frac{\lambda}{2}(u-f)^2 \, dx =: \int_\Omega F(u) \, dx, \tag{4.3.3}$$

where the double-phase functional $\mathcal{H}(x,r) := r^p + a(x)r^q$, $1 < p \leqslant q \leqslant 2$ would serve the purpose of anisotropic diffusion along with isotropic smoothing in the model. The function $0 \leqslant a \in L^\infty(\Omega)$ is assumed to be Lipschitz continuous, which plays as the trade-off between the two types of diffusion. We then propose the following image restoration model:

$$\min_{u \in W^{1,\mathcal{H}}(\Omega) \cap L^2(\Omega)} E(u) = \min_{u \in W^{1,\mathcal{H}}(\Omega) \cap L^2(\Omega)} \int_\Omega \mathcal{H}(x, |\nabla u|) + \frac{\lambda}{2}(u-f)^2 \, dx. \tag{4.3.4}$$

The following result shows the convexity of the functional E.

***Lemma* 4.3.1.** *Let $\Omega \subset \mathbb{R}^n$ and $\mathcal{H} : \Omega \times [0,\infty) \to [0,\infty)$ as defined in* (4.2.1). *Then, $u \mapsto E(u)$ is strictly convex, where $E(u) = \int_\Omega \mathcal{H}(x, |\nabla u|) + \frac{\lambda}{2}(u-f)^2 \, dx$ from* (4.3.3).

Proof. For $u, v \in W^{1,\mathcal{H}}(\Omega) \cap L^2(\Omega)$, $u \neq v$, and $\alpha \in (0,1)$, we have

$$\begin{aligned} &E(\alpha u + (1-\alpha)v) \\ &= \int_\Omega |\nabla(\alpha u + (1-\alpha)v)|^p + a(x)|\nabla(\alpha u + (1-\alpha)v)|^q \\ &\qquad + \frac{\lambda}{2}(\alpha u + (1-\alpha)v - f)^2\, dx \\ &\leqslant \int_\Omega (\alpha|\nabla u| + (1-\alpha)|\nabla v|)^p + a(x)(\alpha|\nabla u| + (1-\alpha)|\nabla v|)^q \\ &\qquad + \frac{\lambda}{2}(\alpha u + (1-\alpha)v - \alpha f - (1-\alpha)f)^2\, dx. \end{aligned}$$

Then using convexity of $r \mapsto r^p, r \mapsto r^q$ and $r \mapsto r^2$, which are strictly convex for $p, q > 1$, we obtain

$$\begin{aligned} &E(\alpha u + (1-\alpha)v) \\ &< \int_\Omega (\alpha|\nabla u|^p + (1-\alpha)|\nabla u|^p) + a(x)(\alpha|\nabla u|^q + (1-\alpha)|\nabla u|^q) \\ &\qquad + \frac{\lambda}{2}(\alpha(u-f)^2 + (1-\alpha)(v-f)^2)\, dx, \end{aligned}$$

that is,

$$E(\alpha u + (1-\alpha)v) < \alpha E(u) + (1-\alpha)E(v), \tag{4.3.5}$$

which proves the strict convexity of the functional E. □

4.3.1 Existence of solution of the minimization problem.

In this section, we discuss the existence and uniqueness of the minimizer to the minimization problem (4.3.4). The approach of the proofs presented here is motivated by those in [23; 39].

Theorem 4.3.2. *Let $\Omega \subset \mathbb{R}^n$ be a bounded open set and $f \in L^2(\Omega)$. The minimization problem*

$$\min_{u \in W^{1,\mathcal{H}}(\Omega) \cap L^2(\Omega)} E(u) \tag{4.3.6}$$

has a unique minimizer $u \in W^{1,\mathcal{H}}(\Omega) \cap L^2(\Omega)$, where $E(u) = \int_\Omega \mathcal{H}(x, |\nabla u|) + \frac{\lambda}{2}(u-f)^2\, dx$, with $\mathcal{H}(x,r) = r^p + a(x)r^q, 1 < p \leqslant q \leqslant 2$ and $0 \leqslant a \in L^\infty(\Omega)$.

Proof. Set $m := \inf_{u \in W^{1,\mathcal{H}}(\Omega) \cap L^2(\Omega)} E(u)$, which is finite, since

$$0 \leqslant m = \inf_u E(u) \leqslant E(0) = \frac{\lambda}{2}\|f\|_2^2 < \infty.$$

The definition of infimum then implies that there exists a minimizing sequence, denoted by $\{u_k\}_{k=1}^\infty$, $u_k \in W^{1,\mathcal{H}}(\Omega) \cap L^2(\Omega)$ such that

$$\inf_u E(u) = \lim_{k\to\infty} E(u_k) = m < \infty.$$

The existence of finite limit implies that the sequence $E(u_k)$ is bounded. Thus,

$$E(u_k) = \int_\Omega \mathcal{H}(x, |\nabla u_k|) + \frac{\lambda}{2}|u_k - f|^2 \, dx \leqslant C,$$

where C denotes a universal strictly positive constant, and so

$$\int_\Omega \mathcal{H}(x, |\nabla u_k|) \, dx \leqslant C \quad \text{and} \quad \int_\Omega (u_k - f)^2 \, dx \leqslant C,$$

for every $k = 1, 2, \ldots$.

Now, by triangle inequality, we have

$$\|u_k\|_2 \leqslant \|f\|_2 + \|u_k - f\|_2 \leqslant C$$

which results in uniform boundedness of $\{u_k\}_{k=1}^\infty$ in $L^2(\Omega)$, that is $\int_\Omega |u_k|^2 \, dx \leqslant C$. Moreover $L^2(\Omega) \hookrightarrow L^{\mathcal{H}}(\Omega)$ [4, Lemma 2.5], which then gives $\int_\Omega \mathcal{H}(x, |u_k|) \, dx \leqslant C$. Thus we get,

$$\varrho_{\mathcal{H}}(u_k) + \varrho_{\mathcal{H}}(|\nabla u_k|) \leqslant C$$

which implies that $\{u_k\}_{k=1}^\infty$ is a bounded sequence in $W^{1,\mathcal{H}}(\Omega)$. Since $W^{1,\mathcal{H}}(\Omega) \cap L^2(\Omega)$ is a reflexive Banach space, there exists a subsequence $\{u_{k_j}\}_{j=1}^\infty \subset \{u_k\}_{k=1}^\infty$ converging weakly to a function u in $W^{1,\mathcal{H}}(\Omega) \cap L^2(\Omega)$. Then Lemma 4.2.3 implies

$$\liminf_{j \to \infty} \int_\Omega \mathcal{H}(x, |\nabla u_{k_j}|) \, dx \geqslant \int_\Omega \mathcal{H}(x, |\nabla u|) \, dx. \tag{4.3.7}$$

Moreover, by the weak lower semicontinuity of the L^2-norm,

$$\liminf_{j \to \infty} \int_\Omega |u_{k_j} - f|^2 \, dx \geqslant \int_\Omega |u - f|^2 \, dx. \tag{4.3.8}$$

Thus, from (4.3.7) and (4.3.8), we can conclude that

$$E(u) \leqslant \liminf_{j \to \infty} E(u_{k_j}) = \inf_u E(u) = m, \tag{4.3.9}$$

and hence u is a minimizer of the problem (4.3.6).

To prove uniqueness of the minimizer, suppose that $u, \tilde{u} \in W^{1,\mathcal{H}}(\Omega) \cap L^2(\Omega)$ are two solutions of the minimization problem (4.3.6). Then $v := \frac{u+\tilde{u}}{2} \in W^{1,\mathcal{H}}(\Omega) \cap L^2(\Omega)$. We claim here that

$$E(v) \leqslant \frac{E(u) + E(\tilde{u})}{2},$$

with a strict inequality, unless $u = \tilde{u}$ a.e.

By Lemma 4.3.1, the functional $E(u)$ is strictly convex. Then, taking $\alpha = \frac{1}{2}$ in (4.3.5), we get $E(v) < \frac{E(u)+E(\tilde{u})}{2} = m = \inf_v E(v)$, which is a contradiction, thus $u = \tilde{u}$ a.e. in Ω. Hence the minimizer of $E(u)$ is unique. □

Theorem 4.3.3. *Let $\Omega \subset \mathbb{R}^n$ be a bounded open set with boundary function $f \in W^{1,\mathcal{H}}(\Omega) \cap L^\infty(\Omega)$. Assume that $C^\infty(\Omega) \cap W^{1,\mathcal{H}}(\Omega)$ is dense in $W^{1,\mathcal{H}}(\Omega)$. Then, for energy functional $E(u) = \int_\Omega \mathcal{H}(x, |\nabla u|) + \frac{\lambda}{2}(u - f)^2\, dx$, where $\mathcal{H}(x, r) = r^p + a(x)r^q, 1 < p \leqslant q \leqslant 2$ as defined in* (4.2.1)*, the minimization problem:*

$$\min_{u \in W^{1,\mathcal{H}}(\Omega) \cap L^2(\Omega), u - f \in W_0^{1,\mathcal{H}}(\Omega)} E(u) \tag{4.3.10}$$

has a unique minimizer $u \in W^{1,\mathcal{H}}(\Omega) \cap L^\infty(\Omega)$, which satisfies $u - f \in W_0^{1,\mathcal{H}}(\Omega) \cap L^\infty(\Omega)$.

Proof. Let U denote the set $\{u : u - f \in W_0^{1,\mathcal{H}}(\Omega)\}$, and $M > 0$ be such that $|f| \leqslant M$ a.e. Let u_M be the function $u \in U$ which has been cut-off at $-M$ and M, that is

$$u_M := \min\{M, \max\{-M, u\}\}.$$

This implies that $u_M - f \in W_0^{1,\mathcal{H}}(\Omega)$ (see the notes in A.1 in the appendix). Moreover, from the lattice property of $W^{1,\mathcal{H}}(\Omega)$ (which follows from [24, Lemma 6.1.6] and [24, Lemma 6.1.7]), we have that

$$\nabla u_M = \begin{cases} \nabla u, & \text{for } |u| \leqslant M, \\ 0, & \text{otherwise} \end{cases}$$

which holds almost everywhere. Hence, $|\nabla u_M| \leqslant |\nabla u|$ a.e., and by increasing property of $\mathcal{H}$, we get $\mathcal{H}(x, |\nabla u_M|) \leqslant \mathcal{H}(x, |\nabla u|)$. For $\{x \in \Omega : |u(x)| \leqslant M\}$, we have $|u_M - f| = |u - f|$, and, for $\{x \in \Omega : u(x) \geqslant M\}$, $|u_M - f| = |M - f| \leqslant |u - f|$ holds. While, for $\{x \in \Omega : u(x) < -M\}$, $|u_M - f| = M + f < -u + f \leqslant |u - f|$ holds. This implies $\|u_M - f\|_{L^2} \leqslant \|u - f\|_{L^2}$. So, we get $E(u_M) \leqslant E(u)$. Thus, we conclude that the possible minimizer satisfies $|u| \leqslant M$ a.e., and belongs to the set $U_M := \{u_M : u \in U\}$.

Since $E(u) \geqslant 0$ for every $u \in U_M$ and $f \in U$, we note that, on taking the same f for an upper bound of the functional E, results in a finite limit of the infimum of E, that is,

$$0 \leqslant m := \inf_{u - f \in W_0^{1,\mathcal{H}}(\Omega)} E(u) \leqslant E(f) = \varrho_{\mathcal{H}}(|\nabla f|) < \infty.$$

Thus, there exists a minimizing sequence $\{u_k\}_{k=1}^\infty \subset U_M$ with $u_k - f \in W_0^{1,\mathcal{H}}(\Omega)$ such that

$$\inf_{u \in U_M} E(u) = \lim_{k \to \infty} E(u_k) = m.$$

This implies that the sequence $E(u_k)$ is bounded, that is

$$E(u_k) = \int_\Omega \mathcal{H}(x, |\nabla u_k|) + \frac{\lambda}{2}(u_k - f)^2\, dx \leqslant C$$

where C denotes a universal strictly positive constant, and so

$$\int_\Omega \mathcal{H}(x, |\nabla u_k|)\, dx \leqslant C, \quad \int_\Omega (u_k - f)^2\, dx \leqslant C.$$

for every $k = 1, 2, \ldots$.

Applying triangle inequality to get

$$\|u_k\|_2 \leqslant \|f\|_2 + \|u_k - f\|_2 \leqslant C,$$

which implies $\{u_k\}_{k=1}^\infty$ is bounded in $L^2(\Omega)$, that is $\int_\Omega |u_k|^2\, dx \leqslant C$, and further since $L^2(\Omega) \hookrightarrow L^{\mathcal{H}}(\Omega)$ by [4, Lemma 2.5], we can conclude that

$$\int_\Omega \mathcal{H}(x, |u_k|)\, dx \leqslant C.$$

Thus, we get

$$\varrho_{\mathcal{H}}(u_k) + \varrho_{\mathcal{H}}(|\nabla u_k|) \leqslant C.$$

Now, the sequence $\{u_k\}_{k=1}^\infty$ being bounded in $W^{1,\mathcal{H}}(\Omega)$ and $L^2(\Omega)$, so similarly using reflexivity, we have a subsequence $\{u_{k_j}\}_{j=1}^\infty \subset \{u_k\}_{k=1}^\infty$, and $u \in W^{1,\mathcal{H}}(\Omega) \cap L^2(\Omega)$ such that

$$u_{k_j} \rightharpoonup u \quad \text{in } W^{1,\mathcal{H}}(\Omega) \cap L^2(\Omega).$$

Moreover, since $\{u_k\}_{k=1}^\infty \subset U_M$, hence $u \in W^{1,\mathcal{H}}(\Omega) \cap L^\infty(\Omega)$. Also note that for $u_k - f \in W_0^{1,\mathcal{H}}(\Omega) \cap L^\infty(\Omega)$, since $W_0^{1,\mathcal{H}}(\Omega) \cap L^\infty(\Omega)$ is a closed subspace of $W^{1,\mathcal{H}}(\Omega) \cap L^\infty(\Omega)$ and $u_k \rightharpoonup u$, then we have $u - f \in W_0^{1,\mathcal{H}}(\Omega) \cap L^\infty(\Omega)$ which is weakly closed. Thereafter, following the same argument in the previous proof to obtain weak lower semicontinuity as in (4.3.9),which implies

$$E(u) \leqslant \liminf_{j\to\infty} E(u_{k_j}) = \inf_u E(u) = m,$$

we can then conclude that u is a minimizer of the minimization problem (4.3.10). The uniqueness of minimizer u follows similarly as proved in the previous result. □

4.4 The associated boundary value problems

In this section, we show that the minimizer of the problem (4.3.4) can be equivalently expressed as solution of a boundary value problem obtained from the minimization functional (4.3.3).

Assume that $f : \overline{\Omega} \to \mathbb{R}$ is continuous, and let

$$K := \{v \in C^2(\overline{\Omega}), \text{ such that } v(x) = f(x) \text{ for } x \in \partial\Omega\},$$

be a non-empty set for the energy functional (4.3.3), which we recall, is of the form,

$$E(v) = \int_\Omega \mathcal{H}(x, |\nabla v|) + \frac{\lambda}{2}(v - f)^2\, dx. \tag{4.4.1}$$

Next, we compute the functional derivative of E with respect to $v \in K$ which generalizes the 'gradient' notion for functions. One version of functional derivative is the Gâteaux derivative for functional E at v in the direction of $h \in C_0^\infty(\Omega)$, denoted by $d(E;h)$, as defined below

$$dE(u;h) := \lim_{\epsilon \to 0} \frac{E(u+\epsilon h) - E(u)}{\epsilon} = \frac{d}{d\epsilon} E(u+\epsilon h) \mid_{\epsilon=0} . \tag{4.4.2}$$

Here, we have $E(u+\varepsilon h) = \int_\Omega \mathcal{H}(x, |\nabla(u+\varepsilon h)|) + \frac{\lambda}{2}(u+\varepsilon h - f)^2\, dx$, where the integrand and its derivative with respect to ε are continuous over Ω. Using Leibniz rule to interchange the order of integration and differentiation, we calculate the derivative of $E(u+\varepsilon w)$ with respect to ε to obtain as follows,

$$\begin{aligned} &\frac{d}{d\varepsilon} E(u+\varepsilon h) \\ &= \int_\Omega (p|\nabla(u+\varepsilon h)|^{p-1} + a(x)\, q|\nabla(u+\varepsilon h)|^{q-1}) \frac{\nabla(u+\varepsilon h)}{|\nabla(u+\varepsilon h)|} \cdot \nabla h \\ &\qquad\qquad\qquad\qquad + \lambda(u+\varepsilon h - f)h\, dx. \end{aligned} \tag{4.4.3}$$

Then (4.4.2) implies,

$$\frac{d}{d\epsilon} E(u+\epsilon h) \mid_{\epsilon=0} = \int_\Omega p|\nabla u|^{p-1} \frac{\nabla u}{|\nabla u|} \cdot \nabla h + a(x)\, q|\nabla u|^{q-1} \frac{\nabla u}{|\nabla u|} \cdot \nabla h + \lambda(u-f)h\, dx,$$

that is,

$$dE(u;h) = \int_\Omega \mathcal{H}_r(x, |\nabla u|) \frac{\nabla u}{|\nabla u|} \cdot \nabla h + \lambda(u-f)h\, dx, \quad h \in C_0^\infty(\Omega). \tag{4.4.4}$$

where $\mathcal{H}_r(x, |\nabla v|) = p|\nabla v|^{p-1} + a(x)\, q|\nabla v|^{q-1}$. Note that when $\nabla u = 0$ the derivative in (4.4.3) is well-defined, since the total degree of ∇u is $p - 1 > 0$.

***Proposition* 4.4.1.** *Let $\Omega \subset \mathbb{R}^n$ be a bounded open set and function $f : \overline{\Omega} \to \mathbb{R}$ be continuous. Consider the following partial differential equation with Dirichlet boundary condition:*

$$-\operatorname{div}\left(\mathcal{H}_r(x, |\nabla u|) \frac{\nabla u}{|\nabla u|}\right) + \lambda(u-f) = 0 \quad \text{in } \Omega, \tag{4.4.5}$$

$$u = f \quad \text{on } \partial\Omega, \tag{4.4.6}$$

where $\mathcal{H}_r(x,r)$ denotes the derivative of $\mathcal{H}(x,r)$ with respect to r. Then the function $u \in K$ is a solution of (4.4.5)–(4.4.6) *if and only if $E(u) = \min_{v \in K} E(v)$, where $E(v)$ is defined as in* (4.4.1).

Proof. Suppose that $u \in K$ is a solution of (4.4.5)–(4.4.6). Let $v \in K$ be arbitrary. We aim to show that $E(u) \leqslant E(v)$ for all $v \in K$, that is, $E(u) = \min_{v \in K} E(v)$.

Now multiplying (4.4.5) by $u-v$ and integrating over Ω, we have

$$-\int_\Omega \operatorname{div}\big(\mathcal{H}_r(x,|\nabla u|)\frac{\nabla u}{|\nabla u|}\big)(u-v)\,dx+\int_\Omega \lambda(u-f)(u-v)\,dx=0,$$

where the derivative $\mathcal{H}_r(x,|\nabla u|)=p|\nabla u|^{p-1}+a(x)\,q|\nabla u|^{q-1}$. Since $u,v\in K$, so $u-v=0$ at the boundary $\partial\Omega$. Then using integration by parts formula for the first integral in the previous expression implies

$$\int_\Omega \mathcal{H}_r(x,|\nabla u|)\frac{\nabla u}{|\nabla u|}\cdot\nabla(u-v)\,dx+\int_\Omega \lambda(u-f)(u-v)\,dx=0.$$

Further rearranging gives

$$\begin{aligned}\int_\Omega \mathcal{H}_r(x,|\nabla u|)\frac{\nabla u}{|\nabla u|}\cdot\nabla(u-v)\,dx&=-\int_\Omega \lambda(u-f)(u-f+f-v)\,dx\\&=-\int_\Omega \lambda(u-f)^2\,dx+\int_\Omega \lambda(u-f)(v-f)\,dx.\end{aligned}$$

Then using the relation $ab\leqslant\frac{1}{2}(a^2+b^2), a,b\in\mathbb{R}$, for the second integral on the right hand side above, we obtain

$$\begin{aligned}\int_\Omega \mathcal{H}_r(x,|\nabla u|)\frac{\nabla u}{|\nabla u|}\cdot\nabla(u-v)\,dx+\int_\Omega \lambda(u-f)^2\,dx\\\leqslant\int_\Omega \frac{\lambda}{2}\Big((u-f)^2+(v-f)^2\Big)\,dx,\end{aligned}$$

and thus we have,

$$\begin{aligned}\int_\Omega \Big(\frac{p}{|\nabla u|^{2-p}}+a(x)\frac{q}{|\nabla u|^{2-q}}\Big)\nabla u\cdot\nabla u\,dx+\int_\Omega \frac{\lambda}{2}(u-f)^2\,dx\\\leqslant\int_\Omega \Big(\frac{p}{|\nabla u|^{2-p}}+a(x)\frac{q}{|\nabla u|^{2-q}}\Big)\nabla u\cdot\nabla v\,dx+\int_\Omega \frac{\lambda}{2}(v-f)^2\,dx.\end{aligned}$$

Now taking the absolute value relation $\nabla u\cdot\nabla v\leqslant|\nabla u\cdot\nabla v|\leqslant|\nabla u||\nabla v|$ in the first integral expression on the right hand side, the above inequality becomes

$$\begin{aligned}\int_\Omega (p|\nabla u|^p+a(x)q|\nabla u|^q)\,dx+\int_\Omega \frac{\lambda}{2}(u-f)^2\,dx\\\leqslant\int_\Omega (p|\nabla u|^{p-1}|\nabla v|+a(x)q|\nabla u|^{q-1}|\nabla v|)\,dx+\int_\Omega \frac{\lambda}{2}(v-f)^2\,dx,\end{aligned}$$

and, thereafter, applying Young's inequality (2.1.7) on each term of the first integral expression on the right hand side above, we obtain the following,

$$\begin{aligned}\int_\Omega (p|\nabla u|^p+a(x)q|\nabla u|^q)\,dx+\int_\Omega \frac{\lambda}{2}(u-f)^2\,dx\\\leqslant\int_\Omega p\left(\tfrac{1}{p}|\nabla v|^p+\left(1-\tfrac{1}{p}\right)|\nabla u|^p\right)+a(x)q\left(\tfrac{1}{q}|\nabla v|^q+\left(1-\tfrac{1}{q}\right)|\nabla u|^q\right)\\+\frac{\lambda}{2}(v-f)^2\,dx,\end{aligned}$$

and therefore, we have

$$\int_\Omega |\nabla u|^p + a(x)|\nabla u|^q + \frac{\lambda}{2}(u-f)^2\,dx \leqslant \int_\Omega |\nabla v|^p + a(x)|\nabla v|^q + \frac{\lambda}{2}(v-f)^2\,dx,$$

which gives,

$$E(u) \leqslant E(v).$$

Since v is any arbitrary function in K, we get $E(u) = \min_{v\in K} E(v)$. Therefore, u is a minimizer of E on K.

Conversely, suppose that $u \in K$ minimizes E. We aim to show that u is a solution of the boundary value problem (4.4.5)–(4.4.6). Let $h \in C_0^\infty(\Omega)$, and for any real number ε, we have $u + \varepsilon h \in K$. Now, consider the functional $\tilde{E} : \mathbb{R} \to \mathbb{R}$, defined as,

$$\tilde{E}(\varepsilon) := E(u + \varepsilon h),$$

such that the function $u + \varepsilon h$ is admissible for the minimization problem. By assumption, u is a minimizer of E, thus the minimizer of $\tilde{E}$ must exist at $\varepsilon = 0$. This implies that the derivative of $\tilde{E}$ vanishes at $\varepsilon = 0$, that is,

$$\frac{d}{d\varepsilon}\tilde{E}(\varepsilon) = \frac{d}{d\varepsilon}E(u + \varepsilon h) = 0 \quad \text{at } \varepsilon = 0. \tag{4.4.7}$$

Using (4.4.4), we have from (4.4.7),

$$\int_\Omega \mathcal{H}_r(x, |\nabla u|)\frac{\nabla u}{|\nabla u|} \cdot \nabla h + \lambda(u-f)h\,dx = 0, \quad h \in C_0^\infty(\Omega), \tag{4.4.8}$$

which indicates that the solution u is weak. As h vanishes on $\partial\Omega$, then using integrating by parts formula for the first integral on the left hand side above, we get

$$\int_\Omega \Big(-\operatorname{div}\Big(\mathcal{H}_r(x, |\nabla u|)\frac{\nabla u}{|\nabla u|}\Big) + \lambda(u-f)\Big)\,h\,dx = 0. \tag{4.4.9}$$

Since h can be chosen arbitrarily, then applying the Fundamental Lemma of Calculus of Variations 2.1.2 for the above equation (4.4.9), we obtain,

$$-\operatorname{div}\Big(\mathcal{H}_r(x, |\nabla u|)\frac{\nabla u}{|\nabla u|}\Big) + \lambda(u-f) = 0, \tag{4.4.10}$$

almost everywhere in Ω. Thus, we conclude that every minimizer of E on K solves (4.4.5)–(4.4.6). □

Next we define weak solution for the above Dirichlet problem (4.4.5)–(4.4.6). Note that, it suffices to require $u \in W^{1,2}(\Omega)$ satisfying (4.4.8), for any $h \in C_0^\infty(\Omega)$, as a solution of (4.4.5) in a weak sense. Additionally, in order to satisfy the boundary value condition (4.4.6), we consider functions in $W_0^{1,2}(\Omega)$ taking zero boundary values in a general sense, so we have $u - f \in W_0^{1,2}(\Omega)$ (refer [15, Section 5.5]).

Thus, a function $u \in W^{1,\mathcal{H}}(\Omega)$ is called weak solution of the Dirichlet problem (4.4.5)–(4.4.6) if the integral identity (4.4.8) holds for every $h \in C_0^\infty(\Omega)$, provided $u - f \in W_0^{1,\mathcal{H}}(\Omega)$. The equation obtained in (4.4.10) is known as the Euler equation or Euler-Lagrange equation, corresponding to the functional E in (4.4.1).

On the other hand, for solving minimization problems where no Dirichlet boundary condition is imposed in the space of admissible functions, as in Theorem 4.3.2, the Neumann boundary problem is considered. The Neumann boundary condition, in general, emerges as a natural consequence of the weak formulation of the problem. In this case, the weak solution formulation is obtained in a similar manner as in the Dirichlet case.

Now, in order to obtain the minimizer of the energy functional (4.4.1), we compute the Euler-Lagrange equation with:

$$dE(u;h) = 0 \text{ in } \Omega. \tag{4.4.11}$$

Then using (4.4.4) we obtain the Gâteaux derivative of E as,

$$dE(u;h) = \int_\Omega \mathcal{H}_r(x, |\nabla u|)\frac{\nabla u}{|\nabla u|} \cdot \nabla h + \lambda(u-f)h\,dx.$$

Further applying the Green's first identity (2.1.5), yields

$$\begin{aligned} dE(u;h) = \int_\Omega \Big(-\operatorname{div}\big(\mathcal{H}_r(x, |\nabla u|)\frac{\nabla u}{|\nabla u|}\big) + \lambda(u-f)\Big)h\,dx \\ + \int_{\partial\Omega} \big(\mathcal{H}_r(x, |\nabla u|)\frac{\nabla u}{|\nabla u|} \cdot N\big)h\,dS \end{aligned}$$

where N is the outer unit normal vector for $x \in \partial\Omega$ and dS denotes the surface measure on $\partial\Omega$. Then, from (4.4.11) and since $h \neq 0$, we obtain the Euler-Lagrange equation,

$$-\operatorname{div}\big(\mathcal{H}_r(x, |\nabla u|)\frac{\nabla u}{|\nabla u|}\big) + \lambda(u-f) = 0, \quad \text{in } \Omega,$$

and a Neumann natural boundary condition,

$$\frac{\partial u}{\partial N} = 0, \quad \text{on } \partial\Omega.$$

Thus, the minimizer of $E(u)$ is obtained by solving the boundary value problem,

$$\begin{cases} -\operatorname{div}\big(\mathcal{H}_r(x, |\nabla u|)\frac{\nabla u}{|\nabla u|}\big) + \lambda(u-f) = 0 & \text{in } \Omega \\ \nabla u \cdot N = 0 & \text{on } \partial\Omega. \end{cases}$$

Note here that, the Neumann boundary condition does not originate from restrictions on the function space where minimization of the functional takes place, but as part of the derivative condition $dE(u) = 0$. Thus, the admissible variations h, in this case, are free to vary on the boundary. Due to such characteristics, boundary problems with Neumann condition are more widely used for numerical purposes, in compared to the Dirichlet boundary condition.

4.4.1 Heat flow and the associated weak solution. In [56], Rudin et al. proposed the use of artificial time to solve Euler-Lagrange equation, which is equivalent to the steepest descent of the energy functional. For such approach, we first consider the image as a function of space and time, say $u \in L^2(0,T;C^2(\Omega))$, having boundary function $f : \overline{\Omega} \to \mathbb{R}$, both continuous up to the boundary of $\Omega_T := \Omega \times (0,T)$ for any $T \in (0,\infty)$. The associated L^2-gradient flow [15] of functional $E(u)$, defined in (4.4.1), is given by $u_t = -\partial E(u)$, where u_t is the partial derivative of $u(x,t)$ with respect to t, and $\partial E(u)$ denotes the functional derivative of E to u which generalizes the 'gradient' notion for functions. One version of functional derivative is the Gâteaux derivative, as in (4.4.2) for functional E where the boundary term vanishes, expressed as follows

$$
\begin{aligned}
dE(u;h) &= \int_\Omega \Big(-\operatorname{div}\big(\mathcal{H}_r(x,|\nabla u|)\frac{\nabla u}{|\nabla u|}\big) + \lambda(u-f)\Big) h\,dx \\
&= \Big\langle -\operatorname{div}\Big(\mathcal{H}_r(x,|\nabla u|)\frac{\nabla u}{|\nabla u|}\Big) + \lambda(u-f), h\Big\rangle_2
\end{aligned}
$$

where $\mathcal{H}_r(x,r)$ is the derivative of $\mathcal{H}(x,r)$ with respect to r and $\langle\cdot,\cdot\rangle$ is the L^2-inner product over Ω. The above can be interpreted as the directional derivative of E in the direction of h, that is,

$$
\langle \partial E(u), h\rangle_2 = \Big\langle -\operatorname{div}\Big(\mathcal{H}_r(x,|\nabla u|)\frac{\nabla u}{|\nabla u|}\Big) + \lambda(u-f), h\Big\rangle_2
$$

and, thus, we get

$$
\partial E(u) = -\operatorname{div}\Big(\mathcal{H}_r(x,|\nabla u|)\frac{\nabla u}{|\nabla u|}\Big) + \lambda(u-f).
$$

Hence, the associated gradient flow of $E(u)$ is of the form,

$$
u_t = \operatorname{div}\Big(\mathcal{H}_r(x,|\nabla u|)\frac{\nabla u}{|\nabla u|}\Big) - \lambda(u-f) \quad \text{in } \Omega \times \mathbb{R}^+.
$$

The associated heat flow to the problem (4.3.4) can then be written as:

$$
u_t = \operatorname{div}\Big(p|\nabla u|^{p-1}\frac{\nabla u}{|\nabla u|} + a(x)\,q|\nabla u|^{q-1}\frac{\nabla u}{|\nabla u|}\Big) - \lambda(u-f), \quad \text{in } \Omega_T \tag{4.4.12}
$$

$$
u(x,t) = f(x), \qquad \text{on } \partial\Omega \times (0,T). \tag{4.4.13}
$$

We also consider an initial condition at $t = 0$, associated with the above boundary problem:

$$
u(x,0) = f(x), \quad \text{in } \Omega \times \{t = 0\}. \tag{4.4.14}
$$

In general, the above boundary problem (4.4.12)–(4.4.13) does not have a classical solution, so we need to introduce a weak formulation in order to satisfy the conditions for the well-posedness of the problem. We now derive the weak solution formulation of the boundary problem as follows.

Suppose that u is a classical solution of (4.4.12)–(4.4.13) belonging to the function space $L^2(0,T;C^2(\Omega)) \cap C^1(\overline{\Omega}_T)$, and $v \in L^2(0,T;W^{1,2}(\Omega)) \cap C^1(\overline{\Omega}_T)$, and both $u(\cdot,t)$ and $v(\cdot,t)$ have the same boundary function f continuous upto $\partial\Omega \times (0,T)$, for every t. Now multiplying (4.4.12) by $v-u$ both sides and integrating over Ω to obtain,

$$\int_\Omega u_t(v-u)\,dx = \int_\Omega \operatorname{div}\Big(\frac{p\nabla u}{|\nabla u|^{2-p}} + a(x)\,\frac{q\nabla u}{|\nabla u|^{2-q}}\Big)(v-u) - \lambda(u-f)(v-u)\,dx.$$

At the boundary $\partial\Omega$, $v-u = f-f = 0$, then using integration by parts formula for the integral on the right hand side above gives,

$$\begin{aligned}
&\int_\Omega u_t(v-u)\,dx \\
&\quad = -\int_\Omega \Big(\frac{p\nabla u}{|\nabla u|^{2-p}} + a(x)\frac{q\nabla u}{|\nabla u|^{2-q}}\Big)\cdot\nabla(v-u)\,dx \\
&\qquad\qquad - \int_\Omega \lambda(u-f)(v-f-u+f)\,dx \\
&\quad = -\int_\Omega \Big(\frac{p}{|\nabla u|^{2-p}} + a(x)\frac{q}{|\nabla u|^{2-q}}\Big)\nabla u\cdot\nabla v\,dx + \int_\Omega (p|\nabla u|^p + a(x)\,q|\nabla u|^q)\,dx \\
&\qquad\qquad + \int_\Omega \lambda(u-f)^2\,dx - \int_\Omega \lambda(u-f)(v-f)\,dx.
\end{aligned}$$

Further using absolute value relation on the first integral on the right implies,

$$\begin{aligned}
\int_\Omega u_t(v-u)\,dx \geqslant & -\int_\Omega (p|\nabla u|^{p-1}|\nabla v| + a(x)q|\nabla u|^{q-1}|\nabla v|)\,dx \\
& + \int_\Omega (p|\nabla u|^p + a(x)q|\nabla u|^q)\,dx + \int_\Omega \lambda(u-f)^2 - \frac{\lambda}{2}(u-f)^2 \\
& \qquad + \frac{\lambda}{2}(v-f)^2\,dx.
\end{aligned}$$

Thereafter, applying Young's inequality (2.1.7) to each of the terms in the first integral on the right above implies,

$$\begin{aligned}
&\int_\Omega u_t(v-u)\,dx \\
&\quad \geqslant -\int_\Omega p\left(\frac{1}{p}|\nabla v|^p + \left(1-\frac{1}{p}\right)|\nabla u|^p\right) + a(x)\,q\left(\frac{1}{q}|\nabla v|^q + \left(1-\frac{1}{q}\right)|\nabla u|^q\right)\,dx \\
&\qquad + \int_\Omega (p|\nabla u|^p + a(x)\,q|\nabla u|^q)\,dx + \int_\Omega \frac{\lambda}{2}(u-f)^2\,dx - \int_\Omega \frac{\lambda}{2}(v-f)^2\,dx,
\end{aligned}$$

that is,

$$\int_\Omega u_t(v-u)\,dx$$
$$\geqslant \int_\Omega \big(|\nabla u|^p + a(x)|\nabla u|^q + \frac{\lambda}{2}(u-f)^2\big) - \big(|\nabla v|^p + a(x)|\nabla v|^q + \frac{\lambda}{2}(v-f)^2\big)\,dx.$$

Hence, we obtain

$$\int_\Omega u_t(v-u)\,dx + E(v) \geqslant E(u).$$

Further integrating both sides with respect to t over $[0,s]$, where $s \in (0,T]$, yields

$$\int_0^s \int_\Omega u_t(v-u)\,dx\,dt + \int_0^s E(v)\,dt \geqslant \int_0^s E(u)\,dt. \tag{4.4.15}$$

Now if we consider that (4.4.15) holds, then for $w \in C_0^\infty(\Omega_T)$, setting $v \coloneqq u + \varepsilon w, \varepsilon \in (0,1)$ in (4.4.15), we have

$$\int_0^s \int_\Omega u_t \varepsilon w\,dx\,dt + \int_0^s E(u+\varepsilon w)\,dt \geqslant \int_0^s E(u)\,dt,$$

which implies that $\int_0^s \int_\Omega u_t \varepsilon w\,dx\,dt + \int_0^s E(u+\varepsilon w)\,dt$ attains a minimum at $\varepsilon = 0$. Thus, at

$$\frac{\partial}{\partial \varepsilon}\left(\int_0^s \int_\Omega u_t \varepsilon w\,dx\,dt + \int_0^s E(u+\varepsilon w)\,dt\right)\bigg|_{\varepsilon=0} = 0,$$

interchanging the order of integration and differentiation using Leibniz's rule, as the integrands: $E(u+\varepsilon w)$ and its derivative with respect to ε are continuous in x and t over Ω_T, we then obtain from above,

$$\int_0^s \int_\Omega u_t w\,dx\,dt + \int_0^s \int_\Omega (p|\nabla u|^{p-1} + a(x)\,q|\nabla u|^{q-1})\frac{\nabla u}{|\nabla u|}\cdot \nabla w$$
$$+ \lambda(u-f)w\,dx\,dt = 0,$$

that is,

$$\int_0^s \int_\Omega u_t w\,dx\,dt + \int_0^s \int_\Omega \big(\mathcal{H}_r(x,|\nabla u|)\frac{\nabla u}{|\nabla u|}\big)\cdot \nabla w + \lambda(u-f)w\,dx\,dt = 0.$$

Since $w = 0$ on $\partial\Omega$, then applying integration by parts formula above to get,

$$\int_0^s \int_\Omega u_t w\,dx\,dt + \int_0^s \int_\Omega \Big(-\operatorname{div}\big(\mathcal{H}_r(x,|\nabla u|)\frac{\nabla u}{|\nabla u|}\big) + \lambda(u-f)\Big)w\,dx\,dt = 0,$$

which holds for all $s \in (0,T]$ and $w \in C_0^\infty(\Omega_T)$. Since w is arbitrary, then by Fundamental Lemma of Calculus of Variations 2.1.2, we obtain from the above equation, $u_t = \operatorname{div}\big(\mathcal{H}_r(x,|\nabla u|)\frac{\nabla u}{|\nabla u|}\big) - \lambda(u-f)$, which is the heat flow equation (4.4.12). This motivates that, if $u \in L^2(0,T;W^{1,\mathcal{H}}(\Omega)\cap L^2(\Omega))$ satisfies (4.4.15) with $u_t \in L^2(\Omega_T)$, then u is a weak solution (pseudosolution) of (4.4.12)–(4.4.13), in the sense of distribution. We, thus, have the following definition.

***Definition* 4.4.2.** *A function $u \in L^2(0,T;W^{1,\mathcal{H}}(\Omega)\cap L^2(\Omega))$ with weak derivative $u_t \in L^2(\Omega_T)$ is called a weak solution of* (4.4.12)–(4.4.13) *if*

(i) $u(\cdot,t)-f \in W_0^{1,\mathcal{H}}(\Omega)$, *for every t, where* $f \in W^{1,\mathcal{H}}(\Omega)$, *and*

(ii) u *satisfies* (4.4.15) *for all* $s \in (0,T]$ *and* $v \in L^2(0,T;W^{1,\mathcal{H}}(\Omega)\cap L^2(\Omega))$ *such that* $v(\cdot,t)-f \in W_0^{1,\mathcal{H}}(\Omega)$, *for every t.*

4.4.2 The approximate functional and the associated boundary problem. Among the various regularization approaches [20; 33] developed so far to achieve specific geometric properties, we consider the ϵ-regularization technique [20] to artificially smooth the problem (4.3.4). For bounded set $\Omega \subset \mathbb{R}^n$ with Lipschitz boundary, we regularize $\mathcal{H}(x,r), r \in \mathbb{R}^+$ by introducing a smoothing parameter $\epsilon \in (0,1)$ such that the approximate functional, $\mathcal{H}^\epsilon : \Omega \times [0,\infty) \to \mathbb{R}$, is defined as:

$$\mathcal{H}^\epsilon(x,r) := \big(\sqrt{r^2+\epsilon^2}\big)^p + a(x)\big(\sqrt{r^2+\epsilon^2}\big)^q, \quad 1 < p \leqslant q \leqslant 2.$$

Here, the function $0 \leqslant a \in L^\infty(\mathbb{R}^n)\cap C^2(\Omega)$ is Lipschitz continuous and additionally assume that a is constant outside some large ball in $\mathbb{R}^n$ containing Ω, so that the decay condition $(A2)$ holds for $\mathcal{H}(x,r)$. Moreover, $\mathcal{H}$ satisfies (A1) condition, provided $\frac{n}{q}(p-q) \leqslant 1$ holds [24, Proposition 7.2.2]. Further, we consider that the function a has bounded second order derivative, which is mainly required to satisfy solvability conditions of the boundary problem.

The derivative of $\mathcal{H}^\epsilon(x,r)$ with respect to r, given by $\mathcal{H}^\epsilon_r : \Omega \times [0,\infty) \to \mathbb{R}$, is of the form,

$$\mathcal{H}^\epsilon_r(x,r) := \frac{\partial}{\partial r}\mathcal{H}^\epsilon(x,r) = \frac{pr}{\big(\sqrt{r^2+\epsilon^2}\big)^{2-p}} + a(x)\frac{qr}{\big(\sqrt{r^2+\epsilon^2}\big)^{2-q}}.$$

Further calculation of second derivative of $\mathcal{H}^\epsilon(x,r)$ with respect to r yields,

$$\frac{\partial^2}{\partial r^2}\mathcal{H}^\epsilon(x,r) = \frac{p(p-1)r^2+p\epsilon^2}{\big(\sqrt{r^2+\epsilon^2}\big)^{4-p}} + a(x)\frac{q(q-1)r^2+q\epsilon^2}{\big(\sqrt{r^2+\epsilon^2}\big)^{4-q}} > 0.$$

This implies that $\mathcal{H}^\epsilon(x,r)$ is (strictly) convex with respect to the second variable. While, as $\epsilon \to 0$, it is clear that $\mathcal{H}^\epsilon(x,r) \to \mathcal{H}(x,r)$ and $\mathcal{H}^\epsilon_r(x,r) \to \mathcal{H}_r(x,r)$.

For any $\epsilon \in (0,1)$, we have the estimates $(r^2+\epsilon^2)^{\frac{p}{2}} \geqslant (r^2)^{\frac{p}{2}}$ and $(r^2+\epsilon^2)^{\frac{q}{2}} \geqslant (r^2)^{\frac{q}{2}}$, where $p,q \in (1,2]$, from which $(\sqrt{r^2+\epsilon^2})^p - (\sqrt{r^2})^p + a(x)((\sqrt{r^2+\epsilon^2})^q - (\sqrt{r^2})^q) \geqslant 0$ holds. Thus, we have the property,

$$\mathcal{H}^\epsilon(x,r) \geqslant \mathcal{H}(x,r). \tag{4.4.16}$$

On the other hand, for $\mathcal{H}^\epsilon$, we use the estimate $\sqrt{a^2+b^2} \leqslant a+b, (a,b \in \mathbb{R}^+)$, to have, for any $r > 0$,

$$\mathcal{H}^\epsilon(x,r) = \big(\sqrt{r^2+\epsilon^2}\big)^p + a(x)\big(\sqrt{r^2+\epsilon^2}\big)^q \leqslant (r+\epsilon)^p + a(x)(r+\epsilon)^q,$$

also expressed as,

$$\mathcal{H}^\epsilon(x,r) \leqslant 2^p\Big(\frac{r+\epsilon}{2}\Big)^p + a(x)\,2^q\Big(\frac{r+\epsilon}{2}\Big)^q.$$

Further due to the convexity of $t \mapsto t^p$ and $t \mapsto t^q$ for the right-hand side terms,

$$\begin{aligned}\mathcal{H}^\epsilon(x,r) &\leqslant 2^{p-1}(r^p+\epsilon^p) + a(x)\,2^{q-1}(r^q+\epsilon^q)\\ &\leqslant 2(r^p + a(x)r^q + (|a(x)|+1)\epsilon^p),\end{aligned}$$

and since $|a| \leqslant M$, where $M > 0$ is constant, we obtain

$$\mathcal{H}^\epsilon(x,r) \leqslant 2\mathcal{H}(x,r) + 2(M+1)\epsilon.$$

Hence, we have the estimate,

$$\mathcal{H}^\epsilon(x,r) \leqslant 2\mathcal{H}(x,r) + C, \tag{4.4.17}$$

where the constant $C = 2(M+1)\epsilon > 0$.

Next, we regularize the boundary function f through mollification. We consider f in the whole of $\mathbb{R}^n$ instead of Ω, such that $f \in W^{1,\mathcal{H}}(\mathbb{R}^n) \cap L^\infty(\mathbb{R}^n)$. First consider a standard mollifier $\zeta \in L^1(\mathbb{R}^n)$, defined as,

$$\zeta(x) := \begin{cases} C\exp\left(\frac{1}{|x|^2-1}\right), & |x| < 1\\ 0, & \text{else,}\end{cases}$$

where constant $C > 0$ is selected so that $\|\zeta\|_1 = \int_{\mathbb{R}^n} \zeta\, dx = 1$. Then, for $\delta > 0$ and $x \in \Omega$, we define the convolution,

$$f_\delta(x) := (f * \zeta_\delta)(x) := \int_{\mathbb{R}^n} f(y)\,\zeta_\delta(x-y)\,dy = \int_{\mathbb{R}^n} f(x-y)\,\zeta_\delta(y)\,dy. \tag{4.4.18}$$

where $\zeta_\delta(x) := \frac{1}{\delta^n}\zeta\big(\frac{x}{\delta}\big)$. It holds that $\zeta_\delta \in C_0^\infty(\mathbb{R}^n)$ and satisfy $\int_{\mathbb{R}^n} \zeta_\delta\, dx = 1$. Thus, by standard mollification properties [15, C.4., Theorem 6] and Young's convolution inequality, we can conclude that $f_\delta \in C^\infty(\overline{\Omega})$ and $\nabla f_\delta := \nabla f * \zeta_\delta$ in Ω. Further, since the set Ω is compactly inside $\mathbb{R}^n$ and f_δ is continuous upto the boundary of Ω, so f_δ is bounded in $\overline{\Omega}$.

Since $\mathcal{H}$ satisfies (aDec), (A0), (A1) and (A2) conditions, then by Lemma 2.2.11, we obtain the following convergence property, as $\delta \to 0$,

$$f_\delta \to f \quad \text{in } W^{1,\mathcal{H}}(\Omega) \tag{4.4.19}$$

which implies $f_\delta \in W^{1,\mathcal{H}}(\Omega)$.

We note another estimate,

$$\|f_\delta\|_{L^\infty(\overline{\Omega})} \leqslant \|f\|_{L^\infty(\mathbb{R}^n)}, \tag{4.4.20}$$

which follows from Young's inequality for convolution (2.1.8).

Now, to prove the existence of solution to the heat flow problem (4.4.12)–(4.4.13), we first consider the following approximated problem:

$$u_t = \epsilon\, \Delta u + \operatorname{div}\left(\mathcal{H}_r^\epsilon(x, |\nabla u|)\frac{\nabla u}{|\nabla u|}\right) - \lambda(u - f_\delta), \qquad \text{in } \Omega_T \qquad (4.4.21)$$

$$u(x,t) = f_\delta(x), \qquad \text{on } \partial\Omega \times [0,T] \qquad (4.4.22)$$

where $f_\delta \in C^\infty(\overline{\Omega})$ and $\Delta u = \sum\limits_{i=1}^{n} \frac{\partial^2 u}{\partial x_i^2}$ is the Laplacian of u. The initial condition associated with the above boundary problem at $t = 0$ is taken as:

$$u(x,0) = f_\delta(x), \quad \text{in } \Omega \times \{t = 0\}. \qquad (4.4.23)$$

Note that in the divergence term in (4.4.21), if $\nabla u = 0$ then $\mathcal{H}_r^\epsilon(x, |\nabla u|)\dfrac{\nabla u}{|\nabla u|} = 0$.

We aim to prove that, under suitable conditions, as $\epsilon \to 0$ and $\delta \to 0$, the solution of the approximated problem (4.4.21)–(4.4.22) converges to the solution of the boundary problem (4.4.12)–(4.4.13). In order for that, we first prove the existence of solution of the approximated problem (4.4.21)–(4.4.22). Taking (4.4.21) as a quasilinear equation in the form,

$$\mathcal{L}u := u_t - \sum_{i,j=1}^{n} g_{ij}(x,t,u,\nabla u)\, u_{x_i x_j} + g(x,t,u,\nabla u) = 0, \qquad (4.4.24)$$

where the coefficient functions g_{ij} and g are continuous functions in the domain. Then calculating the divergence term in (4.4.21) to obtain

$$\operatorname{div}\left(\mathcal{H}_r^\epsilon(x, |\nabla u|)\frac{\nabla u}{|\nabla u|}\right) = \operatorname{div}\left(\frac{p\nabla u}{\left(\sqrt{|\nabla u|^2 + \epsilon^2}\right)^{2-p}} + a(x)\frac{q\nabla u}{\left(\sqrt{|\nabla u|^2 + \epsilon^2}\right)^{2-q}}\right)$$

which gives,

$$\begin{aligned}
&\operatorname{div}\left(\mathcal{H}_r^\epsilon(x, |\nabla u|)\frac{\nabla u}{|\nabla u|}\right) \\
&= \left(\frac{p}{\left(\sqrt{|\nabla u|^2 + \epsilon^2}\right)^{2-p}} + \frac{a(x)\, q}{\left(\sqrt{|\nabla u|^2 + \epsilon^2}\right)^{2-q}}\right)\Delta u + \left(\frac{p(p-2)}{\left(\sqrt{|\nabla u|^2 + \epsilon^2}\right)^{4-p}}\right. \\
&\qquad \left. + \frac{a(x)\, q(q-2)}{\left(\sqrt{|\nabla u|^2 + \epsilon^2}\right)^{4-q}}\right) \sum_{i,j=1}^{n} \frac{\partial u}{\partial x_i}\frac{\partial u}{\partial x_j}\frac{\partial^2 u}{\partial x_i \partial x_j} + \frac{q}{\left(\sqrt{|\nabla u|^2 + \epsilon^2}\right)^{2-q}}(\nabla a \cdot \nabla u).
\end{aligned}$$

Then (4.4.21) is of the form,

$$u_t = \sum_{i,j=1}^{n}\Bigg(\Big(\epsilon + \frac{p}{\big(\sqrt{|\nabla u|^2+\epsilon^2}\big)^{2-p}} + \frac{a(x)\,q}{\big(\sqrt{|\nabla u|^2+\epsilon^2}\big)^{2-q}}\Big)\delta_{ij} + \Big(\frac{p(p-2)|\nabla u|^2}{\big(\sqrt{|\nabla u|^2+\epsilon^2}\big)^{4-p}} + a(x)\frac{q(q-2)|\nabla u|^2}{\big(\sqrt{|\nabla u|^2+\epsilon^2}\big)^{4-q}}\Big)\frac{\partial u}{\partial x_i}\frac{\partial u}{\partial x_j}\Bigg)\frac{\partial^2 u}{\partial x_i \partial x_j} + \frac{q}{\big(\sqrt{|\nabla u|^2+\epsilon^2}\big)^{2-q}}\nabla a\cdot\nabla u - \lambda(u-f_\delta). \tag{4.4.25}$$

Comparing the above expression (4.4.25) with (4.4.24), the coefficient functions g_{ij} and g can be expressed as,

$$\begin{aligned} g_{ij}(x,t,u,\nabla u) = {}&\Big(\epsilon + \frac{p}{\big(\sqrt{|\nabla u|^2+\epsilon^2}\big)^{2-p}} + \frac{a(x)\,q}{\big(\sqrt{|\nabla u|^2+\epsilon^2}\big)^{2-q}}\Big)\delta_{ij} \\ &+ \Big(\frac{p(p-2)|\nabla u|^2}{\big(\sqrt{|\nabla u|^2+\epsilon^2}\big)^{4-p}} + a(x)\frac{q(q-2)|\nabla u|^2}{\big(\sqrt{|\nabla u|^2+\epsilon^2}\big)^{4-q}}\Big)\frac{\partial u}{\partial x_i}\frac{\partial u}{\partial x_j} \end{aligned} \tag{4.4.26}$$

and,

$$g(x,t,u,\nabla u) = -\frac{q}{\big(\sqrt{|\nabla u|^2+\epsilon^2}\big)^{2-q}}(\nabla a\cdot\nabla u) + \lambda(u-f_\delta). \tag{4.4.27}$$

Here, the partial differential equation (4.4.25) is uniformly parabolic, as proved in the appendix A.2, which satisfies the following parabolicity condition, for any $s \in \mathbb{R}$ and $\kappa \in \mathbb{R}^n$,

$$\nu\,|\xi|^2 \leqslant \sum_{i,j=1}^{n} g_{ij}(x,t,s,\kappa)\,\xi_i\xi_j \leqslant \mu\,|\xi|^2, \quad (\nu,\mu>0) \tag{4.4.28}$$

where $\xi = (\xi_1,\ldots,\xi_n)$ is an arbitrary real vector. Due to the parabolic nature, the solvability of the approximated problem (4.4.21)–(4.4.22) can be proved from the theory of general quasilinear parabolic equations, referred in the book of O. A. Ladyzhenskaya, V. A. Solonnikov and N. N. Uralt́seva [38]. We can then conclude the following result on the existence of solution of the quasilinear boundary value problem.

***Lemma* 4.4.3.** *Let $\Omega \subset \mathbb{R}^n$ be a bounded open set with Lipschitz boundary and $w \in C^\infty(\overline{\Omega})$. Consider the following quasilinear boundary problem,*

$$\left.\begin{aligned} &u_t(x,t) - \sum_{i,j=1}^{n} g_{ij}(x,t,u,\nabla u)\,u_{x_ix_j} + g(x,t,u,\nabla u) = 0, && (x,t)\in\Omega_T \\ &u(x,t) = w(x), && (x,t)\in\partial\Omega\times(0,T) \end{aligned}\right\} \tag{4.4.29}$$

where g_{ij} and g are given by (4.4.26) *and* (4.4.27) *respectively. Then there exists a solution of the boundary problem* (4.4.29) *in $C(\overline{\Omega}_T)\cap L^2(0,T;W^{1,2}(\Omega))$. Moreover, $u(\cdot,t)-w \in W_0^{1,2}(\Omega)$ holds for every t.*

Proof. Here, g_{ij} and g in (4.4.26) and (4.4.27), satisfy the solvability conditions in [24, p. 560, Theorem 4.4, Chapter 6], as proved in Proposition A.3.1 in the appendix. Then, for $\gamma \in (0,1)$, we have that $u \in C(\overline{\Omega}_T) \cap C^{2+\gamma,1+\frac{\gamma}{2}}(\Omega_T)$ is a solution of the quasilinear problem (4.4.29), where $C^{2+\gamma,1+\frac{\gamma}{2}}(\Omega_T)$ is the Hölder space consisting of derivatives of $u(x,t)$ with respect to x upto order 2 and with respect to t upto order 1. Moreover, we have $u \in C(\overline{\Omega}_T) \cap L^2(0,T;W^{1,2}(\Omega))$ and $u(\cdot,t) - w \in W_0^{1,2}(\Omega)$ for every t (refer [28, Lemma 1.26]). □

Remark 4.4.30. The result in Lemma 4.4.3 further implies that the solution of the parabolic boundary problem (4.4.29), satisfies C^∞-regularity in Ω_T by [15, Theorem 8, Section 2.3, Chapter 2], irrespective of smoothness of the boundary values on $\partial\Omega \times (0,T)$. This regularity assertion is required for estimating mixed order derivative of the solution with respect to both x and t. Thus, for fixed $\delta, \epsilon > 0$, considering $u_\delta^\epsilon(x,t)$ as the solution of the quasilinear boundary problem (4.4.29), it follows from the previous Lemma 4.4.3 that, the approximated problem (4.4.21)–(4.4.22) has weak solution u_δ^ϵ in $L^2(0,T;W^{1,2}(\Omega)) \cap C^\infty(\Omega_T)$, such that $u_\delta^\epsilon(\cdot,t) - f_\delta \in W_0^{1,2}(\Omega)$ for each t, where $f_\delta \in C^\infty(\overline{\Omega})$.

Next, in order to prove the existence of solution of the original boundary problem (4.4.12)–(4.4.13), we first produce the following *a priori* estimates.

***Lemma* 4.4.4.** *Let $\Omega \subset \mathbb{R}^n$ be a bounded open set with Lipschitz boundary and $f \in W^{1,\mathcal{H}}(\mathbb{R}^n) \cap L^\infty(\mathbb{R}^n)$. For fixed $\delta, \epsilon > 0$ and $f_\delta \in C^\infty(\overline{\Omega})$, if function $u_\delta^\epsilon \in L^2(0,T;W^{1,2}(\Omega)) \cap C^\infty(\Omega_T)$ is a solution of the problem* (4.4.21)–(4.4.22) *having initial condition* (4.4.23)*, then $(u_\delta^\epsilon)_t \in L^2(0,T;L^2(\Omega))$ satisfies*

$$\begin{aligned}
&\int_0^T \int_\Omega |(u_\delta^\epsilon)_t|^2 \, dx \, dt \\
&+ \sup_{s\in(0,T]} \int_\Omega \Big(\frac{\epsilon}{2}|\nabla u_\delta^\epsilon(x,s)|^2 + \mathcal{H}^\epsilon(x,|\nabla u_\delta^\epsilon(x,s)|) + \frac{\lambda}{2}(u_\delta^\epsilon(x,s) - f_\delta)^2\Big) \, dx \leqslant C,
\end{aligned} \tag{4.4.31}$$

where $C > 0$ is a constant depending only on Ω and $\epsilon\|\nabla f_\delta\|^2_{L^2(\Omega)}$.

Proof. Since $u_\delta^\epsilon \in L^2(0,T;W^{1,2}(\Omega)) \cap C^\infty(\Omega_T)$ satisfies the approximated problem (4.4.21)–(4.4.22), then multiplying (4.4.21) by $(u_\delta^\epsilon)_t$ and integrating over Ω, we obtain,

$$\begin{aligned}
\int_\Omega |(u_\delta^\epsilon)_t|^2 \, dx = \int_\Omega \epsilon (u_\delta^\epsilon)_t \Delta u_\delta^\epsilon \, dx + \int_\Omega (u_\delta^\epsilon)_t \operatorname{div}\big(\mathcal{H}_r^\epsilon(x,|\nabla u_\delta^\epsilon|)\frac{\nabla u_\delta^\epsilon}{|\nabla u_\delta^\epsilon|}\big) \, dx \\
- \lambda \int_\Omega (u_\delta^\epsilon)_t (u_\delta^\epsilon - f_\delta) \, dx.
\end{aligned}$$

At the boundary $\partial\Omega$, we have $(u_\delta^\epsilon)_t = 0$, since f_δ is independent of t. Then, for fixed t, applying integration by parts formula over Ω for the first two integrals on the right-hand side above to get,

$$\int_\Omega |(u_\delta^\epsilon)_t|^2\,dx = -\epsilon\int_\Omega \nabla(u_\delta^\epsilon)_t\cdot\nabla u_\delta^\epsilon\,dx - \int_\Omega \nabla(u_\delta^\epsilon)_t\cdot\frac{\nabla u_\delta^\epsilon}{|\nabla u_\delta^\epsilon|}\mathcal{H}_r^\epsilon(x,|\nabla u_\delta^\epsilon|)\,dx \\ -\lambda\int_\Omega (u_\delta^\epsilon)_t(u_\delta^\epsilon - f_\delta)\,dx.$$

Since $u_\delta^\epsilon \in C^\infty(\Omega_T)$, then interchanging the differentiation order of the mixed derivatives of x and t using Schwarz's Theorem, in the previous equation to obtain

$$\int_\Omega |(u_\delta^\epsilon)_t|^2\,dx = -\epsilon\int_\Omega \frac{\partial}{\partial t}(\nabla u_\delta^\epsilon)\cdot\nabla u_\delta^\epsilon\,dx - \int_\Omega \frac{\partial}{\partial t}(\nabla u_\delta^\epsilon)\cdot\frac{\nabla u_\delta^\epsilon}{|\nabla u_\delta^\epsilon|}\mathcal{H}_r^\epsilon(x,|\nabla u_\delta^\epsilon|)\,dx \\ -\lambda\int_\Omega \frac{\partial}{\partial t}(u_\delta^\epsilon - f_\delta)^2\,dx. \tag{4.4.32}$$

For the terms on the right-hand side above, note that $\frac{\partial}{\partial t}(\nabla u_\delta^\epsilon)\cdot\nabla u_\delta^\epsilon = \frac{1}{2}\frac{\partial}{\partial t}|\nabla u_\delta^\epsilon|^2$, and, on the other hand, we compute,

$$\begin{aligned}\frac{\partial}{\partial t}(\mathcal{H}^\epsilon(x,|\nabla u_\delta^\epsilon|)) &= \left(\frac{p|\nabla u_\delta^\epsilon|}{\left(\sqrt{|\nabla u|^2+\epsilon^2}\right)^{2-p}} + a(x)\frac{q|\nabla u_\delta^\epsilon|}{\left(\sqrt{|\nabla u|^2+\epsilon^2}\right)^{2-q}}\right)\frac{\partial}{\partial t}|\nabla u_\delta^\epsilon| \\ &= \mathcal{H}_r^\epsilon(x,|\nabla u_\delta^\epsilon|)\frac{\nabla u_\delta^\epsilon}{|\nabla u_\delta^\epsilon|}\cdot\frac{\partial}{\partial t}(\nabla u_\delta^\epsilon).\end{aligned}$$

Then plugging the above derivatives in (4.4.32) implies,

$$\begin{aligned}&\int_\Omega |(u_\delta^\epsilon)_t|^2\,dx \\ &\quad = -\int_\Omega \frac{\epsilon}{2}\frac{\partial}{\partial t}|\nabla u_\delta^\epsilon|^2\,dx - \int_\Omega \frac{\partial}{\partial t}(\mathcal{H}^\epsilon(x,|\nabla u_\delta^\epsilon|))\,dx - \lambda\int_\Omega \frac{\partial}{\partial t}(u_\delta^\epsilon - f_\delta)^2\,dx,\end{aligned}$$

that is,

$$\int_\Omega |(u_\delta^\epsilon)_t|^2\,dx + \int_\Omega \frac{\partial}{\partial t}\big(\tfrac{\epsilon}{2}|\nabla u_\delta^\epsilon|^2 + \mathcal{H}^\epsilon(x,|\nabla u_\delta^\epsilon|) + \lambda(u_\delta^\epsilon - f_\delta)^2\big)\,dx = 0.$$

Interchanging the order of integration and differentiation above using Leibniz's integral rule, and further integrating with respect to t over the interval $[0,s]$ where $s\in(0,T]$, we have

$$\int_0^s\int_\Omega |(u_\delta^\epsilon)_t|^2\,dx\,dt + \int_0^s\frac{\partial}{\partial t}\int_\Omega\big(\tfrac{\epsilon}{2}|\nabla u_\delta^\epsilon|^2 + \mathcal{H}^\epsilon(x,|\nabla u_\delta^\epsilon|) + \lambda(u_\delta^\epsilon - f_\delta)^2\big)\,dx\,dt = 0.$$

At $t=0$, we have $u_\delta^\epsilon(x,0) = f_\delta(x)$, then the above equation implies,

$$\int_0^s\int_\Omega |(u_\delta^\epsilon)_t|^2\,dx\,dt + \int_\Omega\big(\tfrac{\epsilon}{2}|\nabla u_\delta^\epsilon(x,s)|^2 + \mathcal{H}^\epsilon(x,|\nabla u_\delta^\epsilon(x,s)|) \\ + \lambda(u_\delta^\epsilon(x,s) - f_\delta)^2\big)\,dx = \int_\Omega\big(\tfrac{\epsilon}{2}|\nabla f_\delta|^2 + \mathcal{H}^\epsilon(x,|\nabla f_\delta|)\big)\,dx.$$

Applying the estimate $\mathcal{H}^\epsilon(x,|\nabla f_\delta|) \leqslant 2\mathcal{H}(x,|\nabla f_\delta|)+C$ from (4.4.17), on the right-hand side above, to have,

$$\int_0^s \int_\Omega |(u_\delta^\epsilon)_t|^2 \, dx \, dt + \int_\Omega \Big(\frac{\epsilon}{2} |\nabla u_\delta^\epsilon(x,s)|^2 + \mathcal{H}^\epsilon(x, |\nabla u_\delta^\epsilon(x,s)|) + \lambda (u_\delta^\epsilon(x,s) - f_\delta)^2 \Big) \, dx \leqslant \int_\Omega \Big(\frac{\epsilon}{2} |\nabla f_\delta|^2 + 2\mathcal{H}(x,|\nabla f_\delta|) + C \Big) \, dx. \tag{4.4.33}$$

Further, since $\int_\Omega \mathcal{H}(x,|\nabla f_\delta|)\,dx \leqslant C$(constant), then the previous expression implies,

$$\int_0^s \int_\Omega |(u_\delta^\epsilon)_t|^2 \, dx \, dt + \int_\Omega \Big(\frac{\epsilon}{2} |\nabla u_\delta^\epsilon(x,s)|^2 + \mathcal{H}^\epsilon(x, |\nabla u_\delta^\epsilon(x,s)|) + \lambda (u_\delta^\epsilon(x,s) - f_\delta)^2 \Big) \, dx \leqslant C, \tag{4.4.34}$$

where the constant C depends on Ω and $\epsilon\|\nabla f_\delta\|^2_{L^2(\Omega)}$. Thus, as $s \to T$, then (4.4.34) implies,

$$\int_0^T \int_\Omega |(u_\delta^\epsilon)_t|^2 \, dx \, dt + \sup_{s\in(0,T]} \int_\Omega \Big(\frac{\epsilon}{2} |\nabla u_\delta^\epsilon(x,s)|^2 + \mathcal{H}^\epsilon(x, |\nabla u_\delta^\epsilon(x,s)|) + \lambda (u_\delta^\epsilon(x,s) - f_\delta)^2 \Big) \, dx \leqslant C,$$

with constant C depending only on Ω and $\epsilon\|\nabla f_\delta\|^2_{L^2(\Omega)}$, as required. □

***Lemma* 4.4.5.** *Let $\Omega \subset \mathbb{R}^n$ be a bounded open set with Lipschitz boundary and $f \in W^{1,\mathcal{H}}(\mathbb{R}^n) \cap L^\infty(\mathbb{R}^n)$. For fixed $\delta, \epsilon > 0$ and $f_\delta \in C^\infty(\overline{\Omega})$, if $u_\delta^\epsilon \in L^2(0,T;W^{1,2}(\Omega)) \cap C^\infty(\Omega_T)$ is the weak solution of the approximated problem* (4.4.21)–(4.4.22) *having initial condition* (4.4.23)*, then*

$$\|u_\delta^\epsilon\|_{L^\infty(\Omega_T)} \leqslant \|f\|_{L^\infty(\mathbb{R}^n)}. \tag{4.4.35}$$

Proof. Let $k \coloneqq \|f\|_{L^\infty(\mathbb{R}^n)}$ and G be a C^0 truncation function defined as $G(v) \coloneqq \max\{v, 0\}$. Choosing $v \coloneqq u_\delta^\epsilon - k$, we have

$$G(u_\delta^\epsilon - k) = \begin{cases} u_\delta^\epsilon - k, & \text{if } u_\delta^\epsilon - k \geqslant 0, \\ 0, & \text{otherwise.} \end{cases}$$

By Lemma 2.2.7, we have, for fixed $t > 0$, $G(u_\delta^\epsilon(\cdot,t) - k) \in W^{1,2}(\Omega)$. Let G' denote the derivative of $G(u_\delta^\epsilon(\cdot,t) - k)$ with respect to x, having value $G' = \nabla u_\delta^\epsilon(x,t)$ for almost every $x \in \{u_\delta^\epsilon(\cdot,t) - k \geqslant 0\}$, otherwise $G' = 0$ for all other x. Since u_δ^ϵ

is weak solution of (4.4.21), then multiplying (4.4.21) by $G(u_\delta^\epsilon - k)$ and integrating over Ω gives

$$\begin{aligned}\int_\Omega (u_\delta^\epsilon)_t\, G(u_\delta^\epsilon - k)\, dx \\ = \int_\Omega \epsilon\Delta u_\delta^\epsilon\, G(u_\delta^\epsilon - k) + \operatorname{div}\left(\mathcal{H}_r^\epsilon(x, |\nabla u_\delta^\epsilon|)\frac{\nabla u_\delta^\epsilon}{|\nabla u_\delta^\epsilon|}\right) G(u_\delta^\epsilon - k) \\ - \lambda(u_\delta^\epsilon - f_\delta)\, G(u_\delta^\epsilon - k)\, dx.\end{aligned}$$

Using (4.4.20), we have at the boundary $\partial\Omega$, $u_\delta^\epsilon = f_\delta \leqslant |f_\delta| \leqslant \|f_\delta\|_{L^\infty(\partial\Omega)} \leqslant \|f\|_{L^\infty(\mathbb{R}^n)} = k$, that is $u_\delta^\epsilon - k \leqslant 0$, and hence, by definition, $G(u_\delta^\epsilon - k) = 0$. Then using integration by parts formula for the right-hand side integral above gives,

$$\begin{aligned}\int_\Omega (u_\delta^\epsilon)_t\, G(u_\delta^\epsilon - k)\, dx \\ = -\int_\Omega \epsilon\nabla u_\delta^\epsilon \cdot G'\, dx - \int_\Omega \mathcal{H}_r^\epsilon(x, |\nabla u_\delta^\epsilon|)\frac{\nabla u_\delta^\epsilon}{|\nabla u_\delta^\epsilon|}\cdot G'\, dx \\ - \int_\Omega \lambda(u_\delta^\epsilon - f_\delta)\, G(u_\delta^\epsilon - k)\, dx.\end{aligned} \tag{4.4.36}$$

To evaluate the right-hand expression in (4.4.36), we consider two cases: the first case is for $x \in \{u_\delta^\epsilon(\cdot,t) < k\}$. This yields $G(u_\delta^\epsilon - k) = 0$ and also $G' = 0$. Hence, (4.4.36) implies $\int_{\{u_\delta^\epsilon(\cdot,t)<k\}} (u_\delta^\epsilon)_t\, G(u_\delta^\epsilon - k)\, dx = 0$. The second case is for $x \in \{u_\delta^\epsilon(\cdot,t) \geqslant k\}$. Then, we have $u_\delta^\epsilon(\cdot,t) \geqslant k = \|f\|_{L^\infty(\mathbb{R}^n)} \geqslant \|f_\delta\|_{L^\infty(\overline{\Omega})} \geqslant f_\delta$, that is, $u_\delta^\epsilon(\cdot,t) \geqslant f_\delta$ for every $x \in \{u_\delta^\epsilon(\cdot,t) \geqslant k\}$. Moreover, for $u_\delta^\epsilon(\cdot,t) \geqslant k$, we have $G(u_\delta^\epsilon - k) = u_\delta^\epsilon - k \geqslant 0$. Additionally, $G' = \nabla u_\delta^\epsilon$ in this case.

Now, for the right-hand side integrals in (4.4.36), we have

$$\begin{aligned}&\int_\Omega \lambda(u_\delta^\epsilon - f_\delta)\, G(u_\delta^\epsilon - k)\, dx \\ &= \int_\Omega \lambda(u_\delta^\epsilon - f_\delta)\, G(u_\delta^\epsilon - k)\chi_{\{u_\delta^\epsilon(\cdot,t)<k\}}\, dx \\ &\qquad + \int_\Omega \lambda(u_\delta^\epsilon - f_\delta)\, G(u_\delta^\epsilon - k)\chi_{\{u_\delta^\epsilon(\cdot,t)\geqslant k\}}\, dx \\ &\geqslant 0 + \int_\Omega \lambda(u_\delta^\epsilon - f_\delta)^2\, \chi_{\{u_\delta^\epsilon(\cdot,t)\geqslant k\}}\, dx \geqslant 0\end{aligned}$$

and, for

$$\begin{aligned}\int_\Omega \epsilon\nabla u_\delta^\epsilon \cdot G'\, dx &= \int_\Omega \epsilon\nabla u_\delta^\epsilon \cdot G'\chi_{\{u_\delta^\epsilon(\cdot,t)<k\}}\, dx + \int_\Omega \epsilon\nabla u_\delta^\epsilon \cdot G'\chi_{\{u_\delta^\epsilon(\cdot,t)\geqslant k\}}\, dx \\ &= 0 + \int_\Omega \epsilon|\nabla u_\delta^\epsilon|^2\chi_{\{u_\delta^\epsilon(\cdot,t)\geqslant k\}}\, dx \geqslant 0\end{aligned}$$

and, similarly for,

$$\int_\Omega \mathcal{H}_r^\epsilon(x, |\nabla u_\delta^\epsilon|)\frac{\nabla u_\delta^\epsilon}{|\nabla u_\delta^\epsilon|}\cdot G'\,dx = \int_\Omega \mathcal{H}_r^\epsilon(x, |\nabla u_\delta^\epsilon|)|\nabla u_\delta^\epsilon|\chi_{\{u_\delta^\epsilon(\cdot,t)\geqslant k\}}\,dx \geqslant 0$$

Since the value of the right-hand side expression in (4.4.36) is non-positive for every x, which implies $\int_\Omega (u_\delta^\epsilon)_t G(u_\delta^\epsilon - k)\,dx \leqslant 0$, that is,

$$\frac{1}{2}\int_\Omega \frac{d}{dt}(G(u_\delta^\epsilon - k))^2\,dx \leqslant 0.$$

Interchanging the order of integration and differentiation through Leibniz integral rule, we get $\frac{d}{dt}\int_\Omega \frac{1}{2}(G(u_\delta^\epsilon - k))^2\,dx \leqslant 0$, which implies that $\int_\Omega \frac{1}{2}(G(u_\delta^\epsilon - k))^2\,dx$ is monotonically decreasing in t. Moreover, at the initial condition $\Omega\times\{t=0\}$, we have $u_\delta^\epsilon = f_\delta \leqslant k$, which implies $G(u_\delta^\epsilon - k) = 0$. So we can conclude that,

$$\int_\Omega \frac{1}{2}(G(u_\delta^\epsilon(x,t) - k))^2\,dx \leqslant \int_\Omega \frac{1}{2}(G(u_\delta^\epsilon(x,0) - k))^2\,dx = 0.$$

Thus, we have $\frac{1}{2}\int_\Omega(G(u_\delta^\epsilon(\cdot,s) - k))^2\,dx = 0$, for every $s\in(0,T]$. This yields $G(u_\delta^\epsilon(\cdot,s) - k) = 0$, which implies $u_\delta^\epsilon(\cdot,s)\leqslant k$, for every $s\in(0,T]$.

On the other hand, considering $v := u_\delta^\epsilon + k$ and multiplying (4.4.21) by $G(u_\delta^\epsilon + k)$, we follow the same process as above to obtain $u_\delta^\epsilon(\cdot,s)\geqslant -k$ for every $s\in(0,T]$. Hence, we can conclude that $-k\leqslant u_\delta^\epsilon(\cdot,s)\leqslant k$, that is $|u_\delta^\epsilon|\leqslant k$, which implies,

$$\|u_\delta^\epsilon\|_{L^\infty(\Omega_T)} \leqslant k = \|f\|_{L^\infty(\mathbb{R}^n)}$$

as required. □

4.4.3 Existence and uniqueness of solution of the boundary value problem

***Theorem* 4.4.6.** *Let* $\Omega\subset\mathbb{R}^2$ *be a bounded open set with Lipschitz boundary and* $\mathcal{H}:\Omega\times[0,\infty)\to[0,\infty)$*, as defined in* (4.2.1) *where* $1 < p \leqslant q \leqslant 2$ *such that* $q < \frac{3}{2}p$*. Suppose* $f\in W^{1,\mathcal{H}}(\mathbb{R}^2)\cap L^\infty(\mathbb{R}^2)$*. Then there exists a unique weak solution* $u\in L^\infty(0,T;W^{1,\mathcal{H}}(\Omega)\cap L^\infty(\Omega))$ *of the boundary problem* (4.4.12)–(4.4.13) *having initial condition* (4.4.14)*, with* $u_t\in L^2(\Omega_T)$.

Proof. Existence: For fixed $\delta,\epsilon > 0$, there exists a strong solution of the approximated problem (4.4.21)–(4.4.22) in $L^2(0,T;W^{1,2}(\Omega))\cap C^\infty(\Omega_T)$, say $u_\delta^\epsilon(x,t)$, by Lemma 4.4.3, such that $u_\delta^\epsilon(\cdot,t) - f_\delta\in W_0^{1,2}(\Omega)$ for each t, where $f_\delta\in C^\infty(\overline{\Omega})$ is the convolution function of f as defined in (4.4.18). Note that, on passing $\delta,\epsilon\to 0$ in the approximated problem (4.4.21)–(4.4.22), we would obtain the heat flow problem (4.4.12)–(4.4.13), under suitable convergence conditions. So, to prove the existence of solution for the boundary problem (4.4.12)–(4.4.13), the approach involves first fixing $\delta > 0$ for the solution of the boundary problem (4.4.21)–(4.4.22) $u_\delta^\epsilon(x,t)$ and pass to the limit as $\epsilon \to 0$, and thereafter passing $\delta \to 0$.

Step 1: With fixed $\delta > 0$, passing $\epsilon \to 0$. Now, for fixed $\delta > 0$, we have $\{u_\delta^{\epsilon_i}\}_{i=1}^\infty$ as the sequence of solutions to the approximated problem (4.4.21)–(4.4.22). Then, for each $\epsilon_i \in (0,1)$, $\{u_\delta^{\epsilon_i}\}$ satisfies the estimates in Lemma 4.4.5 and Lemma 4.4.4, which implies that $\{u_\delta^{\epsilon_i}\}$ has uniformly bounded $L^\infty(\Omega_T)$ norm, and $\{(u_\delta^{\epsilon_i})_t\}$ has uniformly bounded $L^2(\Omega_T)$ norm with bounded constant depending on Ω and $\epsilon_i \|\nabla f_\delta\|^2_{L^2(\Omega)}$. Hence, by sequential Banach-Alaoglu Theorem [57, Section 3.15, p 68], there exists a subsequence $\{u_\delta^{\epsilon_{ij}}\}_{j=1}^\infty \subset \{u_\delta^{\epsilon_i}\}$ such that, for $u_\delta \in L^\infty(\Omega_T)$ and $w \in L^2(\Omega_T)$, as $\epsilon_i \to 0$,

$$u_\delta^{\epsilon_{ij}} \rightharpoonup u_\delta \quad \text{weakly* in } L^\infty(\Omega_T) \tag{4.4.37}$$

$$(u_\delta^{\epsilon_{ij}})_t \rightharpoonup w \quad \text{weakly in } L^2(\Omega_T), \tag{4.4.38}$$

Then, from A.4 in the appendix,

$$w(x,t) = (u_\delta(x,t))_t, \quad u_\delta(x,0) = f_\delta(x), \tag{4.4.39}$$

where $(u_\delta)_t \in L^2(\Omega_T)$.

We have $\int_0^s (u_\delta^{\epsilon_{ij}}(x,t))_t \, dt = u_\delta^{\epsilon_{ij}}(x,s) - f_\delta(x)$, where $s \in (0,T], T > 0$, then for $\psi \in L^2(\Omega)$, the following holds,

$$\int_\Omega \int_0^s (u_\delta^{\epsilon_{ij}}(x,t))_t \psi(x) \, dt \, dx = \int_\Omega (u_\delta^{\epsilon_{ij}}(x,s) - f_\delta(x))\psi(x) \, dx.$$

Thereafter, taking $\epsilon_i \to 0$, we apply weak L^2-convergence of $(u_\delta^{\epsilon_{ij}})_t$ from (4.4.38), to obtain,

$$\begin{aligned}
\int_\Omega (u_\delta^{\epsilon_{ij}}(x,s) - f_\delta(x))\psi(x) \, dx &= \int_\Omega \int_0^s (u_\delta^{\epsilon_{ij}}(x,t))_t \psi(x) \, dt \, dx \\
&\to \int_\Omega \int_0^s (u_\delta(x,t))_t \psi(x) \, dt \, dx \\
&= \int_\Omega \int_0^s (u_\delta(x,t))_t \psi(x) \, dt \, dx \\
&= \int_\Omega (u_\delta(x,s) - f_\delta(x))\psi(x) \, dx.
\end{aligned}$$

This implies that, for each $s > 0$, as $\epsilon_i \to 0$,

$$u_\delta^{\epsilon_{ij}}(\cdot,s) \rightharpoonup u_\delta(\cdot,s) \quad \text{weakly in } L^2(\Omega).$$

Since the sequence $\{u_\delta^{\epsilon_i}\}$ satisfies the estimate (4.4.31) from Lemma 4.4.4, we have $\|u_\delta^{\epsilon_i}(\cdot,s) - f_\delta\|_2^2 \leqslant C$ for each $s > 0$, where $C > 0$ is constant depending on Ω and $\epsilon_i\|\nabla f_\delta\|^2_{L^2(\Omega)}$. Then, applying the embedding $L^2(\Omega) \hookrightarrow L^{\mathcal{H}}(\Omega)$, by [4, Lemma 2.5], we obtain $\|u_\delta^{\epsilon_i}(\cdot,s) - f_\delta\|_{\mathcal{H}} \leqslant C$. Also by Lemma 2.2.11, $\|f_\delta\|_{\mathcal{H}} \leqslant C$ where constant $C > 0$ depends on $\|f\|_{\mathcal{H}}$. Then, for each $s > 0$, $\|u_\delta^{\epsilon_i}(\cdot,s)\|_{\mathcal{H}} \leqslant \|u_\delta^{\epsilon_i}(\cdot,s) - f_\delta\|_{\mathcal{H}} + \|f_\delta\|_{\mathcal{H}} \leqslant C$. By definition of the norm of $L^{\mathcal{H}}$ space and its unit ball property (Lemma 2.2.2), we have $\varrho_{\mathcal{H}}(u_\delta^{\epsilon_i}(\cdot,s)) \leqslant C$ and $\varrho_{\mathcal{H}}(u_\delta^{\epsilon_i}(\cdot,s) - f_\delta) \leqslant C$. Moreover, since $\int_\Omega \mathcal{H}^\epsilon(x, |\nabla u_\delta^{\epsilon_i}(\cdot,s)|)\,dx \leqslant C$ from the estimate (4.4.31), and $\mathcal{H} \leqslant \mathcal{H}^\epsilon$ from (4.4.16), we get $\int_\Omega \mathcal{H}(x, |\nabla u_\delta^{\epsilon_i}(\cdot,s)|)\,dx \leqslant C$, that is $\varrho_{\mathcal{H}}(\nabla u_\delta^{\epsilon_i}(\cdot,s)) \leqslant C$. Thus, we have, for each s,

$$\varrho_{\mathcal{H}}(u_\delta^{\epsilon_i}(\cdot,s)) + \varrho_{\mathcal{H}}(\nabla u_\delta^{\epsilon_i}(\cdot,s)) \leqslant C, \tag{4.4.40}$$

which implies $\{u_\delta^{\epsilon_i}(\cdot,s)\}$ is a bounded sequence in $W^{1,\mathcal{H}}(\Omega)$. And since $\|\nabla f_\delta\|_{\mathcal{H}} \leqslant C$ from 2.2.11, we have $\|\nabla(u_\delta^{\epsilon_i}(\cdot,s) - f_\delta)\|_{\mathcal{H}} \leqslant \|\nabla u_\delta^{\epsilon_i}(\cdot,s)\|_{\mathcal{H}} + \|\nabla f_\delta\|_{\mathcal{H}} \leqslant C$, where $C > 0$ is a constant depending on $\Omega, \epsilon\|\nabla f_\delta\|^2_{L^2(\Omega)}$ and $\|\nabla f\|_{\mathcal{H}}$. Using the estimate $\varrho_{\mathcal{H}}(u_\delta^{\epsilon_i}(\cdot,s) - f_\delta) \leqslant C$, we similarly obtain as above,

$$\varrho_{1,\mathcal{H}}(u_\delta^{\epsilon_i}(\cdot,s) - f_\delta) \leqslant C. \tag{4.4.41}$$

Since $\mathcal{H}$ satisfies (A0) and (aInc)$_p$, $(p > 1)$, conditions, then by [24, Lemma 6.1.6], we have $W^{1,\mathcal{H}}(\Omega) \hookrightarrow W^{1,p}(\Omega)$, where p is the index from (aInc)$_p$ condition. Thus, (4.4.40) yields that, for each $t > 0$, the sequence $\{u_\delta^{\epsilon_i}(\cdot,t)\}$ is bounded in $W^{1,p}(\Omega)$. Further, due to the boundedness of Ω in $\mathbb{R}^2$ and applying Rellich-Kondrachov theorem, we have that $W^{1,p}(\Omega)$ compactly embeds into L^q for all $q < \frac{2p}{2-p}$, which implies compact embedding of $W^{1,p}(\Omega)$ into $L^2(\Omega)$ space. Thus, for each $t > 0$, the subsequence $\{u_\delta^{\epsilon_{i_j}}(\cdot,t)\}_{j=1}^\infty \subset \{u_\delta^{\epsilon_i}(\cdot,t)\}$ converges strongly in $L^2(\Omega)$, as $\epsilon_i \to 0$, that is,

$$u_\delta^{\epsilon_{i_j}}(\cdot,t) \to u_\delta(\cdot,t) \quad \text{strongly in } L^2(\Omega). \tag{4.4.42}$$

Next, we have $(u_\delta^{\epsilon_{i_j}})_t u_\delta^{\epsilon_{i_j}} - (u_\delta)_t u_\delta = (u_\delta^{\epsilon_{i_j}})_t(u_\delta^{\epsilon_{i_j}} - u_\delta) + u_\delta((u_\delta^{\epsilon_{i_j}})_t - (u_\delta)_t)$, then integrating both sides over Ω, for each t, implies,

$$\begin{aligned}\int_\Omega &((u_\delta^{\epsilon_{i_j}})_t u_\delta^{\epsilon_{i_j}} - (u_\delta)_t u_\delta)\,dx \\ &\leqslant \int_\Omega |(u_\delta^{\epsilon_{i_j}})_t||u_\delta^{\epsilon_{i_j}} - u_\delta|\,dx + \int_\Omega u_\delta((u_\delta^{\epsilon_{i_j}})_t - (u_\delta)_t)\,dx.\end{aligned}$$

Further, applying Hölder's inequality for the first integral on the right gives,

$$\begin{aligned}\int_\Omega &((u_\delta^{\epsilon_{i_j}})_t u_\delta^{\epsilon_{i_j}} - (u_\delta)_t u_\delta)\,dx \\ &\leqslant \|(u_\delta^{\epsilon_{i_j}})_t\|_2 \|u_\delta^{\epsilon_{i_j}} - u_\delta\|_2 + \int_\Omega u_\delta((u_\delta^{\epsilon_{i_j}})_t - (u_\delta)_t)\,dx.\end{aligned}$$

Then, for each $t > 0$, as $\epsilon_{ij} \to 0$, we apply strong-L^2 convergence of $u_\delta^{\epsilon_{ij}}(\cdot,t)$ from (4.4.42) and weak-L^2 convergence of $(u_\delta^{\epsilon_{ij}}(\cdot,t))_t$ from (4.4.38) in the above expression, which implies,

$$\int_\Omega ((u_\delta^{\epsilon_{ij}}(\cdot,t))_t u_\delta^{\epsilon_{ij}}(\cdot,t) - (u_\delta(\cdot,t))_t u_\delta(\cdot,t))\,dx \to 0, \quad \text{as } \epsilon_{ij} \to 0. \tag{4.4.43}$$

Now consider a test function, say, $v \in L^2(0,T;W^{1,2}(\Omega)) \cap C^\infty(\Omega_T)$ such that $v(\cdot,t) - f_\delta \in W_0^{1,2}(\Omega)$, for every t. Since the solution $\{u_\delta^{\epsilon_{ij}}\}$ satisfies the PDE (4.4.21), then multiplying (4.4.21) by $(v - u_\delta^{\epsilon_{ij}})$, and integrating with respect to x over Ω as well as with respect to t over $[0,s]$, where $s \in (0,T], T > 0$, we have,

$$\begin{aligned}&\int_0^s \int_\Omega (u_\delta^{\epsilon_{ij}})_t (v - u_\delta^{\epsilon_{ij}}) - \epsilon_{ij}\Delta u_\delta^{\epsilon_{ij}}(v - u_\delta^{\epsilon_{ij}})\,dx\,dt\\ &= \int_0^s \int_\Omega \operatorname{div}\big(\mathcal{H}_r^\epsilon(x,|\nabla u_\delta^{\epsilon_{ij}}|)\frac{\nabla u_\delta^{\epsilon_{ij}}}{|\nabla u_\delta^{\epsilon_{ij}}|}\big)(v - u_\delta^{\epsilon_{ij}}) - \lambda(u_\delta^{\epsilon_{ij}} - f_\delta)(v - u_\delta^{\epsilon_{ij}})\,dx\,dt.\end{aligned}$$

Note that, from proof of Lemma 4.4.3, $u_\delta^{\epsilon_{ij}}$ and its derivatives of x upto order 2 are Hölder continuous in Ω_T, which implies they are continuous in $\overline{\Omega}$, and thus we have $\big|\mathcal{H}_r^\epsilon(x,|\nabla u_\delta^{\epsilon_{ij}}(\cdot,t)|)\frac{\nabla u_\delta^{\epsilon_{ij}}(\cdot,t)}{|\nabla u_\delta^{\epsilon_{ij}}(\cdot,t)|}\big| \in W^{1,2}(\Omega)$ for each t. Moreover, since $v(\cdot,t) - u_\delta^{\epsilon_{ij}}(\cdot,t) \in W_0^{1,2}(\Omega)$, then, for fixed t, applying integration by parts formula for Sobolev functions, from (2.1.6), to the integral on the right above, we obtain

$$\begin{aligned}&\int_0^s \int_\Omega (u_\delta^{\epsilon_{ij}})_t (v - u_\delta^{\epsilon_{ij}}) - \epsilon_{i_j}\Delta u_\delta^{\epsilon_{ij}}(v - u_\delta^{\epsilon_{ij}})\,dx\,dt\\ &\qquad= \int_0^s \int_\Omega \big(- \mathcal{H}_r^\epsilon(x,|\nabla u_\delta^{\epsilon_{ij}}|)\frac{\nabla u_\delta^{\epsilon_{ij}}}{|\nabla u_\delta^{\epsilon_{ij}}|} \cdot (\nabla v - \nabla u_\delta^{\epsilon_{ij}})\big)\,dx\,dt \\ &\qquad\qquad - \int_0^s \int_\Omega \lambda(u_\delta^{\epsilon_{ij}} - f_\delta)(v - u_\delta^{\epsilon_{ij}})\,dx\,dt.\end{aligned} \tag{4.4.44}$$

Since $\mathcal{H}^\epsilon(x,|\nabla u_\delta^\epsilon|)$ is convex with respect to the second variable, that is $|\nabla u_\delta^\epsilon|$, then the relation $\mathcal{H}_r^\epsilon(x,|\nabla u_\delta^{\epsilon_{ij}}|)(|\nabla v| - |\nabla u_\delta^{\epsilon_{ij}}|) \leqslant \mathcal{H}^\epsilon(x,|\nabla v|) - \mathcal{H}^\epsilon(x,|\nabla u_\delta^{\epsilon_{ij}}|)$ holds true. First, applying absolute value both sides, we compute

$$\begin{aligned}\mathcal{H}_r^\epsilon(x,|\nabla u_\delta^{\epsilon_{ij}}|)\frac{\nabla u_\delta^{\epsilon_{ij}}}{|\nabla u_\delta^{\epsilon_{ij}}|} \cdot (\nabla v - \nabla u_\delta^{\epsilon_{ij}}) &\leqslant \mathcal{H}_r^\epsilon(x,|\nabla u_\delta^{\epsilon_{ij}}|)\Big(\frac{|\nabla u_\delta^{\epsilon_{ij}} \cdot \nabla v|}{|\nabla u_\delta^{\epsilon_{ij}}|} - \frac{|\nabla u_\delta^{\epsilon_{ij}}|^2}{|\nabla u_\delta^{\epsilon_{ij}}|}\Big)\\ &\leqslant \mathcal{H}_r^\epsilon(x,|\nabla u_\delta^{\epsilon_{ij}}|)(|\nabla v| - |\nabla u_\delta^{\epsilon_{ij}}|),\end{aligned}$$

and then applying the convexity relation of $\mathcal{H}^\epsilon$ to obtain,

$$\mathcal{H}_r^\epsilon(x,|\nabla u_\delta^{\epsilon_{ij}}|)\frac{\nabla u_\delta^{\epsilon_{ij}}}{|\nabla u_\delta^{\epsilon_{ij}}|} \cdot (\nabla v - \nabla u_\delta^{\epsilon_{ij}}) \leqslant \mathcal{H}^\epsilon(x,|\nabla v|) - \mathcal{H}^\epsilon(x,|\nabla u_\delta^{\epsilon_{ij}}|). \tag{4.4.45}$$

On the other hand, we have

$$\begin{aligned}\lambda(u_\delta^{\epsilon_{ij}} - f_\delta)(v - u_\delta^{\epsilon_{ij}}) &= \lambda(u_\delta^{\epsilon_{ij}} - f_\delta)(-u_\delta^{\epsilon_{ij}} + f_\delta + v - f_\delta) \\ &= -\lambda(u_\delta^{\epsilon_{ij}} - f_\delta)^2 + \lambda(u_\delta^{\epsilon_{ij}} - f_\delta)(v - f_\delta) \\ &\leqslant -\lambda(u_\delta^{\epsilon_{ij}} - f_\delta)^2 + \frac{\lambda}{2}(u_\delta^{\epsilon_{ij}} - f_\delta)^2 + \frac{\lambda}{2}(v - f_\delta)^2,\end{aligned}$$

and thus,

$$-\lambda(u_\delta^{\epsilon_{ij}} - f_\delta)(v - u_\delta^{\epsilon_{ij}}) \geqslant \frac{\lambda}{2}(u_\delta^{\epsilon_{ij}} - f_\delta)^2 - \frac{\lambda}{2}(v - f_\delta)^2. \tag{4.4.46}$$

So applying the inequalities (4.4.45) and (4.4.46) in (4.4.44) to obtain,

$$\begin{aligned}\int_0^s \int_\Omega (u_\delta^{\epsilon_{ij}})_t(v - u_\delta^{\epsilon_{ij}}) - \epsilon_{ij}\Delta u_\delta^{\epsilon_{ij}}(v - u_\delta^{\epsilon_{ij}})\,dx\,dt \geqslant \int_0^s \int_\Omega \big(-\mathcal{H}^\epsilon(x, |\nabla v|) \\ + \mathcal{H}^\epsilon(x, |\nabla u_\delta^\epsilon|) + \frac{\lambda}{2}(u_\delta^{\epsilon_{ij}} - f_\delta)^2 - \frac{\lambda}{2}(v - f_\delta)^2\big)\,dx\,dt,\end{aligned}$$

that is,

$$\begin{aligned}\int_0^s \int_\Omega (u_\delta^\epsilon)_t(v - u_\delta^{\epsilon_{ij}}) - \epsilon_{ij}\Delta u_\delta^{\epsilon_{ij}}(v - u_\delta^{\epsilon_{ij}}) + \mathcal{H}^\epsilon(x, |\nabla v|) + \frac{\lambda}{2}(v - f_\delta)^2\,dx\,dt \\ \geqslant \int_0^s \int_\Omega \mathcal{H}^\epsilon(x, |\nabla u_\delta^{\epsilon_{ij}}|) + \frac{\lambda}{2}(u_\delta^{\epsilon_{ij}} - f_\delta)^2\,dx\,dt.\end{aligned}$$

Further, since $\mathcal{H}^\epsilon \geqslant \mathcal{H}$ from (4.4.16), then the previous inequality implies,

$$\begin{aligned}\int_0^s \int_\Omega (u_\delta^{\epsilon_{ij}})_t(v - u_\delta^{\epsilon_{ij}}) - \epsilon_{ij}\Delta u_\delta^{\epsilon_{ij}}(v - u_\delta^{\epsilon_{ij}}) + \mathcal{H}^\epsilon(x, |\nabla v|) + \frac{\lambda}{2}(v - f_\delta)^2\,dx\,dt \\ \geqslant \int_0^s \int_\Omega \mathcal{H}(x, |\nabla u_\delta^{\epsilon_{ij}}|) + \frac{\lambda}{2}(u_\delta^{\epsilon_{ij}} - f_\delta)^2\,dx\,dt.\end{aligned} \tag{4.4.47}$$

Next, for all $v \in L^2(0, T; W^{1,2}(\Omega)) \cap C^\infty(\Omega_T)$, applying the dominated convergence theorem, we obtain,

$$\begin{aligned}\lim_{\epsilon\to 0} \int_0^s \int_\Omega \mathcal{H}^\epsilon(x, |\nabla v|)\,dx\,dt &= \int_0^s \int_\Omega \lim_{\epsilon\to 0} \mathcal{H}^\epsilon(x, |\nabla v|)\,dx\,dt \\ &= \int_0^s \int_\Omega \mathcal{H}(x, |\nabla v|)\,dx\,dt,\end{aligned} \tag{4.4.48}$$

since the integral $\int_\Omega \mathcal{H}^\epsilon(x, |\nabla v|)\,dx$ is bounded with respect to $\epsilon < 1$, and $\mathcal{H}^\epsilon$ is bounded by $2\mathcal{H}$ and a constant, from (4.4.17).

On the other hand, from weak-L^2 convergence of $(u_\delta^{\epsilon_{ij}})_t$ in (4.4.38), we have

$$\int_0^s \int_\Omega (u_\delta^{\epsilon_{ij}})_t v\,dx\,dt \to \int_0^s \int_\Omega (u_\delta)_t v\,dx\,dt,$$

as $\epsilon_i \to 0$.

Due to the uniform boundedness of $\{u_\delta^{\epsilon_i}(\cdot,t)\}$ in $W^{1,\mathcal{H}}(\Omega)$ with respect to t from (4.4.40), there exists a subsequence of $\{u_\delta^{\epsilon_{ij}}(\cdot,t)\}$, denoted by same $\{u_\delta^{\epsilon_{ij}}(\cdot,t)\}_{j=1}^{\infty}$, converging weakly in $W^{1,\mathcal{H}}(\Omega)$ and weakly* in $L^\infty(\Omega)$, to the limit $u_\delta(\cdot,t) \in W^{1,\mathcal{H}}(\Omega)\cap L^\infty(\Omega)$, as $\epsilon_{ij} \to 0$. Then using weak lower semicontinuity from Lemma 4.2.3, we have,

$$\liminf_{\epsilon_{ij}\to 0} \int_\Omega \mathcal{H}(x, |\nabla u_\delta^{\epsilon_{ij}}(\cdot,t)|)\,dx \geqslant \int_\Omega \mathcal{H}(x,|\nabla u_\delta(\cdot,t)|)\,dx. \tag{4.4.49}$$

On the other hand, applying the embedding property $L^2(\Omega) \hookrightarrow L^{\mathcal{H}}(\Omega)$ [4, Lemma 2.5] for $u_\delta^{\epsilon_i}(\cdot,t) - f_\delta \in W_0^{1,2}(\Omega)$, yields $u_\delta^{\epsilon_i}(\cdot,t) - f_\delta \in W_0^{1,\mathcal{H}}(\Omega)$ for every t. Further, since $\{u_\delta^{\epsilon_i}(\cdot,t) - f_\delta\}$ is uniformly bounded in $W^{1,\mathcal{H}}(\Omega)$ with respect to t, by (4.4.41), then taking a subsequence of $\{u_\delta^{\epsilon_{ij}}(\cdot,t) - f_\delta\}$ converging weakly to $u_\delta(\cdot,t) - f_\delta$ in $W^{1,\mathcal{H}}(\Omega)$ as $i_j \to \infty$, we apply Lemma 4.2.2 to obtain that the limit $u_\delta(\cdot,t) - f_\delta \in W_0^{1,\mathcal{H}}(\Omega)$, for each $t > 0$.

Now taking the limit as $\epsilon_i \to 0$ in (4.4.47), we apply the convergence results (4.4.38), (4.4.43) and (4.4.48) to obtain,

$$\begin{aligned}\int_0^s \int_\Omega (u_\delta)_t(v-u_\delta) + \mathcal{H}(x,|\nabla v|) + \frac{\lambda}{2}(v-f_\delta)^2\,dx\,dt \\ \geqslant \lim_{\epsilon_{ij}\to 0}\Big(\int_0^s\int_\Omega \mathcal{H}(x,|\nabla u_\delta^{\epsilon_{ij}}|)\,dx\,dt + \int_0^s\int_\Omega \frac{\lambda}{2}(u_\delta^{\epsilon_{ij}} - f_\delta)^2\,dx\,dt\Big).\end{aligned} \tag{4.4.50}$$

Applying boundedness property of $u_\delta^{\epsilon_{ij}}$ in $W^{1,\mathcal{H}}(\Omega)$ from (4.4.40), and boundedness of $\int_\Omega \frac{\lambda}{2}(u_\delta^{\epsilon_{ij}} - f_\delta)^2\,dx$ with respect to $\epsilon_{ij}\|\nabla f_\delta\|^2_{L^2(\Omega)}$, from the estimate (4.4.31), we have

$$\int_\Omega \mathcal{H}(x,|\nabla u_\delta^{\epsilon_{ij}}|)\,dx\,dt + \int_\Omega \frac{\lambda}{2}(u_\delta^{\epsilon_{ij}} - f_\delta)^2\,dx \leqslant \epsilon_{ij}\|\nabla f_\delta\|^2_{L^2(\Omega)} + C,$$

where constant C depends on Ω. Then taking limit on both sides as $\epsilon_{ij} \to 0$, the above implies,

$$\lim_{\epsilon_{ij}\to 0}\Big(\int_\Omega \mathcal{H}(x,|\nabla u_\delta^{\epsilon_{ij}}|)\,dx\,dt + \int_\Omega \frac{\lambda}{2}(u_\delta^{\epsilon_{ij}} - f_\delta)^2\,dx\Big) \leqslant C.$$

With constant C as the majorant of the right-hand side integral above, we apply the dominated convergence theorem to interchange the order of limit and integral in (4.4.50), and further apply (4.4.49) and (4.4.42) to get,

$$\begin{aligned}\int_0^s \int_\Omega (u_\delta)_t(v-u_\delta) + \mathcal{H}(x,|\nabla v|) + \frac{\lambda}{2}(v-f_\delta)^2\,dx\,dt \\ \geqslant \int_0^s\int_\Omega \mathcal{H}(x,|\nabla u_\delta|) + \frac{\lambda}{2}(u_\delta - f_\delta)^2\,dx\,dt,\end{aligned}$$

that is,

$$\int_0^s \int_\Omega (u_\delta)_t (v - u_\delta)\, dx\, dt + \int_0^s E(v)\, dt \geqslant \int_0^s E(u_\delta)\, dt, \tag{4.4.51}$$

holds for all $v \in L^2(0,T; W^{1,2}(\Omega)) \cap C^\infty(\Omega_T)$, such that $v(\cdot,t) - f_\delta \in W_0^{1,2}(\Omega)$, for each $t > 0$.
Next we prove that (4.4.51) holds for any test function in $L^2(0,T; W^{1,\mathcal{H}}(\Omega))$ with similar boundary condition as u_δ.

First, let $\tilde{f}_\delta(x,t) := f_\delta(x)$ for every t, then we have $\int_0^T \|\tilde{f}_\delta(\cdot,t)\|_{W^{1,2}(\Omega)}\, dt = T\|f_\delta\|_{W^{1,2}(\Omega)}$. So for $v(\cdot,t) - f_\delta \in W_0^{1,2}(\Omega)$, we have $v - \tilde{f}_\delta \in L^2(0,T; W_0^{1,2}(\Omega))$. Assume that $\bar{v} \in L^2(0,T; W^{1,\mathcal{H}}(\Omega))$, such that $\bar{v}(\cdot,t) - \tilde{f}_\delta(\cdot,t) \in W_0^{1,\mathcal{H}}(\Omega)$ for every t, and let $\xi := \bar{v} - \tilde{f}_\delta$.

Since $C_0^\infty(\Omega_T)$ is dense in $L^2(0,T; W_0^{1,\mathcal{H}}(\Omega))$ by Lemma 4.2.5, which implies that, any function in $L^2(0,T; W_0^{1,\mathcal{H}}(\Omega))$ can be approximated by a sequence of $C_0^\infty(\Omega_T)$ functions, say $\{\xi_j\}_{j=1}^\infty$. Thus, there exists a subsequence of ξ_j, denoted by the same ξ_j, such that, as $j \to \infty$, $\xi_j \to \xi$ in $L^2(0,T; W_0^{1,\mathcal{H}}(\Omega))$, that is,

$$\int_0^s \|\xi_j - \xi\|_{W^{1,\mathcal{H}}}^2\, dt \to 0. \tag{4.4.52}$$

Since $\mathcal{H}$ satisfies (aInc)$_p, p > 1$, and (aDec) properties, then using Lemma 2.2.4 to have

$$\int_0^s \|\xi_j - \xi\|_{W^{1,\mathcal{H}}}^2\, dt \geqslant \int_0^s (\min\{\big(\frac{1}{c}\varrho_{1,\mathcal{H}}(\xi_j - \xi)\big)^{\frac{1}{p}}, \big(\frac{1}{c}\varrho_{1,\mathcal{H}}(\xi_j - \xi)\big)^{\frac{1}{q}}\})^2\, dt, \tag{4.4.53}$$

where p and q come from the (aInc) and (aDec) conditions, while c is the maximum of the constants from these conditions. Further, the right-hand side of (4.4.53) implies,

$$\begin{aligned}
&\int_0^s (\min\{\big(\frac{1}{c}\varrho_{1,\mathcal{H}}(\xi_j - \xi)\big)^{\frac{1}{p}}, \big(\frac{1}{c}\varrho_{1,\mathcal{H}}(\xi_j - \xi)\big)^{\frac{1}{q}}\})^2\, dt \\
&\quad = \int_0^s \big(\frac{1}{c}\varrho_{1,\mathcal{H}}(\xi_j - \xi)\big)^{\frac{2}{p}} \chi_{\{\varrho_{1,\mathcal{H}}(\xi_j - \xi) < c\}} + \big(\frac{1}{c}\varrho_{1,\mathcal{H}}(\xi_j - \xi)\big)^{\frac{2}{q}} \chi_{\{\varrho_{1,\mathcal{H}}(\xi_j - \xi) \geqslant c\}}\, dt \\
&\quad \geqslant \frac{1}{c^{\frac{2}{p}}} \Big(\int_0^s \big(\varrho_{1,\mathcal{H}}(\xi_j - \xi)\big)^{\frac{2}{p}} \chi_{\{\varrho_{1,\mathcal{H}}(\xi_j - \xi) < c\}}\, dt\Big)^{\frac{p}{2}} \\
&\qquad\qquad + \frac{1}{c} \int_0^s \varrho_{1,\mathcal{H}}(\xi_j - \xi) \chi_{\{\varrho_{1,\mathcal{H}}(\xi_j - \xi) \geqslant c\}}\, dt.
\end{aligned}$$

Applying Hölder's inequality to have from

$$\begin{aligned}
&\int_0^s (\min\{\big(\tfrac{1}{c}\varrho_{1,\mathcal{H}}(\xi_j-\xi)\big)^{\frac{1}{p}}, \big(\tfrac{1}{c}\varrho_{1,\mathcal{H}}(\xi_j-\xi)\big)^{\frac{1}{q}}\})^2\,dt \\
&\geqslant \frac{1}{c^{\frac{2}{p}}s^{1-\frac{p}{2}}}\Big(\int_0^s \big(\varrho_{1,\mathcal{H}}(\xi_j-\xi)\big)^{\frac{2}{p}}\chi_{\{\varrho_{1,\mathcal{H}}(\xi_j-\xi)<c\}}\,dt\Big)^{\frac{p}{2}}\Big(\int_0^s dt\Big)^{1-\frac{p}{2}} \\
&\qquad\qquad\qquad + \frac{1}{c}\int_0^s \varrho_{1,\mathcal{H}}(\xi_j-\xi)\chi_{\{\varrho_{1,\mathcal{H}}(\xi_j-\xi)\geqslant c\}}\,dt \\
&\geqslant C\int_0^s \varrho_{1,\mathcal{H}}(\xi_j-\xi)\chi_{\{\varrho_{1,\mathcal{H}}(\xi_j-\xi)<c\}}\,dt + C\int_0^s \varrho_{1,\mathcal{H}}(\xi_j-\xi)\chi_{\{\varrho_{1,\mathcal{H}}(\xi_j-\xi)\geqslant c\}}\,dt
\end{aligned}$$

where $C = \min\{\frac{1}{c^{\frac{2}{p}}s^{1-\frac{p}{2}}}, \frac{1}{c}\}$. Thus, we obtain

$$\int_0^s (\min\{\big(\tfrac{1}{c}\varrho_{1,\mathcal{H}}(\xi_j-\xi)\big)^{\frac{1}{p}}, \big(\tfrac{1}{c}\varrho_{1,\mathcal{H}}(\xi_j-\xi)\big)^{\frac{1}{q}}\})^2\,dt \geqslant C\int_0^s \varrho_{1,\mathcal{H}}(\xi_j-\xi)\,dt.$$

Then, from (4.4.53) we get

$$\int_0^s \|\xi_j-\xi\|^2_{W^{1,\mathcal{H}}}\,dt \geqslant C\int_0^s \varrho_{1,\mathcal{H}}(\xi_j-\xi)\,dt. \tag{4.4.54}$$

Then, from (4.4.52), the above estimate (4.4.54) implies that $\int_0^s \varrho_{1,\mathcal{H}}(\xi_j-\xi)\,dt \to 0$, as $j\to\infty$. Since $\varrho_{1,\mathcal{H}}(\xi(\cdot,t))$ is bounded with respect to t, we apply Lemma 4.2.6 to conclude that, as $j\to\infty$,

$$\Big|\int_0^s (\varrho_{1,\mathcal{H}}(\xi_j)-\varrho_{1,\mathcal{H}}(\xi))\,dt\Big| \to 0,$$

from which we have,

$$\int_0^s\int_\Omega \mathcal{H}(x,|\nabla\xi_j|)\,dx\,dt \to \int_0^s\int_\Omega \mathcal{H}(x,|\nabla\xi|)\,dx\,dt, \quad \text{as } j\to\infty. \tag{4.4.55}$$

Using $W^{1,\mathcal{H}}(\Omega)\hookrightarrow L^2(\Omega)$ from Lemma 4.2.4, we obtain the embedding property $L^2(0,T;W^{1,\mathcal{H}}(\Omega))\hookrightarrow L^2(0,T;L^2(\Omega))$ from parabolic space properties in [63, Proposition 23.2(h)], and applying this in (4.4.52) implies

$$\int_0^s\int_\Omega \xi_j^2\,dx\,dt \to \int_0^s\int_\Omega \xi^2\,dx\,dt, \quad \text{as } j\to\infty. \tag{4.4.56}$$

With $(u_\delta)_t\in L^2(0,T;L^2(\Omega))$, we apply the above L^2-convergence to get,

$$\lim_{j\to\infty}\int_0^s\int_\Omega (u_\delta)_t\,\xi_j\,dx\,dt = \int_0^s\int_\Omega (u_\delta)_t\,\xi\,dx\,dt. \tag{4.4.57}$$

We have that (4.4.51) holds for test function $\xi_j + \tilde{f}_\delta$, that is

$$\int_0^s \int_\Omega \mathcal{H}(x,|\nabla u_\delta|) + \frac{\lambda}{2}|u_\delta - \tilde{f}_\delta|^2 \, dx \, dt \leqslant \int_0^s \int_\Omega (u_\delta)_t(\xi_j + \tilde{f}_\delta - u_\delta) + \mathcal{H}(x, |\nabla \xi_j + \nabla \tilde{f}_\delta|) + \frac{\lambda}{2}\xi_j^2 \, dx \, dt. \tag{4.4.58}$$

Now taking the limit as $j \to \infty$ in (4.4.58), we apply (4.4.55), (4.4.56) and (4.4.57), to get

$$\int_0^s \int_\Omega \mathcal{H}(x, |\nabla u_\delta|) + \frac{\lambda}{2}|u_\delta - \tilde{f}_\delta|^2 \, dx \, dt \leqslant \int_0^s \int_\Omega (u_\delta)_t(\xi + \tilde{f}_\delta - u_\delta) + \mathcal{H}(x, |\nabla \xi + \nabla \tilde{f}_\delta|) + \frac{\lambda}{2}\xi^2 \, dx \, dt.$$

Since $\xi = \bar{v} - \tilde{f}_\delta = \bar{v} - f_\delta$, then the above inequality implies

$$\int_0^s \int_\Omega \mathcal{H}(x, |\nabla u_\delta|) + \frac{\lambda}{2}|u_\delta - f_\delta|^2 \, dx \, dt \leqslant \int_0^s \int_\Omega (u_\delta)_t(\bar{v} - u_\delta) + \mathcal{H}(x, |\nabla \bar{v}|) + \frac{\lambda}{2}(\bar{v} - f_\delta)^2 \, dx \, dt, \tag{4.4.59}$$

which holds for all $\bar{v} \in L^2(0,T; W^{1,\mathcal{H}}(\Omega))$, with $\bar{v}(\cdot,t) - f_\delta \in L^2(0,T; W_0^{1,\mathcal{H}}(\Omega))$. Hence, by definition of weak solutions 4.4.2, $u_\delta \in L^2(0,T; W^{1,\mathcal{H}}(\Omega) \cap L^\infty(\Omega))$ is a weak solution of (4.4.12)–(4.4.13), such that, $u_\delta(\cdot,t) - f_\delta \in W_0^{1,\mathcal{H}}(\Omega)$ for each $t > 0$, and $u_\delta(x,0) = f_\delta(x)$.

Step 2: Passing $\delta \to 0$, to obtain the weak solution formulation. Now we pass $\delta \to 0^+$ in (4.4.59), to complete the proof.
First note that Lemma 4.4.4 holds for $u_\delta^{\epsilon_i}$, so fixing $\delta > 0$ and taking $\epsilon_i \to 0$ in (4.4.33), we apply (4.4.42), (4.4.38) to obtain

$$\int_0^T \int_\Omega |(u_\delta)_t|^2 \, dx \, dt + \sup_{s \in (0,T]} \left(\int_\Omega \mathcal{H}(x, |\nabla u_\delta(\cdot,s)|) + \frac{\lambda}{2}(u_\delta(\cdot,s) - f_\delta)^2 \, dx \right) \leqslant C \tag{4.4.60}$$

where the constant $C > 0$ is independent of δ and ϵ_i, and depends only on Ω. This implies that u_δ is in $L^\infty(0,T; W^{1,\mathcal{H}}(\Omega))$, with bounded constant independent of t.

On the other hand, taking $\epsilon_i \to 0$ in (4.4.35), we apply (4.4.37) to obtain,

$$\|u_\delta\|_{L^\infty(\Omega_T)} \leqslant \liminf_{\epsilon_i \to 0} \|u_\delta^{\epsilon_i}\|_{L^\infty(\Omega_T)} \leqslant \|f\|_{L^\infty(\mathbb{R}^n)}, \tag{4.4.61}$$

where the bound is independent of both δ and ϵ_i.

Now taking $\delta = \delta_i, i \in \mathbb{N}$, we have from (4.4.60), that $\{(u_{\delta i})_t\}$ is uniformly bounded in $L^2(\Omega_T)$. While, (4.4.61) implies that $\{u_{\delta i}\}$ is uniformly bounded in $L^\infty(\Omega_T)$. Also, from (4.4.60), for fixed $s \in (0, T]$, $\varrho_{\mathcal{H}}(\nabla u_{\delta i}(\cdot, s)) \leqslant C$ and $\|u_{\delta i}(\cdot, s) - f_{\delta i}\|_2 \leqslant C$. Again, using the embedding $L^2 \hookrightarrow L^{\mathcal{H}}$ [4, Lemma 2.5], we obtain, for fixed s, $\|u_{\delta i}(\cdot, s) - f_{\delta i}\|_{\mathcal{H}} \leqslant C$, and further $\|u_{\delta i}(\cdot, s)\|_{\mathcal{H}} \leqslant C$. Hence, by properties of the norm, we get $\varrho_{\mathcal{H}}(u_{\delta i}(\cdot, s)) \leqslant C$ and $\varrho_{\mathcal{H}}(u_{\delta i}(\cdot, s) - f_{\delta i}) \leqslant C$.

While, with $|\nabla(u_{\delta i} - f_{\delta i})| \leqslant |\nabla u_{\delta i}| + |\nabla f_{\delta i}|$, applying the doubling property of $\mathcal{H}$ and integrating over Ω to get $\varrho_{\mathcal{H}}(\nabla(u_{\delta i}(\cdot, s) - f_{\delta i})) \leqslant C$. Thus, for fixed $s > 0$, $\{u_{\delta i}(\cdot, s)\}$ as well as $\{u_{\delta i}(\cdot, s) - f_{\delta i}\}$ are bounded sequences in $W^{1,\mathcal{H}}(\Omega)$.

Next, we follow the same arguments used to obtain (4.4.37), (4.4.38), (4.4.42) and (4.4.49), to conclude that, for each t, there exists a subsequence $\{u_{\delta ij}(\cdot, t)\}_{j=1}^{\infty} \subset \{u_{\delta i}(\cdot, t)\}$ and a function $u \in L^\infty(0, T; W^{1,\mathcal{H}}(\Omega) \cap L^\infty(\Omega))$ with $u_t \in L^2(\Omega_T)$ such that as $\delta_{ij} \to 0$,

$$u_{\delta ij} \rightharpoonup u \text{ weakly* in } L^\infty(\Omega_T) \tag{4.4.62}$$

$$\left(u_{\delta ij}\right)_t \rightharpoonup u_t \text{ weakly in } L^2(\Omega_T) \tag{4.4.63}$$

$$u_{\delta ij}(\cdot, t) \to u(\cdot, t) \text{ strongly in } L^2(\Omega) \tag{4.4.64}$$

$$\liminf_{\delta_i \to 0} \int_\Omega \mathcal{H}(x, |\nabla u_{\delta ij}|)\, dx \geqslant \int_\Omega \mathcal{H}(x, |\nabla u|)\, dx, \tag{4.4.65}$$

and $u(x, 0) = \lim_{\delta_i \to 0} f_{\delta i}(x)$, follows from A.4 as well. Moreover, since $q < \frac{3}{2}$ which implies $\mathcal{H}$ satisfies $(A1)$ condition [24, Proposition 7.2.2], then using (4.4.19) we have $f_{\delta i} \to f$ in $W^{1,\mathcal{H}}(\Omega)$, as $\delta_i \to 0$. Hence we obtain $u(x, 0) = f(x)$. Also, since $u_{\delta i}(\cdot, t) - f_{\delta i} \in W_0^{1,\mathcal{H}}(\Omega)$ for each t, then by Lemma 4.2.2, we have $u(\cdot, t) - f \in W_0^{1,\mathcal{H}}(\Omega)$.

On the other hand, using the uniform boundedness of $\{u_{\delta_i}(\cdot, t)\}$ in $W^{1,\mathcal{H}}(\Omega)$ with respect to t, and the pointwise estimate (4.4.65), we obtain

$$\begin{aligned} \liminf_{\delta_i \to 0} \int_0^s \int_\Omega \mathcal{H}(x, |\nabla u_{\delta ij}|)\, dx &= \int_0^s \liminf_{\delta_i \to 0} \int_\Omega \mathcal{H}(x, |\nabla u_{\delta ij}|)\, dx\, dt \\ &\geqslant \int_0^s \int_\Omega \mathcal{H}(x, |\nabla u|)\, dx\, dt. \end{aligned} \tag{4.4.66}$$

Now, using weak-L^2 convergence of $(u_{\delta ij})_t$ from (4.4.63) and strong-L^2 convergence of $u_{\delta i}(\cdot, t)$ from (4.4.64), we follow a similar approach used to obtain (4.4.43), to have,

$$\int_\Omega \left(u_{\delta ij}(\cdot, t)\right)_t u_{\delta ij}(\cdot, t)\, dx \to \int_\Omega u_t(\cdot, t) u(\cdot, t)\, dx, \quad \text{as } \delta_{ij} \to 0. \tag{4.4.67}$$

Since $f_{\delta i} \to f$ in $W^{1,\mathcal{H}}(\Omega)$, as $\delta_i \to 0$, by (4.4.19), then we apply the embedding $W^{1,\mathcal{H}}(\Omega) \hookrightarrow L^2(\Omega)$, from Lemma 4.2.4, to obtain

$$\|f_{\delta ij}\|_2 \to \|f\|_2 \quad \text{as } \delta_{ij} \to 0. \tag{4.4.68}$$

Consider $\delta = \delta_{ij}$ in (4.4.59). Then taking the limit as $\delta_i \to 0$ in (4.4.59), we apply the convergence criteria (4.4.62)–(4.4.68), to obtain that, for all $\bar{v} \in L^2(0,T; W^{1,\mathcal{H}}(\Omega))$,

$$\int_0^s \int_\Omega u_t(\bar{v}-u) + \mathcal{H}(x, |\nabla \bar{v}|) + \frac{\lambda}{2}(\bar{v}-f)^2 \, dx \, dt \geqslant \int_0^s \int_\Omega \mathcal{H}(x, |\nabla u|) + \frac{\lambda}{2}(u-f)^2 \, dx \, dt$$

that is,

$$\int_0^s \int_\Omega u_t(\bar{v}-u) \, dx \, dt + \int_0^s E(\bar{v}) \, dt \geqslant \int_0^s E(u) \, dt. \tag{4.4.69}$$

Additionally, since we have $u(\cdot, t) - f \in W_0^{1,\mathcal{H}}(\Omega)$ for each $t > 0$, and $u(x,0) = f(x)$, hence, by definition 4.4.2, we conclude that $u \in L^\infty(0,T; W^{1,\mathcal{H}}(\Omega) \cap L^\infty(\Omega))$ is weak solution of the heat flow problem (4.4.12)–(4.4.13).
Uniqueness: To prove the uniqueness of the solution, suppose that $u_1(x,t)$ and $u_2(x,t)$ are weak solutions of (4.4.12)-(4.4.13). Here we can obtain two inequalities from (4.4.69): the first one by setting $u = u_1$ and $v = u_2$ in (4.4.69) as follows:

$$\int_0^s \int_\Omega (u_1)_t(u_2-u_1) \, dx \, dt + \int_0^s E(u_2) \, dt \geqslant \int_0^s E(u_1) \, dt \tag{4.4.70}$$

and the second one by setting $u = u_2$ and $v = u_1$ in (4.4.69) as follows:

$$\int_0^s \int_\Omega (u_2)_t(u_1-u_2) \, dx \, dt + \int_0^s E(u_1) \, dt \geqslant \int_0^s E(u_2) \, dt. \tag{4.4.71}$$

For all $s > 0$, adding the two inequalities (4.4.70) and (4.4.71) gives,

$$\int_0^s \int_\Omega ((u_1)_t - (u_2)_t)(u_1-u_2) \, dx \, dt \leqslant 0$$

so, we have,

$$\int_0^s \int_\Omega \frac{1}{2} \frac{d}{dt} |u_1 - u_2|^2 \, dx \, dt \leqslant 0.$$

Then applying Fubini's Theorem for interchanging the order of integration above, and using the initial conditions, $u_1(x,0) = u_2(x,0) = f(x)$, to obtain

$$\int_\Omega |u_1(s) - u_2(s)|^2 \, dx \leqslant 0$$

for every $s \in (0,T]$. Thus, for $T > 0$, we have $\int_\Omega |u_1 - u_2|^2 \, dx = 0$, which implies $u_1 = u_2$ almost everywhere in $\Omega \times [0, \infty)$, and hence the solution is unique. $\square$

Next, we discuss steady state of the solution to the boundary value problem (4.4.12)–(4.4.13). A steady state solution for a differential equation is defined as the solution where the value of the function do not change over time, that is, a time independent function. This is obtained by setting the partial derivative(s) with respect to t in the partial differential equation to constant zero, and then solving the equation for a function that depends only on the spatial variable x. With such approach, we now prove the stability of the solution for the heat flow problem in the following result.

***Theorem* 4.4.7.** *Let $\Omega \subset \mathbb{R}^2$ be a bounded open set with Lipschitz boundary, and $\mathcal{H} : \Omega \times [0, \infty) \to [0, \infty)$, as defined from* (4.2.1) *where $1 < p \leqslant q \leqslant 2$ such that $q < \frac{3}{2}p$. Then, for any $T > 0$,*

$$\|u_1 - u_2\|_{L^\infty(\Omega_T)} \leqslant \|f_1 - f_2\|_{L^\infty(\mathbb{R}^2)}$$

where u_1 and u_2 are weak solutions of the boundary problem (4.4.12)–(4.4.13) *with initial values $f_1, f_2 \in W^{1,\mathcal{H}}(\mathbb{R}^2) \cap L^\infty(\mathbb{R}^2)$.*

Proof. Setting $k := \|f_1 - f_2\|_{L^\infty(\mathbb{R}^2)}$, we define,

$$v := u_1 - (u_1 - u_2 - k)_+, \quad w := u_2 + (u_1 - u_2 - k)_+, \tag{4.4.72}$$

where,

$$(u_1 - u_2 - k)_+ := \begin{cases} u_1 - u_2 - k, & \text{if } u_1 - u_2 - k \geqslant 0 \\ 0, & \text{otherwise.} \end{cases} \tag{4.4.73}$$

From Theorem 4.4.6, the weak solutions $u_1, u_2 \in L^\infty(0, T; W^{1,\mathcal{H}}(\Omega) \cap L^\infty(\Omega))$ such that $u_1(\cdot, t) - f_1 \in W_0^{1,\mathcal{H}}(\Omega)$ and $u_2(\cdot, t) - f_2 \in W_0^{1,\mathcal{H}}(\Omega)$ for each t, with $(u_1)_t, (u_2)_t \in L^2(\Omega_T)$. Then for the above expressions, by chain rule we get $v, w \in L^2(0, T; W^{1,\mathcal{H}}(\Omega) \cap L^\infty(\Omega))$. Moreover, we have $(u_1 - u_2 - k)_+ \in W_0^{1,\mathcal{H}}(\Omega)$, then v and w have the same boundary value as u_1 and u_2 in Sobolev sense, that is, f_1 and f_2, respectively. Further, we estimate the following using (4.4.72) and (4.4.73),

$$\nabla v = \begin{cases} \nabla u_2, & u_1 - u_2 \geqslant k \\ \nabla u_1, & u_1 - u_2 < k \end{cases}, \qquad \nabla w = \begin{cases} \nabla u_1, & u_1 - u_2 \geqslant k \\ \nabla u_2, & u_1 - u_2 < k. \end{cases} \tag{4.4.74}$$

Now, u_1 and u_2 being weak solutions, then by definition of weak solution, we have that, for each $t > 0$,

$$\int_0^s \int_\Omega (u_1)_t (v - u_1) + \mathcal{H}(x, |\nabla v|) + \frac{\lambda}{2}(v - f_1)^2 \, dx \, dt$$
$$\geqslant \int_0^s \int_\Omega \mathcal{H}(x, |\nabla u_1|) + \frac{\lambda}{2}(u_1 - f_1)^2 \, dx \, dt,$$

and,

$$\int_0^s \int_\Omega (u_2)_t(w-u_2) + \mathcal{H}(x,|\nabla w|) + \frac{\lambda}{2}(w-f_2)^2\,dx\,dt \geqslant \int_0^s \int_\Omega \mathcal{H}(x,|\nabla u_2|) + \frac{\lambda}{2}(u_2-f_2)^2\,dx\,dt.$$

Summing up the above two inequalities implies,

$$\int_0^s \int_\Omega (u_1)_t(v-u_1) + (u_2)_t(w-u_2) + \mathcal{H}(x,|\nabla v|) + \mathcal{H}(x,|\nabla w|) + \frac{\lambda}{2}(v-f_1)^2 + \frac{\lambda}{2}(w-f_2)^2\,dx\,dt \geqslant \int_0^s \int_\Omega \mathcal{H}(x,|\nabla u_1|) + \mathcal{H}(x,|\nabla u_2|) + \frac{\lambda}{2}(u_1-f_1)^2 + \frac{\lambda}{2}(u_2-f_2)^2\,dx\,dt. \tag{4.4.75}$$

From (4.4.74), we have

$$\mathcal{H}(x,|\nabla v|) + \mathcal{H}(x,|\nabla w|) = \begin{cases} \mathcal{H}(x,|\nabla u_2|) + \mathcal{H}(x,|\nabla u_1|), & \text{for } u_1-u_2 \geqslant k \\ \mathcal{H}(x,|\nabla u_1|) + \mathcal{H}(x,|\nabla u_2|), & \text{for } u_1-u_2 < k \end{cases}$$

and, hence

$$\mathcal{H}(x,|\nabla v|) + \mathcal{H}(x,|\nabla w|) = \mathcal{H}(x,|\nabla u_1|) + \mathcal{H}(x,|\nabla u_2|), \tag{4.4.76}$$

holds almost everywhere. On the other hand, consider,

$$\int_\Omega ((u_1-f_1)^2 - (v-f_1)^2) + ((u_2-f_2)^2 - (w-f_2)^2)\,dx = \int_\Omega (u_1-v)(u_1+v-2f_1) + (u_2-w)(u_2+w-2f_2)\,dx$$

where we substitute the values of v and w from (4.4.72) to obtain,

$$\begin{aligned}\int_\Omega &(u_1-f_1)^2 - (v-f_1)^2 + (u_2-f_2)^2 - (w-f_2)^2\,dx \\ &= \int_\Omega (u_1-u_2-k)_+(u_1+v-2f_1-u_2-w+2f_2)\,dx \\ &= \int_\Omega (u_1-u_2-k)_+(2u_1-2u_2-2(u_1-u_2-k)_+-2f_1+2f_2)\,dx,\end{aligned}$$

that is,

$$\int_\Omega (u_1-f_1)^2 + (u_2-f_2)^2 - (v-f_1)^2 - (w-f_2)^2\,dx = \int_\Omega 2(u_1-u_2-k)_+((u_1-u_2-k) - (u_1-u_2-k)_+ - (f_1-f_2-k))\,dx. \tag{4.4.77}$$

On the other hand, by the definition of k, we have $f_1 - f_2 - k = f_1 - f_2 - \|f_1 - f_2\|_{L^\infty(\Omega)} \leqslant f_1 - f_2 - |f_1 - f_2| \leqslant 0$, that is $f_1 - f_2 - k \leqslant 0$.
So, now, for values of $u_1 - u_2 < k$, since $(u_1 - u_2 - k)_+ = 0$ by (4.4.73), then the right-hand side expression of (4.4.77) implies,

$$\int_{\{u_1-u_2<k\}} 2(u_1 - u_2 - k)_+((u_1-u_2-k)-(u_1 - u_2 - k)_+-(f_1-f_2-k))\, dx = 0. \tag{4.4.78}$$

While, for $u_1 - u_2 \geqslant k$, since $(u_1 - u_2 - k)_+ = u_1 - u_2 - k$ by (4.4.73), then the right-hand side expression of (4.4.77) implies,

$$\begin{aligned}\int_{\{u_1-u_2\geqslant k\}} & 2(u_1 - u_2 - k)_+((u_1 - u_2 - k) - (u_1 - u_2 - k)_+ - (f_1 - f_2 - k))\, dx \\ & \geqslant \int_{\{u_1-u_2\geqslant k\}} 2(u_1 - u_2 - k)((u_1 - u_2 - k) - (u_1 - u_2 - k))\, dx = 0,\end{aligned}$$

that is,

$$\int_{\{u_1-u_2\geqslant k\}} 2(u_1 - u_2 - k)_+((u_1-u_2-k)-(u_1 - u_2 - k)_+-(f_1-f_2-k))\, dx \geqslant 0. \tag{4.4.79}$$

From (4.4.78) and (4.4.79), we conclude

$$\int_\Omega 2(u_1 - u_2 - k)_+((u_1 - u_2 - k) - (u_1 - u_2 - k)_+ - (f_1 - f_2 - k))\, dx \geqslant 0.$$

Then (4.4.77) implies,

$$\int_\Omega (u_1 - f_1)^2 + (u_2 - f_2)^2\, dx \geqslant \int_\Omega (v - f_1)^2 + (w - f_2)^2\, dx. \tag{4.4.80}$$

Now applying (4.4.76) and (4.4.80) in (4.4.75), to obtain

$$\int_0^s \int_\Omega (u_1)_t(v - u_1) + (u_2)_t(w - u_2)\, dx\, dt \geqslant 0.$$

Substituting the values of v and w gives,

$$\int_0^s \int_\Omega (-(u_1)_t(u_1 - u_2 - k)_+ + (u_2)_t(u_1 - u_2 - k)_+)\, dx\, dt \geqslant 0,$$

that is,

$$\int_0^s \int_\Omega ((u_1)_t - (u_2)_t)(u_1 - u_2 - k)_+\, dx\, dt \leqslant 0.$$

The above expression is of the form,

$$\int_0^s \int_\Omega \frac{d}{dt}|(u_1 - u_2 - k)_+|^2\, dx\, dt \leqslant 0.$$

Then, applying Leibniz rule to interchange the order of integration and differentiation above, gives

$$\int_0^s \frac{d}{dt}\left(\int_\Omega |(u_1 - u_2 - k)_+|^2\, dx\right) dt \leqslant 0.$$

and thus, integrating over the interval $[0, s]$ implies,

$$\begin{aligned}\int_\Omega |(u_1(x,s) - u_2(x,s) - k)_+|^2\, dx &\leqslant \int_\Omega |(u_1(x,0) - u_2(x,0) - k)_+|^2\, dx \\ &= \int_\Omega |(f_1 - f_2 - k)_+|^2\, dx.\end{aligned} \tag{4.4.81}$$

Since $f_1 - f_2 - k \leqslant 0$, so we have $(f_1 - f_2 - k)_+ = 0$, almost everywhere, then from (4.4.81), it follows,

$$\int_\Omega |(u_1 - u_2 - k)_+|^2\, dx \leqslant 0.$$

This gives us that $(u_1 - u_2 - k)_+ = 0$ a.e., which implies $u_1 - u_2 \leqslant k$.

Similarly, considering $(k - u_1 + u_2)_+$, we follow the same approach to obtain $u_1 - u_2 \geqslant -k$. Thus, we get $-k \leq u_1 - u_2 \leqslant k$, that is, $|u_1 - u_2| \leqslant k$, which gives,

$$k = \|f_1 - f_2\|_{L^\infty(\mathbb{R}^2)} \geqslant |u_1 - u_2| = \|u_1 - u_2\|_{L^\infty(\Omega_T)},$$

as required.

□

4.4.4 Behavior of the solution

***Theorem* 4.4.8.** *Let $\Omega \subset \mathbb{R}^2$ be a bounded open set with Lipschitz boundary and $\mathcal{H} : \Omega \times [0,\infty) \to [0,\infty)$, as defined in* (4.2.1) *where $1 < p \leqslant q \leqslant 2$ such that $q < \frac{3}{2}p$. Assume that $f \in W^{1,\mathcal{H}}(\mathbb{R}^2) \cap L^\infty(\mathbb{R}^2)$. Then, there exists a sequence of positive numbers $\{s_i\}, i = 1, 2, \ldots,$ and an integral average of weak solution, denoted as $u(x,t)$, of the heat flow problem* (4.4.12)–(4.4.13) *over the interval $[0, s_i]$, defined as*

$$w(x,s) := \frac{1}{s}\int_0^s u(x,t)\, dt, \qquad s = s_1, s_2, \ldots, \tag{4.4.82}$$

such that a subsequence of $w(x, s_i)$ strongly converges to the solution of the minimization problem (4.3.10) *in $L^2(\Omega)$, as $s_i \to \infty$.*

Proof. With $u \in L^\infty(0,T; W^{1,\mathcal{H}}(\Omega) \cap L^\infty(\Omega))$ as weak solution of the problem (4.4.12)–(4.4.13), we have, by definition 4.4.2, u satisfies (4.4.15) such that $u(\cdot,t) -$

$f \in W_0^{1,\mathcal{H}}(\Omega)$ for each $t > 0$. For any $v \in W^{1,\mathcal{H}}(\Omega)$, having the same boundary value as u, we have from (4.4.15) the weak solution formulation,

$$\int_0^s \int_\Omega u_t(x,t)(v(x) - u(x,t))\,dx\,dt + \int_0^s E(v)\,dt \geqslant \int_0^s E(u)\,dt.$$

Applying Fubini's Theorem to interchange the order of integration for the first integral, implies

$$\int_\Omega \int_0^s u_t(x,t)v(x)\,dt\,dx - \int_\Omega \int_0^s u_t(x,t)u(x,t)\,dt\,dx + E(v)\int_0^s dt \geqslant \int_0^s E(u)\,dt.$$

Since $u_t u = \frac{1}{2}\frac{d}{dt}(u^2)$ and $u(x,0) = f(x)$, then integrating with respect to t over the interval $[0,s]$ for the left hand side integrals, it follows,

$$\int_\Omega (u(\cdot,s) - f)v\,dx - \frac{1}{2}\int_\Omega (u^2(\cdot,s) - f^2)\,dx + s\int_\Omega (\mathcal{H}(x,|\nabla v|) + \frac{\lambda}{2}(v-f)^2)\,dx \geqslant \int_0^s \int_\Omega \mathcal{H}(x,|\nabla u|) + \frac{\lambda}{2}(u-f)^2\,dx\,dt. \tag{4.4.83}$$

Since $u \in L^\infty(0,\infty; W^{1,\mathcal{H}}(\Omega) \cap L^\infty(\Omega))$ by Theorem 4.4.6, so for each $s > 0$, using the convexity property of $\mathcal{H}$ and the definition of w from (4.4.82) to get

$$\varrho_{\mathcal{H}}(w(\cdot,s)) = \varrho_{\mathcal{H}}\big(\frac{1}{s}\int_0^s u\,dt\big) = \int_\Omega \mathcal{H}(x, \frac{1}{s}\int_0^s u\,dt)\,dx \leqslant \frac{1}{s}\int_\Omega \int_0^s \mathcal{H}(x,u)\,dt\,dx.$$

Interchanging the order of integration above and using the boundedness of $u(\cdot,t)$ in $W^{1,\mathcal{H}}(\Omega)$ to obtain,

$$\varrho_{\mathcal{H}}(w(\cdot,s)) \leqslant \frac{1}{s}\int_0^s \int_\Omega \mathcal{H}(x,u)\,dx\,dt = \frac{1}{s}\int_0^s \varrho_{\mathcal{H}}(u)\,dt \leqslant \frac{1}{s}\int_0^s C\,dt \leqslant C, \tag{4.4.84}$$

where the constant C depends on $s > 0$. Similarly, for each $s > 0$, we obtain $\varrho_{\mathcal{H}}(\nabla w(\cdot,s)) \leqslant C$, and hence we conclude that $\varrho_{1,\mathcal{H}}(w(\cdot,s)) \leqslant C$. Thus, $\|w(\cdot,s)\|_{W^{1,\mathcal{H}}(\Omega)} \leqslant C$. On the other hand, since $\|u(\cdot,t)\|_\infty \leqslant C$ for every t, we estimate

$$\|w(\cdot,s)\|_\infty = \|\frac{1}{s}\int_0^s u(x,t)\,dt\|_\infty \leqslant \frac{1}{s}\int_0^s \|u\|_\infty\,dt \leqslant \frac{1}{s}\int_0^s C\,dt \leqslant C, \tag{4.4.85}$$

where the constant C depends on $s > 0$.

Next, we compute estimates for $w(\cdot,s)$. Since $|\nabla w(x,s)| \leqslant \frac{1}{s}\int_0^s |\nabla u(x,t)|\,dt$, then increasing and convexity properties of $\mathcal{H}$ implies

$$\mathcal{H}(x,|\nabla w(\cdot,s)|) \leqslant \mathcal{H}(x, \frac{1}{s}\int_0^s |\nabla u|\,dt) \leqslant \frac{1}{s}\int_0^s \mathcal{H}(x,|\nabla u|)\,dt. \tag{4.4.86}$$

On the other hand, using Cauchy-Schwartz inequality, we have

$$(w(x,s)-f(x))^2 = \big(\frac{1}{s}\int_0^s u(x,t)\,dt - f(x)\big)^2 = \big(\frac{1}{s}\int_0^s (u(x,t)-f(x))\,dt\big)^2 \leqslant \frac{1}{s}\int_0^s (u(x,t)-f(x))^2\,dt. \tag{4.4.87}$$

Now, in (4.4.83), interchanging the order of integration by applying Fubini's theorem for the right hand side expression, we divide both the sides by s, to obtain

$$\frac{1}{s}\int_\Omega (u(x,s)-f(x))v(x)\,dx - \frac{1}{2s}\int_\Omega (u^2(x,s)-f^2(x))\,dx + \int_\Omega (\mathcal{H}(x,|\nabla v|) + \frac{\lambda}{2}(v-f)^2)\,dx \geqslant \int_\Omega \frac{1}{s}\big(\int_0^s \mathcal{H}(x,|\nabla u|)\,dt + \frac{\lambda}{2s}\int_0^s (u-f)^2\,dt\big)\,dx,$$

and applying (4.4.86) and (4.4.87) implies,

$$\int_\Omega \mathcal{H}(x,|\nabla w(x,s)|) + \frac{\lambda}{2}(w(x,s)-f(x))^2\,dx \leqslant \frac{1}{s}\int_\Omega (u(x,s)-f(x))v(x)\,dx - \frac{1}{2s}\int_\Omega (u^2(x,s)-f^2(x))\,dx + \int_\Omega \mathcal{H}(x,|\nabla v(x)|) + \frac{\lambda}{2}(v(x)-f(x))^2\,dx.$$

Further, using the estimate $|uv| \leqslant \frac{1}{2}(|u|^2+|v|^2)$ and $|fv| \leqslant \frac{1}{2}(|f|^2+|v|^2)$ on the left-hand side above implies,

$$\int_\Omega \mathcal{H}(x,|\nabla w(x,s)|) + \frac{\lambda}{2}(w(x,s)-f(x))^2\,dx \leqslant \frac{1}{2s}\int_\Omega (|(u(x,s)|^2 + |v(x)|^2 + |f(x)|^2 + |v(x)|^2)\,dx - \frac{1}{2s}\int_\Omega (u^2(x,s)-f^2(x))\,dx + \int_\Omega \mathcal{H}(x,|\nabla v(x)|) + \frac{\lambda}{2}(v(x)-f(x))^2\,dx.$$

that is,

$$\int_\Omega \mathcal{H}(x,|\nabla w(x,s)|) + \frac{\lambda}{2}(w(x,s)-f(x))^2\,dx \leqslant \frac{1}{s}\int_\Omega (|v(x)|^2 + |f(x)|^2)\,dx + \int_\Omega \mathcal{H}(x,|\nabla v(x)|) + \frac{\lambda}{2}(v(x)-f(x))^2\,dx. \tag{4.4.88}$$

Since $W^{1,\mathcal{H}}(\Omega) \hookrightarrow L^2(\Omega)$ from Lemma 4.2.4, we have $f \in L^2(\mathbb{R}^2)$ as square-integrable function, then $\int_\Omega (|v(x)|^2 + |f(x)|^2)\,dx$ is finite in the above expression.

Taking $s = s_i, i \in \mathbb{N}$, we have from (4.4.84) and (4.4.85) that the sequence $\{w(\cdot,s_i)\}_{i\in\mathbb{N}}$ is uniformly bounded in $W^{1,\mathcal{H}}(\Omega)$ and $L^\infty(\Omega)$, for each s_i. Therefore, there exists a subsequence $\{w(\cdot, s_{ij})\}_{j=1}^{\infty} \subset \{w(\cdot, s_i)\}$ which converges weakly in $W^{1,\mathcal{H}}(\Omega)$ and weakly* in $L^{\infty}(\Omega)$ to a function $\tilde{u} \in W^{1,\mathcal{H}}(\Omega) \cap L^{\infty}(\Omega)$, as $s_i \to \infty$. Applying weak lower semicontinuity from Lemma 4.2.3, it follows,

$$\liminf_{s_i \to \infty} \int_{\Omega} \mathcal{H}(x, |\nabla w(\cdot, s_{ij})|)\, dx \geqslant \int_{\Omega} \mathcal{H}(x, |\nabla \tilde{u}|)\, dx. \tag{4.4.89}$$

Moreover, from the embedding $W^{1,\mathcal{H}}(\Omega) \hookrightarrow W^{1,p}(\Omega)$ [24, Lemma 6.1.6], where $p > 1$ is from $(\text{aInc})_p$, it follows that $\{w(\cdot, s_i)\}$ is a bounded sequence in $W^{1,p}(\Omega)$ for each s_i. Further, since $W^{1,p}(\Omega)$ is compactly embedded in $L^{p*}(\Omega)$ for $n = 2$, where $p* < \frac{2p}{2-p}$, we conclude that $\{w(\cdot, s_{ij})\}$ converges strongly in $L^2(\Omega)$ space. Hence, as $s_{ij} \to \infty$,

$$w(\cdot, s_{ij}) \to \tilde{u} \quad \text{strongly in } L^2(\Omega). \tag{4.4.90}$$

Now consider $s = s_i, i \in \mathbb{N}$ in (4.4.88). Then taking limit as $s_i \to \infty$ in (4.4.88), and applying (4.4.89) and (4.4.90), we obtain

$$\int_{\Omega} \mathcal{H}(x, |\nabla \tilde{u}|) + \frac{\lambda}{2}(\tilde{u} - f)^2\, dx \leqslant \int_{\Omega} \mathcal{H}(x, |\nabla v|) + \frac{\lambda}{2}(v - f)^2\, dx,$$

that is,

$$E(\tilde{u}) \leqslant E(v),$$

which shows that $\tilde{u}$ is the minimizer of (4.3.4). □

4.5 Numerical methods and experimental results

Image restoration is an application of minimization problems, as introduced earlier, whose primary objective is removal of noise such that the features of the image are preserved. The approach involves the minimization of nonsmooth energy functionals for yielding quality restored results. Solving such functionals typically requires large number of iterations, and thus an efficient algorithm is preferable where the runtime is crucial. It is a classic problem in the field of image processing and computer vision, for which various methods involving variational and numerical approaches [16; 56], have been developed to address noise reduction, in the recent decades.

Nonsmooth optimization techniques to solve image restoration problems include the bundle methods [42]. The basic idea in bundle methods is to approximate the subdifferential [11] of an objective function with a bundle formed by collecting subgradients from previous iterations. This information is used to construct a model of the objective, which is utilized to determine a solution of the original problem. Among the bundle methods, Limited Memory Bundle Method (LMBM) [22; 34] is used for solving large-scale nonsmooth unconstrained optimization problems that guarantee global convergence. So, it is considered appropriate for image restoration

[33; 41], as digital images contains thousands of pixels to be processed in a short span. We apply this method to produce experimental results by solving the image restoration optimization problem through FORTRAN.

In this section, the primary goal is to present simple numerical results, to understand how the image restoration model works with noisy signal and images. However, we do not discuss the details of the optimization algorithm used, since we mainly deal with the illustration of the behavior of the image restoration model.

First, we give the discretized formulation of the image restoration model along with its discretized gradient, in the following subsection.

4.5.1 Discretization of the model In solving optimization problems, Euler made extensive use of the 'method of finite differences'. By replacing smooth curves by polygonal lines, he reduced the problem of finding extrema of a function to the problem of finding extrema of a function of N variables, and then he obtained exact solutions by passing to the limit as $N \to \infty$. The finite difference discretization method [17] is one of the simplest forms of discretization, which approximates the differential operators constituting an equation locally. In image processing, due to the digital structure of an image as a set of pixels uniformly distributed, finite difference approximation schemes are easy to implement to an image domain.

We assume that Ω is an open and bounded domain in $\mathbb{R}^n$, where $n \geqslant 1$ is the dimension of image. In one dimension, for signals, Ω is an interval in $\mathbb{R}$. In two dimensions, we deal with planar images and Ω is a rectangle in $\mathbb{R}^2$. Given an observed image f, possibly degraded by noise, and a balancing parameter $\lambda > 0$, we solve the corresponding optimization problem of the image restoration model from (4.3.3) as,

$$\min_{u \in W^{1,\mathcal{H}}(\Omega) \cap L^2(\Omega)} E(u) = \min_{u \in W^{1,\mathcal{H}}(\Omega) \cap L^2(\Omega)} \mathcal{F}(u) + \frac{\lambda}{2} \|u - f\|_2^2, \tag{4.5.1}$$

where $\mathcal{F}(u) = \int_\Omega |\nabla u|^p + a(x)|\nabla u|^q \, dx$, with $1 < p \leqslant q \leqslant 2$ and $a \geqslant 0$ is a regulatory component. The weak lower semicontinuity of the cost functional $E(u)$ in $W^{1,\mathcal{H}}(\Omega) \cap L^2(\Omega)$, yields the unique solvability of the above minimization problem in this space, as established in Theorem 4.3.2. The discretization of the above model is approached by finite difference method, using backward difference approximation.

For the numerical experiments, we take the case of one-dimensional signal $f : \Omega \to \mathbb{R}$, where $\Omega = (0, 1) \subset \mathbb{R}$. Consider the number of discretized points in Ω as $N \in \mathbb{N}$. We introduce the equidistributed grid points $(x_i)_{0 \leqslant i \leqslant N}$ given by $x_i := ih$, and $h := \frac{1}{N+1}$ denote the mesh step size. Also, let $u_i := u(x_i)$. For discretization, we fix N and assume u, f, a as N-dimensional vectors, where f is the given noisy signal and a is chosen appropriately. To perform the operation over the entire signal in Ω, we simply iterate over all the points indices x_i for $i = 0, 1, \ldots, N$, and thus determine the numerical value of u_i.

Let D denote the backward difference approximation operator of ∇, defined as

$$(Du)_i := \frac{1}{h}(u_i - u_{i-1}), \quad i = 1, \ldots, N, \tag{4.5.2}$$

with $u_0 = 0$, taken as a Dirichlet boundary condition. Further, we denote

$$|Du|^p := \sum_{i=1}^{N} |(Du)_i|^p \quad \text{and} \quad |Du|_a^q := \sum_{i=1}^{N} a_i |(Du)_i|^q.$$

The discrete approximation of the problem (4.5.1) is then defined as,

$$\min_{u \in \mathbb{R}^N} E(u) = \min_{u \in \mathbb{R}^N} \frac{\lambda}{2}(u - f)^T (u - f) + |Du|^p + |Du|_a^q, \tag{4.5.3}$$

where $(u - f)^T$ denotes the transpose of $(u - f)$.

Note that, for $1 < p \leqslant q \leqslant 2$, $E(u)$ is strictly convex and differentiable in $\mathbb{R}^n$, which implies that E has a global minimum value at the point where the gradient of E vanishes. So, next we arrive at the necessary and sufficient optimality condition for (4.5.3), by computing the gradient of $E(u)$.

We denote,

$$E(u) = e(u) + \mathcal{H}(u),$$

where $e(u) = \frac{\lambda}{2}(u - f)^T (u - f)$ and $\mathcal{H}(u) = |Du|^p + |Du|_a^q$. We then calculate the gradients of $e(u)$ and $\mathcal{H}(u)$ with respect to u. The gradient of $e(u)$ is given by,

$$\nabla e(u) = \lambda(u - f).$$

While, calculating the gradient of $\mathcal{H}(u)$ gives,

$$\frac{\partial}{\partial u_i}\mathcal{H}(u) = \frac{\partial}{\partial u_i}\big(|(Du)_i|^p + a_i |(Du)_i|^q + |(Du)_{i+1}|^p + a_{i+1}|(Du)_{i+1}|^q\big).$$

Plugging in the formula for backward difference operator D from (4.5.2), we have from above

$$\begin{aligned}
\frac{\partial}{\partial u_i}\mathcal{H}(u) &= \frac{\partial}{\partial u_i}\left(\Big|\frac{u_i - u_{i-1}}{h}\Big|^p + \Big|\frac{u_{i+1} - u_i}{h}\Big|^p + a_i\Big|\frac{u_i - u_{i-1}}{h}\Big|^q + a_{i+1}\Big|\frac{u_{i+1} - u_i}{h}\Big|^q\right) \\
&= p\Big|\frac{u_i - u_{i-1}}{h}\Big|^{p-1} \operatorname{sgn}\Big(\frac{u_i - u_{i-1}}{h}\Big)\frac{1}{h} + p\Big|\frac{u_{i+1} - u_i}{h}\Big|^{p-1} \operatorname{sgn}\Big(\frac{u_{i+1} - u_i}{h}\Big)\Big(-\frac{1}{h}\Big) \\
&\quad + a_i\, q\Big|\frac{u_i - u_{i-1}}{h}\Big|^{q-1} \operatorname{sgn}\Big(\frac{u_i - u_{i-1}}{h}\Big)\frac{1}{h} \\
&\qquad + a_{i+1}\, q\Big|\frac{u_{i+1} - u_i}{h}\Big|^{q-1} \operatorname{sgn}\Big(\frac{u_{i+1} - u_i}{h}\Big)\Big(-\frac{1}{h}\Big),
\end{aligned}$$

where $\operatorname{sgn}(x) := \dfrac{\partial}{\partial x}|x|, x \neq 0$ denotes the sign or signum function. Using again $(Du)_i$, the above expression can be expressed as

$$\begin{aligned}
\frac{\partial}{\partial u_i}\mathcal{H}(u) &= p\left[|(Du)_i|^{p-1}\operatorname{sgn}((Du)_i) - |(Du)_{i+1}|^{p-1}\operatorname{sgn}((Du)_{i+1})\right]\frac{1}{h} \\
&\quad + q\left[a_i|(Du)_i|^{q-1}\operatorname{sgn}((Du)_i) - a_{i+1}|(Du)_{i+1}|^{q-1}\operatorname{sgn}((Du)_{i+1})\right]\frac{1}{h}.
\end{aligned} \tag{4.5.4}$$

Let us denote by

$$\mu_i := |(Du)_i|^{p-1}\operatorname{sgn}((Du)_i) = \begin{cases} \frac{(Du)_i}{|(Du)_i|^{2-p}}, & \text{for } (Du)_i \neq 0 \\ 0, & \text{for } (Du)_i = 0 \end{cases} \tag{4.5.5}$$

and,

$$\eta_i := a_i|(Du)_i|^{q-1}\operatorname{sgn}((Du)_i) = \begin{cases} \frac{a_i(Du)_i}{|(Du)_i|^{2-q}}, & \text{for } (Du)_i \neq 0 \\ 0, & \text{for } (Du)_i = 0. \end{cases} \tag{4.5.6}$$

Taking a forward difference operator, denoted by D^T, such that $(D^T\mu)_i := \frac{1}{h}(\mu_{i+1} - \mu_i)$ and $(D^T\eta)_i := \frac{1}{h}(\eta_{i+1} - \eta_i)$, then (4.5.4) implies,

$$\frac{\partial}{\partial u_i}\mathcal{H}(u) = -(p(D^T\mu)_i + q(D^T\eta)_i).$$

So considering that the minimum value of $E(u)$ exists at $u = u^*$, we have

$$\nabla E(u^*) = \lambda(u^* - f) - (pD^T\mu^* + qD^T\eta^*) = 0, \tag{4.5.7}$$

where μ_i^* and η_i^* correspond to μ_i and η_i, defined similarly as in (4.5.5) and (4.5.6) for $(Du^*)_i$. The equation (4.5.7) is the necessary and sufficient optimality condition for (4.5.3).

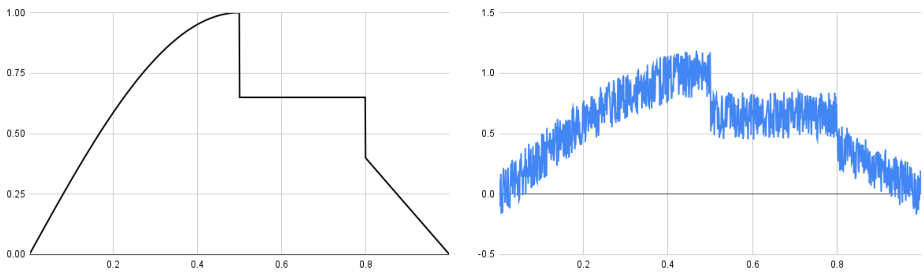
4.5.2 Experimental results. Here, we present a few examples to demonstrate the working of the proposed image restoration model (4.5.3) with one-dimensional signals and numerical results. Also, depending on the original as well as noisy signals, we formulate the function $a(x)$ to examine the behaviour of the image restoration model.

For the 1-dimensional noisy signals [33] considered for image restoration, we implement LMBM algorithm [22] (coded by Napsu Karmitsa) using Fortran 77, presented in the appendix A.5. The quality of image restoration is measured using the reconstruction error, which is the average error between the true signal and the obtained result. The number of discretization points taken for all the cases is $N = 1000$.

We consider the following 1-dimensional signal, say $I_1(x)$, formulated as:

$$I_1(x) := \begin{cases} \sin(2\pi x), & x < 0.5, \\ 0.65, & 0.5 < x < 0.8, \\ -2x + 2, & 0.8 < x < 1. \end{cases}$$

The above signal along with its noisy signal are presented in figure 1.



(a) Original signal **(b)** Noisy signal

Figure 1. 1D signal: first example

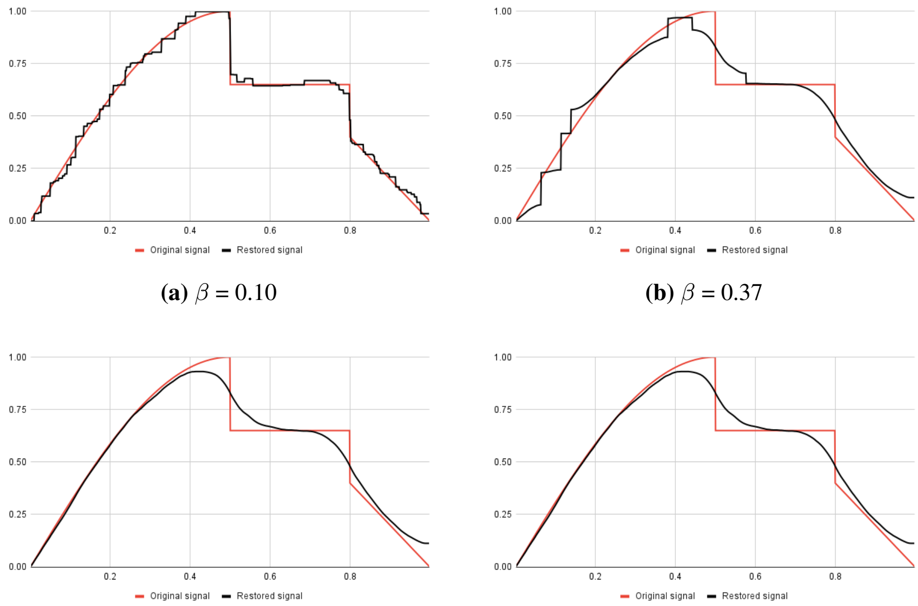
We consider two approaches in formulating the function $a(x)$: first, based on the noisy signal and, second, based on the original signal. In practical cases, it is not always feasible to correctly estimate the original signal, but for theoretical purpose we adopt such approach to examine the efficiency of our image restoration model. The main idea here is to test if we have the best possible $a(x)$ and check how the model works in that case.

For the first approach where the function $a(x)$ is formulated using the gradient of noisy signal, say $f_1(x)$ in figure 1(b), denoting it by $a_1(x)$, we have:

$$a_1(x) := \begin{cases} 0.9, & |Df_1(x)| \leqslant \beta, \\ 0, & |Df_1(x)| > \beta, \end{cases}$$

where $\beta > 0$ is an appropriate cut-off value, which depends on the noise level present at the particular point.

We now compare the restored signals obtained with different values of β in figure 2. Here, we fix $p = 1.0001, q = 2.0$. With $\beta = 0.1$, the restored signal, figure 2(a), gives an error of $7.41 \cdot 10^{-4}$ and computation time being 2.73 seconds. While with $\beta = 0.37$, we obtain the restored signal, figure 2(b), with error $1.81 \cdot 10^{-3}$ and computation time of 1.49 seconds. Further, with $\beta = 0.5$ and $\beta = 1.0$, the obtained restored signals, figures 2(c) and 2(d), both give approximately same error of $1.59 \cdot 10^{-3}$, with computation time of 1.14 seconds and 1.18 seconds respectively.



(a) $\beta = 0.10$ **(b)** $\beta = 0.37$ **(c)** $\beta = 0.50$ **(d)** $\beta = 1.0$

Figure 2. Restored signals of 1(b) with $p = 1.0001, q = 2.0$

Next, taking a fixed β, we consider the extreme values of the parameters p and q to check the behaviour of the model for the same noise level. With fixed $\beta = 0.5$, choosing $p = 1.0001$ and $q = 1.0007$ we obtain the restored signal, figure 3(a), with error $6.81 \cdot 10^{-4}$ and computation time of 18.95 seconds. On the other hand, choosing $p = 1.9$ and $q = 2.0$ we obtain the restored signal, figure 3(b), with error $1.99 \cdot 10^{-3}$ and computation time of 0.16 second.

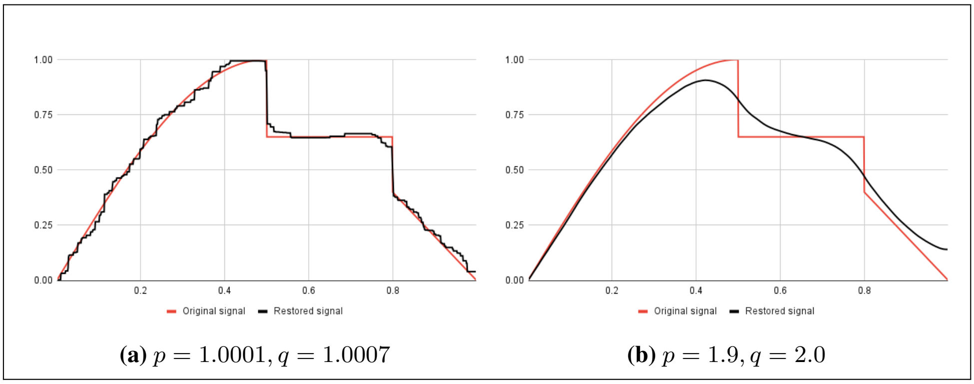


Figure 3. Restored signals of 1(b) with $\beta = 0.5$

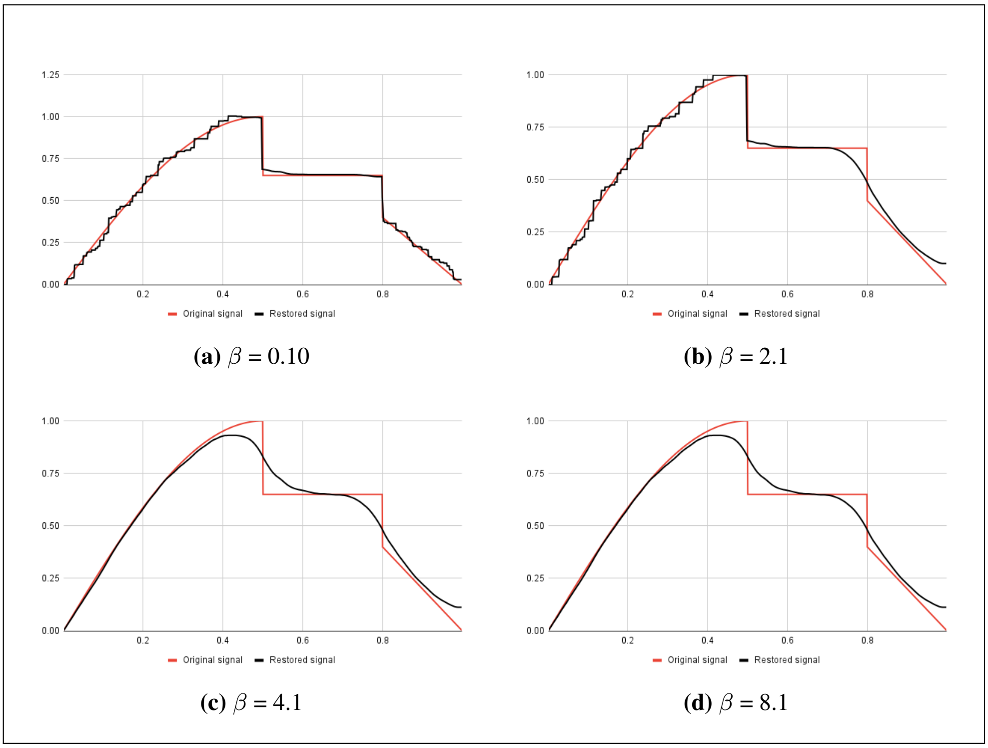
Next, we use the second approach where the function $a(x)$ depends on the gradient of the original signal, $I_1(x)$ in figure 1(a), that is,

$$a_2(x) := \begin{cases} 0.9, & |\nabla I_1(x)| \leqslant \beta, \\ 0, & |\nabla I_1(x)| > \beta. \end{cases}$$

In this case, the (absolutely continuous part of) gradient, $\nabla I_1(x)$, is formulated as:

$$\nabla I_1(x) := \begin{cases} 2\pi \cos(2\pi x), & x < 0.5, \\ 0, & 0.5 \leqslant x < 0.8, \\ -2, & 0.8 \leqslant x < 1. \end{cases}$$

We again compare the restored signals obtained with different values of β, in figure 4. The parameters $p = 1.0001, q = 2.0$ are fixed here. With $\beta = 0.1$, the restored signal, in figure 4(a), gives an error of $7.90 \cdot 10^{-4}$ and computation time being 1.86 seconds. While with $\beta = 2.1$, we obtain the restored signal, in figure 4(b), with error $1.18 \cdot 10^{-3}$ and computation time of 2.54 seconds. Further, with $\beta = 4.1$ and $\beta = 8.1$, the obtained restored signals, figures 4(c) and 4(d), give approximately same error of $1.59 \cdot 10^{-3}$, with computation time of 1.15 seconds and 1.19 seconds respectively.

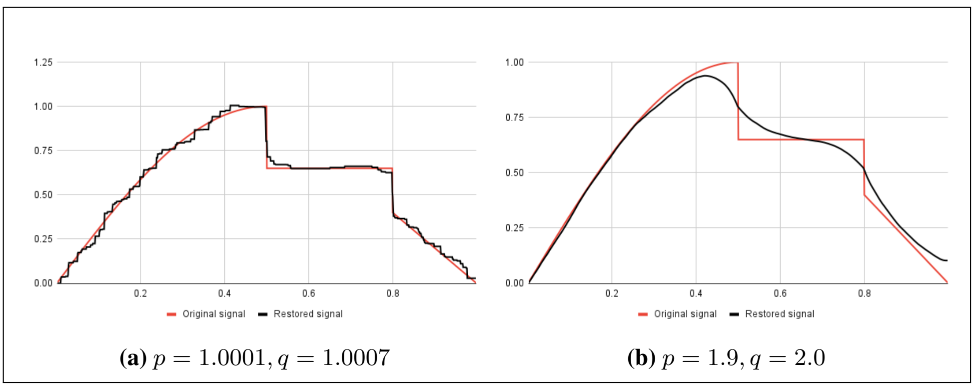


(a) $\beta = 0.10$ **(b)** $\beta = 2.1$

(c) $\beta = 4.1$ **(d)** $\beta = 8.1$

Figure 4. Restored signals: $p = 1.0001, q = 2.0$

Next, taking a fixed β, we consider the extreme values of the parameters p and q. With fixed $\beta = 0.1$, choosing $p = 1.0001$ and $q = 1.0007$ we obtain the restored signal, in figure 5(a), with error $7.42 \cdot 10^{-4}$ and computation time of 35.68 seconds. On the other hand, choosing $p = 1.9$ and $q = 2.0$ we obtain the restored signal, in figure 5(b), with error $1.63 \cdot 10^{-3}$ and computation time of 0.19 second.



(a) $p = 1.0001, q = 1.0007$ **(b)** $p = 1.9, q = 2.0$

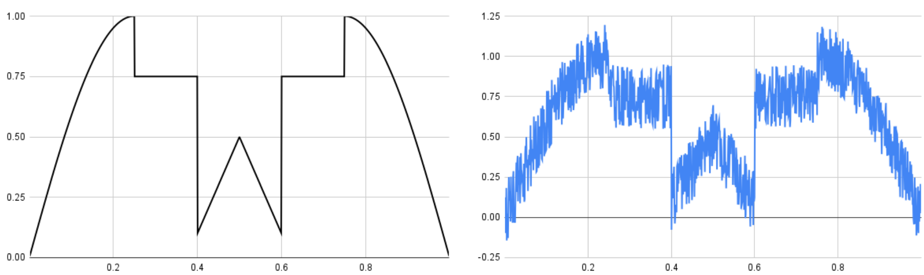
Figure 5. Restored signals with $\beta = 0.1$

Next, we consider another 1-dimensional signal, say $I_2(x)$, formulated as:

$$I_2(x) := \begin{cases} \sin(2\pi x), & x < 0.25, \\ 0.75, & 0.25 < x < 0.4, \\ 4x - 1.56, & 0.4 < x < 0.5, \\ -4x + 2.5, & 0.5 < x < 0.6, \\ 0.75, & 0.6 < x < 0.75, \\ -\sin(2\pi x), & 0.75 < x < 1. \end{cases}$$

We follow a similar process as in the previous signal, to check the working of the image restoration model in this case.

The above signal along with its noisy signal are presented in figure 6.



(a) Original signal **(b)** Noisy signal

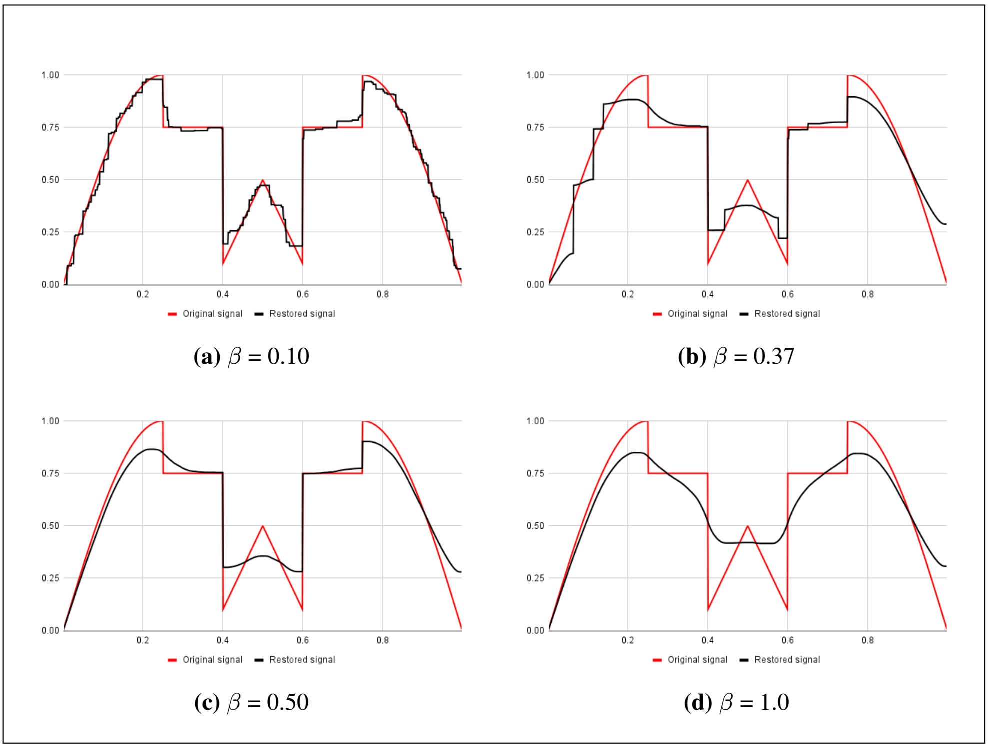
Figure 6. 1D signal: second example

For the first approach where the function $a(x)$ is based on the gradient of noisy signal, say $f_2(x)$ in figure 6(b), denoting it by $a_3(x)$, we have

$$a_3(x) := \begin{cases} 0.9, & |Df_2(x)| \leqslant \beta, \\ 0, & |Df_2(x)| > \beta, \end{cases}$$

where $\beta > 0$ is an appropriate cut-off value.

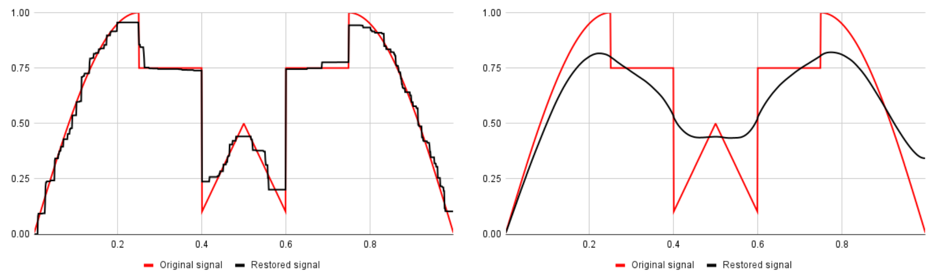
We now compare the restored signals obtained with different values of β, in figure 7. The parameters $p = 1.0001, q = 2.0$ are fixed. With $\beta = 0.1$, the restored signal, figure 7(a), gives an error of $9.67 \cdot 10^{-4}$ and computation time being 4.18 seconds. While with $\beta = 0.37$, we obtain the restored signal, figure 7(b), with error $2.38 \cdot 10^{-3}$ and computation time of 5.60 seconds. Further, with $\beta = 0.5$, the obtained restored signal, figure 7(c), give error of $1.59 \cdot 10^{-3}$ and $3.12 \cdot 10^{-3}$, with computation time of 1.17 seconds. Lastly, with $\beta = 1.0$, the obtained restored signal, figure 7(d), give error of $3.12 \cdot 10^{-3}$, with computation time of 2.78 seconds.



(a) $\beta = 0.10$ (b) $\beta = 0.37$ (c) $\beta = 0.50$ (d) $\beta = 1.0$

Figure 7. Restored signals: $p = 1.0001, q = 2.0$

Next, taking a fixed β, we consider the extreme values of the parameters p and q to check the behaviour of the model. With fixed $\beta = 0.5$, choosing $p = 1.0001$ and $q = 1.0007$ we obtain the restored signal, figure 8(a), with error $1.05 \cdot 10^{-3}$ and computation time of 11.92 seconds. On the other hand, choosing $p = 1.9$ and $q = 2.0$ we obtain the restored signal, figure 8(b), with error $4.31 \cdot 10^{-3}$ and computation time of 0.17 second.



(a) $p = 1.0001, q = 1.0007$ **(b)** $p = 1.9, q = 2.0$

Figure 8. Restored signals with $\beta = 0.5$

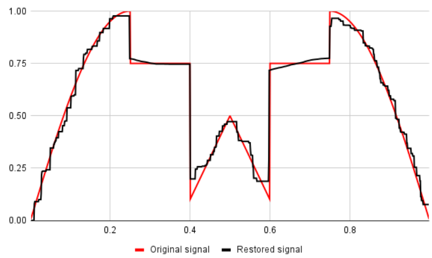
Next, in the second approach, the function $a(x)$ depends on the gradient of the original signal in figure 6(a), that is,

$$a_4(x) := \begin{cases} 0.9, & |\nabla I_2(x)| \leqslant \beta, \\ 0, & |\nabla I_2(x)| > \beta \end{cases}$$

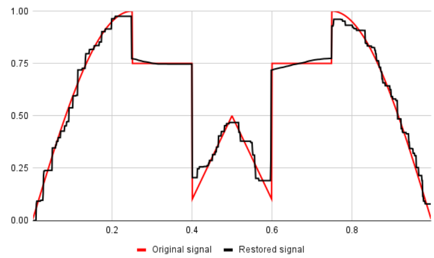
where the (absolutely continuous part of) gradient, $\nabla I_2(x)$, is formulated as:

$$\nabla I_2(x) := \begin{cases} 2\pi \cos(2\pi x), & x < 0.25, \\ 0, & 0.25 \leqslant x < 0.4, \\ 4, & 0.4 \leqslant x < 0.5, \\ -4, & 0.5 \leqslant x < 0.6, \\ 0, & 0.6 \leqslant x < 0.75, \\ -2\pi \cos(2\pi x), & 0.75 \leqslant x < 1. \end{cases}$$

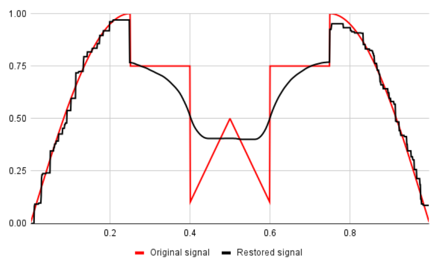
We again compare the restored signals obtained with different values of β, in figure 9. The parameters $p = 1.0001, q = 2.0$ are fixed here. With $\beta = 0.1$, the restored signal, figure 9(a), gives an error of $1.27 \cdot 10^{-3}$ and computation time being 2.16 seconds. While with $\beta = 2.1$, we obtain the restored signal, figure 9(b), with error $1.29 \cdot 10^{-3}$ and computation time of 2.16 seconds. With $\beta = 4.1$, the obtained restored signal, figure 9(c), give error of $3.06 \cdot 10^{-3}$, with computation time of 4.39 seconds. Further with $\beta = 8.1$, the obtained restored signal, figure 9(d), give error of $3.80 \cdot 10^{-3}$, with computation time of 2.62 seconds.



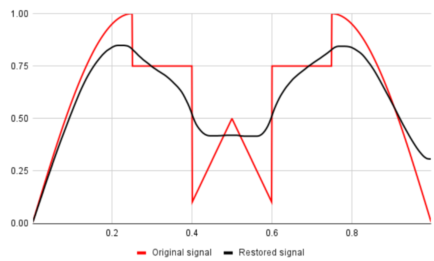
(a) $\beta = 0.10$



(b) $\beta = 2.1$



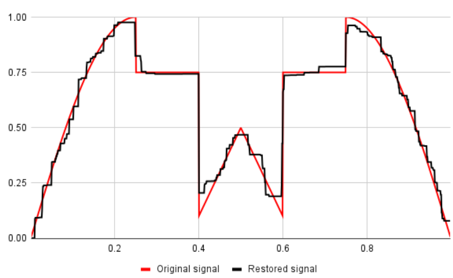
(c) $\beta = 4.1$



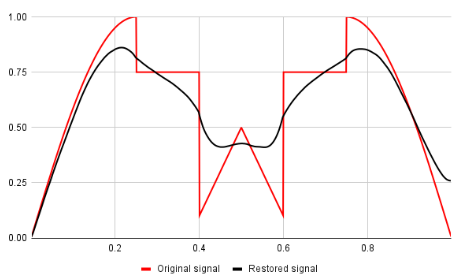
(d) $\beta = 8.1$

Figure 9. Restored signals: $p = 1.0001, q = 2.0$

Finally, taking a fixed β, we consider the extreme values of the parameters p and q. With fixed $\beta = 0.1$, choosing $p = 1.0001$ and $q = 1.0007$ we obtain the restored signal, figure 10(a), with error $1.08 \cdot 10^{-3}$ and computation time of 15.17 seconds. On the other hand, choosing $p = 1.9$ and $q = 2.0$ we obtain the restored signal, figure 10(b), with error $3.68 \cdot 10^{-3}$ and computation time of 0.17 second.



(a) $p = 1.0001, q = 1.0007$



(b) $p = 1.9, q = 2.0$

Figure 10. Restored signals with $\beta = 0.1$

Thus, as the exponents p, q of the gradient function in the model (4.5.3) move from 1 to 2, the sharp edge-like features tend to become smooth, with varying error tolerance depending on the image features and noise level.

4.5.3 Summary and discussion The results presented here are measured using quantitative performance measures such as reconstruction error, as well as in terms of visual quality of the images. The techniques here assume the noise model to be Gaussian, while in reality, this assumption may not always hold true due to the varied nature and sources of noise. An ideal denoising procedure requires *a priori* knowledge of the noise, whereas a practical procedure may not have the required information about the variance of the noise or the noise model. Thus, in most of the cases, Gaussian noise with different variance values is added in the natural images to test the performance of an algorithm and also to compare different algorithms or different denoising models.

The parameters chosen here comply with the image restoration model considered, where the functional $\mathcal{H}(x, |\nabla u|) = |\nabla u|^p + a(x)|\nabla u|^q, 1 < p \leqslant q \leqslant 2$ is used to keep a balance between edge detection and smoothing effect in the denoised image regulated by the value of a. Moreover these parameters are chosen accordingly at best to reduce the reconstruction error mathematically. The difference in the visual quality of the reconstructed images might not be reflected distinctly, as shown mathematically, but this aspect is more significant in the practical applications to obtain clarity of image features to the possible extent.

The message we wish to convey is that to interpret a physical phenomenon, the intuition that leads to certain formulations and the underlying theoretical study are often complementary. However, developing a theoretical justification of a problem is not simply *art for art's sake*, but in fact, a deep understanding of the theoretical difficulties may lead to the development of suitable numerical schemes or different models. Hence, for denoising purpose, the formulation quality of the image restoration model as well as the algorithm solver are crucial to justify a well-established nature of methodology.

Appendix

A.1 Density argument in the proof of Theorem 4.3.3

To prove: $u_M - f \in W_0^{1,\mathcal{H}}(\Omega)$, where $u_M := \min\{M, \max\{-M, u\}\}$ with $u \in U := \{u \in W^{1,\mathcal{H}}(\Omega) \cap L^2(\Omega) \mid u - f \in W_0^{1,\mathcal{H}}(\Omega)\}$ and $M > 0$ such that $|f| \leqslant M$.

Since $C_0^\infty(\Omega)$ is dense in $W_0^{1,\mathcal{H}}(\Omega)$, then there exists a sequence $\{\psi_i\}_{i=1}^\infty \in C_0^\infty(\Omega)$ such that $\psi_i \to u - f$ in $W^{1,\mathcal{H}}(\Omega)$, as $i \to \infty$. In order to establish the claim, we need to prove that there exists an approximating sequence $\tilde{\psi}_{\epsilon_i} \in C_0^\infty(\Omega)$, such that $\tilde{\psi}_{\epsilon_i} \to u_M - f$ in $W^{1,\mathcal{H}}(\Omega)$ as $i \to \infty$.

We have that,

$$\varrho_{1,\mathcal{H}}(\psi_i + f - u) \to 0, \quad \text{as } i \to \infty, \tag{A.1.1}$$

so $\psi_i + f$ converges to u in $W^{1,\mathcal{H}}(\Omega)$. Now cutting-off the functions $\psi_i + f, u$ at $-M$ and M, $M > 0$, we set,

$$(\psi_i + f)_M := \min\{M, \max\{-M, \psi_i + f\}\}, \quad u_M := \min\{M, \max\{-M, u\}\}.$$

Next we prove that $(\psi_i + f)_M \to u_M$ in $W^{1,\mathcal{H}}(\Omega)$ as $i \to \infty$, that is,

$$\varrho_{1,\mathcal{H}}((\psi_i + f)_M - u_M) \to 0, \quad \text{as } i \to \infty.$$

For that, we consider the following subsets of the domain Ω, where the values of the modular $\varrho_{1,\mathcal{H}}((\psi_i + f)_M - u_M)$ are calculated.

Consider $\Omega_1 := \{x \in \Omega : -M \leqslant (\psi_i + f)(x), u(x) \leqslant M\}$, $\Omega_2 := \{x \in \Omega : (\psi_i + f)(x), u(x) < -M\}$ and $\Omega_3 := \{x \in \Omega : (\psi_i + f)(x), u(x) > M\}$. Let χ_1, χ_2 and χ_3 be the corresponding characteristic functions to the sets Ω_1, Ω_2 and Ω_3. Then, at Ω_1, Ω_2 and Ω_3, the modular $\varrho_{1,\mathcal{H}}((\psi_i + f)_M - u_M)$ takes the following forms.

$$\begin{aligned}
&\varrho_{1,\mathcal{H}}(((\psi_i + f)_M - u_M)\chi_1) = \varrho_{1,\mathcal{H}}((\psi_i + f - u)\chi_1),\\
&\varrho_{1,\mathcal{H}}(((\psi_i + f)_M - u_M)\chi_2) = \varrho_{1,\mathcal{H}}((-M + M)\chi_2) = 0,\\
&\varrho_{1,\mathcal{H}}(((\psi_i + f)_M - u_M)\chi_3) = \varrho_{1,\mathcal{H}}((M - M)\chi_3) = 0.
\end{aligned}$$

At $\Omega_4 := \{x \in \Omega : (\psi_i + f)(x) > M\} \cap \{x \in \Omega : u(x) < -M\}$, with χ_4 as the characteristic function to the set Ω_4, we have the modulars

$$\begin{aligned}
&\varrho_{\mathcal{H}}(((\psi_i + f)_M - u_M)\chi_4) = \varrho_{\mathcal{H}}((M - (-M))\chi_4) = \varrho_{\mathcal{H}}((2M)\chi_4),\\
&\varrho_{\mathcal{H}}((\nabla(\psi_i + f)_M - \nabla u_M)\chi_4) = \varrho_{\mathcal{H}}((0)\chi_4) = 0.
\end{aligned}$$

At $\Omega_5 \coloneqq \{x \in \Omega : |(\psi_i + f)(x)| < M\} \cap \{x \in \Omega : u(x) < -M\}$, with χ_5 as the characteristic function to the set Ω_5, we have

$$\varrho_{\mathcal{H}}(((\psi_i + f)_M - u_M)\chi_5) = \varrho_{\mathcal{H}}((\psi_i + f - (-M))\chi_5) = \varrho_{\mathcal{H}}((\psi_i + f + M)\chi_5),$$
$$\varrho_{\mathcal{H}}((\nabla(\psi_i + f)_M - \nabla u_M)\chi_5) = \varrho_{\mathcal{H}}(\nabla(\psi_i + f)\chi_5).$$

At $\Omega_6 \coloneqq \{x \in \Omega : (\psi_i + f)(x) > M\} \cap \{x \in \Omega : |u(x)| < M\}$, with χ_6 as the characteristic function to the set Ω_6, we have

$$\varrho_{\mathcal{H}}(((\psi_i + f)_M - u_M)\chi_6) = \varrho_{\mathcal{H}}((M - u)\chi_6),$$
$$\varrho_{\mathcal{H}}((\nabla(\psi_i + f)_M - \nabla u_M)\chi_6) = \varrho_{\mathcal{H}}((-\nabla u)\chi_6).$$

At $\Omega_7 \coloneqq \{x \in \Omega : (\psi_i + f)(x) < -M\} \cap \{x \in \Omega : |u(x)| < M\}$, with χ_7 as the characteristic function to the set Ω_7, we have

$$\varrho_{\mathcal{H}}(((\psi_i + f)_M - u_M)\chi_7) = \varrho_{\mathcal{H}}((-M - u)\chi_7),$$
$$\varrho_{\mathcal{H}}((\nabla(\psi_i + f)_M - \nabla u_M)\chi_7) = \varrho_{\mathcal{H}}((-\nabla u)\chi_7).$$

At $\Omega_8 \coloneqq \{x \in \Omega : |(\psi_i + f)(x)| < M\} \cap \{x \in \Omega : u(x) > M\}$, with χ_8 as the characteristic function to the set Ω_8, we have

$$\varrho_{\mathcal{H}}(((\psi_i + f)_M - u_M)\chi_8) = \varrho_{\mathcal{H}}((\psi_i + f - M)\chi_8),$$
$$\varrho_{\mathcal{H}}((\nabla(\psi_i + f)_M - \nabla u_M)\chi_8) = \varrho_{\mathcal{H}}((\nabla(\psi_i + f))\chi_8).$$

At $\Omega_9 \coloneqq \{x \in \Omega : (\psi_i + f)(x) < -M\} \cap \{x \in \Omega : u(x) > M\}$, with χ_9 as the characteristic function to the set Ω_9, we have

$$\varrho_{\mathcal{H}}(((\psi_i + f)_M - u_M)\chi_9) = \varrho_{\mathcal{H}}(-M - M)\chi_9 = \varrho_{\mathcal{H}}((-2M)\chi_9),$$
$$\varrho_{\mathcal{H}}((\nabla(\psi_i + f)_M - \nabla u_M)\chi_9) = \varrho_{\mathcal{H}}((0)\chi_9) = 0.$$

Since $\Omega_1, \Omega_2, \ldots, \Omega_9$ form a partition of Ω, then using the triangle-inequality, we obtain

$$\varrho_{\mathcal{H}}((\psi_i + f)_M - u_M) = \sum_{k=1}^{9} \varrho_{\mathcal{H}}(((\psi_i + f)_M - u_M)\chi_k). \tag{A.1.2}$$

As $i \to \infty$, we have $\varrho_{1,\mathcal{H}}(\psi_i + f - u) \to 0$, from (A.1.1), hence we conclude $\varrho_{1,\mathcal{H}}(((\psi_i + f) - u)\chi_1) \to 0$ as $i \to \infty$. Using the fact that each characteristic function χ_k depend on the definition of the corresponding set Ω_k, and since ψ_i converges pointwise to $u - f$ as $i \to \infty$, in Ω_1, hence the characteristics functions $\chi_2, \ldots, \chi_9$ tend to 0, as $i \to \infty$.

On the other hand, we have $|(\psi_i + f)_M - u_M| \leqslant |(\psi_i + f)_M| + |u_M| \leqslant M + M = 2M$ which is integrable, hence using the dominated convergence theorem, we conclude that $\lim_{i\to\infty} \varrho_{1,\mathcal{H}}((((\psi_i + f)_M - u_M)\chi_k) = 0$ for $k = 1, \ldots, 9$. Further, from (A.1.2), we obtain $\lim_{i\to\infty} \varrho_{1,\mathcal{H}}((\psi_i + f)_M - u_M) = 0$, which implies

$$(\psi_i + f)_M - f \to u_M - f \quad \text{in } W^{1,\mathcal{H}}(\Omega),$$

as $i \to \infty$. Since $(\psi_i + f)_M - f$ has a compact support in Ω and $C^\infty(\Omega) \cap W^{1,\mathcal{H}}(\Omega)$ is dense in $W^{1,\mathcal{H}}(\Omega)$ by assumption, then by [24, Lemma 6.1.10] we conclude that $u_M - f \in W_0^{1,\mathcal{H}}(\Omega)$.

A.2 Verification of the parabolicity condition

To prove: For any $s \in \mathbb{R}$ and $\kappa \in \mathbb{R}^n$,

$$\nu|\xi|^2 \leqslant \sum_{i,j=1}^{n} g_{ij}(x,t,s,\kappa)\,\xi_i\xi_j \leqslant \mu|\xi|^2,$$

where $\xi = (\xi_1, \ldots, \xi_n)$ is a real vector and $\nu, \mu > 0$ are constants.
For $(x,t) \in \overline{\Omega}_T$, $g_{ij}(x,t,s,\kappa)$ is defined from (4.4.26) as,

$$\sum_{i,j=1}^{n} g_{ij}(x,t,s,\kappa)\xi_i\xi_j = \left(\epsilon + \frac{p}{\left(\sqrt{|\kappa|^2+\epsilon^2}\right)^{2-p}} + a(x)\frac{q}{\left(\sqrt{|\kappa|^2+\epsilon^2}\right)^{2-q}}\right)\sum_{i,j=1}^{n}\xi_i\xi_j\delta_{ij}$$
$$+ \left(\frac{p(p-2)}{\left(\sqrt{|\kappa|^2+\epsilon^2}\right)^{4-p}} + a(x)\frac{q(q-2)}{\left(\sqrt{|\kappa|^2+\epsilon^2}\right)^{4-q}}\right)\sum_{i,j=1}^{n}(\kappa_i\xi_i)(\kappa_j\xi_j),$$

where $1 < p \leqslant q \leqslant 2$ and $a(x)$ is Lipschitz continuous function.
In order to get the required inequality, we further apply absolute value for each term above to get,

$$\sum_{i,j=1}^{n} g_{ij}(x,t,s,\kappa)\,\xi_i\xi_j \leqslant \left(\epsilon + \frac{p}{\epsilon^{2-p}} + |a(x)|\frac{q}{\epsilon^{2-q}}\right)|\xi|^2 + \left(\frac{|p(p-2)|}{|\kappa|^2\epsilon^{2-p}} + |a(x)|\frac{|q(q-2)|}{|\kappa|^2\epsilon^{2-q}}\right)|\kappa|^2|\xi|^2$$
$$\leqslant \left(\epsilon + \frac{2C}{\epsilon^{2-p}} + \frac{2C}{\epsilon^{2-q}}\right)|\xi|^2, \quad \text{where } C \geqslant 0 \text{ is a constant.}$$

Thus, $\sum_{i,j=1}^{n} g_{ij}(x,t,s,\kappa)\,\xi_i\xi_j \leqslant \mu|\xi|^2$, where $\mu = \epsilon + \frac{2C}{\epsilon^{2-p}} + \frac{2C}{\epsilon^{2-q}} > 0$.
On the other hand,

$$\sum_{i,j=1}^{n} g_{ij}(x,t,s,\kappa)\,\xi_i\xi_j \geqslant \left(\epsilon + \frac{p}{(\sqrt{|\kappa|^2+\epsilon^2})^{2-p}} + \frac{a(x)q}{(\sqrt{|\kappa|^2+\epsilon^2})^{2-q}}\right)|\xi|^2$$
$$+ \left(\frac{p(p-2)}{|\kappa|^2\left(\sqrt{|\kappa|^2+\epsilon^2}\right)^{2-p}} + a(x)\frac{q(q-2)}{|\kappa|^2\left(\sqrt{|\kappa|^2+\epsilon^2}\right)^{2-q}}\right)|\kappa|^2|\xi|^2$$
$$\geqslant \left(\epsilon + \frac{p(p-1)}{\left(\sqrt{|\kappa|^2+\epsilon^2}\right)^{2-p}} + a(x)\frac{q(q-1)}{\left(\sqrt{|\kappa|^2+\epsilon^2}\right)^{2-q}}\right)|\xi|^2$$

Since each term on the right-hand side above is non-negative, hence we have

$$\sum_{i,j=1}^{n} g_{ij}(x,t,s,\kappa)\,\xi_i\xi_j \geqslant \nu|\xi|^2, \quad \text{where } \nu > \epsilon > 0,$$

holds a.e. Therefore, the condition (4.4.28): $\nu|\xi|^2 \leqslant \sum_{i,j=1}^{n} g_{ij}(x,t,s,\kappa)\,\xi_i\xi_j \leqslant \mu|\xi|^2$ is satisfied for constants $\nu, \mu > 0$.

A.3 Existence of solution of quasilinear boundary problem

***Proposition* A.3.1.** [38, p. 560, Theorem 4.4, Chapter 6] *Let $\Omega \subset \mathbb{R}^n$ be a bounded open set with Lipschitz boundary. Consider the quasi-linear equation with Dirichlet boundary,*

$$\left.\begin{aligned} &u_t - \sum_{i,j=1}^{n} g_{ij}(x,t,u,\nabla u)\,u_{x_ix_j} + g(x,t,u,\nabla u) = 0, && (x,t) \in \Omega_T \\ &u(x,t) = f_\delta(x), && (x,t) \in \partial\Omega \times (0,T). \end{aligned}\right\} \tag{A.3.1}$$

Suppose that the following conditions hold.

(a) For $(x,t) \in \overline{\Omega}_T$ and arbitrary $s \in \mathbb{R}$, the following conditions are fulfilled,

$$\sum_{i,j=1}^{n} g_{ij}(x,t,s,0)\,\xi_i\xi_j \geqslant 0 \quad \textit{and} \quad s\,g(x,t,s,0) \geqslant -b_1 s^2 - b_2,$$

where b_1 and b_2 are non-negative constants, and $\xi = (\xi_1,\ldots,\xi_n)$ is an arbitrary real vector.

(b) For $(x,t) \in \overline{\Omega}_T$, $|s| \leqslant M$ (where $M > 0$ is constant) and, $\kappa \in \mathbb{R}^n$, the functions $g_{ij}(x,t,s,\kappa)$ and $g(x,t,s,\kappa)$ are continuous and differentiable with respect to x, s and κ, and satisfy the following inequalities with $m \geqslant 1$,

$$\nu(1+|\kappa|)^{m-2}\xi^2 \leqslant \sum_{i,j=1}^{n} g_{ij}(x,t,s,\kappa)\,\xi_i\xi_j \leqslant \mu(1+|\kappa|)^{m-2}\xi^2, \quad \nu,\mu > 0$$

$$\left|\frac{\partial g_{ij}}{\partial \kappa_k}\right|(1+|\kappa|)^3 + |g| + \left|\frac{\partial g}{\partial \kappa_k}\right|(1+|\kappa|) \leqslant \mu_1(1+|\kappa|)^m,$$

$$\left|\frac{\partial g_{ij}}{\partial x_k}\right|(1+|\kappa|)^2 + \left|\frac{\partial g}{\partial x_k}\right| \leqslant [\beta + P(|\kappa|)](1+|\kappa|)^{m+1}, \quad \beta > 0$$

$$\left|\frac{\partial g_{ij}}{\partial s}\right| \leqslant [\beta + P(|\kappa|)](1+|\kappa|)^{m-2},$$

$$-\frac{\partial g}{\partial s} \leqslant [\beta + P(|\kappa|)](1+|\kappa|)^m.$$

where $P(\rho)$ is a non-negative continuous function that tends to zero for $\rho \to \infty$ and β is sufficiently small determined by the numbers M, ν, μ, μ_1 and $\tilde{P} = \max_{\rho \geqslant 0} P(\rho)$.

(c) *For $(x,t) \in \overline{\Omega}_T$, $|s| \leqslant M$ and $|\kappa| \leqslant M_1$ (where $M, M_1 > 0$ are constants), the functions $g_{ij}(x,t,s,\kappa)$ and $g(x,t,s,\kappa)$ are continuously differentiable with respect to all of their arguments.*

(d) $f_\delta \in C(\overline{\Omega}_T) \cap H^{2+\beta, 1+\frac{\beta}{2}}(\Omega_T)$.

(e) *Each point of the boundary of Ω can be touched from without by a ball (or cone) of fixed size in such a way that the ball (cone) does not have any points in common with Ω.*

Then there exists a unique solution of the problem (A.3.1) *in* $C(\overline{\Omega}_T) \cap H^{2+\beta,1+\frac{\beta}{2}}(\Omega_T)$.

For $(x,t) \in \overline{\Omega}_T$ and arbitrary $\kappa \in \mathbb{R}^n$, $s \in \mathbb{R}$, the functions g_{ij} and g are defined from (4.4.26) and (4.4.27), as follows

$$\begin{aligned} g_{ij}(x,t,s,\kappa) = &\left(\epsilon + \frac{p}{\left(\sqrt{|\kappa|^2+\epsilon^2}\right)^{2-p}} + a(x)\frac{q}{\left(\sqrt{|\kappa|^2+\epsilon^2}\right)^{2-q}}\right)\delta_{ij} \\ &+ \left(\frac{p(p-2)}{\left(\sqrt{|\kappa|^2+\epsilon^2}\right)^{4-p}} + a(x)\frac{q(q-2)}{\left(\sqrt{|\kappa|^2+\epsilon^2}\right)^{4-q}}\right)\kappa_i\kappa_j, \end{aligned} \tag{A.3.2}$$

and,

$$g(x,t,s,\kappa) = -\frac{q}{(\sqrt{|\kappa|^2+\epsilon^2})^{2-q}}\sum_{i=1}^{n}\kappa_i\frac{\partial a}{\partial x_i} + \lambda(s - f_\delta), \tag{A.3.3}$$

(a) To prove: For $(x,t) \in \overline{\Omega}_T$ and arbitrary $s \in \mathbb{R}$,

$$\sum_{i,j=1}^{n} g_{ij}(x,t,s,0)\,\xi_i\xi_j \geqslant 0 \quad \text{and} \quad sg(x,t,s,0) \geqslant -b_1 s^2 - b_2,$$

where b_1 and b_2 are non-negative constants.
For real-valued vectors ξ_i, ξ_j, we have

$$\begin{aligned} \sum_{i,j=1}^{n} g_{ij}(x,t,s,0)\,\xi_i\xi_j &= \left(\epsilon + \tfrac{p}{\epsilon^{2-p}} + a(x)\tfrac{q}{\epsilon^{2-q}}\right)\sum_{i,j=1}^{n}\xi_i\xi_j\,\delta_{ij} \\ &\geqslant \left(\epsilon + \tfrac{p}{\epsilon^{2-p}}\right)|\xi|^2 \geqslant 0. \end{aligned}$$

On the other hand,

$$\begin{aligned} s\,g(x,t,s,0) &= s\lambda(s-f_\delta) \geqslant \lambda(s^2-|sf_\delta|) \\ &\geqslant \lambda\big(s^2-\tfrac{|s|^2}{2}-\tfrac{|f_\delta|^2}{2}\big) \quad \text{(by Young's inequality)} \\ &= \lambda\big(\tfrac{s^2}{2}-\tfrac{|f_\delta|^2}{2}\big). \end{aligned}$$

and, since $|f_\delta| \leqslant C$ ($C>0$ is constant), hence we have,

$$s\,g(x,t,s,0) \geqslant \lambda\big(\frac{s^2}{2}-\frac{C^2}{2}\big) \geqslant -\lambda\frac{C^2}{2}.$$

Thus, $g_{ij}(x,t,s,0)\,\xi_i\xi_j \geqslant 0$ and $sg(x,t,s,0) \geqslant -b_1s^2-b_2$, where $b_1=0$ and $b_2=\frac{\lambda C^2}{2}$ are non-negative constants.

(b) To prove: We prove the inequalities here for $m=2$, that is, the following estimates:

$$\nu|\xi|^2 \leqslant \sum_{i,j=1}^{n} g_{ij}(x,t,s,\kappa)\,\xi_i\xi_j \leqslant \mu|\xi|^2, \quad \nu,\mu = \text{constant} > 0 \tag{A.3.4}$$

$$\Big|\frac{\partial g_{ij}}{\partial\kappa_k}\Big|(1+|\kappa|)^3+|g|+\Big|\frac{\partial g}{\partial\kappa_k}\Big|(1+|\kappa|) \leqslant \mu_1(1+|\kappa|)^2, \tag{A.3.5}$$

$$\Big|\frac{\partial g_{ij}}{\partial x_k}\Big|(1+|\kappa|)^2+\Big|\frac{\partial g}{\partial x_k}\Big| \leqslant [\beta+P(|\kappa|)](1+|\kappa|)^3, \quad (\beta>0) \tag{A.3.6}$$

$$\Big|\frac{\partial g_{ij}}{\partial s}\Big| \leqslant [\varepsilon+P(|\kappa|)], \tag{A.3.7}$$

$$-\frac{\partial g}{\partial s} \leqslant [\varepsilon+P(|\kappa|)](1+|\kappa|)^2 \tag{A.3.8}$$

where $P(\rho)\to 0$ for $\rho\to\infty$.

First, we compute the following estimates, for arbitrary $s\in\mathbb{R}$ and $\kappa\in\mathbb{R}^n$,

Now differentiating g_{ij} in (A.3.2) with respect to κ_k,

$$\begin{aligned} \frac{\partial}{\partial\kappa_k}g_{ij}(x,t,s,\kappa) = \frac{\partial}{\partial\kappa_k}\Bigg(&\Big(\epsilon+\frac{p}{(\sqrt{|\kappa|^2+\epsilon^2})^{2-p}}+a(x)\frac{q}{(\sqrt{|\kappa|^2+\epsilon^2})^{2-q}}\Big)\delta_{ij} \\ &+\Big(\frac{p(p-2)}{(\sqrt{|\kappa|^2+\epsilon^2})^{4-p}}+a(x)\frac{q(q-2)}{(\sqrt{|\kappa|^2+\epsilon^2})^{4-q}}\Big)\kappa_i\kappa_j\Bigg), \end{aligned}$$

$$\begin{aligned} \frac{\partial g_{ij}}{\partial\kappa_k} = &\left(\frac{p(p-2)|\kappa|}{(\sqrt{|\kappa|^2+\epsilon^2})^{4-p}}+a(x)\frac{q(q-2)|\kappa|}{(\sqrt{|\kappa|^2+\epsilon^2})^{4-q}}\right)\delta_{ij}\frac{\partial}{\partial\kappa_k}(|\kappa|) \\ &+\left(\frac{p(p-2)(p-4)|\kappa|}{(\sqrt{|\kappa|^2+\epsilon^2})^{6-p}}+a(x)\frac{q(q-2)(q-4)|\kappa|}{(\sqrt{|\kappa|^2+\epsilon^2})^{6-q}}\right)\kappa_i\kappa_j\frac{\partial}{\partial\kappa_k}(|\kappa|) \\ &+\left(\frac{p(p-2)}{(\sqrt{|\kappa|^2+\epsilon^2})^{4-p}}+a(x)\frac{q(q-2)}{(\sqrt{|\kappa|^2+\epsilon^2})^{4-q}}\right)\frac{\partial}{\partial\kappa_k}(\kappa_i\kappa_j) \end{aligned}$$

which further gives,

$$\frac{\partial g_{ij}}{\partial \kappa_k} = \left(\frac{p(p-2)}{(\sqrt{|\kappa|^2+\epsilon^2})^{4-p}} + a(x)\frac{q(q-2)}{(\sqrt{|\kappa|^2+\epsilon^2})^{4-q}} \right)(\kappa_i\delta_{jk} + \kappa_j\delta_{ik} + \kappa_k\delta_{ij}) + \left(\frac{p(p-2)(p-4)}{(\sqrt{|\kappa|^2+\epsilon^2})^{6-p}} + a(x)\frac{q(q-2)(q-4)}{(\sqrt{|\kappa|^2+\epsilon^2})^{6-q}} \right)\kappa_i\kappa_j\kappa_k$$

Then,

$$\begin{aligned}\frac{\partial g_{ij}}{\partial \kappa_k} &\leqslant \left(\frac{p(p-2)}{(|\kappa|+1)^{4-p}} + a(x)\frac{q(q-2)}{(|\kappa|+1)^{4-q}} \right)(\kappa_i\delta_{jk} + \kappa_j\delta_{ik} + \kappa_k\delta_{ij}) \\ &\quad + \left(\frac{p(p-2)(p-4)}{(\sqrt{|\kappa|^2+\epsilon^2})^{6-p}} + a(x)\frac{q(q-2)(q-4)}{(\sqrt{|\kappa|^2+\epsilon^2})^{6-q}} \right)\kappa_i\kappa_j\kappa_k \\ &\leqslant \left(\frac{|p(p-2)|}{(|\kappa|+1)^{4-p}} + a(x)\frac{|q(q-2)|}{(|\kappa|+1)^{4-q}} \right)(|\kappa_i| + |\kappa_j| + |\kappa_k|) \\ &\quad + \left(\frac{p(p-2)(p-4)}{(\sqrt{|\kappa|^2\epsilon^2+\epsilon^2})^{6-p}} + a(x)\frac{q(q-2)(q-4)}{(\sqrt{|\kappa|^2\epsilon^2+\epsilon^2})^{6-q}} \right)|\kappa_i||\kappa_j||\kappa_k|\end{aligned}$$

which implies,

$$\begin{aligned}\frac{\partial g_{ij}}{\partial \kappa_k} &\leqslant \left(\frac{|p(p-2)|}{|\kappa|((|\kappa|+1)^{3-p}} + a(x)\frac{|q(q-2)|}{|\kappa|((|\kappa|+1)^{3-q}} \right)3|\kappa| \\ &\quad + \left(\frac{p(p-2)(p-4)}{|\kappa|^3\epsilon(\sqrt{|\kappa|^2+1})^{3-p}} + a(x)\frac{q(q-2)(q-4)}{|\kappa|^3\epsilon(\sqrt{|\kappa|^2+1})^{3-q}} \right)|\kappa|^3 \\ &\leqslant \frac{3|p(p-2)|}{(|\kappa|+1)^{3-p}} + |a(x)|\frac{3|q(q-2)|}{(|\kappa|+1)^{3-q}} + \frac{p(p-2)(p-4)}{\epsilon^2(\sqrt{|\kappa|^2+1})^{3-p}} \\ &\qquad + a(x)\frac{q(q-2)(q-4)}{\epsilon^2(\sqrt{|\kappa|^2+1})^{3-q}}\end{aligned}$$

Thus, we obtain

$$\frac{\partial g_{ij}}{\partial \kappa_k} \leqslant C\big(1+\tfrac{1}{\epsilon^2}\big)\left(\frac{1}{(|\kappa|+1)^{3-p}} + \frac{1}{(|\kappa|+1)^{3-q}} \right) \leqslant 2C\big(1+\tfrac{1}{\epsilon^2}\big)\frac{1}{(|\kappa|+1)^{3-q}}$$

where constant $C = p(p-2)(p-4) \in [0,3)$. that is,

$$\left|\frac{\partial g_{ij}}{\partial \kappa_k}\right| \leqslant 2C\big(1+\tfrac{1}{\epsilon^2}\big)\frac{1}{(|\kappa|+1)^{3-q}}. \tag{A.3.9}$$

Also, differentiating g_{ij} with respect to x_k, we have

$$\frac{\partial}{\partial x_k} g_{ij}(x,t,s,\kappa) = \frac{\partial}{\partial x_k}\Bigg(\big(\epsilon + \frac{p}{(\sqrt{|\kappa|^2+\epsilon^2})^{2-p}} + a(x)\frac{q}{(\sqrt{|\kappa|^2+\epsilon^2})^{2-q}}\big)\delta_{ij} + \big(\frac{p(p-2)}{(\sqrt{|\kappa|^2+\epsilon^2})^{4-p}} + a(x)\frac{q(q-2)}{(\sqrt{|\kappa|^2+\epsilon^2})^{4-q}} \big)\kappa_i\kappa_j \Bigg)$$

$$
\begin{aligned}
\frac{\partial g_{ij}}{\partial x_k} &= \left(\frac{q}{\left(\sqrt{|\kappa|^2+\epsilon^2}\right)^{2-q}}\delta_{ij} + \frac{q(q-2)}{\left(\sqrt{|\kappa|^2+\epsilon^2}\right)^{4-q}}\kappa_i\kappa_j \right)\frac{\partial}{\partial x_k}a(x) \\
&\leqslant \left(\frac{q}{|\kappa|^{2-q}} + \frac{q(q-2)}{(|\kappa|+1)^{4-q}}\kappa_i\kappa_j \right)\frac{\partial}{\partial x_k}a(x) \\
&\leqslant \left(\frac{q}{|\kappa|^{2-q}} + \frac{|q(q-2)|}{|\kappa|^2(|\kappa|+1)^{2-q}}|\kappa_i||\kappa_j| \right)\Big|\frac{\partial}{\partial x_k}a(x)\Big|.
\end{aligned}
$$

Thus,

$$
\Big|\frac{\partial g_{ij}}{\partial x_k}\Big| \leqslant \left(\frac{2}{|\kappa|^{2-q}} + \frac{2}{|\kappa|^2(|\kappa|+1)^{2-q}} \right)C, \tag{A.3.10}
$$

where $|\frac{\partial}{\partial x_k}a(x)| \leqslant C$, since $a(x)$ is Lipschitz continuous.

Next, we have for $g(x,t,s,\kappa)$ in (A.3.3),

$$
|g| = \left| -\frac{q}{(\sqrt{|\kappa|^2+\epsilon^2})^{2-q}}\sum_{i=1}^{n}\kappa_i\frac{\partial a}{\partial x_i} + \lambda(s-f_\delta) \right| \leqslant C(|\kappa|+1), \tag{A.3.11}
$$

and differentiating $g(x,t,s,\kappa)$ with respect to κ_k gives,

$$
\begin{aligned}
\frac{\partial g}{\partial \kappa_k} &= \frac{\partial}{\partial \kappa_k}\big(-\frac{q}{(\sqrt{|\kappa|^2+\epsilon^2})^{2-q}}\sum_{i=1}^{n}\kappa_i\frac{\partial a}{\partial x_i} + \lambda(s-f_\delta)\big) \\
&= \frac{-q\kappa_k a_{x_k}}{(\sqrt{|\kappa|^2+\epsilon^2})^{2-q}} - \frac{q(q-2)\kappa_k}{(\sqrt{|\kappa|^2+\epsilon^2})^{4-q}}\sum_{i=1}^{n}\kappa_i a_{x_i}.
\end{aligned}
$$

The above implies

$$
\begin{aligned}
&\frac{\partial g}{\partial \kappa_k} \leqslant \frac{q|\kappa_k a_{x_k}|}{(\sqrt{|\kappa|^2+\epsilon^2})^{2-q}} + \frac{|q(q-2)||\kappa_k|}{(\sqrt{|\kappa|^2+\epsilon^2})^{4-q}}\Big|\sum_{i=1}^{n}\kappa_i a_{x_i}\Big| \\
&\leqslant \frac{C|\kappa|}{(\sqrt{|\kappa|^2+\epsilon^2})^{2-q}} + \frac{C|\kappa|^2}{|\kappa|^2(\sqrt{|\kappa|^2+\epsilon^2})^{2-q}} \leqslant \frac{C|\kappa|}{\epsilon(\sqrt{|\kappa|^2+1})^{2-q}} + \frac{C}{\epsilon(\sqrt{|\kappa|^2+1})^{2-q}} \\
&\leqslant \frac{C|\kappa|}{\epsilon^2(|\kappa|+1)^{2-q}} + \frac{C}{\epsilon^2(|\kappa|+1)^{2-q}} \leqslant \frac{C(|\kappa|+1)}{\epsilon^2(|\kappa|+1)^{2-q}}
\end{aligned}
$$

Thus,

$$
\Big|\frac{\partial g}{\partial \kappa_k}\Big| \leqslant \frac{C}{\epsilon^2}(|\kappa|+1)^{q-1}. \tag{A.3.12}
$$

On the other hand, differentiating g with respect to x_k,

$$
\begin{aligned}
\frac{\partial g}{\partial x_k} &= \frac{\partial}{\partial x_k}\big(-\frac{q}{(\sqrt{|\kappa|^2+\epsilon^2})^{2-q}}\sum_{i=1}^{n}\kappa_i\frac{\partial a}{\partial x_i} + \lambda(s-f_\delta)\big) \\
&= -\frac{q}{(\sqrt{|\kappa|^2+\epsilon^2})^{2-q}}\sum_{i=1}^{n}\kappa_i\frac{\partial^2 a}{\partial x_i\partial x_k} + \lambda\frac{\partial}{\partial x_k}f_\delta(x),
\end{aligned}
$$

and, further, since $a(x)$ has bounded second order derivative and also the derivative of $f_\delta(x)$ is bounded, we have

$$\begin{aligned}\frac{\partial g}{\partial x_k} &\leqslant \frac{q}{(\sqrt{|\kappa|^2+\epsilon^2})^{2-q}}\sum_{i=1}^{n}|\kappa_i a_{x_i x_k}| + \lambda|\frac{\partial}{\partial x_k}f_\delta(x)| \\ &\leqslant \frac{q|\kappa|C}{|\kappa|^{2-q}} + C \leqslant C|\kappa|^{q-1} + C \leqslant C(|\kappa|+1)^{q-1} + C.\end{aligned}$$

Thus,

$$\left|\frac{\partial g}{\partial x_k}\right| \leqslant C(|\kappa|+1). \tag{A.3.13}$$

Now, to prove (A.3.4), we refer the parabolicity condition proved in section A.2,

$$\nu|\xi|^2 \leqslant \sum_{i,j=1}^{n} g_{ij}(x,t,s,\kappa)\,\xi_i\xi_j \leqslant \mu|\xi|^2,$$

where $\xi = (\xi_1, \ldots, \xi_n)$ is a real vector and $\nu, \mu > 0$ are constants.
This implies that (A.3.4) holds for the case $m = 2$.

Next to prove (A.3.5), we apply (A.3.9), (A.3.11) and (A.3.12) on the left-hand side of (A.3.5), which gives,

$$\begin{aligned}&\left|\frac{\partial g_{ij}}{\partial \kappa_k}\right|(1+|\kappa|)^3 + |g| + \left|\sum_{k=1}^{n}\frac{\partial g}{\partial \kappa_k}\right|(1+|\kappa|) \\ &\leqslant 2C\big(1+\tfrac{1}{\epsilon^2}\big)\frac{1}{(1+|\kappa|)^{3-q}}(1+|\kappa|)^3 + C(1+|\kappa|) + \frac{C}{\epsilon^2}(1+|\kappa|)^{q-1}(1+|\kappa|) \\ &\leqslant 2C\big(1+\tfrac{1}{\epsilon^2}\big)(1+|\kappa|)^q + C(1+|\kappa|) + \frac{C}{\epsilon^2}(1+|\kappa|)^q \\ &\leqslant 3C\big(1+\tfrac{1}{\epsilon^2}\big)(1+|\kappa|)^2\end{aligned}$$

Hence, we get,

$$\left|\sum_{k=1}^{n}\frac{\partial g_{ij}}{\partial \kappa_k}\right|(1+|\kappa|)^3 + |g| + \left|\sum_{k=1}^{n}\frac{\partial g}{\partial \kappa_k}\right|(1+|\kappa|) \leqslant \mu_1(1+|\kappa|)^2,$$

where $\mu_1 \geqslant 3C\big(1+\frac{1}{\epsilon^2}\big) > 0$ is constant. This implies that (A.3.5) holds true.

Next, in case of (A.3.6), we apply (A.3.10) and (A.3.13) on the left-hand side of (A.3.6), which gives,

$$\begin{aligned}
&\left|\frac{\partial g_{ij}}{\partial x_k}\right|(1+|\kappa|)^2+\left|\frac{\partial g}{\partial x_k}\right| \\
&\qquad\leqslant \frac{2C}{|\kappa|^{2-q}}(1+|\kappa|)^2+\frac{C|q(q-2)|}{(1+|\kappa|)^{2-q}}(1+|\kappa|)^2+C(1+|\kappa|) \\
&\qquad\leqslant \frac{2C}{|\kappa|^{2-q}}(1+|\kappa|)^2+C|q(q-2)|(1+|\kappa|)^3+C(1+|\kappa|) \\
&\qquad\leqslant (1+|\kappa|)^3\left(\frac{2C}{|\kappa|^{2-q}(1+|\kappa|)}+C|q(q-2)|+\frac{C}{(1+|\kappa|)^2}\right)
\end{aligned}$$

Thus,

$$\left|\frac{\partial g_{ij}}{\partial x_k}\right|(1+|\kappa|)^2+\left|\frac{\partial g}{\partial x_k}\right| \leqslant [\beta+P(|\kappa|)](1+|\kappa|)^3,$$

where the constant $\beta = C|q(q-2)|+\epsilon > 0$ depending on ϵ is sufficiently small, and $P(|\kappa|) = \frac{2C}{|\kappa|^{2-q}(1+|\kappa|)}+\frac{C}{(1+|\kappa|)^2} \to 0$ as $|\kappa| \to \infty$.
Hence, (A.3.6) holds true for $m = 2$.

The remaining conditions (A.3.7) and (A.3.8) also hold true, since $\dfrac{\partial g_{ij}}{\partial s} = 0$ and $-\dfrac{\partial g}{\partial s} = -\lambda < 0$.

(c) The condition that, for $(x,t) \in \overline{\Omega}_T$, $|s| \leqslant M$ and $|r| \leqslant M_1$, the functions $g_{ij}(x,t,s,\kappa)$ and $g(x,t,s,\kappa)$ are continuously differentiable with respect to all of their arguments, holds true for being smooth up to order 2.

(d) Since $f_\delta \in C^\infty(\overline{\Omega}_T)$, we can conclude that $f_\delta \in C(\overline{\Omega}_T) \cap H^{2+\beta,1+\frac{\beta}{2}}(\Omega_T)$ is satisfied.

(e) The condition that, each point of the boundary $\partial\Omega$ can be touched from without by a ball (or cone) of fixed size in such a way that the ball (cone) does not have any points in common with Ω, holds true because the boundary of domain Ω satisfies Lipschitz continuity.

A.4 Arguments for initial and boundary values in proof of Theorem 4.4.6

In the proof of Theorem 4.4.6, for $w \in L^2(\Omega_T)$ and $f_\delta \in C^\infty(\overline{\Omega})$, we need to prove that $w(x,t) = (u_\delta)_t(x,t)$ and $u_\delta(x,0) = f_\delta(x)$ hold.
It is established that $\{u_\delta^{\epsilon_i}\}_{i\in\mathbb{N}} \in L^2(0,T;W^{1,2}(\Omega)) \cap C^\infty(\Omega_T)$ is the sequence of solution to the approximated problem (4.4.21)–(4.4.22), and we have the following

convergence results for the subsequence $\{u_\delta^{\epsilon_{ij}}\}_{j\in\mathbb{N}} \subset \{u_\delta^{\epsilon_i}\}$, from the proof of Theorem 4.4.6, as $\epsilon_i \to 0$,

$$u_\delta^{\epsilon_{ij}} \rightharpoonup u_\delta \text{ weakly* in } L^\infty(\Omega_T) \tag{A.4.1}$$

$$(u_\delta^{\epsilon_{ij}})_t \rightharpoonup w \text{ weakly in } L^2(\Omega_T), \tag{A.4.2}$$

as $\epsilon_i \to 0$, for some $u_\delta \in L^\infty(\Omega_T)$ and $w \in L^2(\Omega_T)$.

Let z be a function in $C^{2,1}(\Omega_T)$, with compact support in Ω_T. From weak convergence of $(u_\delta^{\epsilon_{ij}})_t$ in $L^2(\Omega_T)$ from (A.4.2), we have, as $\epsilon_i \to 0$

$$\int_0^s \int_\Omega (u_\delta^{\epsilon_{ij}})_t z \, dx \, dt \to \int_0^s \int_\Omega wz \, dx \, dt. \tag{A.4.3}$$

On the other hand, because we have a uniform bound for $u_\delta^{\epsilon_{ij}}$ in $L^\infty(\Omega_T)$ and the bounded constant is integrable on Ω_T, then by dominated convergence theorem, as $\epsilon_i \to 0$,

$$\int_0^s \int_\Omega u_\delta^{\epsilon_{ij}} z_t \, dx \, dt \to \int_0^s \int_\Omega u_\delta z_t \, dx \, dt.$$

Then interchanging the order of integration in the above expression and applying integration by parts formula with respect to t for the left-hand side integral, we obtain

$$-\int_\Omega \int_0^s (u_\delta^{\epsilon_{ij}})_t z \, dt \, dx \to \int_\Omega \int_0^s u_\delta z_t \, dt \, dx. \tag{A.4.4}$$

From (A.4.3) and (A.4.4), the uniqueness of limit then implies

$$\int_\Omega \int_0^s wz \, dt \, dx = -\int_\Omega \int_0^s u_\delta z_t \, dt \, dx$$

which by definition of weak derivative implies that w is the derivative of u_δ with respect to t that exists in a weak sense, in $L^2(\Omega_T)$.

By the First fundamental theorem of Calculus, we then consider that u is a.e. equal to an antiderivative of w as follows,

$$u_\delta(\cdot, t) := \xi + \int_0^t w(\cdot, s) \, ds,$$

for $\xi \in L^2(\Omega)$, a.e. $t \in (0, T]$. Clearly, the above gives us that $u_\delta(\cdot, 0) = \xi$.

We next prove that $u_\delta(x, 0) = f_\delta(x)$. Let $\Psi \in C^\infty(0, T; L^2(\Omega))$ such that $0 < \Psi(x, 0) \in L^2(\Omega)$ and $\Psi(x, T) = 0$, for almost every $x \in \Omega$. Moreover, $\Psi_t \in L^2(\Omega_T)$.

Consider the integral $\int_0^T \int_\Omega u_\delta^{\epsilon_{ij}} \Psi_t \, dx \, dt$, and applying Fubini's Theorem to interchange the integration order, we use integration by parts formula with respect to t

to obtain

$$\int_0^T \int_\Omega u_\delta^{\epsilon_{ij}} \Psi_t \, dx \, dt = -\int_\Omega \int_0^T (u_\delta^{\epsilon_{ij}})_t \Psi \, dt \, dx - \int_\Omega u_\delta^{\epsilon_{ij}}(x,0)\Psi(x,0)\, dx,$$
$$= -\int_0^T \int_\Omega (u_\delta^{\epsilon_{ij}})_t \Psi \, dx \, dt - \int_\Omega u_\delta^{\epsilon_{ij}}(x,0)\Psi(x,0)\, dx.$$

Since $u_\delta^{\epsilon_{ij}}(x,0) = f_\delta(x)$, then the above expression implies

$$-\int_\Omega f_\delta(x)\Psi(x,0)\, dx - \int_0^T \int_\Omega (u_\delta^{\epsilon_{ij}})_t \Psi \, dx \, dt = \int_0^T \int_\Omega u_\delta^{\epsilon_{ij}} \Psi_t \, dx \, dt.$$

Taking limit as $\epsilon_i \to 0$ in the above equation, we apply the convergence conditions (A.4.1) and (A.4.2) to get,

$$-\int_\Omega f_\delta(x)\Psi(x,0)\, dx - \int_0^T \int_\Omega w\Psi \, dx \, dt = \int_0^T \int_\Omega u_\delta \Psi_t \, dx \, dt.$$

Again interchanging the integration order through Fubini's theorem on the right hand side above, we apply integration by parts formula with respect to t to obtain,

$$-\int_\Omega f_\delta(x)\Psi(x,0)\, dx - \int_0^T \int_\Omega w\Psi \, dx \, dt = \int_\Omega \int_0^T u_\delta \Psi_t \, dt \, dx$$
$$= -\int_\Omega \int_0^T w\Psi \, dt \, dx - \int_\Omega u_\delta(x,0)\Psi(x,0)\, dx.$$

Further interchanging the order of integration on the right hand side above,

$$-\int_\Omega f_\delta(x)\Psi(x,0)\, dx = -\int_\Omega u_\delta(x,0)\Psi(x,0)\, dx,$$

we thus obtain, for $\Psi(x,0) \in L^2(\Omega)$ and,

$$u_\delta(x,0) = f_\delta(x) \quad \text{for every } x \in \Omega,$$

as required.

A.5 FORTRAN code for the image restoration model

```
*       * PROGRAM TLMBM *
*       * Purpose *
*       Test program for limited memory bundle subroutine for
    large-scale
*       unconstrained nonsmooth optimization with noise
    reduction problem.
*       * Parameters *
*       I    N               Number of variables.
*       I    NA              Maximum bundle dimension, NA >= 2.
*       I    MCU             Upper limit for maximum number of
    stored
*                              corrections, MCU >= 3.
*       I    NW              Dimension of the work vector W:
*                              NW >= N*(9+2*NA+2*(MCU+1)) + 3*NA
*                               + 3*(MCU+1)*(MCU+2)/2 + 9*(MCU+1))
*                               + N*MG + N*MH.
*       * Variables *
*       I    MC              Maximum number of stored corrections,
*                              MCU >= MC >= 3.
*       R    X(N)            Vector of variables.
*       R    F               Value of the objective function.
*       R    RPAR(6)         Real parameters:
*              RPAR(1)         Tolerance for change of function
    values.
*              RPAR(2)         Tolerance for the function value.
*              RPAR(3)         Tolerance for the termination
    criterion.
*              RPAR(4)         Distance measure parameter, 0 <=
    RPAR(4).
*              RPAR(5)         Line search parameter, 0 < RPAR(5)
    < 0.25.
*              RPAR(6)         Maximum stepsize, 1 < RPAR(6).
*       I    IPAR(7)         Integer paremeters:
*              IPAR(1)         Exponent for distance measure.
*              IPAR(2)         Maximum number of iterations.
*              IPAR(3)         Maximum number of function
    evaluations.
*              IPAR(4)         Maximum number of iterations with
    changes of
*                                function values smaller than RPAR
    (1).
*              IPAR(5)         Printout specification:
*                                  -1  - No printout.
*                                   0  - Only the error messages.
*                                   1  - The final values of the
```

```
   objective
*                                   function.
*                                2 - The final values of the
   objective
*                                   function and the most
   serious
*                                   warning messages.
*                                3 - The whole final solution.
*                                4 - At each iteration values
   of the
*                                   objective function.
*                                5 - At each iteration the
   whole
*                                   solution
*            IPAR(6)        Selection of the method:
*                                0 - Limited memory bundle
   method.
*                                1 - L-BFGS bundle method.
*            IPAR(7)        Selection of the scaling:
*                                0 - Interval scaling with STU
   /UTU.
*                                1 - Interval scaling with STS
   /STU.
*                                2 - Scaling at every
   iteration with STU/UTU.
*                                3 - Scaling at every
   iteration with STS/STU.
*                                4 - Preliminary scaling with
   STU/UTU.
*                                5 - Preliminary scaling with
   STS/STU.
*                                6 - No scaling.
*      I    IOUT(3)         Integer parameters:
*            IOUT(1)         Number of used iterations.
*            IOUT(2)         Number of used function evaluations
   .
*            IOUT(3)         Cause of termination:
*                                1 - The problem has been
   solved.
*                                   with desired accuracy.
*                                2 - Number of function calls
   > IPAR(3).
*                                3 - Number of iterations >
   IPAR(2).
*                                4 - Changes in function
   values < RPAR(1)
*                                   in IPAR(4) subsequent
```

```
    iterations.
*                                  5   - F(1) < RPAR(2).
*                                 -1   - Two consecutive restarts
    or number
*                                        of restarts > maximum
    number of
*                                        restarts.
*                                 -2   - TMAX < TMIN in two
    subsequent
*                                        iterations.
*                                 -3   - Failure in function or
    subgradient
*                                        calculations (assigned by
     the user).
*                                 -4   - Failure in attaining the
    demanded
*                                        accuracy.
*                                 -5   - Invalid input parameters.
*                                 -6   - Not enough working space.
*      R    W(NW)           Work vector.
*      * Variables in COMMON /CPUTIME/ *
*      R    STARTU          Initial CPU-time.
*      R    TIME            Maximum CPU-time in seconds.
*      * Subprograms used *
*      S    LMBMU           Initialization of limited memory
    bundle method
*                           for nonsmooth optimization.
*      RF   ETIME           Execution time.
*       Initial code by Napsu Karmitsa (nee Haarala) (2002 -
    2004)

            PROGRAM TLMBM

*       Parameters
       INTEGER N,NA,MCU,NW
       PARAMETER(
!      N = number of variables (i.e. dimension of u).
     &      N = 1000,
c       For the rest of these do not touch
     &      NA = 2,
     &      MCU = 7,
     &      NW = 1 + 9*N + 2*N*NA + 3*NA + 2*N*(MCU+1) +
     &       3*(MCU+2)*(MCU+1)/2 + 9*(MCU+1))

*       Scalar Arguments
       INTEGER MC
       DOUBLE PRECISION F
```

```
*       Array Arguments
        INTEGER IPAR(7),IOUT(3)
        DOUBLE PRECISION W(NW),X(N),RPAR(6)

*       Local Scalars
        INTEGER MCINIT,I,J,IMODEL

        DOUBLE PRECISION Z(N),ZZ(N),APU(N)
        DOUBLE PRECISION L,H,MY,G,DELTA,P,Q
        DOUBLE PRECISION TEMP,S,EXACT,TEMP2,VIRHE,JJ,VEKT(N)

*       CPU-time
        REAL START, FINI, ETIME, TARRAY(2)
        REAL STARTU, FINIU, TIME
        COMMON /CPUTIME/STARTU,TIME

*       Common blocks
        COMMON /KIMPPUM/ Z,H,MY,G,P,Q,APU,IMODEL

*       External Subroutines
        EXTERNAL LMBMU

        TIME=1800.0E+00

*       Initial number of stored corrections
*
        MC = 7
        MCINIT = MC

*       Choice of integer and real parameters
*
        DO 10 I = 1,7
           IPAR(I) = 0
 10     CONTINUE

*       Printout specification
*
        IPAR(5) = 2

*       Selection of the method
*
        IPAR(6) = 0

*       Selection of the scaling
*
        IPAR(7) = 2
```

```
      DO 20 I = 1,6
         RPAR(I) = 0.0D0
 20   CONTINUE

*     Desired accuracy
*
c This can be changed. Smaller value should produce more
   accurate results
c from optimization view point.
c
      RPAR(3) = 1.0D-4

*     Locality measure
*
c For convex problems 0 should be ok.

      RPAR(4) = 0.0D+00

*     Stepsize
*
c This parameter has large effect on results, for values in
   (1,1000]

      RPAR(6) = 2.50D+00

*     Line search parameter
*
c This parameter affects how we accept the new point in
   optimization procedure
c preferable values that are smaller than 0.25

      RPAR(5) = 0.2499D+00

*     Maximum numbers of iterations and function evaluations
*
      IPAR(2) =500000000
      IPAR(3) =500000000

c     These are parameters for models:
c     MY, G = regularization parameters
c     P, Q = exponents
c     MY = lambda and g=a for the image restoration model

!     Selection of model:
c     IMODEL = 3 -> model with a(x) depending on noisy signal
   's gradient
```

```
c      IMODEL = 4 -> model with a(x) depending on non noisy
    signal's gradient
c       Just for testing the second example (Change ISIGNA=1 in
     EXACT below).
c      IMODEL = 5 -> model with a(x) depending on non noisy
    signal's gradient
c       Just for testing the first example (Change ISIGNA=2 in
    EXACT below).
c

       IMODEL=4
       L = 1.0D0
c Try different values for P and Q.
       MY = 1000.0D0
!      MY = 5.D-2
!        P = 1.9D0
       P = 1.0001D0
       Q = 2.0D0 ! for IMODEL 3 and 4
!        Q = 1.0007D0


c      DELTA = noise
       DELTA = 0.2D0

c       Discretization step
       H = L/(N+1)

C      read N random numbers from  NOISE.DAT
C      S = randomnumber in [-1.0,1.0]

       OPEN(10,FILE='noise.dat')
       READ (10,*) I
       IF (I.LT.N) THEN
          PRINT *,I,' RANDOM NUMBERS IN FILE NOISE.DAT, BUT N
              =',N,'.'
          PRINT *,'CANNOT CONTINUE. BYE.'
          STOP
       END IF
       TEMP = H
       DO I=1,N
           READ(10,*) S
           Z(I) = EXACT(I,N,TEMP) + DELTA*S
           ZZ(I)= EXACT(I,N,TEMP)
           TEMP = TEMP + H
       END DO
       CLOSE(10)
```

```
      CALL DCOPY (N,Z,1,X,1)

*
*     CPU-time
*
      START = ETIME(TARRAY)
      STARTU = TARRAY(1)

*     Solution
*
      CALL LMBMU(N,NA,MC,MCU,NW,X,F,IPAR,IOUT,RPAR,W)

*     CPU-time
*
      FINI = ETIME(TARRAY)
      FINIU = TARRAY(1)

      PRINT*
      PRINT*,'ITERM    = ',IOUT(3)
      PRINT*
      PRINT*,'F(X)     = ',F
      PRINT*,'N        = ',N
      PRINT*,'NA       = ',NA
      PRINT*,'MCINIT   = ',MCINIT
      PRINT*,'MC       = ',MC
      PRINT*,'MCU      = ',MCU
      PRINT*,'NIT      = ',IOUT(1)
      PRINT*,'NFV      = ',IOUT(2)
      PRINT*,'XMAX     = ',RPAR(6)
      PRINT*,'GAM      = ',RPAR(4)
      PRINT*,'EPSL     = ',RPAR(5)
      PRINT*,'EPS      = ',RPAR(3)
      PRINT*,'METHOD   = ',IPAR(6)
      PRINT*,'SCALING  = ',IPAR(7)

C     Error:

      TEMP2 = H
      TEMP = 0.D0
      DO I=1,N
         TEMP = TEMP + (X(I) -EXACT(I,N,TEMP2))**2
         TEMP2=TEMP2+H
      END DO
!      TEMP = TEMP/N
      VIRHE = SQRT(TEMP)/DBLE(N)
      PRINT*
      PRINT*,'ERROR=',VIRHE
```

```
        !PRINT*,'JJ  =',JJ(N,X,H,MY,G,P,Q,Z,VEKT,IMODEL)
        !print*,'TARKKA=',JJ(N,ZZ,H,MY,G,P,Q,Z,VEKT,IMODEL)

        PRINT*
        PRINT*,'Used time = ',FINIU-STARTU
        PRINT*

C       Printing the results to files

200     CONTINUE

        OPEN(10,FILE='U.DAT')
        OPEN(30,FILE='Z.DAT')
        OPEN(40,FILE='EU.DAT')
c
        TEMP = 0.D0
        DO I=1,N
           TEMP = TEMP + H
           WRITE(10,'(2F10.4)') TEMP,X(I)
           WRITE(30,'(2F10.4)') TEMP,Z(I)
           WRITE(40,'(2F10.4)') TEMP,EXACT (I,N,TEMP)
        END DO

        CLOSE(10)
        CLOSE(30)
        CLOSE(40)

c         END DO
 999    CONTINUE

        STOP
        END


*********************************
*
*       * SUBROUTINE FUNDER *

*       * PURPOSE *
*
*       COMPUTATION OF THE VALUE AND THE SUBGRADIENT OF THE
    OBJECTIVE
*       FUNCTION.

*       * CALLING SEQUENCE *
```

```
*
*      CALL FUNDER(N,X,F,G)

*      * PARAMETERS *
*
*      II  N           NUMBER OF VARIABLES.
*      RI  X(N)        A VECTOR OF VARIABLES.
*      RO  F           THE VALUE OF THE OBJECTIVE FUNCTION.
*      RO  G(N)        THE SUBGRADIENT OF THE OBJECTIVE
    FUNCTION.
*
      SUBROUTINE FUNDER(NF,U,F,D)
      INTEGER NF,I,IMODEL
      DOUBLE PRECISION F,U(NF),D(NF),H, MY, G, Z(1000),EPS,P,
         Q,APU(1000)
      DOUBLE PRECISION H2,H3,TMP,TMP2,BETA,PI,GN(1000)

!       DOUBLE PRECISION pw(NF),Gs(NF),CNV(NF)
C
      parameter (eps = 1.d-12)
      COMMON /KIMPPUM/ Z,H,MY,G,P,Q,APU,IMODEL

!        BETA = 8.10D0 ! This goes well with IMODEL 4 and 5.
                       ! In fact anything below 4.0 will give
                          almost the same result with imodel 4
                          and
                       ! Anything below 2.0 will give almost
                          the same result with imodel 5.
                       ! Interesting values with imodel=5 are
                          at least 4.1, 2.1, and 0.1.
!        BETA = 0.370D0 ! This goes approximately well with
    IMODEL 3.
                       ! Smaller values will make more
                          staircase effect.
                       ! The suitable value debends on noise
                          level.
        BETA = 4.10D0
      H2=1.D0/(H*H)
      H3=1.D0/H
      F = 0.D0

      PI=DACOS(-1.D0)

      IF (IMODEL == 3) THEN
c      this is the model with a(x) computed from noisy signals
     subgradient.
!      The noisy signal's gradient is computed here at the
```

```
    beginning.
!     Note that it would be enough to compute it once at the
    beginning of the program.

          GN(1) = Z(1) !/DBLE(NF) !Z(0)=0 and we take number
               of discretization points as step
          IF (ABS(GN(1)) .GE. BETA) THEN
               GN(1) = 0.00D0
          ELSE
               GN(1) = 0.9D0
          END IF

          DO I=2,NF
               GN(I) = (Z(I)-Z(I-1)) !/DBLE(NF)

               ! Computing the piecewise function a(x) in the
                    thesis.
               IF (ABS(GN(I)) .GE. BETA) THEN
                    GN(I) = 0.00D0
               ELSE
                    GN(I) = 0.9D0
               END IF
          END DO

          TMP = U(1)*H3
          F = ABS(TMP)**P + GN(1)*ABS(TMP)**Q
          IF (ABS(TMP) .GE. eps) THEN
               D(1)=P*TMP/(ABS(TMP)**(2.0D0-P))
               APU(1)=GN(1)*Q*TMP/(ABS(TMP)**(2.0D0-Q))

          ELSE
               D(1)=0.0D0
               APU(1)=0.0D0
          END IF

          DO I=2,NF
                    !PRINT*,'GN=',I,GN(I)
               TMP = (U(I)-U(I-1))*H3
               F = F + ABS(TMP)**P + GN(I)*ABS(TMP)**Q
               IF (ABS(TMP) .GE. eps) THEN
                    D(I)=P*TMP/(ABS(TMP)**(2.0D0-P))
                    APU(I)=GN(I)*Q*TMP/(ABS(TMP)**(2.0D0-Q))
               ELSE
                    D(I)=0.0D0
                    APU(I)=0.0D0
               END IF
          END DO
```

```
        TMP = 0.D0
        DO I=1,NF-1
            TMP2 = (D(I)-D(I+1))*H3
            D(I) = TMP2 + MY*( U(I) - Z(I))
            TMP2 =  (APU(I)-APU(I+1))*H3
            D(I) = D(I)+TMP2
            TMP = TMP  + (U(I)-Z(I))**2
        END DO

        TMP2 = D(NF)*H3
        D(NF) = D(NF)*H3 + GN(NF)*TMP2 + MY*(U(NF) - Z(NF))
        TMP = TMP + (U(NF) - Z(NF))**2
        F = F + 0.5D0*MY*TMP

    ELSE IF (IMODEL == 4) THEN
c     this is the model with function a(x) computed from
   original (not noisy) signal's gradient.
c     Just for testing the model with the one dimensional
   signal given in thesis (ISIGNAL=1 in EXACT below).
c     Note! G is "hard" coded. It cannot be used but to this
   one signal.
c     In addition, as the original signal is not changing, G
   could be coded as
c     a vector at the beginning and not computed at every
   iteration.

       TMP = U(1)*H3
       ! G is a function depending on original signals
           gradient
       G = 2.0D00*PI              ! cos(0)=1

       ! Computing the piecewise function a(x) in the
           thesis.
       IF (ABS(G) .GE. BETA) THEN
           G = 0.00D0
       ELSE
           G = 0.9D0
       END IF

       F = ABS(TMP)**P + G*ABS(TMP)**Q
       IF (ABS(TMP) .GE. eps) THEN
           D(1)=P*TMP/(ABS(TMP)**(2.0D0-P))
           APU(1)=G*Q*TMP/(ABS(TMP)**(2.0D0-Q))
       ELSE
           D(1)=0.0D0
           APU(1)=0.0D0
```

```
END IF

DO I=2,NF
    TMP = (U(I)-U(I-1))*H3

    ! Computing the (sub)gradient of the original
        signal.
    IF (I < 0.25*NF) THEN ! The signal is piecewise
        defined.
        G = 2.0*PI*COS(PI*2.0D+00/DBLE(I))
    ELSE IF (I<0.40*NF) THEN
        G = 0.0D00
    ELSE IF (I<0.50*NF) THEN
        G = 4.0D00
    ELSE IF (I<0.60*NF) THEN
        G = -4.0D00
    ELSE IF (I<0.75*NF) THEN
        G = 0.0D00
    ELSE
        G = -2.0*PI*COS(PI*2.0D+00/DBLE(I))
    END IF

    ! Computing the piecewise function a(x) in the
        thesis.
    IF (ABS(G) .GE. BETA) THEN
        G = 0.00D0
    ELSE
        G = 0.9D0
    END IF

    F = F + ABS(TMP)**P + G*ABS(TMP)**Q
    IF (ABS(TMP) .GE. eps) THEN
        D(I)=P*TMP/(ABS(TMP)**(2.0D0-P))
        APU(I)=G*Q*TMP/(ABS(TMP)**(2.0D0-Q))
    ELSE
        D(I)=0.0D0
        APU(I)=0.0D0
    END IF
 END DO

 TMP = 0.D0
 DO I=1,NF-1
     TMP2 = (D(I)-D(I+1))*H3
     D(I) = TMP2 + MY*( U(I) - Z(I))
     TMP2 =  (APU(I)-APU(I+1))*H3
     D(I) = D(I)+TMP2
     TMP = TMP  + (U(I)-Z(I))**2
```

```
          END DO

          TMP2 = D(NF)*H3
          D(NF) = D(NF)*H3 + G*TMP2 + MY*(U(NF) - Z(NF))
          TMP = TMP + (U(NF) - Z(NF))**2
          F = F + 0.5D0*MY*TMP

        ELSE ! IMODEL == 5
c         this is the model with function a(x) computed from
    original (not noisy) signal's gradient.
c         Just for testing the first example (Change ISIGNAL=2 in
    EXACT below).
c         Note! G is "hard" coded. It cannot be used but to this
    one signal.
c         In addition, as the original signal is not changing, G
    could be coded as
c         a vector at the beginning and not computed at every
    iteration.

          TMP = U(1)*H3
          ! G is a function debending on original signals
              gradient
          G = PI             ! cos(0)=1

          ! Computing the piecewise function a(x) in the
              thesis.
          IF (ABS(G) .GE. BETA) THEN
              G = 0.00D0
          ELSE
              G = 0.9D0
          END IF

          F = ABS(TMP)**P + G*ABS(TMP)**Q
          IF (ABS(TMP) .GE. eps) THEN
              D(1)=P*TMP/(ABS(TMP)**(2.0D0-P))
              APU(1)=G*Q*TMP/(ABS(TMP)**(2.0D0-Q))
          ELSE
              D(1)=0.0D0
              APU(1)=0.0D0
          END IF

          DO I=2,NF
              TMP = (U(I)-U(I-1))*H3

              ! Computing the (sub)gradient of the original
                  signal.
              IF (I < NF/2) THEN ! The signal is piecewise
```

```
            defined.
            G = PI*COS(PI/DBLE(I))
        ELSE IF (I<4*NF/5) THEN
            G = 0.0D00
        ELSE
            G = -2.0D00
        END IF

        ! Computing the piecewise function a(x) in the
            thesis.
        IF (ABS(G) .GE. BETA) THEN
            G = 0.00D0
        ELSE
            G = 0.9D0
        END IF

        F = F + ABS(TMP)**P + G*ABS(TMP)**Q
        IF (ABS(TMP) .GE. eps) THEN
            D(I)=P*TMP/(ABS(TMP)**(2.0D0-P))
            APU(I)=G*Q*TMP/(ABS(TMP)**(2.0D0-Q))
        ELSE
            D(I)=0.0D0
            APU(I)=0.0D0
        END IF
    END DO

    TMP = 0.D0
    DO I=1,NF-1
        TMP2 = (D(I)-D(I+1))*H3
        D(I) = TMP2 + MY*( U(I) - Z(I))
        TMP2 =  (APU(I)-APU(I+1))*H3
        D(I) = D(I)+TMP2
        TMP = TMP  + (U(I)-Z(I))**2
    END DO

    TMP2 = D(NF)*H3
    D(NF) = D(NF)*H3 + G*TMP2 + MY*(U(NF) - Z(NF))
    TMP = TMP + (U(NF) - Z(NF))**2
    F = F + 0.5D0*MY*TMP

END IF
RETURN

END

*       EMPTY SUBROUTINES
*
```

```
      SUBROUTINE FUN(NF,KA,X,FA)
C       .. Scalar Arguments ..
      DOUBLE PRECISION FA
      INTEGER KA,NF
C       ..
C       .. Array Arguments ..
      DOUBLE PRECISION X(*)
C       ..
      RETURN

      END
      SUBROUTINE DER(NF,KA,X,GA)
C       .. Scalar Arguments ..
      INTEGER KA,NF
C       ..
C       .. Array Arguments ..
      DOUBLE PRECISION GA(*) ,X(*)
C       ..
      RETURN

      END
      SUBROUTINE HES(NF,X,H)
C       .. Scalar Arguments ..
      INTEGER NF
C       ..
C       .. Array Arguments ..
      DOUBLE PRECISION H(*) ,X(*)
C       ..
      RETURN

      END

C**********************************
C
C     ORIGINAL SIGNAL
C     comment one of these.
C     You can add here some more different signals
************************************
      DOUBLE PRECISION FUNCTION EXACT (I,N,TEMP)
      INTEGER I, N, ISIGNA
      DOUBLE PRECISION PI,TEMP

      PI = ACOS(-1.D0)

c Selection of signal Set this 1 or 2.
      ISIGNA = 1
```

```
      IF (ISIGNA .EQ. 1) THEN
C
C Example 2:
C
      IF (I .LE. N/4) THEN
         EXACT = SIN(PI*TEMP*2.d0)
      ELSE IF (I .GT. 3*N/4 ) THEN
         EXACT = -SIN(PI*TEMP*2.d0)
      ELSE IF (I.GT.N/4 .AND. I.LE.4*N/10 .OR.
     &        I.GT.6*N/10 .AND. I.LE.3*N/4) THEN
         EXACT =  0.75d0
      ELSE IF (I.GT.4*N/10 .AND. I.LE.N/2 ) THEN
         EXACT =  4.d0*TEMP-1.5d0
      ELSE IF (I.GT.N/2 .AND. I.LE.6*N/10 ) THEN
         EXACT = 2.5d0-4.d0*TEMP
      END IF

      ELSE
C
C Example 1:
c
      IF (I .LE. N/2) THEN
         EXACT = SIN(PI*TEMP)
      ELSE IF (I.GT.N/2 .AND. I.LE.4*N/5) THEN
         EXACT =  0.65d0
      ELSE
         EXACT =  2.d0-2.d0*TEMP
      END IF

      END IF

      RETURN
      END

C***********************************
c The value of cost function (Same as in FUNDER but without
   derivatives)
c note that not all parameters are used in this example
C**********************************
      DOUBLE PRECISION FUNCTION JJ(N,U,H,MY,G,P,Q,Z,APU,
         IMODEL)
      INTEGER N,I
      DOUBLE PRECISION U(N), H, MY, G, P, Q, Z(N), APU(N)
      DOUBLE PRECISION H2,H3,TMP
c     Here U is the optimized variable, z is the noisy data,
c     h is the step size in discretization
```

```
      H3=1.D0/H
      JJ = 0.D0
c       PRINT*,'hihu',IMODEL

      IF (IMODEL .EQ. 1) THEN
c       PRINT*,'hihu'
C
      H2=1.D0/(H*H)

      TMP = U(1)*H3
      JJ = JJ + G*ABS(TMP)

      DO I=2,N
         TMP = (U(I)-U(I-1))*H3
         JJ = JJ + G*ABS(TMP)
      END DO

      TMP = (2.D0*U(1) - U(2))*H2
      JJ = JJ + 0.5D0*(MY*U(1)*TMP    + (U(1) - Z(1))**2)

      DO I=2,N-1
         TMP = (2.D0*U(I) - U(I-1) - U(I+1))*H2
         JJ = JJ + 0.5D0*(MY*U(I)*TMP + (U(I) - Z(I))**2)
      END DO
      TMP = (2.D0*U(N) - U(N-1))*H2
      JJ = JJ + 0.5D0*(MY*U(N)*TMP + (U(N) - Z(N))**2)
c
      ELSE
      TMP = U(1)*H3
      JJ = ABS(TMP)**P + G*ABS(TMP)**Q

      DO I=2,N
         TMP = (U(I)-U(I-1))*H3
         JJ = JJ + ABS(TMP)**P + G*ABS(TMP)**Q
      END DO

      TMP = 0.D0
      DO I=1,N
         TMP = TMP + (U(I)-Z(I))**2
      END DO
      JJ = JJ + 0.5D0*MY*TMP

      END IF

      RETURN
      END
```

List of References

[1] Emilio Acerbi and Giuseppe Mingione. Regularity results for a class of functionals with non-standard growth. *Arch. Ration. Mech. Anal.*, 156(2):121–140, 2001.

[2] Gilles Aubert and Pierre Kornprobst. *Mathematical Problems in Image Processing*, volume 147 of *Applied Mathematical Sciences*. Springer, New York, 2nd edition, 2006.

[3] Paolo Baroni, Maria Colombo, and Giuseppe Mingione. Harnack inequalities for double phase functionals. *Nonlinear Anal.*, 121:206–222, 2015.

[4] Debangana Baruah, Petteri Harjulehto, and Peter Hästö. Capacities in generalized Orlicz spaces. *J. Funct. Spaces*, pages 10, Art. ID 8459874, 2018.

[5] Tony Chan, Antonio Marquina, and Pep Mulet. High-order total variation-based image restoration. *SIAM J. Sci. Comput.*, 22(2):503–516, 2000.

[6] Yunmei Chen, Stacey Levine, and Murali Rao. Variable exponent, linear growth functionals in image restoration. *SIAM J. Appl. Math.*, 66(4):1383–1406, 2006.

[7] Yunmei Chen and Murali Rao. Minimization problems and associated flows related to weighted p energy and total variation. *SIAM J. Math. Anal.*, 34(5):1084–1104, 2003.

[8] Iwona Chlebicka. A pocket guide to nonlinear differential equations in Musielak-Orlicz spaces. *Nonlinear Anal.*, 175:1–27, 2018.

[9] Gustave Choquet. Theory of capacities. *Ann. Inst. Fourier (Grenoble)*, 5:131–295, 1955.

[10] Gustave Choquet. La naissance de la théorie des capacités: réflexion sur une expérience personnelle. *C. R. Acad. Sci. Sér. Gén. Vie Sci.*, 3(4):385–397, 1986.

[11] Frank H. Clarke. *Optimization and Nonsmooth Analysis*, volume 5 of *Classics in Applied Mathematics*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, second edition, 1990.

[12] Maria Colombo and Giuseppe Mingione. Bounded minimisers of double phase variational integrals. *Arch. Ration. Mech. Anal.*, 218(1):219–273, 2015.

[13] Maria Colombo and Giuseppe Mingione. Regularity for double phase variational problems. *Arch. Ration. Mech. Anal.*, 215(2):443–496, 2015.

[14] Lars Diening, Petteri Harjulehto, Peter Hästö, and Michael Růžička. *Lebesgue and Sobolev Spaces with Variable Exponents*, volume 2017 of *Lecture Notes in Mathematics*. Springer, Heidelberg, 2011.

[15] Lawrence C. Evans. *Partial Differential Equations*, volume 19 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, second edition, 2010.

[16] Linwei Fan, Fan Zhang, Hui Fan, and Caiming Zhang. Brief review of image denoising techniques. *Vis. Comput. Ind. Biomed. Art*, 2(1):7, 2019.

[17] George E. Forsythe and Wolfgang R. Wasow. *Finite-difference Methods for Partial Differential Equations*. Dover Phoenix Editions. Dover Publications, Inc., Mineola, NY, 2004. Reprint of the 1960 original.

[18] Isarel M. Gelfand and Sergei V. Fomin. *Calculus of Variations*. Revised English edition translated and edited by Richard A. Silverman. Prentice-Hall, Inc., Englewood Cliffs, N.J., 1963.

[19] Guy Gilboa. Variational methods in image processing. In *Nonlinear Eigenproblems in Image Processing and Computer Vision*, Advances in Computer Vision and Pattern Recognition, pages 15–37. Springer, Cham, 2018.

[20] Roland Glowinski and Jinchao Xu, editors. *Handbook of Numerical Analysis. Vol. XVI. Special volume: numerical methods for non-Newtonian fluids*, volume 16 of *Handbook of Numerical Analysis*. Elsevier/North-Holland, Amsterdam, 2011.

[21] Piotr Gwiazda, Petra Wittbold, Aneta Wróblewska, and Aleksandra Zimmermann. Renormalized solutions of nonlinear elliptic problems in generalized Orlicz spaces. *J. Differential Equations*, 253(2):635–666, 2012.

[22] Napsu Haarala, Kaisa Miettinen, and Marko M. Mäkelä. Globally convergent limited memory bundle method for large-scale nonsmooth optimization. *Math. Program.*, 109(1, Ser. A):181–205, 2007.

[23] Petteri Harjulehto and Peter Hästö. Boundary regularity under generalized growth conditions. *Z. Anal. Anwend.*, 38(1):73–96, 2019.

[24] Petteri Harjulehto and Peter Hästö. *Orlicz Spaces and Generalized Orlicz Spaces*, volume 2236 of *Lecture Notes in Mathematics*. Springer, Cham, 2019.

[25] Petteri Harjulehto and Peter Hästö. Double phase image restoration. *J. Math. Anal. Appl.*, 501(1):Paper No. 123832, 2021.

[26] Petteri Harjulehto, Peter Hästö, and Riku Klén. Generalized Orlicz spaces and related PDE. *Nonlinear Anal.*, 143:155–173, 2016.

[27] Petteri Harjulehto, Peter Hästö, Visa Latvala, and Olli Toivanen. Critical variable exponent functionals in image restoration. *Appl. Math. Lett.*, 26(1):56–60, 2013.

[28] Juha Heinonen, Tero Kilpeläinen, and Olli Martio. *Nonlinear potential theory of degenerate elliptic equations*. Oxford Mathematical Monographs. The Clarendon Press, Oxford University Press, New York, 1993. Oxford Science Publications.

[29] Juha Heinonen, Tero Kilpeläinen, and Olli Martio. *Nonlinear Potential Theory of Degenerate Elliptic Equations*. Dover Publications, Inc., Mineola, NY, 2006. Unabridged republication of the 1993 original.

[30] Anil K. Jain. *Fundamentals of Digital Image Processing*. Prentice-Hall, Inc., USA, 1989.

[31] Anna Kamińska. Flat Orlicz-Musielak sequence spaces. *Bull. Acad. Polon. Sci. Sér. Sci. Math.*, 30(7-8):347–352, 1982.

[32] Anna Kamińska and Henryk Hudzik. Some remarks on convergence in Orlicz space. *Comment. Math. Prace Mat.*, 21(1):81–88, 1980.

[33] Tommi Kärkkäinen, Kirsi Majava, and Marko M. Mäkelä. Comparison of formulations and solution methods for image restoration problems. *Inverse Problems*, 17(6):1977–1995, 2001.

[34] Napsu Karmitsa. Limited memory bundle method and its variations for large-scale nonsmooth optimization. In Adil M. Bagirov, Manlio Gaudioso, Napsu Karmitsa, Marko M. Mäkelä, and Sona Taheri, editors, *Numerical Nonsmooth Optimization*, pages 167–199. Springer Publishing, Switzerland, 2020.

[35] Tero Kilpeläinen. A remark on the uniqueness of quasi continuous functions. *Ann. Acad. Sci. Fenn. Math.*, 23(1):261–262, 1998.

[36] Ondrej Kováčik and Jiří Rákosník. On spaces $L^{p(x)}$ and $W^{k,p(x)}$. *Czechoslovak Math. J.*, 41(116)(4):592–618, 1991.

[37] Nicolai V. Krylov. *Lectures on Elliptic and Parabolic Equations in Hölder Spaces*, volume 12 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 1996.

[38] Olga A. Ladyženskaia, Vsevolod A. Solonnikov, and Nina N. Ural´seva. *Linear and Quasi-linear Equations of Parabolic Type*, volume 23 of *Translations of Mathematical Monographs*. American Mathematical Society, Providence Rhode Island, 1988.

[39] Fang Li, Zhibin Li, and Ling Pi. Variable exponent functionals in image restoration. *Appl. Math. Comput.*, 216(3):870–882, 2010.

[40] Marius Lysaker and Xue-Cheng Tai. Iterative image restoration combining total variation minimization and a second-order functional. *Int. J. Comput. Vis.*, 66(1):5–18, 2006.

[41] Kirsi Majava, Napsu Haarala, and Tommi Kärkkäinen. Solving variational image denoising problems using limited memory bundle method. In Wenbin Proceedings of SCPDE05, Liu, Michael

Ng, and Zhong-Ci Shi, editors, *Recent Progress in Scientific Computing*, pages 319–332, Beijing, 2007. Science Press.

[42] Marko M. Mäkelä. Survey of bundle methods for nonsmooth optimization. *Optim. Methods Softw.*, 17(1):1–29, 2002.

[43] Paolo Marcellini. Regularity of minimizers of integrals of the calculus of variations with nonstandard growth conditions. *Arch. Rational Mech. Anal.*, 105(3):267–284, 1989.

[44] Paolo Marcellini. Regularity and existence of solutions of elliptic equations with p, q-growth conditions. *J. Differential Equations*, 90(1):1–30, 1991.

[45] Paolo Marcellini. Regularity for elliptic equations with general growth conditions. *J. Differential Equations*, 105(2):296–333, 1993.

[46] Julian Musielak. *Orlicz Spaces and Modular Spaces*, volume 1034 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 1983.

[47] Hidegorô Nakano. *Modulared Semi-Ordered Linear Spaces*. Maruzen Co., Ltd., Tokyo, 1950.

[48] Takao Ohno and Tetsu Shimomura. Musielak-Orlicz-Sobolev spaces with zero boundary values on metric measure spaces. *Czechoslovak Math. J.*, 66(141)(2):371–394, 2016.

[49] Władysław Orlicz. Über konjugierte Exponentenfolgen. *Stud. Math.*, 3(1):200–211, 1931.

[50] Władysław Orlicz. Über eine gewisse Klasse von Raumen vom Typus B. *Bull. Internat. Acad. Polon. Ser A Cracovie*, 8:207–220, 1932.

[51] Mikko Parviainen. Lecture note for the course Partial Differential Equations 2 (MATS340), University of Jyväkylä. https://www.jyu.fi/science/en/maths/research/nonlinear-pde/pde2/lecturenote2017.pdf. Online; accessed 14 September 2022.

[52] Pietro Perona and Jitendra Malik. Scale-space and edge detection using anisotropic diffusion. *IEEE Trans. Pattern Anal. Mach. Intell.*, 12(7):629–639, 1990.

[53] William K. Pratt and Faramarz Davarian. Fast computational techniques for pseudoinverse and Wiener image restoration. *IEEE Trans. Comput.*, C-26(6):571–580, 1977.

[54] Michael Růžička. *Flow of Shear Dependent Electrorheological Fluids: Unsteady Space Periodic Case*. Kluwer/Plenum, New York, 1999.

[55] Michael Růžička. *Electrorheological Fluids: Modeling and Mathematical Theory*, volume 1748 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 2000.

[56] Leonid I. Rudin, Stanley Osher, and Emad Fatemi. Nonlinear total variation based noise removal algorithms. *Phys. D*, 60(1-4):259–268, 1992.

[57] Walter Rudin. *Functional Analysis*. International Series in Pure and Applied Mathematics. McGraw-Hill, Inc., New York, second edition, 1991.

[58] David M. Strong, Jean-François Aujol, and Tony F. Chan. Scale recognition, regularization parameter selection, and Meyer's G norm in total variation regularization. *Multiscale Model. Simul.*, 5(1):273–303, 2006.

[59] Andrei N. Tikhonov. On the solution of ill-posed problems and the method of regularization. *Dokl. Akad. Nauk SSSR*, 151:501–504, 1963.

[60] Beibei Wang, Duchao Liu, and Peihao Zhao. Hölder continuity for nonlinear elliptic problem in Musielak-Orlicz-Sobolev space. *J. Differential Equations*, 266(8):4835–4863, 2019.

[61] Aneta Wróblewska-Kamińska. Existence result for the motion of several rigid bodies in an incompressible non-Newtonian fluid with growth conditions in Orlicz spaces. *Nonlinearity*, 27(4):685–716, 2014.

[62] Zhuoqun Wu, Jingxue Yin, and Chunpeng Wang. *Elliptic & Parabolic Equations*. World Scientific Publishing Co. Pte. Ltd., Hackensack, NJ, 2006.

[63] Eberhard Zeidler. *Nonlinear Functional Analysis and Its Applications. II/A*. Springer-Verlag, New York, 1990. Linear monotone operators, Translated from the German by the author and Leo F. Boron.

[64] Vasilii V. Zhikov. Averaging of functionals of the calculus of variations and elasticity theory. *Izv. Akad. Nauk SSSR Ser. Mat.*, 50(4):675–710, 877, 1986.

[65] Vasilii V. Zhikov. On the Lavrentiev's effect. *Dokl. Akad. Nauk*, 345(1):10–14, 1995.

TURUN
YLIOPISTO
UNIVERSITY
OF TURKU

Painosalama Oy, Turku, Finland 2022

ISBN 978-951-29-9035-1 (PRINT)
ISBN 978-951-29-9036-8 (PDF)
ISSN 0082-7002 (PRINT)
ISSN 2343-3175 (ONLINE)