The Dissertation Committee for Mauricio Santillana
certifies that this is the approved version of the following dissertation:

Analysis and Numerical Simulation of the Diffusive Wave Approximation of the Shallow Water Equations

Committee:

_____
Clint N. Dawson, Supervisor

_____
Mary F. Wheeler

_____
Bjorn Engquist

_____
Irene M. Gamba

_____
Bayani Cardenas

Analysis and Numerical Simulation of the Diffusive Wave Approximation of the Shallow Water Equations

by

Mauricio Santillana, B.S., M.S.

Dissertation

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

Doctor of Philosophy

The University of Texas at Austin

August 2008

To my mother Luz Patricia, *in memoriam...*

Acknowledgments

I would like to thank my advisor, Clint Dawson, for his support and encouragement to pursue the current investigation. I deeply appreciate his patience and help during difficult times of my life in my years at UT Austin. I would like to thank my colleague and friend, Ricardo Alonso, for all the support and interest he showed when I tried to drag him into studying the DSW equation. He was very patient with me and taught me many concepts and techniques applied throughout this work.

I would like to thank Irene M. Gamba, for her guidance and support. In fact, it was Irene who convinced me to pursue my graduate work at UT Austin and she was very kind to support me throughout all my years at UT, both as an academic advisor and as a friend. I would also like to thank Dr. Mary F. Wheeler for her support and advise. She represents an example of leadership and scholarship to me. I would like to thank Dr. Bjorn Engquist for the eye opening lectures on mathematical modeling and numerical analysis from which I benefited greatly. I would like to thank Dr. Bayani Cardenas for accepting to become a member of my committee and reading through my work.

I would like to thank Juan Luis Vázquez for the patient guidance he offered me in the field of nonlinear diffusion equations and for the many insightful conversations we had during his visit to UT Austin in the Fall of 2007. I would also like to thank Luis Caffarelli for having the patience to hear many of the disconnected ideas I presented to him, and giving me in return meaningful insights about my work. I would also like to thank Rafael de la Llave for his kind guidance and support throughout all my years at UT, both as a friend

and as academic advisor. I would like to thank Dr. Tinsley Oden, for having supported me during my first years at UT.

I would like to thank many of my fellow ICES graduate students, Ross Heath, Marco A. Iglesias, Sri Harsha T. , Victor Calo, Konstantine Mardanov, Yuri Bazilevs, Shweta Bansal, among many others that I might have not mentioned, for their enthusiasm and support. I would like to thank both Connie Baxter and Lorraine Sanchez for their support and kind friendship.

I would like to thank the Director of CentroGeo, Dr. Carmen Reyes for all her kind advise and support. She represents an example of leadership and achievement to me. I would like to thank CentroGeo for partially supporting my graduate stay at UT.

I would like to thank my beloved wife, Monica Yudron, for her unconditional love, support, and for the extreme patience she showed to me during the development of this work. She was very kind to read many of the drafts that lead to the final version of this work and she offered me extremely meaningful suggestions in numerous occasions.

Finally, I would like to thank my brother Humberto, my father, my mother (to whom I dedicate this thesis), my sister Maru, my brother Sebastian, Fabiola, Maru and Erik for their love and support.

MAURICIO SANTILLANA

The University of Texas at Austin
August 2008

Analysis and Numerical Simulation of the Diffusive Wave Approximation of the Shallow Water Equations

Publication No. _____

Mauricio Santillana, Ph.D.
The University of Texas at Austin, 2008

Supervisor: Clint N. Dawson

In this dissertation, the quantitative and qualitative aspects of modeling shallow water flow driven mainly by gravitational forces and dominated by shear stress, using an effective equation often referred to in the literature as the diffusive wave approximation of the shallow water equations (DSW) are presented. These flow conditions arise for example in overland flow and water flow in vegetated areas such as wetlands. The DSW equation arises in shallow water flow models when special assumptions are used to simplify the shallow water equations and contains as particular cases: the Porous Medium equation and the time evolution of the p-Laplacian. It has been successfully applied as a suitable model to simulate overland flow and water flow in vegetated areas such as wetlands; yet, no formal mathematical analysis has been carried out addressing, for example, conditions for which weak solutions may exist, and conditions for which a numerical scheme can be successful in approximating them. This

thesis represents a first step in that direction. The outline of the thesis is as follows. First, a survey of relevant results coming from the studies of doubly nonlinear diffusion equations that can be applied to the DSW equation when topographic effects are ignored, is presented. Furthermore, an original proof of existence of weak solutions using constructive techniques that directly lead to the implementation of numerical algorithms to obtain approximate solutions is shown. Some regularity results about weak solutions are presented as well. Second, a numerical approach is proposed as a means to understand some properties of solutions to the DSW equation, when topographic effects are considered, and conditions for which the continuous and discontinuous Galerkin methods will succeed in approximating these weak solutions are established.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

1.1 Motivation

The purpose of this thesis is to study, analytically and numerically, an efective equation often referred to in the literature as the diffusive wave approximation of the shallow water system of equations. This equation has been used to simulate overland flow in wetlands and open channels.

Wetlands are some of the most important ecosystems on earth. Historically, they have been called swamps, marshes, bogs, fens, or sloughs. Since wetlands provide an abundant supply of water, they are able to host a variety of plant and animal species including mammals, birds, reptiles, amphibians, and fishes. During the past decades, the increased degradation of wetlands has been connected to damage to the overall biodiversity of our planet. Consequently, natural wetlands are increasingly protected and construction of artificial wetlands is being encouraged.

Of particular interest is the relevant role that wetlands can play during storm surges and flooding events. Indeed, coastal marshes and swamps act as a buffer zone, for example, between the Gulf of Mexico and inhabited inland areas in Louisiana, where an estimated 60-75 % of residents live within 50 miles of the coast (1993) and where between 1899 and 1995 over a dozen major hurricanes (class 3-5) have hit[1] (with the two most recent hits of category 5 hurricanes Katrina and Rita in 2005). Wetlands also buffer the tidal increase and wave intensity of hurricanes, since they act as energy dissipators; they provide natural flood control by detaining and slowing flood waters, and they help protect areas prone to serious

---

[1]Source: USGS National Wetlands Research Center. http://www.lacoast.gov/

erosion. Furthermore, hurricane winds subside substantially once they reach the wetland buffer. Another fact about wetlands that make them an important subject of study is that due to the high rate of biological activity, wetlands can transform many common pollutants that occur in waste waters into harmless byproducts or into nutrients that can be used for biological productivity. Thus, it is clear that understanding the dynamics of the flows in such environments is essential, in particular for management and designing purposes. In this work, we intend to provide insight to improve the ways in which we model flow conditions in vegetated areas, such as wetlands, in order to deepen our understanding of such ecologically rich environments.

Wetlands are frequently transitional areas between uplands (terrestrial systems) and continuously or deeply flooded (aquatic) systems. They are also found at topographic lows or in areas with high slopes and low permeability soils such as seepage slopes. *Overland flow* is a term used to describe water flow regimes in these environments, where uniform and fully developed turbulent flow conditions arise since the water flow is driven mainly by gravitational forces and dominated by shear stress. More precisely, overland flow is a term that is used mostly to describe the shallow movement of water across land surfaces both when rainfall has exceeded the infiltration rate of the ground's surface (Horton Overland Flow), and when the entire soil column becomes completely saturated and water exfiltrates at the surface (Dunne Overland Flow). An effective equation that has been used to simulate these particular water flow regimes is a *doubly nonlinear and degenerate diffusion equation* often referred to in the literature as the diffusive wave approximation of the shallow water equations (DSW). This equation is obtained from the shallow water system of equations, by approximating the 2-D depth-averaged continuity equations by empirical laws commonly used in open channel flow theory, such as Manning's or Chézy's formulas, and combining the resulting expression with the free surface boundary condition. Formally speaking, the DSW equation is given by

$$\frac{\partial u}{\partial t} - \nabla \cdot \left(\frac{(u-z)^\alpha}{c_f \, |\nabla u|^{1-\gamma}} \nabla u \right) = f$$

where $u(x, y; t)$ is the free surface water elevation, $z(x, y)$ the bed surface or bathymetry, $u - z$ the water depth, $f(x, y, t)$ a source/sink, and c_f is a friction coefficient. The values of α and γ correspond to, $\alpha = 5/3$ and $\gamma = 1/2$ for Manning's formula, and $\alpha = 3/2$ and $\gamma = 1/2$ for Chézy's formula.

Diverse numerical schemes have been implemented to approximately solve the DSW

equation and have been successfully applied as suitable models to simulate overland flow and water flow in vegetated areas such as wetlands ([69], [44], [17], [38], [39]); yet, to the best of our knowledge, no formal mathematical analysis has been carried out in order to study existence, uniqueness, and regularity of weak solutions to this equation; and as a consequence, the approximation properties of numerical methods, such as error estimates between the true solution (if uniqueness holds) and the numerical approximant, rates of convergence, and so on, have not been developed for this equation.

To motivate more deeply this study, it is important to mention that solving the DSW equation computationally requires significantly less work than solving the shallow water system of equations (SWE), see [44]. However, despite the obvious appeal to use the DSW *in lieu* of the SWE, analysing the mathematical properties of the DSW equation is not a simple task. Note that the DSW equation contains as particular cases two complicated nonlinear diffusion equations: the Porous Medium equation (when $z = 0$ and $\gamma = 1$) and the p-Laplacian for $1 < p < 2$ (when $\alpha = 0$ and $p = \gamma + 1$, this case is not considered in this work).

In this thesis, we will set up the appropriate initial boundary value problem arising from the DSW equation. This will be done by combining the knowledge from the modeling applications of the DSW ([69], [44], [17], [38], [39]) and the knowledge coming from experimental works (such as [61]) where the parameters α and γ are considered to be flexible (generally in the ranges $\alpha > 1$ and $0 < \gamma \leq 1$) as a way to account for more general circumstances when flow changes back and forth between turbulent and laminar conditions in vegetated regions. We also carry out an investigation of the mathematical properties of the DSW equation, addressing the mathematical issues discussed in the previous paragraphs.

The DSW equation can be characterized as a *doubly nonlinear parabolic equation* and gives rise to the initial boundary value problem (2.1) stated in Chapter 2. Its doubly nonlinear nature comes from the fact that the nonlinear behavior appears inside the divergence term as a product of two nonlinearities involving u and ∇u, namely $(u - z)^\alpha$ and $\nabla u / |\nabla u|^{1-\gamma}$. Furthermore, when rewriting it as (for $c_f \equiv 1$):

$$\frac{\partial u}{\partial t} - \nabla \cdot (a(u, \nabla u) \, \nabla u) = f \qquad \text{with} \qquad a(u, \nabla u) = \frac{(u - z)^\alpha}{|\nabla u|^{1-\gamma}}, \tag{1.1}$$

where a is the diffusion coefficient, one can be more specific and characterize it as a doubly nonlinear and degenerate-singular equation. This is the case since $a \to 0$ when $(u - z) \to 0$

3

and $a \to \infty$ when $\nabla u \to 0$ for the given choice of $0 < \gamma \le 1$, and $1 < \alpha < 2$. Important works addressing existence, uniqueness and regularity of solutions to doubly nonlinear parabolic equations related to the DSW equations include: [56], [41], [6], [45], and [35]. These works will provide the starting point of this thesis.

Previous work aimed to analyse approximate solutions to nonlinear diffusion equations using the Galerkin finite element method, such as the work of Wheeler [67] and Douglas and Dupont [32], deal with equations with nonlinear diffusion coefficients that only depend on the function u itself and not on ∇u, *i.e.* diffusion coefficients of the form $a = a(u)$. The analysis carried out in such cases requires roughly two assumptions,

$$0 < \mu \le a(u) \le M, \quad \text{and} \quad |a'(u)| \le B \quad \text{for } u \in \mathbb{R} \tag{1.2}$$

so that a is uniformly Lipschitz with respect to u and bounded below by a small constant μ. These assumptions ensure in particular, that one can construct a weak formulation such that, for some Sobolev space V, one has two fundamental conditions:

$$\mu\|u\|_V^2 \le (a(u)\nabla u, \nabla u) \quad \text{and} \quad (a(u)\nabla u, \nabla w) \le M\|u\|_V\|w\|_V \quad \text{for } u, w \in V, \tag{1.3}$$

where $(\cdot, \cdot)$ represents the appropriate duality pairing. The doubly nonlinear nature of the DSW equation poses new challenges that come from the possible degeneracy of the diffusion coefficient a in (1.1) when $(u - z) = 0$, and the possible singular nonlinear dependency of a with respect to ∇u. In fact, with the condition that $0 < \gamma \le 1$, one can only expect that the diffusion coefficient be uniformly Lipschitz with respect to ∇u if $\gamma = 1$, that is when the dependency with respect to ∇u disappears and the DSW equation becomes the Porous Medium equation (for $z \equiv 0$). In general, the diffusion coefficient given by (1.1) is at most Hölder continuous with respect to ∇u and possibly degenerate (i.e. $a(u, \nabla u) = 0$) in subsets of Ω, thus, one cannot expect that similar expressions such as those shown in (1.3) will hold. This fact motivates the need of further assumptions or properties on the type of solutions to be approximated, such as physical consistency, if one is to produce a meaningful numerical method.

Some solutions of the DSW equation on flat topographies ($z \equiv 0$) evolving from compactly supported initial conditions and without a forcing term, are themselves compactly supported, at least locally in time, and exhibit a finite speed of propagation of information. This fact gives rise to free boundaries (interfaces between $u = 0$ and $u > 0$) and local discontinuities in the gradient of the solution, and it is shown to persist (based on numerical

experiments) even for bounded and regular nonflat topographies. Many solutions of this class, violate *globally* some of the main assumptions used to derive the DSW in the context of shallow water modeling and should be carefully identified as incorrect models for water flow. Even though in this thesis work we do not elaborate on the accurate approximation of free boundaries, the discontinuous Galerkin method will be explored as a natural choice to better approximate solutions of the DSW equation in 2D. Furthermore, the *local discontinuous Galerkin* (LDG) method will be introduced as an appropriate setting to handle complex geometries as well as a natural technique that allows the easy coupling of the DSW model with other surface water models (such as the SWE) and other subsurface flow models (such as Darcy's flow or other infiltration models).

The overall mathematical strategy[2] of this work can be summarized as follows. First, we present an original, simple, and constructive proof of existence of weak solutions to the DSW equation using the Faedo Galerkin method when *topographic effects* are ignored. Most of the techniques presented in this proof were originally introduced by Lions [50] and further developed for quasilinear and doubly nonlinear parabolic equations by Alt and Luckhaus [3] and Bernis[14], respectively. We proceed to show proofs of basic regularity results such as boundedness of solutions and uniqueness. This is done using both *a priori* estimates, and through a comparison principle first introduced in the context of doubly nonlinear equation by Bamberger [6], respectively. The latter is also used to find nonnegativity of solutions. The constructive method presented in the proof of existence will be used as a natural setting for a computational method to find approximate solutions even in the case when *topographic effects* are incorporated. Then, we will present a numerical approach in 1-D as a means to understand some properties of solutions to the DSW equation and thus to provide conditions for which the use of the DSW equation may be inappropriate as a model for shallow water flow in vegetated areas, both from the physical and the mathematical points of view. Error estimates and rates of convergence for both the continuous Galerkin finite element method, and the local discontinuous Galerkin method, will be established in 2-D. The general analytic study of the DSW equation, when *topographic effects* are incorporated, remains an open problem. Some of the difficulties of this problem will be discussed.

In terms of the application of the numerical method to real life situations,[3] we will present evidence to support the fact that the DSW can be utilized as a model to simulate the

---

[2]This aspect of the dissertation fulfills the requirements of areas A and B of the Computational and Applied Mathematics Ph. D. program.
[3]This portion of the dissertation fulfills the requirements of area C of the Computational and Applied Mathematics Ph. D. program.

main aspects of overland flow in a controlled environment. On one hand, we will show how numerically simulated hydrographs and water depth profiles in 1-D compare successfully to the measurements of a set of laboratory experiments conducted by Iwagaki [46]. On the other hand, we will present 2-D simulations of the evolution of water depth profiles that capture the salient features of: *(i)* an ideal dam break problem and *(ii)* water flow in a channel containing vegetation.

1.2 Literature review

1.2.1 Modeling overland flow

Models for overland flow are derived both from the three-dimensional incompressible Navier-Stokes (NS) equations and from empirical observations such as Manning's formula and Chézy's formula. Generally, depending on the physics of the flow, scaling arguments are used in the NS system of equations in order to obtain effective equations capable of reproducing the phenomena under study. In practice, the Raynolds Averaged Navier-Stokes (RANS) equations are used as a point of departure in the derivation of effective equations. In shallow water flows, the main scaling assumption consists in considering that the vertical scales are small relative to the horizontal ones. This approximation reduces the vertical momentum equation to the hydrostatic pressure relation

$$\frac{\partial p}{\partial x_3} = \rho g,$$

where p is the hydrostatic pressure, ρ is the density of the fluid, g the gravitational acceleration constant, and x_3 the vertical coordinate. Integrating the two horizontal momentum equations and the continuity equation over the depth, and using appropriate kinematic boundary conditions at the free surface and bottom leads to the two-dimensional shallow water system of equations (SWE). For a detailed description of shallow water hydrodynamics see [65] and [63]. If, in addition, horizontal shear stresses are assumed to be small, and Coriolis effects and surface wind shear stress are neglected, the two dimensional SWE can be written as a depth-averaged mass continuity equation

$$\frac{\partial H}{\partial t} + \frac{\partial (hu)}{\partial x} + \frac{\partial (hv)}{\partial y} = f(x, y, t), \tag{1.4}$$
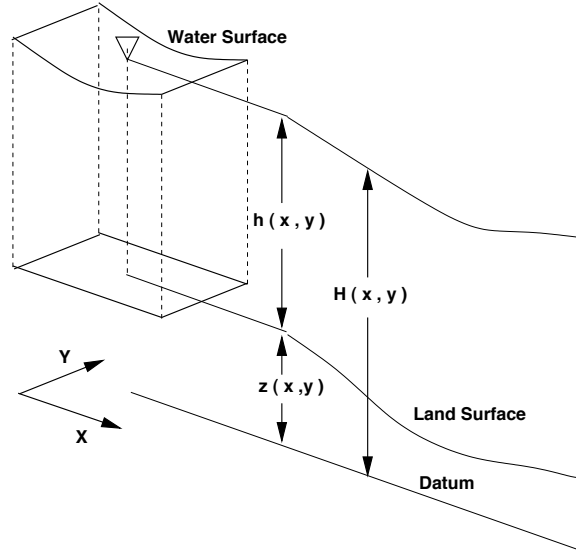
FIGURE 1.1: *Land surface elevation is measured from the datum and represented by function z(x,y), free water surface elevation by H(x,y) and the water depth by h(x,y).*

and the 2-D depth-averaged momentum equations

$$\frac{\partial(hu)}{\partial t} + \frac{\partial(hu^2)}{\partial x} + \frac{\partial(huv)}{\partial y} + \quad gh\frac{\partial H}{\partial x} \quad + \quad \frac{\tau_{bx}}{\rho_0} = 0, \tag{1.5}$$

$$\underbrace{\frac{\partial(hv)}{\partial t} + \frac{\partial(huv)}{\partial x} + \frac{\partial(hv^2)}{\partial y}}_{\text{inertial terms}} + \underbrace{gh\frac{\partial H}{\partial y}}_{\text{pressure}} + \underbrace{\frac{\tau_{by}}{\rho_0}}_{\text{friction}} = 0, \tag{1.6}$$

where $H(x, y; t)$ is the free surface elevation or hydraulic head, z is the bed surface, bathymetry or land elevation, (See figure 1.1), $h(x, y; t) = H(x, y; t) - z(x, y)$ is the water depth, f is a source/sink (such as rain or infiltration), $u(x, y; t)$ and $v(x, y; t)$ are the x and y components of the depth averaged horizontal velocity vector $\boldsymbol{V}$, and τ_{bx} and τ_{by} are the x and y components of the averaged boundary shear stress at the bed surface z, and ρ_0 the reference density of the fluid.

Further assumptions and simplifications of these equations lead to different overland flow models. These models are generalizations of the one-dimensional open channel flow approximations, which have been studied based on their capacity to simulate wave propagation. See [55]. Mainly, two approximations are relevant to overland flow: the *kinematic wave* and the *diffusive wave* approaches. The *kinematic wave approach* assumes that the

inertial terms as well as the pressure terms in the momentum equations (1.5) and (1.6) are negligible when compared to the friction terms. The *diffusive wave approach* assumes that only the inertial terms are negligible. Their applicability to one dimensional overland flow was studied, for example by Ponce et al. in [54], who found conditions where both can be used to simulate the physical phenomenon instead of the full SWE (also known as St. Venant Equations).

Research teams have explored different models in order to capture the main features of overland flow and water flow in vegetated areas in a computationally efficient manner. Some recent studies addressing modeling techniques for these flow regimes, particularly those arising in vegetated areas, can be found in [12] and the references therein. In this section, however, we will mainly discuss those studies in which the use of the DSW equation was successful in modeling overland flow and water flow in wetlands. For example, Xanthopoulos and Koutitas [69], as well as Hromadka et al. [44] validated a two-dimensional dam-break model for flood wave propagation and flood plain study, respectively, based on the diffusive wave approach. They concluded that the inertial terms are indeed negligible in situations where the bed surface is flat and derived the DSW equation under this conditions. Hromadka et al. also concluded that the computational cost required to solve the full St. Venant equations increases by 50% the cost required to solve the DSW equation. Xanthopoulos and Koutitas [69] solved the DSW equation using an explicit finite difference (FD) scheme in Eulerian space and their results compared successfully with laboratory experiments. They eventually used this approximation to model a plain in Northern Greece. Hromadka et al. [44] solved the DSW equation using an integrated finite difference (IFD) formulation for regular and irregular triangle elements. They showed that the two-dimensional diffusive approach could be used to predict a hypothetical dam break more accurately than a one-dimensional model. Giammarco et al. [39] used a control volume finite element (CVFE) formulation to solve the DSW equation and showed that this approach improves the classical Finite Element (FE) formulation since it is locally mass conservative. They also showed that the CVFE represents the gradients better than the IFD. Feng and Molz [38] derived an effective equation similar to the DSW equation using the diffusive wave approach in order to model flow in wetlands. They solved such an effective equation using a fully implicit finite difference scheme in a regular mesh with square elements, and using a Picard Iteration scheme to resolve the nonlinear terms in a fixed rectangular domain with irregular wet boundaries. They also proposed an alternative friction formula deduced empirically by Turner and Chanmeesri [61], see equation (2.5), which relates the one-dimensional flow velocity to the one-dimensional friction bed slope for shallow flow of

8

water through non-submerged vegetation. Observations by Turner and Chanmeesri [61] show that $\alpha > 1$ and, in general $0 < \gamma \leq 1$. Bolster and Saiers [16] found that the value of $\gamma = 1$ was the most suitable choice when calibrated with field measurements of hydraulic heads. They used a predictor-corrector finite-difference scheme to solve the DSW equation when modeling wetlands in the Shark River Slough in the Everglades in Florida. As a last example of successful application of the DSW equation in order to model wetlands we mention the work of Bauer et al. [11], who developed a numerical model using a finite difference scheme to simulate the flow in a wetland system coupled with groundwater flow in the Okavango Delta, in Botswana.

A related work using the SWE as a model for overland flow is the work by Zhang and Cundy. In [73], they developed a fully dynamical model solving equations (1.4)-(1.6) assuming particular forms of the friction shear stresses and using a MacCormack finite difference scheme. Their model allows spatial variations of hillslope features, including surface roughness, infiltration, and microtopography. Their main conclusion was that microtopography is the dominant factor causing spatial variation in overland flow depth , velocity, and direction.

1.2.2 Doubly nonlinear diffusion equations

To the best of our knowledge, the DSW equation has not been studied in its general form as presented in the initial boundary value problem (2.1) of Chapter 2. However, when topographic effects are neglected ($z \equiv 0$) and zero-Dirichlet initial/boundary conditions are assumed ($\partial\Omega = \Gamma_D$), one can find a fairly extensive number of works that study doubly nonlinear equations that are relevant to the DSW equation. See for example [50], [56],[41], [6], [45]. Most of these works study alternative formulations of problem (2.1). These will be explained in the subsequent paragraphs.

Two important particular cases of the DSW equation, that do not fall into the category of doubly nonlinear diffusion equations that deserve to be mentioned are, the Porous Medium Equation (PME) and the p-Laplacian. A comprehensive study of the PME can be found in the book by Vázquez [62] and the references therein. For the time evolution of the p-Laplacian the reader is referred to the book by DiBenedetto [31] and the references therein. An interesting reference addressing a class of nonlinear diffusion equations in the context of image processing and edge detection is [24]. The diffusion coefficients studied in [24] depend purely on the gradient of the solution.

We proceed to summarize the most relevant results from studies of doubly nonlinear diffusion equations existing in the literature that can be applied to the DSW equation when topographic effects are ignored. In order to do so, we will introduce an alternative formulation that has been studied before. This formulation will be used throughout this section and in Chapter 3 of this thesis and is given by

$$
\begin{cases}
\dfrac{\partial \phi(v)}{\partial t} - \eta^\gamma \, \nabla \cdot \left(\dfrac{\nabla v}{|\nabla v|^{1-\gamma}} \right) = f & \text{on } \Omega \times (0, T] \\[2mm]
v = 0 & \text{on } \partial\Omega \times [0, T] \\[2mm]
v = v_0 & \text{on } \Omega \times \{t = 0\}
\end{cases}
\tag{1.7}
$$

where Ω is either $\mathbb{R}^n$ or an open (and in most cases bounded) subset of $\mathbb{R}^n$, η is a positive constant, and the function $\phi(s) \in C^{0,\eta}(\mathbb{R})$ is an *odd* function satisfying the following properties:

(i) $|\phi(s)| \leq |s|^\eta$ for $0 < \eta \leq \gamma < 1$, with equality for $|s| \geq R$ for some $R \geq 0$

(ii) $\phi(s)$ is a concave increasing function for $s \geq 0$.

Note that with the change of variables defined by $u = \phi(v)$, problem (1.7) is transformed into

$$
\begin{cases}
\dfrac{\partial u}{\partial t} - \eta^\gamma \, \nabla \cdot \left(\left((\phi^{-1})'(u)\right)^\gamma \dfrac{\nabla u}{|\nabla u|^{1-\gamma}} \right) = f & \text{on } \Omega \times (0, T] \\[2mm]
u = 0 & \text{on } \partial\Omega \times [0, T] \\[2mm]
u = u_0 & \text{on } \Omega \times \{t = 0\}
\end{cases}
\tag{1.8}
$$

Now, choosing

$$
0 < \eta = \frac{\gamma}{\alpha + \gamma} < 1, \qquad \text{and} \qquad \phi(s) = \frac{s}{|s|^{1-\eta}}
\tag{1.9}
$$

we can obtain the explicit expression for

$$
(\phi^{-1})'(s) = (1 + \theta)|s|^\theta \qquad \text{where} \qquad \theta = \frac{1 - \eta}{\eta} = \frac{\alpha}{\gamma}
\tag{1.10}
$$

which yields the following equation

$$
\frac{\partial u}{\partial t} - \nabla \cdot \left(|u|^\alpha \frac{\nabla u}{|\nabla u|^{1-\gamma}} \right) = f.
\tag{1.11}
$$

The previous manipulations show, at least formally, that nonnegative solutions of problem (1.7) are solutions of the DSW equation for flat topographies. Most of the results found in

10

the literature address the IBVP (1.7).

1.2.2.1   Existence of solutions

Lions [50] introduced the techniques of compactness and monotonicity later utilized in the subsequent works in the proofs of existence for problem (1.7). Raviart [56], and Grange and Mignot [41] prove the existence of weak solutions to problem (1.7), provided Ω is an open and bounded subset of $\mathbb{R}^n$, constructing approximate solutions using implicit finite differences in time and passing to the limit by means of compactness and monotonicity. In [56] Raviart worked directly with problem (1.7), and in [41] Grange and Mignot extended such results to the abstract setting of equations of the type:

$$\frac{\partial Bu}{\partial t} + Au = f$$

where A and B denote the subdifferentials of convex functionals. Their analysis is based on the essential restriction that these functionals must be continuous on appropriate Banach spaces. Bernis further extends these results to the case when Ω is any open set of $\mathbb{R}^n$ in [14]. Another relevant reference is [15], where Blanchard and Francfort address the semi-abstract problem

$$\frac{\partial}{\partial t}b(u) - \nabla \cdot (D\Phi(\nabla u)) = f$$

where b is a locally Lipschitz function and may grow faster than any power function at infinity, and Φ is a C^1 convex functional with specific coercivity assumptions. They obtain existence and comparison results with the aid of a Galerkin approximation technique which uses truncation-penalization of the time nonlinearity and *a priori* estimates through convex conjugate functions. An important work addressing quasilinear and doubly nonlinear parabolic equations is found in Alt and Luckhaus [3].

1.2.2.2   Comparison principles and uniqueness

In [6], Bamberger studies the existence of particular solutions to problem (1.7) which are the limit of *solutions fortes i.e.* solutions that have the property $\phi(u)_t \in L^1(0, T, L^1(\Omega))$. Bamberger refers to this kind of solutions as *limite de solutions fortes*. In addition, he presents a very concise exposition of a comparison principle between solutions that are *limite de solutions fortes* and uses this result to find uniqueness. See Section 3.3.

1.2.2.3 Regularity

When topographic effects are neglected ($z \equiv 0$) and zero-Dirichlet initial/boundary conditions are assumed ($\partial\Omega = \Gamma_D$) Problem (2.1) can also be re-written in the form:

$$\frac{\partial u}{\partial t} - \nabla \cdot \left(|\nabla u^m|^{\gamma-1} \nabla u^m \right) = f \tag{1.12}$$

with $m = 1 + \alpha/\gamma$. Esteban and Vázquez [35] studied this equation in 1-D for the Cauchy problem ($\Omega = \mathbb{R}$). They study the *local velocity of propagation*

$$V(x,t) = -v_x |v_x|^{\gamma-1}$$

where v is the nonlinear potential defined as:

$$v = \begin{cases} \dfrac{m\gamma}{m\gamma - 1} \, u^{\frac{m\gamma-1}{\gamma}} & \text{if} \quad m\gamma \neq 1 \\[2ex] \dfrac{1}{\gamma} \, \log u & \text{if} \quad m\gamma = 1 \end{cases}$$

Recall that in the DSW equation, it is assumed that $m\gamma = \alpha + \gamma > 1$. In their work, they base their approach on the existing theory for the Porous Medium Equation and find the estimate

$$V_x \leq \frac{1}{\gamma(m+1)\,t}.$$

Using the previous estimate as the main tool, they construct a theory for the Cauchy problem with nonnegative, integrable initial data. In particular, they address the following questions:

- Existence, uniqueness and regularity of strong solutions,

- Existence and regularity of free boundaries,

- Asymptotic behaviour of solutions and free boundaries.

In [45], Ishige gives a sufficient condition for the growth order of the initial data at infinity for the existence of weak solutions of the Cauchy problem ($\Omega = \mathbb{R}$) (1.7).

1.2.2.4 Additional properties of solutions

Some interesting facts about nonnegative solutions to problem (1.7) are:

- *Finite speed of propagation.* Indeed, Barenblatt constructed a class of self-similar source type solutions for the Cauchy problem ($\Omega = \mathbb{R}$) which have the property that

12

their supports propagate in time with finite speed, when $\alpha + \gamma > 1$. See [7].

- *Extinction property.* In [6], using simple arguments, Bamberger exhibits that for $f = 0$, nonnegative solutions to the zero-Dirichlet boundary value problem ($\Omega \subset \mathbb{R}$, bounded) become zero in finite time.

- *Traveling waves.* It is worthwhile mentioning that an interesting example of traveling wave type solutions

$$u(x,t) = U(t - n \cdot x) \qquad \text{with} \quad U(s) = 0 \qquad \text{for} \quad s > 0,$$

to the zero-Dirichlet boundary value problem ($\Omega \subset \mathbb{R}$, bounded) is shown in [6] for the case when $\eta > \gamma$ (equivalently $\alpha < 1 - \gamma$). In the DSW equation this case does not arise since $\alpha > 1$ and $0 < \gamma \leq 1$.

Other properties of solutions including *nonexistence of global nonnegative solutions* and *blow up solutions* can be found in [14] and [45] respectively, for particular choices of the parameters η and γ that do not happen in the DSW equation case.

1.2.3 Numerical methods for nonlinear parabolic problems

As mentioned before, finite difference schemes and finite element techniques have been implemented to approximate the solution of the DSW equation and have been used successfully to simulate water flow in shallow systems in [69], [44], [39], [38], [61], [11]. However, no formal numerical analysis has been carried out in order to show that the proposed methods converge in some sense to the *true* solution of the IBVP (2.1) . This is not surprising given the complexity of the general formulation of the IBVP (2.1) and the lack of analytical techniques to prove for example uniqueness of solutions in the presence of topographic effects.

Generally speaking, numerical schemes to solve parabolic problems have been widely explored. The most important methodologies utilized in the past are two: finite difference schemes and finite element methods. The early works of Courant, Friedrichs and Lewy [29] served as a starting point in the development of finite difference schemes, where the approximate solution of the equation is constructed by solving algebraic difference equations at given mesh points in the domain at each time step. See, for example, [59]. Later on, the works of Wheeler [67] and Douglas and Dupont [32] played a key role in the study of approximate solutions (globally throughout the domain) to nonlinear diffusion equations using the Galerkin finite element method. In the finite element method an approximate solution is constructed as a linear combination of so-called basis functions of a linear space

by solving a discrete variational formulation derived directly from the partial differential equation at hand. Loosely speaking, the Galerkin finite element has proved to be superior when numerically solving partial differential equations in complex domains or when the solution lacks smoothness. References such as [18] and [34] describe the mathematical background and applications of this method. Of particular interest is the book by Thomée [60] and the references therein, where a comprehensive account of the mathematical theory of Galerkin finite element methods as applied to parabolic partial differential equations is presented. The Galerkin finite element method, both in its continuous and discontinuous versions, will be the basis for this thesis work.

Within the context of the continuous Galerkin finite element method, relevant works approximating degenerate parabolic equations include for example: [53], [72], [57], [42], [37], and [4]. Even though the DSW equation (as presented in Chapter 2) does not fall into the set of equations solved in these references, they provide important techniques that will be explored in this thesis. In particular, in [53], Nochetto and Verdi present a numerical method to approximate degenerate parabolic problems, based on the finite element method, similar to the one used in this thesis work. In their study they analyze equations of the form

$$\frac{\partial u}{\partial t} - \nabla \cdot (\nabla v + b(r(v))) + f(r(v)) = 0, \qquad u \in m(v), \tag{1.13}$$

where $m(v)$ is a maximal monotone graph in $\mathbb{R} \times \mathbb{R}$ possibly with a singularity at the origin $(m'(0) = \infty)$. Stefan type, nonstationary filtration type, and porous-medium type degenerate parabolic equations can be written in the form (1.13). One may replace $r(v)$ by v for intuition purposes. In cases when singularities in m appear, Nochetto and Verdi use a smoothing procedure similar to the one that will be used in Chapter 3. In their approximation strategy, they construct a numerical scheme that approximates a regularized problem obtained by replacing m in (1.13) by a smooth function m_ϵ with maximal slope equal to $1/\epsilon$, for some regularization parameter $\epsilon > 0$. Then they discretize this regularized problem in space and time to compute the regularized numerical approximation $U_\epsilon^h(t)$. Finally, roughly speaking, they show global error estimates between the solution $u(t)$ of (1.13) and the regularized numerical approximation $U_\epsilon^h(t)$. Their method will be discussed again in Chapter 4 when we introduce the numerical strategy proposed in this thesis work.

In [72], Yong and Pop present a methodology to numerically solve porous medium-type equations based on the maximum principle. In their approach they locally perturb the (initial and boundary) data instead of the nonlinear diffusion coefficients to keep solutions away from degeneracy. In [57], Rulla and Walkington obtain optimal *spatial* error estimates

for degenerate parabolic problems by using semigroup theory. However, in their approach, they do not discuss the numerical techniques to solve these degenerate equations. This is the case as well in [42], where Hansen and Ostermann focus their study in improving the *temporal* error estimates of porous medium-type equations by using high order implicit Runge-Kutta methods. In [37], Evje and Karlsen analyse discrete approximations of weak solutions of bounded variation (BV), in space and time, to doubly nonlinear parabolic equations. Their approach is carried out in the framework of BV because weak solutions of the equations they study are, in general, not uniquely determined by their data. The kind of doubly nonlinear parabolic equations presented in [37] differ from the DSW in the fact that topographic effects cannot be incorporated in their formulation. Finally, in [4], Arbogast et al. develop and analyse two mixed finite element approximations to obtain approximate solutions of a nonlinear, degenerate advection diffusion equation arising in petroleum reservoir and groundwater aquifer simulation. The diffusion coefficient of the equation they study does not contain nonlinearities with respect to the gradient of the solution. For completeness, we refer the reader to [8], [48] and the reference therein for numerical studies addressing equations with diffusion coefficients that depend purely on the gradient of the solution, such as the time evolution of the p-Laplacian.

In Chapter 5, the *local discontinuous Galerkin* (LDG) method will be introduced as an appropriate setting to handle complex geometries as well as a natural technique that allows the easy coupling of the DSW model with other surface water models (such as the SWE) and other subsurface flow models (such as Darcy's flow or other infiltration models). The LDG method is one of many *discontinuous Galerkin* (DG) methods. In these methods, continuity across elements is not enforced in the linear space where the basis functions live, thus giving rise to "broken" or discontinuous approximate solutions. This is a major difference with the continuous Galerkin finite element method. For a review on the development of discontinuous Galerkin methods see the book by Cockburn et al. [27].

In the context of elliptic and parabolic problems, the DG methods emerged from the *interior penalty* (IP) methods. The IP methods were originally devised as a means to impose Dirichlet boundary conditions *weakly* rather than incorporating the boundary conditions into the finite element space [52]. This idea was further generalized, at the element level, by enforcing continuity across elements weakly in [68]. The development of these ideas in the context of elliptic equations can be found in the Chapter: "Discontinuous Galerkin Methods for Elliptic Problems" in [27].

The LDG method was introduced by Cockburn and Shu in [28] as an extension, to general convection-diffusion problems, from the numerical techniques introduced by Bassi and Rebay in [10] to solve the compressible Navier-Stokes equations. One of the basic ideas in the LDG method is to rewrite, say the parabolic equation at hand, as a degenerate first order system of equations where both u and ∇u $(= \boldsymbol{q})$ are now considered as *independent unknowns*. This strategy is also present in methods based on a *mixed* formulation. In the LDG method, one further discretizes the resulting first order system using particular DG techniques. Details about the method set up in the context of elliptic problems can be found in [22]. Some general characteristics of the DG methods are the following:

- They can easily handle various shapes in different elements across the domain, as well as local spaces of different types (orders). This is the case since continuity is not enforced *strongly* across elements.

- The previous property makes these methods suitable to handle structured and unstructured meshes in domains with general geometries.

- Their high degree of locality makes them highly parallelizable.

- They are element-wise conservative (This statement is meaningful when modeling (nonlinear) conservation laws).

- They are ideally suited for *hp*-refinement (or *hp*-adaptivity).

Particular to the LDG method studied in this thesis work, the approximation to u, and the approximation to each of the components of $\boldsymbol{q}$ belong to the same approximation spaces, making the coding simpler than in the standard mixed methods. Also in the so-called *numerical flux*, introduced to properly define the values of the fluxes across all element boundaries, $\widehat{u}$ will not depend on $\boldsymbol{q}$ making it possible for the local variable $\boldsymbol{q}$ to be solved in terms of u. These characteristics will be described in detail in Chapter 5.

Works addressing the properties of the LDG method in the context of convection diffusion problems include for example: [28], [26], [23], and [1]. The applicability of the LDG method has been explored for example, for elliptic problems in [5], for nonlinear diffusion problems in [19], for Richard's equation (a nonlinear parabolic equation) in [49], and for PDE's with higher order derivatives in [70] and [71].

1.3   Accomplishments

The main results of this dissertation can be summarized as follows:

- In section 1.2 of this Chapter, we present a collection of studies where the DSW has been used as a suitable model for shallow water flow in particular flow regimes; we describe the most relevant results in studies of doubly nonlinear diffusion equations existing in the literature that can be applied to the DSW equation when *topographic effects* are ignored (*i.e.*, when $z \neq 0$); and finally, we give an overview of some relevant works addressing the numerical approximation of solutions to nonlinear parabolic equations.

- In Chapter 2, we state the appropriate initial boundary value problem arising from the DSW equation. This is done by combining the knowledge from the modeling applications of the DSW and the knowledge coming from experimental works where the parameters α and γ are considered to be flexible (generally in the ranges $\alpha > 1$ and $0 < \gamma \leq 1$) as a way to account for more general circumstances when flow changes back and forth between turbulent and laminar conditions in vegetated regions. An alternative, simple, and intuitive derivation of the DSW equation in the context of shallow water flows is obtained.

- In Chapter 3, an original and constructive proof of existence of weak solutions to the zero-Dirichlet initial/boundary value problem (2.1) using the Faedo Galerkin method is presented, for situations when *topographic effects* are ignored. Proofs for some regularity results are also shown, including conditions for which a comparison principle for solutions can be established and thus conditions for which uniqueness of solutions can be ensured. The ideas presented in this Chapter are the result of a joint collaboration with Ricardo Alonso.

- In Chapter 4, a numerical method based on the continuous Galerkin finite element method is proposed. Conditions, based on physical consistency, for which *a priori* error estimates and convergence rates can be ensured, are described for this method. The results of numerical experiments in 1-D, using a lumped mass finite element code (in Matlab), are reported. The main objectives of these experiments include:

  - to verify the convergence rates of the method,
  - to numerically simulate water depth profiles and hydrographs in unsteady flow conditions. The numerical results matched the real measurements obtained in an experiment conducted by Iwagaki [46].

17

– to investigate qualitative properties of solutions to the DSW equation when topographic effects are incorporated.

• In Chapter 5, the local discontinuous Galerkin method is introduced along with conditions for which *a priori* error estimates and convergence rates can be ensured. The results of numerical experiments in 2-D, using an LDG finite element code (in FORTRAN) implemented to solve the DSW equation and aiming at verifying the modeling qualities of the DSW equation in the context of shallow water modeling, are presented. These include: the simulation of a dam break problem, and the simulation of water flow in an inclined and vegetated region.

Chapter 2

The DSW Equation and Preliminaries

The outline of this Chapter is the following. In section 2.1, we will introduce the initial boundary value problem associated with the DSW equation, following the notation found in the mathematical literature on Partial Differential Equations (PDE), see for example [36]. Throughout this dissertation, we will consistently use this notation except for two sections: section 1.2.1 and section 2.2, where the notation commonly used in the context of shallow water hydrodynamics will be utilized, see for example [65]. In 2.2, a derivation of the DSW equation from the shallow water system of equations will be provided. For consistency, we provide further details about the mathematical notation used in this dissertation in section 2.3. Well known results in interpolation theory will be listed for completeness in section 2.4.

2.1 The initial-boundary value problem

The DSW equation gives rise to the following initial/boundary-value problem prescribed for any fixed $T > 0$

$$
\begin{cases}
\dfrac{\partial u}{\partial t} - \nabla \cdot \left(\dfrac{(u-z)^\alpha}{|\nabla u|^{1-\gamma}} \nabla u \right) = f & \text{on } \Omega \times (0, T] \\[2ex]
u = u_0 & \text{on } \Omega \times \{t = 0\} \\[2ex]
\left(\dfrac{(u-z)^\alpha}{|\nabla u|^{1-\gamma}} \nabla u \right) \cdot n = g_N & \text{on } \partial\Omega \cap \Gamma_N \times (0, T] \\[2ex]
u = g_D & \text{on } \partial\Omega \cap \Gamma_D \times (0, T]
\end{cases}
\tag{2.1}
$$

where Ω is an open, bounded subset of $\mathbb{R}^n$ ($n = 1, 2$) and Γ_N and Γ_D are subsets of $\partial\Omega \in C^1$ such that $\partial\Omega = \Gamma_N + \Gamma_D$. Also $f : \Omega \times (0, T] \to \mathbb{R}$, $u_0 : \Omega \to \mathbb{R}$, $g_N : \Gamma_N \times (0, T] \to \mathbb{R}$, $g_D : \Gamma_D \times (0, T] \to \mathbb{R}$ are given functions, $z : \overline{\Omega} \to \mathbb{R}^+$ is a positive time independent function, $0 < \gamma \leq 1$, $1 < \alpha < 2$, and $u : \overline{\Omega} \times [0, T] \to \mathbb{R}$ is the unknown. Here $|\cdot| : \mathbb{R}^n \to \mathbb{R}$ refers to the Euclidean norm in $\mathbb{R}^n$.

As mentioned in Chapter 1, problem (2.1) is characterized as doubly nonlinear since the nonlinear behaviour appears inside the divergence term as a product of two nonlinearities involving $u - z$ and ∇u, namely $(u - z)^\alpha$ and $\nabla u / |\nabla u|^{1-\gamma}$. In fact, when the DSW equation is re-written as

$$\frac{\partial u}{\partial t} - \nabla \cdot (a(u, \nabla u) \, \nabla u) = f \qquad \text{with} \qquad a(u, \nabla u) = \frac{(u - z)^\alpha}{|\nabla u|^{1-\gamma}},$$

where a is the diffusion coefficient, one can be more specific and characterize it as a doubly nonlinear and degenerate-singular equation for the given choice of $0 < \gamma \leq 1$, and $1 < \alpha < 2$. This is the case since $a \to 0$ when $(u - z) \to 0$ and $a \to \infty$ when $\nabla u \to 0$. In the context of shallow water modeling, $u(x, t)$ represents the surface water elevation in the position x at time t, the positive time independent function $z(x)$ describes the bathymetry of the bed surface throughout the domain and introduces the commonly called *topographic effects* into the model.

The types of physical boundary conditions appropriate for this model are two, a prescribed water depth g_D on Γ_D, and/or a prescribed water flux g_N on Γ_N. The first one corresponds to a Dirichlet type boundary condition and it is mostly used to model an infinite source of water on the boundary Γ_D. The second one corresponds to a Neumann type boundary condition and it is the most natural choice to model water flux through a boundary Γ_N.

2.2 Derivation from the shallow water equations

Models for surface water flows are derived from the incompressible, three-dimensional Navier-Stokes (NS) equations, which consist of momentum equations for the three velocity components and a continuity equation. Depending on the physics of the flow, scaling arguments are used in order to obtain effective equations for the problem at hand. Equation (2.1) is a simplified version of the two-dimensional shallow water equations called the dif-

fusive wave or zero-inertia approach. This equation is commonly derived by neglecting the inertial terms in the horizontal momentum equations and substituting the bottom slope in Manning's formula by the water surface slope. This approach is shown in [38], [69], [44], [17], and [39]. In the following paragraphs we provide an intuitive, concise and equally valid derivation following a more empirical approach, such as the one used to derive the Porous Medium Equation in section 2 of [62].

Recall that in shallow water theory, the main scaling assumption is that the vertical scales are small relative to the horizontal ones. This approximation reduces the vertical momentum equation to the hydrostatic pressure relation

$$\frac{\partial p}{\partial x_3} = \rho g \tag{2.2}$$

where g is the gravitational constant, x_3 the vertical coordinate and p the pressure, and leaves us with two effective momentum equations in the horizontal direction. Upon vertical integration of the NS equations, we obtain two depth-averaged momentum equations and a depth-averaged continuity equation. These resulting equations are called the 2-D shallow water equations. For a detailed description of shallow water hydrodynamics see [65] and [63]. When combining the depth-averaged continuity equation with the free surface boundary condition, one obtains the mass balance equation

$$\frac{\partial h}{\partial t} + \nabla \cdot (h\boldsymbol{V}) = f, \tag{2.3}$$

where $h(x,t) = H(x,t) - z(x)$ is the water depth, $H(x,t)$ is the free water surface elevation or hydraulic head, $z(x)$ is the bed surface, bathymetry, or land elevation, $\boldsymbol{V}(x,t)$ is the depth-averaged velocity, and $f(x,t)$ is a source/sink (such as rainfall or infiltration).
In open channel flow theory, empirical laws such as Manning's formula or Chézy's formula have been observed to successfully describe the dynamics of water flow in regimes when fluid motion is dominated by gravity and balanced by the bottom boundary shear stress. See [47] or chapter 11 in [30]. Examples of open channel flow include water flow in rivers, in partially full drains and surface runoff. Manning's and Chézy's formulas relate the mean velocity of the flow V with the so-called *hydraulic radius*[1] R and the *bottom slope S* through

---

[1] The hydraulic radius for open channels is calculated as $R = A/\omega$, where A is the cross section of the channel, and ω is the wetted perimeter. Note that for a rectangular cross section with base L and depth h, $R = hL/(L + 2h) \sim h$ when $L >> h$.

a friction coefficient c_f in the following way:

$$V = \frac{1}{c_f} R^{\alpha-1} S^{\gamma}, \tag{2.4}$$

for particular choices of α and γ. For Manning's formula[2] $\alpha = 5/3$ and $\gamma = 1/2$, and for Chézy's formula $\alpha = 3/2$ and $\gamma = 1/2$. When one multiplies equation (2.4) by the hydraulic radius R, one obtains an equivalent relation in terms of the water discharge Q

$$Q = RV = \frac{1}{c_f} R^{\alpha} S^{\gamma}. \tag{2.5}$$

The discharge-depth equation (2.5) is a generalization of both Manning's formula or Chézy's formulas and was proposed in [61] as a way to account for more general circumstances when flow changes back and forth between turbulent and laminar conditions. This is the case, for example, in water flow in vegetated areas. In [61], the authors study equation (2.5) as a prediction model for shallow water flow in vegetated areas based entirely on empirical procedures. In their study they conclude that equation (2.5) with flexible coefficients α and γ results in a broader and better model than the particular Manning's formula. Experimentally, they reported values in the ranges $1 \leq \alpha \leq 2$ and $0 < \gamma < 1$. These values motivate the ranges of α and γ in the present work. Further assumptions in open channel theory that justify the application of velocity-depth equations like (2.4) include:

- the approximation of the hydraulic radius R by the water depth h in (2.4) and (2.5),

- the assumption that the slope of the bathymetry is small, and

- the assumption that the bottom slope is comparable to the free water surface slope.

In the diffusive wave approximation, one makes use of the previous assumptions, and extends the scaling of the mean flow velocity V with respect to R and S in (2.4), to the depth-averaged velocity $\boldsymbol{V}(x,t)$ in (2.3) along the direction of the flow. This is done in the following way: since the flow is assumed to be dominated by gravity, the direction of the flow will be along the unitary vector $\nabla H/|\nabla H|$ (recall equation (2.2)), and thus, equation (2.4) is transformed into

$$\boldsymbol{V} = -\frac{h^{\alpha-1}}{c_f} \frac{\nabla H}{|\nabla H|} |\nabla H|^{\gamma} = -\frac{(H-z)^{\alpha-1}}{c_f} \frac{\nabla H}{|\nabla H|^{1-\gamma}}, \tag{2.6}$$

The DSW equation is obtained from substituting the particular form of the depth-averaged

---

[2]For a derivation of Manning's formula based on the phenomenological theory of turbulence see [40].

horizontal velocity given by (2.6), into equation (2.3)

$$\frac{\partial H}{\partial t} - \nabla \cdot \left(\frac{(H-z)^\alpha}{c_f} \frac{\nabla H}{|\nabla H|^{1-\gamma}} \right) = f(t,x), \qquad \text{for } (t,x) \in \mathbb{R}^+ \times \mathbb{R}^2, \qquad (2.7)$$

The assumptions made to obtain the DSW equation suggest that it may be suitable to serve as a model in low-to-moderate velocity-flow regimes. See section 2 of [39] and the references therein.

Remark 2.2.1. In hydrological systems, z describes the bed surface over which water flows, thus, in physically meaningful situations one assumes that ∇z must be bounded. This in turn implies, in physically meaningful solutions, the boundedness of ∇H. This is an extra assumption that will be used in the numerical error analysis that aligns well with the physics of the associated problem.

Remark 2.2.2. In this context, equation (2.7) makes sense physically only if $H - z \geq 0$. It is with this in mind that we will not pay attention to the approximation of negative solutions of (2.1). Note that in writing (2.1) we have assumed that $c_f(x) \equiv 1$.

Remark 2.2.3. Note that if one identifies the water elevation H with the hydrostatic pressure p, the expression that relates the velocity and the water elevation gradient (2.6) becomes a nonlinear version of the empirical *Darcy's law* for gas flow through a porous medium. Indeed, flow in vegetated areas such as wetlands can be understood as a flow through a porous medium.

One interesting fact about solutions of the DSW is the following. When one sees the DSW equation as a conservation law with respect to the depth $u^* = u - z$, it becomes

$$\frac{\partial u^*}{\partial t} - \nabla \cdot (u^* V) = f,$$

where the horizontal velocity V is given by (2.6), and its magnitude is

$$|V| = \frac{|u^*|^{\alpha-1}}{c_f} |\nabla u|^\gamma,$$

which indicates that at the free boundary (interface between regions where $u^* > 0$ and $u^* = 0$) or any place in the domain where the depth of the water u^* is zero, the magnitude of the velocity is zero since $\alpha > 1$.

Remark 2.2.4. For studies addressing the applicability of the DSW equation as a model to simulate shallow water flow, instead of the full Saint Venant (or Shallow Water) equations in experimental and real life settings we refer the reader, for example, to: in the 1 -D case

the works of Ponce et al. [55] and [54], and in 2-D cases the references mentioned in section 1.2.1.

2.3 Notation

We will use the standard notation introduced in [36]. Let X be a real Banach space, with norm $\|\cdot\|$. The symbol $L^p(0,T;X)$ will denote the Banach space of all measurable functions $u:[0,T] \to X$ such that

(i) $\|u\|_{L^p(0,T;X)} := \left(\int_0^T \|u(t)\|^p \right)^{1/p} < \infty,$ for $1 \le p < \infty$, and

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

We will denote with $C([0,T];X)$ the space of all continuous functions $u:[0,T] \to X$ such that

$$\|u\|_{C(0,T;X)} := \max_{0 \le t \le T} \|u(t)\| < \infty.$$

Let $u \in L^1(0,T;X)$, we say $v \in L^1(0,T;X)$ is the weak time derivative of u, denoted $u_t = v$, provided

$$\int_0^T \psi_t(t)\, u(t) = - \int_0^T \psi(t)\, v(t)$$

for all scalar test functions $\psi \in C_0^\infty(0,T)$. Throughout the paper, $W^{1,p}(0,T;X)$ will denote the space of all functions $u \in L^p(0,T;X)$ such that u_t exists in the weak sense and $u_t \in L^p(0,T;X)$ with the norm

$$\|u\|_{W^{1,p}(0,T;X)} := \begin{cases} \left(\int_0^T \|u(t)\|^p + \|u_t(t)\|^p \right)^{1/p} & (1 \le p < \infty), \\[2mm] \operatorname{ess\,sup}_{0 \le t \le T} \left(\|u(t)\| + \|u_t(t)\| \right) & (p = \infty). \end{cases}$$

For $1 \le p \le +\infty$, we will denote its conjugate as p^* i.e., $1/p + 1/p^* = 1$. For any measurable set $E \subset \Omega$ and real valued vector functions $u \in L^p(E)$ and $v \in L^{p^*}(E)$ we will denote the duality pairing between u and v as

$$(u,v)_E := \int_E u \cdot v.$$

For simplicity, we use $(u,v) := (u,v)_\Omega$. Similarly, we will denote the duality pairing between $u \in W^{-1,p^*}(\Omega)$ and $v \in W_0^{1,p}(\Omega)$ as $\langle u,v \rangle$. Recall that the elements of $W^{-1,p^*}(\Omega)$ are the distributions that have continuous extension to $W_0^{1,p}(\Omega)$. These spaces are characterized in

24

the following way: if $u \in W^{-1,p^*}(\Omega)$, then there exists functions $f^0, f^1, \cdots, f^n$ in $L^{p^*}(\Omega)$ such that

$$\langle u, v \rangle = (f^0, v) + \sum_{i=1}^{n}(f^i, v_{x_i}).$$

Throughout the paper, C will be a generic constant with different values and the explicit dependence with respect to parameters will be written inside parenthesis.

2.4 Interpolation theory results. Continuous case

For Lemmas 2.4.1, 2.4.2, 2.5.1, and 2.5.2, we will consider τ to be a quasi-uniform triangulation of Ω into elements E_i, $i = 1, ..., m$, with $\text{diam}(E_i) = h_i$ and $h = \max_i(h_i)$. $\mathcal{M}(= \mathcal{P}^k)$ will denote a finite dimensional subspace of $H_0^1(\Omega)$ defined on this triangulation consisting of piecewise polynomials of degree at most k and K_0 will denote a constant independent of h and v.

Lemma 2.4.1 (Interpolation error). *Let* $u \in H^{k+1}(\Omega)$, *then there exists* $\hat{u} \in \mathcal{M}$, *projection of* u, *defined by*

$$\int (\hat{u} - u)v = 0 \quad \forall\, v \in \mathcal{M} \tag{2.8}$$

with the following property:

$$\|\hat{u} - u\|_{H^s(\Omega)} \le C\, h^{k+1-s}\, \|u\|_{H^{k+1}(\Omega)}$$

where $0 \le s \le k$.

Proof. See section 4.4 in [18]. $\qquad\square$

Lemma 2.4.2 (Inverse inequalities). *Let* $v \in \mathcal{M}$ *then, there exists a constant* K_0 *independent of* h *and* v *such that*

$$\|v\|_{L^\infty(\Omega)} \le K_0 h^{-1}\|v\|_{L^2(\Omega)} \qquad and \qquad \|\nabla v\|_{L^\infty(\Omega)} \le K_0 h^{-1}\|\nabla v\|_{L^2(\Omega)}$$

Proof. See section 4.5 in [18]. $\qquad\square$

Remark 2.4.1. Lemma 2.4.1 implies that for a subspace $\mathcal{M} = \mathcal{P}^1$ consisting of piecewise linear polynomials,

$$\|\hat{u} - u\|_{L^2(\Omega)} \le C\, h^2\, \|u\|_{H^2(\Omega)} \qquad \text{and} \qquad \|\nabla\hat{u} - \nabla u\|_{L^2(\Omega)} \le C\|\hat{u} - u\|_{H^1(\Omega)} \le C\, h\, \|u\|_{H^2(\Omega)} \tag{2.9}$$

These inequalities will be useful in section 4.4.

2.5 Interpolation theory results. Discontinuous case

In Chapter 5, a discontinuous Galerkin method will be introduced where the need to have a possibly broken approximate solution across elements may arise. In such cases, the results of Lemmas 2.4.1 and 2.4.2 will be valid when applied to each element Ω_e separately, and as a consequence, they will also be valid when applied to the union (or the sum over elements). However, we require two extra Lemmas in order to bound the errors associated to the element boundary integrals. The following lemmas address this issue.

Lemma 2.5.1 (Trace interpolation error). *Let $u \in H^{k+1}(\Omega)$, then there exists $\hat{u} \in \mathcal{M}$, interpolant of u with the following property:*

$$\|\hat{u} - u\|_{H^s(\partial\Omega_e)} \leq C\, h^{k+\frac{1}{2}-s}\, \|u\|_{H^{k+1}(\Omega_e)}$$

where $0 \leq s \leq k$.

Proof. See [25]. $\qquad\square$

Lemma 2.5.2 (Trace inequality). *Provided $\boldsymbol{s} \in (\mathcal{M})^n$, where $n = 1, 2$, is the space dimension, then there exists a constant C independent of the mesh size h such that*

$$\|\boldsymbol{s}\|_{L^2(\partial\Omega_e)} \leq C\, h^{-\frac{1}{2}}\, \|\boldsymbol{s}\|_{L^2(\Omega_e)}$$

Proof. See [25]. $\qquad\square$

Chapter 3

Existence, Regularity and Uniqueness of Weak Solutions

As mentioned in Chapter 1, the study of existence, uniqueness, and regularity of weak solutions to the IBVP (2.1) remains an open problem. However, when topographic effects are ignored ($z \equiv 0$) and for nonnegative solutions, the IBVP (2.1) can be transformed into the IBVP (1.7). The overall strategy of this Chapter is to study in depth the IBVP (1.7) and then interpret the findings in terms of the IBVP (1.8) through corollaries and observations. Recall that the IBVP (1.8) is a generalization (when $z \equiv 0$) of the IBVP (2.1) and that nonnegative solutions of the IBVP (1.8) coincide with nonnegative solutions of the IBVP (2.1) when $\phi(s)$ is chosen as in (1.9). The ideas presented in this Chapter were conceived in joint collaboration with Ricardo Alonso and resulted in the publication of [2].

The outline of this Chapter is the following. In section 3.0.1, we describe the relevance of the analysis carried out in this Chapter in the context of shallow water modeling. In section 3.0.2, we provide some intuition about the regularization technique used in the proof of existence. The notion of weak solution and its precise meaning in this analysis is introduced in section 3.0.3. In section 3.1, we present a concise and constructive proof of existence of weak solution to the alternative problem (1.7) using techniques originally introduced by Lions [50] and further developed for quasi-linear parabolic equations by Alt and Luckhaus [3] and for doubly nonlinear equations by Bernis[14]. This constructive method provides a natural setting for a computational method to find approximate solutions to problem (1.7), further described in Chapters 4 and 5, within the framework of finite element techniques. In this proof of existence, instead of following the time discretization approach established in [56] and [41], we take advantage of the continuous in time evolu-

tion of the appropriate Banach space norms of the approximate solutions and find *a priori* estimates for them. This is a standard technique proposed in [50] that does not require any truncation-penalization technique as the one used in [15]. The approximate solutions constructed in this proof are *solutions fortes* in the sense of Bamberger [6], and thus, they and their limit will satisfy all the results presented in [6]. In particular, the result on uniqueness of *limite de solutions fortes* in [6] will ensure that the numerical schemes analysed in Chapters 4 and 5 will converge to a unique solution. It is important to note that in this study we do not require the nonlinearity in time to be locally Lipschitz as in [15].

In addition, we include a concise argument to prove the L^∞ control and integrability properties of the time derivative of solutions in section 3.2. Although these results have been studied, the regularity arguments we present in this thesis are hard to find in the literature and provide insight on the complexities of the equation.

For completeness, in section 3.3, we include the proof of a comparison result mentioned in [6], and use it to prove uniqueness, nonnegativity and stability of the proposed approximation scheme. In the last section, we present possible avenues of research as well as a brief discussion of this Chapter.

3.0.1   The obstacle problem

It is important to mention at this point that, within the shallow water modeling context, a complete analysis of the DSW equation and thus problem (2.1), should be posed as an obstacle problem, in other words, any physical solution u of problem (2.1) should be greater or equal than the topography z (in this Chapter considered flat) regardless of the sign of the input f (possibly negative when modeling physical processes such as *infiltration* or *evaporation*). In other words, one should study the time evolution of the positivity set of $u - z$, denoted by $(u - z)_+$, and as a consequence, one should also characterize the properties of the *free boundary* (interface between regions where $u = z$ and $u > z$). As stated in the previous paragraphs, this is not the way we will analyse problem (2.1). Furthermore, note that in the approach followed in this Chapter, a solution u of problem (1.7) (recall equation (1.11)) could be negative, and thus physically inconsistent. As a note in favor of our approach, however, we will show in section 3.3 that the nonnegativity of f will imply the nonnegativity of a solution u of problem (1.7), for any physically consistent initial condition $u_0 \geq 0$. Furthermore, the analysis presented in this Chapter will be physically relevant for all cases when the combination of inputs f (*infiltration, evaporation,* and *rainfall*) are such that $u \geq 0$.

A classical example of an obstacle problem approach can be found in [62] for the Porous

28

Medium Equation, where free boundary issues need to be explicitly addressed. A closely related one dimensional obstacle problem formulation can be found in [20] for a doubly nonlinear parabolic equation arising in ice sheet dynamics. In general, the theory of free boundaries is an important and difficult subject of mathematical investigation. In particular, the free boundary theory for doubly nonlinear equations is an area of research far from being complete.

3.0.2 Regularization technique

Since the *regularization technique* used in the the proof of existence inspires the formulation of the numerical methods used to approximate solutions of the DSW equation in Chapters 4 and 5, we will elaborate on it briefly. The key idea of this *regularization technique* is to find a solution to problem (1.7) when ϕ is replaced by a Lipschitz function ϕ_{reg} approximating ϕ uniformly, such that $|\phi_{reg}| \leq |\phi|$. Then, the goal is to show that a solution to problem (1.7) can be found as a limit of these *regularized* solutions. Similar regularization techniques have been used in the approximation of solutions of degenerate parabolic equations as described in section 1.2.3. These include: [53], [57], [64], [9], [72].

This regularization process can be interpreted as enforcing some sort of ellipticity condition for the original problem (2.1) (by means of problem (1.8)), since it is a regularization for small values of u where the degenerate character of problem (2.1) arises. To support such interpretation, we plotted in Figure 3.1, functions $\phi(x)$, $\phi^{-1}(x)$ and $(\phi^{-1})'(x)$ without the Lipschitz property, and functions $\phi_{reg}(x)$, $\phi_{reg}^{-1}(x)$ and $(\phi_{reg}^{-1})'(x)$ with the Lipschitz property. In particular, the plot of $(\phi_{reg}^{-1})'(x)$ shows that the replacement of ϕ by a Lipschitz function ϕ_{reg} in the IBVP (1.7) implies naturally the enforcing of ellipticity in the IBVP (1.8) (and thus in the IBVP (2.1)).

3.0.3 Definitions of Weak Solutions

From now on, we will assume that $\phi(s)$ and η are given by (1.9), and $0 < \gamma \leq 1$, $1 < \alpha < 2$.

Definition 3.0.1. we say a function

$$v \in L^{1+\gamma}(0, T; W_0^{1,(1+\gamma)}(\Omega)), \text{ with } \phi(v)_t \in L^{(1+\gamma)^*}(0, T; W^{-1,(1+\gamma)^*}(\Omega)),$$

is a weak solution of the initial/boundary-value problem (1.7) provided

$$\langle \phi(v)_t, w \rangle + \eta^\gamma \left(\frac{\nabla v}{|\nabla v|^{1-\gamma}}, \nabla w \right) = (f, w) \qquad \text{a.e in time } 0 \leq t \leq T, \qquad (3.1)$$

29

FIGURE 3.1: *Figures showing the difference between ϕ and ϕ_{reg}.* **(top left)** $\phi(x)$, **(top right)** $\phi_{reg}(x)$, **(center left)** $\phi^{-1}(x)$, **(center right)** $\phi_{reg}^{-1}(x)$, **(bottom left)** $(\phi^{-1})'(x)$, **(bottom right)** $(\phi_{reg}^{-1})'(x)$.

for any $w \in W_0^{1,(1+\gamma)}(\Omega))$ and

$$v(0) = v_0. \tag{3.2}$$

Definition 3.0.2. we say a function u, with the properties

$$\phi^{-1}(u) \in L^{1+\gamma}(0,T; W_0^{1,1+\gamma}(\Omega)), \text{ and } u_t \in L^{(1+\gamma)^*}(0,T; W^{-1,(1+\gamma)^*}(\Omega)),$$

is a weak solution of the initial/boundary-value problem (1.8) provided

$$\langle u_t, w \rangle + \eta^\gamma \left(\left((\phi^{-1})'(u) \right)^\gamma \frac{\nabla u}{|\nabla u|^{1-\gamma}}, \nabla w \right) = (f, w) \qquad \text{a.e in time } 0 \le t \le T, \tag{3.3}$$
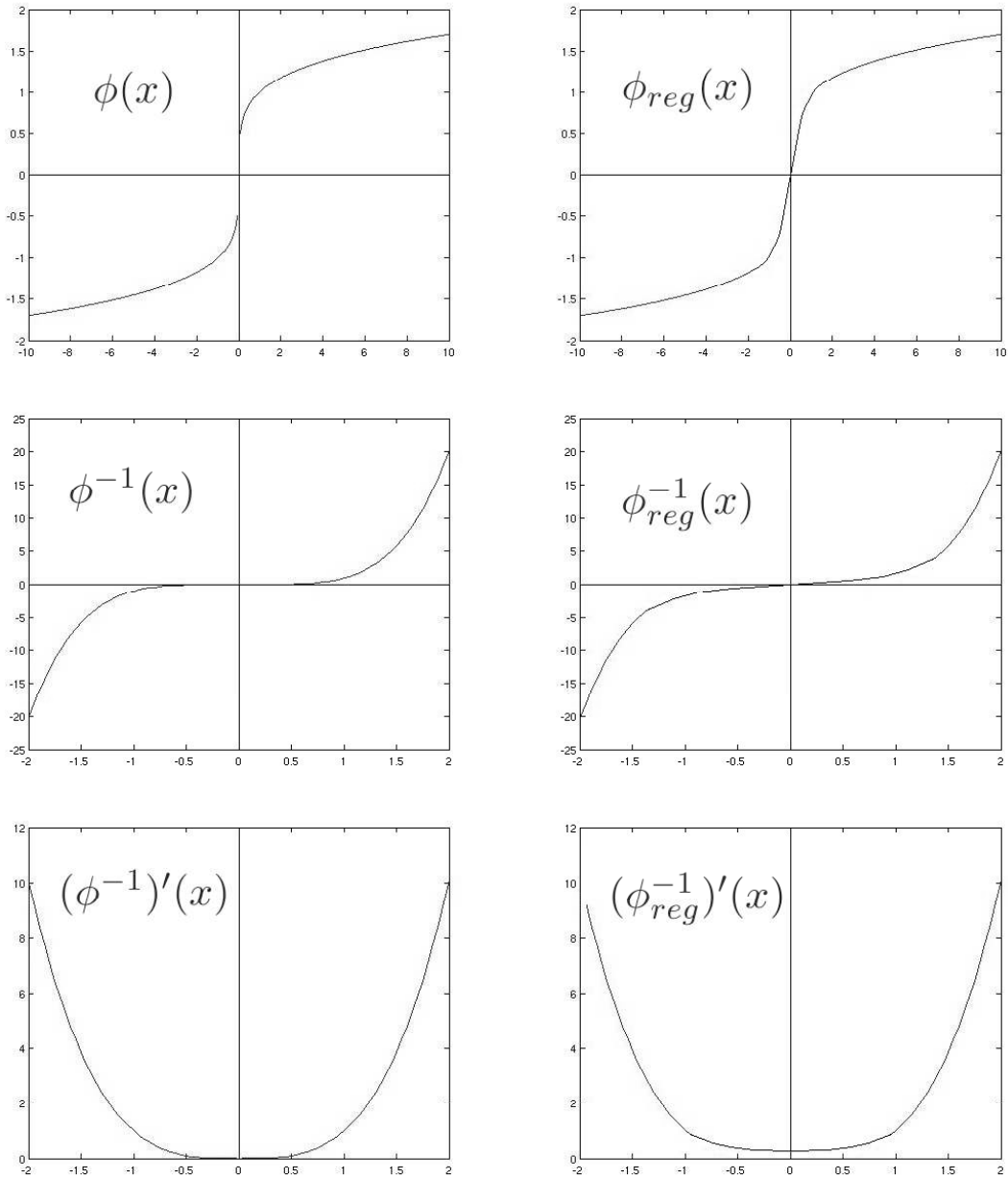
for any $w \in W_0^{1,(1+\gamma)}(\Omega))$ and

$$u(0) = u_0. \tag{3.4}$$

Remark 3.0.1. A consequence of Definition 3.0.1 (resp. Definition 3.0.2) is that

$$\phi(v) \in C([0,T]; W^{-1,(1+\gamma)^*}(\Omega)) \qquad (\text{ resp. } u \in C([0,T]; W^{-1,(1+\gamma)^*}(\Omega)))$$

thus condition (3.2) (resp. (3.4)) makes sense.

Remark 3.0.2. In Definition 3.0.2, we understand the pointwise gradient of u, denoted as ∇u, as the function

$$\nabla u = \begin{cases} \phi'(v)\nabla v & \text{if} \quad |v| > 0 \\ 0 & \text{if} \quad v = 0. \end{cases}$$

where $v \in L^{1+\gamma}(0,T; W_0^{1,(1+\gamma)}(\Omega))$ is a weak solution of the initial/boundary-value problem (1.7).

3.1 Existence

In order to prove the existence of a weak solution of problem (1.7) we will use the Faedo-Galerkin method using compactness and monotonicity arguments as explained in [50]. The method consists of five main steps:

Step 1. Constructing approximate solutions by the method of Faedo-Galerkin.

Step 2. Finding *a priori* estimates on such approximate solutions.

Step 3. Using the properties of compactness to extract a converging sub-sequence to pass to the limit.

Step 4 and Step 5. Using the monotonicity of the nonlinear operator $\mathcal{A}(\mathbf{x})$ (See Appendix A) to prove that the limit process indeed leads to a weak solution.

The key idea of the proof is to find a solution to problem (1.7) when ϕ is replaced by a Lipschitz function ϕ_{reg} approximating ϕ uniformly, such that $|\phi_{reg}| \leq |\phi|$. Throughout this Chapter we will refer to any of these approximations as *regular ϕ* or ϕ_{reg}, indistinctively. Then, we will show that a solution to problem (1.7) can be found as a limit of these *regularized* solutions. For the sake of clarity, the reader can think of the following family of regularized Lipschitz functions,

$$\phi_{reg} = \phi_\epsilon(s) = \begin{cases} \dfrac{\phi(\epsilon)}{\epsilon} s & \text{if} \quad |s| \leq \epsilon, \\ \phi(s) & \text{if} \quad |s| > \epsilon, \end{cases}$$

Clearly, $\phi_\epsilon(s) \to \phi(s)$ uniformly as $\epsilon \to 0$. In fact,

$$\sup_s |\phi(s) - \phi_\epsilon(s)| \leq \epsilon^\eta.$$

With the previous ideas in mind, *Step 1*, *Step 3* and *Step 4* will be performed for any regular ϕ_{reg}, the *a priori* estimates obtained in *Step 2* will be computed for ϕ and thus, they will hold uniformly for any ϕ_{reg}. The latter fact will allow us to find in *Step 5* a subsequence of *regularized* solutions that will converge to a solution of problem (1.7).

Theorem 3.1.1. *Let f and v_0 satisfy*

$$v_0 \in L^{1+\eta}(\Omega) \qquad and \qquad f \in L^{(1+\gamma)^*}(0, T; L^{(1+\gamma)^*}(\Omega)), \tag{3.5}$$

then, there exist a function v with the properties

$$v \in L^{(1+\gamma)}(0, T; W_0^{1,(1+\gamma)}(\Omega)), \tag{3.6}$$

and

$$\phi(v)_t \in L^{(1+\gamma)^*}(0, T; W^{-1,(1+\gamma)^*}(\Omega)), \tag{3.7}$$

such that it solves problem (1.7).

Proof. For clarity we organize the proof in the steps previously described.

Step 1. Approximate Solutions

Let $\{w_j\}_{j=1}^{\infty}$ be a basis of $V = W_0^{1,(1+\gamma)}(\Omega)$. Construct the Faedo-Galerkin approximate solution of problem (1.7), $v_m(t)$, the following way. For any fixed t

$$v_m(t) = \sum_{j=1}^{m} \zeta_j(t) w_j(x) \in [w_1, \ldots, w_m] = \text{the space generated by } \{w_j\}_{j=0}^{m}$$

and satisfying

$$(\phi(v_m)_t, w_j) + \eta^{\gamma} \left(\frac{\nabla v_m}{|\nabla v_m|^{1-\gamma}}, \nabla w_j \right) = (f, w_j) \qquad 1 \le j \le m, \tag{3.8}$$

$$v_m(0) = v_{0,m} \in [w_1, \ldots, w_m],$$

where $v_{0,m} \to v_0$ in $L^{1+\eta}(\Omega)$.

Step 2. *A priori* Estimates

Lemma 3.1.1. *Set $\phi(s) = s/|s|^{1-\eta}$. Let v_m be a Faedo-Galerkin approximate solution of problem (1.7), then the following estimates hold.*

$$\sup_{0 \le t \le T} \|\phi(v_m)(t)\|_{L^{(1+\eta)^*}(\Omega)}^{(1+\eta)^*} \le C \left(\|v_0\|_{L^{1+\eta}(\Omega)}, \|f\|_{L^{(1+\eta)^*}(0,T;L^{(1+\eta)^*}(\Omega))}, T \right) \tag{3.9}$$

and

$$\|\nabla v_m\|_{L^{1+\gamma}(0,T;L^{1+\gamma}(\Omega))}^{1+\gamma} \le C \left(\|v_0\|_{L^{1+\eta}(\Omega)}, \|f\|_{L^{(1+\eta)^*}(0,T;L^{(1+\eta)^*}(\Omega))}, T \right) \tag{3.10}$$

where $(1+\eta)^* = (1+\eta)/\eta$.

Proof. Multiply equation (3.8) by $\zeta_j(t)$ and sum for $1 \le j \le m$ to obtain

$$\frac{d}{dt} \|\phi(v_m)(t)\|_{L^{(1+\eta)^*}(\Omega)}^{(1+\eta)^*} + \frac{1+\eta}{\eta^{1-\gamma}} \int_{\Omega} |\nabla v_m|^{1+\gamma} = \frac{1+\eta}{\eta}(f, v_m) \tag{3.11}$$

and from Young's inequality

$$(f, v_m) \le \frac{\eta}{1+\eta} \|f\|_{L^{(1+\eta)^*}(\Omega)}^{(1+\eta)^*} + \frac{1}{1+\eta} \|\phi(v_m)(t)\|_{L^{(1+\eta)^*}(\Omega)}^{(1+\eta)^*}. \tag{3.12}$$

33

Now, since

$$\int_\Omega |\nabla v_m|^{1+\gamma} \geq 0$$

we obtain the inequality

$$\frac{d}{dt} \|\phi(v_m)(t)\|_{L^{(1+\eta)^*}(\Omega)}^{(1+\eta)^*} \leq \|f\|_{L^{(1+\eta)^*}(\Omega)}^{(1+\eta)^*} + \frac{1}{\eta} \|\phi(v_m)(t)\|_{L^{(1+\eta)^*}(\Omega)}^{(1+\eta)^*}.$$

Using Gronwall's lemma we get that for all $t \in [0, T]$

$$\|\phi(v_m)(t)\|_{L^{(1+\eta)^*}(\Omega)}^{(1+\eta)^*} \leq C\left(\|v_0\|_{L^{1+\eta}(\Omega)}, \|f\|_{L^{(1+\eta)^*}(0,T;L^{(1+\eta)^*}(\Omega))}, T \right)$$

which leads to the first estimate stated in (3.9).

Note: we have assumed, without loss of generality, that

$$\|v_{0,m}\|_{L^{1+\eta}(\Omega)} \leq \|v_0\|_{L^{1+\eta}(\Omega)}.$$

Integrating equation (3.11) in time

$$\|\phi(v_m)(T)\|_{L^{(1+\eta)^*}(\Omega)}^{(1+\eta)^*} + \frac{1+\eta}{\eta^{1-\gamma}} \int_0^T \int_\Omega |\nabla v_m|^{1+\gamma} = \frac{1+\eta}{\eta} \int_0^T (f, v_m) + \|v_{0,m}\|_{L^{1+\eta}(\Omega)}^{1+\eta}.$$

The above expression and inequality (3.12) imply that

$$\|\nabla v_m\|_{L^{1+\gamma}(0,T;L^{1+\gamma}(\Omega))}^{1+\gamma} \leq C\left(\|v_0\|_{L^{1+\eta}(\Omega)}, \|f\|_{L^{(1+\eta)^*}(0,T;L^{(1+\eta)^*}(\Omega))}, T \right)$$

which finishes the proof. $\qquad\square$

Remark 3.1.1. Note that by the Poincaré inequality

$$\|v_m\|_{L^{1+\gamma}(0,T;L^{1+\gamma}(\Omega))} \leq C(\Omega) \|\nabla v_m\|_{L^{1+\gamma}(0,T;L^{1+\gamma}(\Omega))}$$

therefore the sequence $\{v_m\} \subset L^{1+\gamma}(0,T;W_0^{1,1+\gamma}(\Omega))$ and it is uniformly bounded.

Step 3. Passing to the limit

Let $v_m(t)$ be the Faedo-Galerkin sequence of approximate solutions of problem (1.7) defined by (3.8). Estimates (3.9) and (3.10) in Lemma 3.1.1 imply that there exists a convergent

subsequence $\{v_\mu\}$ of $\{v_m\}$ such that

$$v_\mu \rightharpoonup v \qquad \text{in} \quad L^{1+\gamma}(0,T;W_0^{1,1+\gamma}(\Omega)) \quad \text{weakly,} \tag{3.13}$$

$$\phi(v_\mu)(T) \rightharpoonup \xi \qquad \text{in} \qquad L^{(1+\eta)^*}(\Omega) \qquad \text{weakly.} \tag{3.14}$$

In addition, inequality (3.10) implies

$$\frac{\nabla v_\mu}{|\nabla v_\mu|^{1-\gamma}} \rightharpoonup \chi \qquad \text{in} \qquad L^{(1+\gamma)^*}(0,T;L^{(1+\gamma)^*}(\Omega)) \text{ weakly.} \tag{3.15}$$

Integrating equation (3.8) in time and using the aforementioned convergence results, we can take the limit as $\mu \to \infty$ to find that for any $w \in L^{1+\gamma}(0,T;W_0^{1,1+\gamma}(\Omega))$

$$\lim_{\mu \to \infty} \int_0^T (\phi(v_\mu)_t, w) = -\eta^\gamma \int_0^T (\chi, \nabla w) + \int_0^T (f, w). \tag{3.16}$$

We can conclude that

$$\phi(v_\mu)_t \rightharpoonup \vartheta \qquad \text{in} \qquad L^{(1+\gamma)^*}(0,T;W^{-1,(1+\gamma)^*}(\Omega)) \qquad \text{weakly,} \tag{3.17}$$

where the functional ϑ is defined by the right hand side of equation (3.16). Using (3.13) and Theorem A.0.2 in Appendix A we can conclude that

$$\phi(v)_t = \vartheta. \tag{3.18}$$

Therefore, for any $w \in L^{1+\gamma}(0,T;W_0^{1,1+\gamma}(\Omega))$

$$\int_0^T \langle \phi(v)_t, w \rangle = -\eta^\gamma \int_0^T (\chi, \nabla w) + \int_0^T (f, w). \tag{3.19}$$

Note also that $L^{(1+\gamma)/\eta}(\Omega) \subset W^{-1,(1+\gamma)^*}(\Omega)$, hence we have

$$\phi(v) \in L^{(1+\gamma)^*}(0,T;L^{(1+\gamma)/\eta}(\Omega)) \subset L^{(1+\gamma)^*}(0,T;W^{-1,(1+\gamma)^*}(\Omega)).$$

Using the previous fact, together with (3.17) and (3.18)

$$\phi(v) \in W^{1,(1+\gamma)^*}(0,T;W^{-1,(1+\gamma)^*}(\Omega)).$$

So by Theorem A.0.1 in Appendix A we conclude that

$$\phi(v) \in C([0,T];W^{-1,(1+\gamma)^*}(\Omega))$$

and

$$\phi(v)(t) - \phi(v)(s) = \int_s^t \phi(v)_t \ \text{ for all } \ 0 \le s \le t \le T. \tag{3.20}$$

Multiply equation (3.20) by $w \in W^{1,1+\gamma}(\Omega)$ and integrate in Ω to obtain

$$
\begin{aligned}
\langle \phi(v)(T) - \phi(v_0), w \rangle &= \int_0^T \langle \phi(v)_t, w \rangle \\
&= \lim_{\mu \to \infty} \int_0^T (\phi(v_\mu)_t, w) \\
&= \lim_{\mu \to \infty} (\phi(v_\mu)(T) - \phi(v_{0,\mu}), w) \\
&= \langle \xi - \phi(v_0), w \rangle .
\end{aligned}
$$

Since w is arbitrary, we conclude that

$$\phi(v)(T) = \xi. \tag{3.21}$$

Step 4. Monotonicity argument

It only remains to show that

$$\chi = \frac{\nabla v}{|\nabla v|^{1-\gamma}}$$

in equation (3.19). For that purpose, recall by the monotonicity Lemma A.0.1 in Appendix A that for any $w \in L^{1+\gamma}(0, T; W_0^{1,1+\gamma}(\Omega))$

$$X_\mu \equiv \eta^\gamma \int_0^T \left(\frac{\nabla v_\mu}{|\nabla v_\mu|^{1-\gamma}} - \frac{\nabla w}{|\nabla w|^{1-\gamma}}, \nabla v_\mu - \nabla w \right) \ge 0$$

which we can rewrite as

$$X_\mu = T_{1,\mu} + T_{2,\mu}$$

where

$$T_{1,\mu} = \eta^\gamma \int_0^T \left(\frac{\nabla v_\mu}{|\nabla v_\mu|^{1-\gamma}}, \nabla v_\mu \right)$$

and

$$T_{2,\mu} = -\eta^\gamma \int_0^T \left(\frac{\nabla v_\mu}{|\nabla v_\mu|^{1-\gamma}}, \nabla w \right) - \eta^\gamma \int_0^T \left(\frac{\nabla w}{|\nabla w|^{1-\gamma}}, \nabla v_\mu - \nabla w \right).$$

Note that

$$\limsup_\mu X_\mu = \limsup_\mu T_{1,\mu} + \limsup_\mu T_{2,\mu} \ge 0. \tag{3.22}$$

From (3.13) and (3.15) one can easily see that

$$\limsup_{\mu} T_{2,\mu} = \lim_{\mu} T_{2,\mu} = -\eta^\gamma \int_0^T (\chi, \nabla w) - \eta^\gamma \int_0^T \left(\frac{\nabla w}{|\nabla w|^{1-\gamma}}, \nabla v - \nabla w \right). \tag{3.23}$$

For the term $T_{1,\mu}$ one needs to be more careful. Using equation (3.8)

$$
\begin{aligned}
T_{1,\mu} &= -\int_0^T (\phi(v_\mu)_t, v_\mu) + \int_0^T (f, v_\mu) \\
&= -\frac{\eta}{\eta+1} \int_0^T \frac{d}{dt} \|\phi(v_\mu)\|_{L^{(1+\eta)^*}(\Omega)}^{(1+\eta)^*} + \int_0^T (f, v_\mu) \\
&= \frac{\eta}{\eta+1} \|\phi(v_{0,\mu})\|_{L^{(1+\eta)^*}(\Omega)}^{(1+\eta)^*} - \frac{\eta}{\eta+1} \|\phi(v_\mu)(T)\|_{L^{(1+\eta)^*}(\Omega)}^{(1+\eta)^*} + \int_0^T (f, v_\mu).
\end{aligned}
$$

Since by (3.21) and a well know property of weak limits

$$\|\phi(v)(T)\|_{L^{(1+\eta)^*}(\Omega)} = \|\xi\|_{L^{(1+\eta)^*}(\Omega)} \leq \liminf_{\mu} \|\phi(v_\mu)(T)\|_{L^{(1+\eta)^*}(\Omega)}.$$

Thus, we are lead to

$$\limsup_{\mu} T_{1,\mu} \leq \frac{\eta}{\eta+1} \left(\|\phi(v_0)\|_{L^{(1+\eta)^*}(\Omega)}^{(1+\eta)^*} - \|\phi(v)(T)\|_{L^{(1+\eta)^*}(\Omega)}^{(1+\eta)^*} \right) + \int_0^T (f, v).$$

Now, substitute v for w in (3.19). Perform the integration in time to find that

$$\eta^\gamma \int_0^T (\chi, \nabla v) = \frac{\eta}{\eta+1} \left(\|\phi(v_0)\|_{L^{(1+\eta)^*}(\Omega)}^{(1+\eta)^*} - \|\phi(v)(T)\|_{L^{(1+\eta)^*}(\Omega)}^{(1+\eta)^*} \right) + \int_0^T (f, v). \tag{3.24}$$

Thus, from (3.22), (3.23) and (3.24) we observe that

$$\int_0^T \left(\chi - \frac{\nabla w}{|\nabla w|^{1-\gamma}}, \nabla v - \nabla w \right) \geq 0$$

if we choose $w = v - \lambda\psi$ for $\lambda > 0$ and $\psi \in L^{1+\gamma}(0, T; W_0^{1,1+\gamma}(\Omega))$ in the previous equation, then

$$\int_0^T \left(\chi - \frac{\nabla(v - \lambda\psi)}{|\nabla(v - \lambda\psi)|^{1-\gamma}}, \nabla\psi \right) \geq 0.$$

Taking the limit as $\lambda \to 0$ we finally obtain that

$$\int_0^T \left(\chi - \frac{\nabla v}{|\nabla v|^{1-\gamma}}, \nabla\psi \right) \geq 0$$

which implies by Lebesgue's lemma that

$$\chi = \frac{\nabla v}{|\nabla v|^{1-\gamma}}.$$

The previous fact completes the proof of Theorem 3.1.1 for any ϕ_{reg}.

Step 5. Going from ϕ_{reg} to ϕ

Next, take $\{\phi_k\}_{k=1}^{\infty}$ to be a sequence of regularized functions converging uniformly to $\phi(s) = s/|s|^{1-\eta}$. Then, *a priori* estimates (3.9) and (3.10), which are independent of k, hold for the sequences $\{\phi_k(v_k)\}$ and $\{v_k\}$. Whence, *Step 3* and *Step 4* can be identically performed to find that v defined as

$$v = \lim_{k \to \infty} v_k$$

is a weak solution of the problem for the non regular ϕ. $\qquad\qquad\square$

Corollary 3.1.1. *There exists a weak solution to problem (1.8), where the gradient of u is understood as the pointwise gradient.*

Proof. Let v be a weak solution of problem (1.7) with initial condition $v_0 = \phi^{-1}(u_0)$ and let $u = \phi(v)$. Immediately, the following holds:

(i) $u = 0$ in $(0, T) \times \partial\Omega$,

(ii) $u(0) = \phi(v(0)) = \phi(v_0) = \phi(\phi^{-1}(u_0)) = u_0$,

(iii) $\phi(v)_t = u_t$.

It only remains to show that the weak gradient of v and the pointwise gradient of u are related by

(iv) $\nabla v = (\phi^{-1})'(u)\nabla u \qquad$ a.e. in $(0, T) \times \Omega$.

For this purpose, observe that since $v \in L^{1+\gamma}(0, T; W_0^{1,(1+\gamma)}(\Omega))$ there exists a sequence $v_m \in L^{1+\gamma}(0, T; C^{\infty}(\Omega))$ such that

$$v_m \to v \quad \text{strongly in} \quad L^{1+\gamma}(0, T; L^{1+\gamma}(\Omega)) \quad \text{and a.e. in } (0, T) \times \Omega.$$

Define the sequence $u_m = \phi(v_m)$. Since $v_m \in L^{1+\gamma}(0, T; C^{\infty}(\Omega))$, we have that the following relation holds true a.e.

$$\nabla u_m = \begin{cases} \phi'(v_m)\nabla v_m & \text{if} \quad |v_m| > 0 \\ 0 & \text{if} \quad v_m = 0. \end{cases}$$

38

Therefore,

$$u_m \to u \qquad \text{and} \qquad \nabla u_m \to \nabla u \qquad \text{a.e. in } (0, T) \times \Omega.$$

In addition, $v_m = \phi^{-1}(u_m)$, thus

$$\nabla v_m = (\phi^{-1})'(u_m) \, \nabla u_m \qquad \text{a.e. in } (0, T) \times \Omega.$$

Sending $m \to \infty$ in the previous expression, we find that

$$\nabla v = (\phi^{-1})'(u) \, \nabla u \qquad \text{in } L^{1+\gamma}(0, T; L^{1+\gamma}(\Omega)).$$

To conclude the proof, substitute *(iii)* and *(iv)* in equation (3.1) to obtain equation (3.3). $\quad \square$

Remark 3.1.2. As pointed out in equation (1.9) and (1.11) an immediate consequence of Corollary 3.1.1 is that if u is a nonnegative solution of problem (1.7) then it solves problem (1.8) in the sense of Definition 3.0.2.

Corollary 3.1.2. *Let v a weak solution of the initial/boundary value problem (1.7). Then for any $w \in L^{1+\gamma}(0, T; W_0^{1,(1+\gamma)}(\Omega))$*

$$\langle \phi(v)_t, w \rangle + \left(\frac{\nabla v}{|\nabla v|^{1-\gamma}}, \nabla w \right) = (f, w) \quad \text{*a.e. in*} \ [0, T].$$

Proof. Fix $w \in L^{1+\gamma}(0, T; W_0^{1,(1+\gamma)}(\Omega))$ and let $\{w_j\}$ be a basis for $W_0^{1,(1+\gamma)}(\Omega)$. Take a sequence $\{\psi_m\}$ of the form

$$\psi_m = \sum_{j=1}^m d_j^m(t) w_j \ \text{ with } \ d_j^m(t) \in L^\infty([0, T])$$

such that $\psi_m \to w$ strongly in $L^{1+\gamma}(0, T; W_0^{1,(1+\gamma)}(\Omega))$. This is possible by density of such finite sums in the mentioned space.

Since v is weak solution of problem (1.7) we get

$$\langle \phi(v)_t, \psi_m \rangle + \left(\frac{\nabla v}{|\nabla v|^{1-\gamma}}, \nabla \psi_m \right) = (f, \psi_m) \quad \text{a.e. in} \ [0, T].$$

Send $m \to +\infty$ to conclude. $\hfill \square$

3.2 Regularity

In this section we investigate basic regularity properties of solutions found in the existence Theorem 3.1.1. It is desirable to find more information on the time derivative of the function $\phi(v)$, in particular, it is worthwhile to find that it is a regular distribution.

Theorem 3.2.1. *Assume*

$$v_0 \in W_0^{1,1+\gamma}(\Omega), \quad and \quad f \in L^{(1+\eta)*}(0,T;L^{(1+\eta)*}(\Omega)).$$

Let v a solution of problem (1.7) constructed as in Theorem 3.1.1, then

(i) $v \in L^\infty(0,T;W_0^{1+\gamma}(\Omega))$,

(ii) v_t *exists as a regular distribution that lies in $L^{1+\eta}(0,T;L^{1+\eta}(\Omega))$*

with the estimate

$$\int_{\{|v|>0\}} \left(\phi'(v)^{1/2} v_t \right)^2 + \sup_{[0,T]} \|\nabla v(t)\|_{L^{1+\gamma}(\Omega)}^{1+\gamma} \leq$$

$$C\left(T,\ \|f\|_{L^{(1+\eta)*}(0,T;L^{(1+\eta)*}(\Omega))},\ \|v_0\|_{W_0^{1,1+\gamma}(\Omega)}\right). \quad (3.25)$$

Moreover, when ϕ is regular then $\phi(v)_t$ also lies in $L^{1+\eta}(0,T;L^{1+\eta}(\Omega))$ and

$$\phi(v)_t = \phi'(v)v_t. \quad (3.26)$$

Proof. Let $\{\phi_k\}_{k=1}^\infty$ be a sequence of regularized functions converging uniformly to $\phi(s) = s/|s|^{1-\eta}$ and let $v_k(t)$ be the solution associated each ϕ_k. Then

$$(\phi_k(v_k)_t, (v_k)_t) + \eta^\gamma \left(\frac{\nabla v_k}{|\nabla v_k|^{1-\gamma}}, \nabla(v_k)_t \right) = (f, (v_k)_t).$$

Hence,

$$\left\| \frac{\phi_k(v_k)_t}{\phi_k'(v_k)^{1/2}} \right\|_{L^2(\Omega)}^2 + \frac{\eta^\gamma}{1+\gamma} \frac{d}{dt} \|\nabla v_k\|_{L^{1+\gamma}(\Omega)}^{1+\gamma} = (f, (v_k)_t).$$

In addition, note that

$$(f, (v_k)_t) \leq 1/2 \left\| \frac{f}{\phi_k'(v_k)^{1/2}} \right\|_{L^2(\Omega)}^2 + 1/2 \left\| \frac{\phi_k(v_k)_t}{\phi_k'(v_k)^{1/2}} \right\|_{L^2(\Omega)}^2.$$

Thus, combining the last two relations we get

$$1/2 \left\| \frac{\phi_k(v_k)_t}{\phi_k'(v_k)^{1/2}} \right\|_{L^2(\Omega)}^2 + \frac{\eta^\gamma}{1+\gamma} \frac{d}{dt} \|\nabla v_k\|_{L^{1+\gamma}(\Omega)}^{1+\gamma} \leq 1/2 \left\| \frac{f}{\phi_k'(v_k)^{1/2}} \right\|_{L^2(\Omega)}^2. \tag{3.27}$$

Integrating (3.27) in time from 0 to T, we obtain that

$$1/2 \int_0^T \left\| \frac{\phi_k(v_k)_t}{\phi_k'(v_k)^{1/2}} \right\|_{L^2(\Omega)}^2 + \frac{\eta^\gamma}{1+\gamma} \sup_{[0,T]} \|\nabla v_k(t)\|_{L^{1+\gamma}(\Omega)}^{1+\gamma} \leq$$
$$1/2 \int_0^T \left\| \frac{f}{\phi_k'(v_k)^{1/2}} \right\|_{L^2(\Omega)}^2 + \|\nabla v_0\|_{L^{1+\gamma}(\Omega)}^{1+\gamma}. \tag{3.28}$$

By the hypothesis imposed on f, the right hand side of (3.28) converges to

$$1/2 \int_0^T \left\| \frac{f}{\phi'(v)^{1/2}} \right\|_{L^2(\Omega)}^2 + \|\nabla v_0\|_{L^{1+\gamma}(\Omega)}^{1+\gamma} \quad \text{as} \quad k \to \infty.$$

This immediately implies that the right hand side is bounded. Because of the nonlinearities that occur in the left hand side of (3.28), it is not straightforward to send $k \to \infty$ to establish estimate (3.25). For this purpose, we will first establish a weak convergence result for the sequence $\{(v_k)_t\}$ in the following way.

Observe that since $\phi(v_k)_t = \phi_k'(v_k)(v_k)_t$ then

$$\int_0^T \|(v_k)_t\|_{L^{1+\eta}(\Omega)}^{1+\eta} \leq \frac{1+\eta}{2} \int_0^T \left\| \frac{\phi_k(v_k)_t}{\phi_k'(v_k)^{1/2}} \right\|_{L^2(\Omega)}^2 + \frac{1-\eta}{2} \int_0^T \left\| 1/\phi_k'(v_k) \right\|_{L^q(\Omega)}^q \tag{3.29}$$

where $q = (1+\eta)/(1-\eta)$. Note that

$$\phi'(s) = \frac{\eta}{|s|^{1-\eta}}, \tag{3.30}$$

therefore

$$\frac{1}{\phi'(s)} = \frac{|\phi(s)|^{\frac{1-\eta}{\eta}}}{\eta} \quad \text{and} \quad \int_0^T \left\| 1/\phi_k'(v_k) \right\|_{L^q(\Omega)}^q = \frac{1}{\eta^q} \|\phi_k(v_k)\|_{L^{(1+\eta)^*}(0,T;L^{(1+\eta)^*}(\Omega))}^{(1+\eta)^*}. \tag{3.31}$$

Hence, as a consequence of (3.28), (3.29), and (3.31) the sequence $\{(v_k)_t\}$ is bounded in $L^{1+\eta}(0,T;L^{1+\eta}(\Omega))$. Thus, there exists a subsequence of $\{(v_k)_t\}$, labeled with the index μ such that

$$(v_\mu)_t \rightharpoonup v_t \quad \text{weakly in } L^{1+\eta}(0,T;L^{1+\eta}(\Omega)) \text{ as } \mu \to +\infty. \tag{3.32}$$

41

Second, define for all $\epsilon > 0$ and $m \geq 1$ the set

$$\Omega_{m,\epsilon} := \bigcap_{j \geq m}^{+\infty} \{[0,T] \times \Omega : |v_j| \geq \epsilon\} .$$

Thus,

$$\int_0^T \left\| \frac{\phi_\mu(v_\mu)_t}{\phi_\mu'(v_\mu)^{1/2}} \right\|_{L^2(\Omega)}^2 = \int_0^T \left\| \phi_\mu'(v_\mu)^{1/2} \, (v_\mu)_t \right\|_{L^2(\Omega)}^2$$

$$\geq \int_{\Omega_{m,\epsilon}} \left(\phi_\mu'(v_\mu)^{1/2} \, (v_\mu)_t \right)^2 . \tag{3.33}$$

Now, in $\Omega_{m,\epsilon}$ we have the bound $\phi_\mu'(v_\mu)^{1/2} \leq \eta \, \epsilon^{\eta-1}$ for $\mu \geq m$ and clearly,

$$\phi_\mu'(v_\mu)^{1/2} \to \phi'(v)^{1/2} \text{ a.e. in } \Omega_{m,\epsilon}.$$

Using this fact with (3.32) we obtain

$$\phi_\mu'(v_\mu)^{1/2} \, (v_\mu)_t \rightharpoonup \phi'(v)^{1/2} \, v_t \text{ weakly in } L^{1+\eta}(\Omega_{m,\epsilon}).$$

Therefore, taking $\liminf_{\mu \to +\infty}$ in (3.33) and using the weakly lower semicontinuity property of convex functionals on L^p it follows that

$$\int_{\Omega_{m,\epsilon}} \left(\phi'(v)^{1/2} \, v_t \right)^2 \leq \liminf_{\mu \to +\infty} \int_0^T \left\| \frac{\phi_\mu(v_\mu)_t}{\phi_\mu'(v_\mu)^{1/2}} \right\|_{L^2(\Omega)}^2 . \tag{3.34}$$

As $v_j \to v$ a.e. in $[0,T] \times \Omega$, it follows that

$$\lim_{m \to \infty, \epsilon \to 0} \Omega_{m,\epsilon} = \{|v| > 0\} .$$

Hence, taking these limits in (3.34) we obtain

$$\int_{\{|v|>0\}} \left(\phi'(v)^{1/2} v_t \right)^2 \leq \liminf_{\mu \to +\infty} \int_0^T \left\| \frac{\phi_\mu(v_\mu)_t}{\phi_\mu'(v_\mu)^{1/2}} \right\|_{L^2(\Omega)}^2 . \tag{3.35}$$

This takes care of the first term in (3.28). The second term of the left hand side is simpler to deal with. Note that by (3.28) there exist a subsequence of $\{v_k\}$, labeled again with the index μ, such that

$$v_\mu \rightharpoonup \xi \text{ in } L^\infty(0,T; W_0^{1,1+\gamma}(\Omega)) \text{ weak}^*.$$

Since the sequence already converged weakly in $L^{1+\gamma}(0,T;W_0^{1,1+\gamma}(\Omega))$ to v, we conclude that $\xi = v$. Therefore, we can take $\liminf_{\mu \to +\infty}$ in (3.28) to obtain

$$1/2 \int_{\{|v|>0\}} \left(\phi'(v)^{1/2} v_t\right)^2 + \frac{\eta^\gamma}{1+\gamma} \sup_{[0,T]} \|\nabla v(t)\|_{L^{1+\gamma}(\Omega)}^{1+\gamma}$$
$$\leq 1/2 \int_0^T \left\|\frac{f}{\phi'(v)^{1/2}}\right\|_{L^2(\Omega)}^2 + \|\nabla v_0\|_{L^{1+\gamma}(\Omega)}^{1+\gamma}. \quad (3.36)$$

To get estimate (3.25), observe that using the first expression in (3.31) one can prove, using Hölder's inequality, that

$$\int_0^T \left\|\frac{f}{\phi'(v)^{1/2}}\right\|_{L^2(\Omega)}^2 \leq \|f\|_{L^{(1+\eta)^*}(0,T;L^{(1+\eta)^*}(\Omega))}^2 \|\phi(v)\|_{L^{(1+\eta)^*}(0,T;L^{(1+\eta)^*}(\Omega))}^{\frac{1-\eta}{\eta}}$$

which together with estimate (3.9) prove *(i)*, *(ii)* and estimate (3.25). Finally when ϕ is regular, it is Lipschitz, then the chain rule formula in (3.26) follows by a standard result for Sobolev functions. $\square$

Corollary 3.2.1. *Assume the conditions of Theorem 3.2.1. Then for any regular ϕ,*

$$\phi(v)_t \in L^2(0,T;L^2(\Omega)),$$

and the following estimate holds

$$\|\phi(v)_t\|_{L^2(0,T;L^2(\Omega))}^2 \leq C\left(T,\ \phi'(0),\ \|f\|_{L^{(1+\eta)^*}(0,T;L^{(1+\eta)^*}(\Omega))},\ \|v_0\|_{W_0^{1,1+\gamma}(\Omega)}\right).$$

Proof. The conditions on any regular ϕ imply that for any $s \in \mathbb{R}$

$$1 \leq \frac{\phi'(0)}{\phi'(s)}$$

thus, after applying the chain rule (3.26) in Theorem 3.2.1, it follows that

$$\int_{\{|v|>0\}} \phi(v)_t^2 = \int_{\{|v|>0\}} \left(\phi'(v) v_t\right)^2$$
$$\leq \phi'(0) \int_{\{|v|>0\}} \left(\phi'(v)^{1/2} v_t\right)^2.$$

In addition, observe that in the set $\{v = 0\}$ we have that $\phi(v) = 0$. Hence, a direct calculation shows that $\phi(v)_t = 0$ in the interior of this set. But $\phi(v)_t$ is measurable,

43

therefore
$$\int_{\{v=0\}} \phi(v)_t^2 = 0.$$

Consequently,
$$\|\phi(v)_t\|_{L^2(0,T;L^2(\Omega))}^2 \le \phi'(0) \int_{\{|v|>0\}} \left(\phi'(v)^{1/2} v_t \right)^2.$$

Using estimate (3.25) in Theorem 3.2.1 we conclude the proof. $\square$

Remark 3.2.1. Corollary 3.2.1 shows that the solutions of problem (1.7) constructed as in Theorem 3.1.1 are *solutions fortes* in the sense of [6].

Theorem 3.2.2. *Assume v is a solution of problem (1.7) constructed as in Theorem 3.1.1, and additionally assume that*
$$v_0 \in L^\infty(\Omega) \qquad and \qquad f \in L^\infty(0,T;L^\infty(\Omega)),$$
then
$$\sup_{t\in[0,T]} \|v(t)\|_{L^\infty(\Omega)} \le C \left(\|v_0\|_{L^\infty(\Omega)}, \|f\|_{L^\infty(0,T;L^\infty(\Omega))}, T \right). \tag{3.37}$$

Proof. In order to find an L^∞ bound on v, we would like to uniformly control its L^p norms. For this purpose, the key idea would be to multiply equation (1.7) by the test function $v/|v|^{1-a}$ for any $a \ge 1$ and use Gronwall's Lemma to establish the result. However, for a fixed time t, the test function $v/|v|^{1-a}$ does not necessarily belong to $W_0^{1,1+\gamma}(\Omega)$, so that we need to regularize it. For this end, let us introduce the family $\{\rho_\delta(s)\}_{\delta>0}$ approximating the function $s/|s|^{1-a}$
$$\rho_\delta(s) = \frac{1}{(1+\delta|s|)^a} \frac{s}{|s|^{1-a}}.$$

Note that $\rho_\delta(v)(t) \in L^{1+\gamma}(0,T;W_0^{1,(1+\gamma)}(\Omega))$ since $\rho_\delta(s)$ is a $C^1([0,\infty))$ function with bounded derivative.

Using Corollary (3.1.2) we can chose $\rho_\delta(v)$ as a test function in equation (1.7). Observe that for any regular ϕ, the solution v has time derivative $v_t \in L^{1+\eta}(0,T;L^{1+\eta}(\Omega))$ by Theorem (3.2.1), whence the chain rules applies,
$$\phi(v)_t = \phi'(v)v_t.$$

Therefore, the following relation holds immediately
$$\frac{d}{dt}\|\Phi_\delta(v)(t)\|_{L^1(\Omega)} = \langle \phi(v)_t, \Phi_\delta(v) \rangle$$

44

where
$$\Phi_\delta(s) = \int_0^s \phi'(z)\rho_\delta(z).$$

Thus, we obtain

$$\frac{d}{dt}\|\Phi_\delta(v)(t)\|_{L^1(\Omega)} + \eta^\gamma(|\nabla v|^{1+\gamma}, \rho'_\delta(v)) = (f, \rho_\delta(v)). \tag{3.38}$$

The second term in the left hand side of (3.38) is nonnegative, thus the following inequality holds

$$\frac{d}{dt}\|\Phi_\delta(v)(t)\|_{L^1(\Omega)} \le \|f\|_{L^\infty(\Omega)}\|\rho_\delta(v)\|_{L^1(\Omega)}.$$

Using the fact that

$$|\rho_\delta(s)| \le 1 + \frac{\eta+a}{\eta}\Phi_\delta(s),$$

we obtain from the previous relation that

$$\frac{d}{dt}X_\delta(t) \le \|f(t)\|_{L^\infty(\Omega)}\left(|\Omega| + \frac{\eta+a}{\eta}X_\delta(t)\right),$$

where

$$X_\delta(t) = \|\Phi_\delta(v)(t)\|_{L^1(\Omega)}.$$

Using Gronwall's lemma we get

$$X_\delta(t) \le \exp\left(\frac{\eta+a}{\eta}\|f\|_{L^\infty(0,T;L^\infty(\Omega))}T\right)\{X_\delta(0) + \|f\|_{L^\infty(0,T;L^\infty(\Omega))}T\}. \tag{3.39}$$

Inequality (3.39) is valid for any ϕ_{reg}. Similarly, observe that

$$\Phi_\delta(v)(t) \longrightarrow \frac{\eta}{\eta+a}|v|^{\eta+a}(t) \text{ pointwise as } \delta \to 0 \text{ in } [0,T] \times \Omega.$$

Thus, sending $\delta \to 0$ in (3.39) and using Fatou's Lemma it follows that

$$\frac{\eta}{\eta+a}\|v(t)\|_{L^{\eta+a}(\Omega)}^{\eta+a} \le \exp\left(\frac{\eta+a}{\eta}\|f\|_{L^\infty(0,T;L^\infty(\Omega))}T\right)\left\{\frac{\eta}{\eta+a}\|v_0\|_{L^{\eta+a}(\Omega)}^{\eta+a} + |\Omega|\,\|f\|_{L^\infty(0,T;L^\infty(\Omega))}T\right\}. \tag{3.40}$$

Taking the $\eta + a$ root in (3.40) and letting $a \to \infty$ we find that for $0 \le t \le T$

$$\|v(t)\|_{L^\infty(\Omega)} \le \exp\left(\eta^{-1}\|f\|_{L^\infty(0,T;L^\infty(\Omega))}T\right)\max\left(1, \|v_0\|_{L^\infty(\Omega)}\right) \tag{3.41}$$

which proves the result for any regular ϕ. Next, take $\{\phi_k\}_{k=1}^\infty$ to be a sequence of regularized functions converging uniformly to $\phi(s) = s/|s|^{1-\eta}$. Let v_k be the solution associated to each

ϕ_k, then, as in the proof of existence,

$$v = \lim_{k \to \infty} v_k \qquad \text{pointwise in} \quad (0, T) \times \Omega$$

Thus, estimate (3.41) holds for v. This concludes the proof. $\qquad\square$

Corollary 3.2.2. *Assume u is a solution of problem (1.8) found as in Corollary 3.1.1, and additionally assume that*

$$u_0 \in L^\infty(\Omega) \qquad and \qquad f \in L^\infty(0, T; L^\infty(\Omega)),$$

then

$$\sup_{t \in [0,T]} \|u\|_{L^\infty(\Omega)} \leq C \left(\|u_0\|_{L^\infty(\Omega)}, \|f\|_{L^\infty(0,T;L^\infty(\Omega))}, T \right).$$

Proof. The solution for problem (1.8) found as in Corollary 3.1.1 is given by $u = \phi(v)$, thus by estimate (3.37) we have

$$\sup_{t \in [0,T]} \|\phi^{-1}(u)\|_{L^\infty(\Omega)} \leq C \left(\|\phi^{-1}(u_0)\|_{L^\infty(\Omega)}, \|f\|_{L^\infty(0,T;L^\infty(\Omega))}, T \right). \tag{3.42}$$

Since ϕ^{-1} is a monotonically increasing function, we have the property that

$$\sup_{t \in [0,T]} \|\phi^{-1}(u)\|_{L^\infty(\Omega)} = \phi^{-1} \left(\sup_{t \in [0,T]} \|u\|_{L^\infty(\Omega)} \right).$$

Substituting the previous fact and applying ϕ on both sides of (3.42) one obtains

$$
\begin{aligned}
\sup_{t \in [0,T]} \|u\|_{L^\infty(\Omega)} &\leq \phi \left(C \left(\phi^{-1} \left(\|(u_0)\|_{L^\infty(\Omega)} \right), \|f\|_{L^\infty(0,T;L^\infty(\Omega))}, T \right) \right), \\
&\leq C \left(\|(u_0)\|_{L^\infty(\Omega)}, \|f\|_{L^\infty(0,T;L^\infty(\Omega))}, T \right)
\end{aligned}
$$

which finishes the proof. $\qquad\square$

3.3 Comparison result, uniqueness and nonnegativity

Generally speaking, if v is a weak solution of problem (1.7) some basic regularity on $\phi(v)_t$ must be obtained for pursuing a uniqueness result, otherwise this task can be very complex. Moreover, uniqueness may not be true. In Theorem (3.3.1) we will prove a comparison

result due to Bamberger [6], that will lead to a uniqueness result under the assumption that

$$\phi(u)_t \in L^1(0, T; L^1(\Omega)). \tag{3.43}$$

In a hydrologic context, the previous assumption can be interpreted in the following way. Condition (3.43) implies that $u_t \in L^1(0, T; L^1(\Omega))$ in problem (1.8). Hence

$$u \in C(0, T; L^1(\Omega)) \subseteq W^{1,1}(0, T; L^1(\Omega)).$$

Recall that when u is nonnegative, u represents the free water surface elevation, or the column of water at a given point in the domain Ω in a physical system. Thus the volume $\mathcal{V}$ of water in Ω may be represented as

$$\mathcal{V}(\Omega, t) = \int_\Omega u(t).$$

Condition (3.43) implies that the the volume in the domain Ω changes continuously in time. This is a natural condition when modeling hydrologic systems. The fact that the volume is a time continuous function follows when integrating expression *(ii)* of Theorem A.0.1 to obtain

$$\mathcal{V}(t_1) - \mathcal{V}(t_0) = \int_{t_0}^{t_1} \mathcal{V}_t(t),$$

where

$$\mathcal{V}_t(\Omega, t) = \int_\Omega u_t(t) \quad \in L^1(0, T).$$

Therefore $\mathcal{V}$ is an absolutely continuous function in $[0, T]$.

In the current section we will use the standard notation f^+ and f^- to denote the positive and negative part of the function f respectively.

Theorem 3.3.1 (Bamberger[1]). *Assume u and v are weak solutions of problem (1.7) associated to the initial data u_0 and v_0, and the forcing terms f and g respectively. Assume the additional property that*

$$\phi(u)_t, \ \phi(v)_t \in L^1(0, T; L^1(\Omega)) \quad and \quad f - g \in L^1(0, T; L^1(\Omega)), \tag{3.44}$$

then

$$\int_\Omega \lambda(\phi(u) - \phi(v)) \le \int_\Omega \lambda(\phi(u_0) - \phi(v_0)) + \int_0^t \int_\Omega \lambda(f - g), \tag{3.45}$$

---

[1]See [6]

where $\lambda(s)$ is any of the following three functions, $|s|$, s^+, or s^-.

Proof. Since u and v are weak solutions of problem (1.7) then

$$\langle \phi(u)_t - \phi(v)_t, w \rangle + \left(\frac{\nabla u}{|\nabla u|^{1-\gamma}} - \frac{\nabla v}{|\nabla v|^{1-\gamma}}, \nabla w \right) = (f - g, w) \qquad (3.46)$$

for any $w \in W_0^{1,(1+\gamma)}(\Omega))$. Let $\{\beta_\delta(s)\}_{\delta>0}$ be the family of $C^1(\mathbb{R})$ increasing functions such that,

(i) $|\beta_\delta(s)| \leq 1$, and

(ii) $\beta_\delta(s) \longrightarrow \lambda'(s)$ as $\delta \to \infty$.

Substituting $w = \beta_\delta(u - v)$ in (3.46) we find that

$$\langle \phi(u)_t - \phi(v)_t, \beta_\delta(u - v) \rangle + \left(\frac{\nabla u}{|\nabla u|^{1-\gamma}} - \frac{\nabla v}{|\nabla v|^{1-\gamma}}, \beta_\delta'(u - v)\nabla(u - v) \right) = (f - g, \beta_\delta(u - v)).$$

Since $\beta_\delta'(u - v) \geq 0$, by Lemma A.0.1 in Appendix A, the second term in the previous expression is nonnegative, thus

$$\int_0^t \langle \phi(u)_t - \phi(v)_t, \beta_\delta(u - v) \rangle \leq \int_0^t (f - g, \beta_\delta(u - v)).$$

Note that $\{\beta_\delta(u - v)\} \subset L^\infty(0, T; L^\infty(\Omega))$. But $\phi(u)_t$ and $\phi(v)_t$ lie in $L^\infty(0, T; L^\infty(\Omega))^*$ by assumption, thus

$$\langle \phi(u)_t - \phi(v)_t, \beta_\delta(u - v) \rangle = (\phi(u)_t - \phi(v)_t, \beta_\delta(u - v)).$$

Using Lebesgue's Dominated Convergence Theorem we can take the limit as $\delta \to \infty$ in the above inequality to find that

$$\int_0^t (\phi(u)_t - \phi(v)_t, \lambda'(u - v)) \leq \int_0^t \int_\Omega \lambda(f - g).$$

Observe that since $\lambda'(u - v) = \lambda'(\phi(u) - \phi(v))$, then for $0 \leq t \leq T$,

$$
\begin{aligned}
\int_0^t (\phi(u)_t - \phi(v)_t, \lambda'(u - v)) &= \int_0^t ((\phi(u) - \phi(v))_t, \lambda'(\phi(u) - \phi(v))) \\
&= \int_0^t \frac{d}{dt} \int_\Omega \lambda(\phi(u) - \phi(v)) \qquad (3.47) \\
&= \int_\Omega \lambda(\phi(u)(t) - \phi(v)(t)) - \int_\Omega \lambda(\phi(u_0) - \phi(v_0)),
\end{aligned}
$$

48

from which (3.45) follows. ▢

Remark 3.3.1. By hypothesis $\phi(v) \in C([0,T]; L^1(\Omega))$ since $\phi(v) \in W^{1,1}([0,T]; L^1(\Omega))$. See Theorem A.0.1 in Appendix A. Thus, the last step in (3.47) can be safely performed.

Remark 3.3.2. Note that if ϕ is regular we know from Corollary (3.2.1) that solutions of problem (1.7) constructed as in Theorem 3.1.1 satisfy

$$\phi(u)_t, \ \phi(v)_t \in L^2(0,T; L^2(\Omega)) \subset L^1(0,T; L^1(\Omega)).$$

Hence, the previous result applies for them.

Corollary 3.3.1 (Uniqueness). *Assume u and v are weak solutions of problem (1.7) satisfying*

$$\phi(u)_t, \ \phi(v)_t \in L^1(0,T; L^1(\Omega)),$$

then $u = v$.

Proof. Use Theorem 3.3.1 with $\lambda(s) = |s|$, $u_0 = v_0$ and $f = g$. ▢

Corollary 3.3.2. *Assume u and v are weak solutions of problem (1.7) associated to the initial data u_0 and v_0, and the forcing terms f and g respectively. Assume also*

$$\phi(u)_t, \ \phi(v)_t \in L^1(0,T; L^1(\Omega)) \quad and \quad f - g \in L^1(0,T; L^1(\Omega)), \tag{3.48}$$

Additionally assume that

$$
\begin{aligned}
v_0 &\leq u_0 \quad a.e.\ in\ \Omega, \\
g &\leq f \quad a.e.\ in\ (0,T) \times \Omega
\end{aligned}
$$

$$\tag{3.49}$$

then $v \leq u$ a.e. in $(0,T) \times \Omega$.

Proof. Use Theorem 3.3.1 with $\lambda(s) = s^-$ to deduce that

$$\int_\Omega (\phi(u) - \phi(v))^- \leq 0,$$

thus, $\phi(u) - \phi(v) \geq 0$ *a.e. in $(0,T) \times \Omega$. Since $\phi(s)$ is strictly increasing the result of the corollary follows.* ▢

Remark 3.3.3 (Nonnegativity for ϕ_{reg} and ϕ). Note that any solution u of problem (1.7) associated to a ϕ_{reg}, with $u_0 \geq 0$ and $f \geq 0$, is unique and nonnegative. The previous observations are consequences of Corollary 3.3.1 and Corollary 3.3.2, and the fact that $\phi_{reg}(u)_t \in L^1(0, T; L^1(\Omega))$. Furthermore, the solution constructed in *Step 5* of the proof of existence will be nonnegative as well since it is a pointwise limit of solutions associated to regularized problems.

Corollary 3.3.3. *Assume u and v are weak solutions of problem (1.8) found as in Corollary 3.1.1, associated to the initial data u_0 and v_0, and the forcing terms f and g respectively. Assume the additional property that*

$$u_t, \ v_t \in L^1(0, T; L^1(\Omega)) \quad and \quad f - g \in L^1(0, T; L^1(\Omega)), \tag{3.50}$$

then

$$\int_\Omega \lambda(u - v) \leq \int_\Omega \lambda(u_0 - v_0) + \int_0^t \int_\Omega \lambda(f - g), \tag{3.51}$$

where $\lambda(s)$ is any of the following three functions, $|s|$, s^+, or s^-.

The proof of Corollary 3.3.3 is an immediate consequence of Theorem 3.3.1 and it is an equivalent comparison result for problem (1.8). From this corollary, one obtains equivalent uniqueness and nonnegativity results for problem (1.8).

3.4 Open problem: topographic effects

To the best of our knowledge, existence, uniqueness, and regularity of solutions of the DSW equation in its general form (2.1), *i.e.* when *topographic* effects are considered, have not been studied. Observe that when one formally carries out the spatial differentiation inside the divergence term in the first equation,

$$\frac{\partial u}{\partial t} - h_1(u, z) \nabla(u - z) \cdot \nabla u - h_2(u, z) \nabla \cdot \left(\frac{\nabla u}{|\nabla u|^{1 - \gamma}} \right) = f,$$

where $h_1(u, z) = \alpha(u - z)^{\alpha - 1} / |\nabla u|^{1 - \gamma}$ and $h_2(u, z) = (u - z)^\alpha$, one can see the appearance of a nonlinear advection term, and a nonlinear diffusive term involving the bathymetry. The topographic effects change qualitatively the direction of the advection $\nabla(u - z)$, and scale both the advection and the diffusion terms. Some of the difficulties that arise when one introduces a non flat bathymetry z are:

- The aforementioned techniques to prove existence of solutions (used when $z = 0$) fail, since one cannot send the nonlinearity $(u - z)^\alpha$ to the time derivative term. In other

50

words the change of variables described at the beginning of section 1.2.2 does not work correctly. This situation introduces further difficulties when trying to prove the validity of the Galerkin method as a suitable way to obtain approximate solutions.

- In general one expects the regularity of solutions of problem (2.1) to depend on the properties of z. Technically speaking, it is not clear how to proceed in order to incorporate such properties in the analysis and relate them directly with the properties of u.

- Presumably, in order to prove uniqueness of solutions for problem (2.1) one may need to impose an entropy condition as described in [21]. This condition may provide means to identify unique physically consistent solutions.

Chapter 4

A Continuous Galerkin Approach

This Chapter may be thought of as the numerical analysis counterpart of Chapter 3. The main contribution presented here is the implementation of a methodology to numerically approximate the solution of the DSW equation using the continuous Galerkin finite element method, in the context of shallow water modeling. This is done by means of the regularization techniques proposed in Chapter 3. Most of the ideas in this Chapter were submitted for publication and can be found in [58].

The outline of this Chapter is the following. we begin by highlighting the main ideas of the Chapter in section 4.1. In sections 4.2, we state the regularized version of problem (2.1) to be numerically approximated and present some auxiliary results that will be key in proving the error estimates in the subsequent sections. In section 4.3, the Galerkin finite element method is introduced and studied. In sections 4.3.3 and 4.3.4, we obtain *a priori* error estimates between the regularized solution of the DSW and semi discrete and fully discrete solutions constructed using the Galerkin finite element method. Special conditions, based on the regularity of solutions of the DSW equation and physical consistency are required to establish the aforementioned results. In section 4.4, we present numerical experiments that provide relevant information about the numerical accuracy of the method and the applicability of the DSW equation as a model to simulate observed quantities in a real life experimental setting. Furthermore, we investigate numerically the qualitative behavior of solution to problem (2.1), including *topographic effects*, and in particular, we investigate whether some of the qualitative properties studied by Esteban and Vázquez in [35] for the solutions of equation (1.12) persist in the more general case for a nonzero and regular topography z in 1-D. Rainfall is considered in this Chapter, however, neither evaporation nor infiltration are investigated.

4.1 Preliminaries

For intuition purposes, we will briefly recapitulate the ideas that were introduced in Chapter 3 that will be utilized in this Chapter. We will also discuss some of the characteristics of the kind of solutions of the DSW equation for which our numerical scheme will succeed to approximate.

The main idea of Chapter 3 was to study an alternative formulation of the IBVP (2.1), namely the IBVP (1.7). It was shown that for nonnegative solutions and $z = 0$, one can, roughly speaking, use the change of variables $u = v^{1/m}$ to transform the DSW into:

$$\frac{\partial v^{1/m}}{\partial t} - m^{-\gamma} \nabla \cdot \left(\frac{\nabla v}{|\nabla v|^{1-\gamma}} \right) = f \tag{4.1}$$

where $0 < 1/m = \gamma/(\alpha + \gamma) < 1$. This change of variables allowed sending the nonlinearity u^α, inside the divergence term in the DSW equation, to the time derivative term in (4.1). This, in turn, moved the difficulty of dealing with the possible degeneracy of the diffusion coefficient (1.1) when $u = 0$, to dealing with differentiability issues of the function $v^{1/m}$ at $v = 0$. As described in section 3.0.2, the regularization technique used in the proof of existence of solutions for problem (1.7) makes use of a sequence of regularized functions $\{\phi_\epsilon(v_\epsilon)\}$ converging uniformly to $\phi(v) = v^{1/m}$ with the property that $\phi'_\epsilon(0) < +\infty$ in order to construct approximate solutions to (4.1). If one transforms back (4.1), using the change of variables $u^m = v$, into

$$\frac{\partial u}{\partial t} - m^{-\gamma} \nabla \cdot \left(\left((\phi^{-1})'(u)\right)^\gamma \frac{\nabla u}{|\nabla u|^{1-\gamma}} \right) = f , \tag{4.2}$$

then, this *regularized function strategy* suggests naturally the use of a sequence of non-degenerate diffusion coefficients a_ϵ, with the property that $0 < \epsilon \leq a_\epsilon(u, \nabla u)$, and given by

$$a_\epsilon(u, \nabla u) = \frac{(\phi_\epsilon^{-1})'(u)}{|\nabla u|^{1-\gamma}}, \tag{4.3}$$

in order to approximate the solutions of problem (2.1) when $z = 0$ and for small ϵ. This is the case since the property that $\phi'_\epsilon(0) < +\infty$ implies directly that $(\phi_\epsilon^{-1})'(0) > 0$. Recall Figure 3.1 in section 3.0.2. When proceeding this way, one then needs to show that the solution of these alternative non-degenerate problems are indeed close, in some sense, to

the original (possibly degenerate) problem (2.1). This was done indirectly in *Step 5* of the proof of Theorem 3.1.1 in Chapter 3 for the case when $z = 0$ and for nonnegative solutions. See also Corollary 3.1.1. We will use this approach in the analysis of the numerical scheme presented in this Chapter even in the case when $z \neq 0$, *i.e.*, even when topographic effects are incorporated, without further justification.

Numerically speaking, we will follow a similar strategy to the one presented in [53], where Nochetto and Verdi construct a numerical scheme that approximates a regularized problem obtained by replacing m in (1.13) by a smooth function m_ϵ with maximal slope equal to $1/\epsilon$, for some regularization parameter $\epsilon > 0$. Then they discretize this regularized problem in space and time to compute the regularized numerical approximation U_ϵ^h. Finally, roughly speaking, in order to obtain global error estimates between the solution $u(t)$ of (1.13) and the regularized numerical approximation $U_\epsilon^h(t)$, they obtain bounds for the quantities $\|u(t) - u_\epsilon(t)\|_{L^\infty}$ and $\|u_\epsilon(t) - U_\epsilon^h(t)\|_{L^\infty}$ for $0 \leq t \leq T$, where $u_\epsilon(t)$ is the true solution of the regularized problem (solving (1.13) with m_ϵ instead of m). From the two L^∞-estimates they can obtain a global estimate of $\|u(t) - U_\epsilon^h(t)\|_{L^\infty}$ using the triangle inequality. Even though our strategy is similar, the analysis we present will not be as complete. We will find bounds similarly, for the quantity $\|u_\epsilon(t) - U_\epsilon^h(t)\|_{L^2}$ associated to the nondegenerate problem (when $u - z > \epsilon > 0$ throughout the domain), however, estimating the difference between the solution u to the IBVP (2.1) and u_ϵ, the solutions to the nondegenerate problems is not yet completely understood in the general setting when topographic effects are considered ($z \neq 0$). This is so, since an appropriate proof of uniqueness of solutions to problem (2.1) has not been developed yet. On a positive note, from the modeling point of view, our analysis is relevant for all situation when there is at least a thin layer of water of depth ϵ in all regions of the domain. Note that this condition is physically relevant and in most cases, one can model even completely dry regions of the domain associating a thin depth of order ϵ to such regions.

It is interesting to mention that, according to Bamberger [6], when $z = 0$, a sufficient condition for uniqueness of solutions to the IBVP (1.7) is that $u_t \in L^1(0, T; L^1(\Omega))$, see Theorem 3.3.1. This condition implies, since $u \in L^\infty(0, T; L^\infty(\Omega))$, that $u \in C^0(0, T; L^1(\Omega))$. If one identifies u (or H as described in section 2.2) with the free water surface elevation in a hydrological context, this condition implies that there will be a unique solution if the volume of water in the domain $\int_\Omega u$ changes continuously in time. This is a natural and physically-consistent condition when modeling hydrologic systems.

To conclude this section, we will describe the assumptions required on the *true* weak solutions of problem 4.4 (and thus implicitly on solutions of problem (2.1)) in order for the analysis presented in sections 4.3.3 and 4.3.4 to apply. These assumptions are based on two criteria, the first one being the physical relevance of solutions to the DSW in the context of shallow water models, and the second one coming from the analytical results shown in Chapter 3. Regarding the first criterion, the error analysis will be restricted to approximating the set of nondegenerate solutions to problem (4.4) (and thus (2.1)) whose gradients are bounded. Despite the fact that these solutions may seem very particular, they play an important role in physical applications. See Remark 4.1.1 below and Remark 2.2.1. The second criterion, consisting of assuming that $u \in C^0(0,T;L^1(\Omega))$, will ensure uniqueness of solutions from a purely mathematical point of view, as described in the previous paragraph.

Remark 4.1.1. It is important to mention that the condition that ∇u be bounded in the L^∞ sense does not necessarily hold for all solutions of (2.1). In fact, even in the particular case when (2.1) becomes the PME ($z = 0$ and $\gamma = 1$) in two or higher dimensions, there exists a class of solutions called *focusing solutions* that exhibit no local regularity on the gradient in subsets of Ω. See Chapter 19 in [62] . Even though we ignore this class of solutions in our analysis by assuming boundedness of ∇u, it is justified to do so in the context of shallow water modeling where for small beds z, one expects *small* values of ∇z and thus *small* values of ∇u as explained in section 2.2.

Remark 4.1.2. Note that the natural norm induced by multiplying the DSW equation by u and integrating by parts, in the nondegenerate case, is the $W^{1,1+\gamma}(\Omega)$ norm. Indeed

$$\int_\Omega \frac{\partial u}{\partial t} u - \int_\Omega \left(\frac{(u-z)^\alpha}{|\nabla u|^{1-\gamma}} \nabla u \right) \cdot \nabla u = \int_\Omega fu$$

implies that for a sufficiently regular u and say zero Newmann boundary conditions

$$\frac{1}{2}\frac{\partial}{\partial t}\|u\|_{L^2(\Omega)} + \int_\Omega (u-z)^\alpha |\nabla u|^{1+\gamma} = \int_\Omega fu.$$

After some manipulations on the previous expression along with assumptions on the nondegeneracy $u - z > \epsilon > 0$, and $u_0, f \in L^2(\Omega)$ for all $t \in [0,T]$ one obtains the analytic stability result,

$$\|u\|_{L^2(\Omega)} \leq C \left(\|u_0\|_{L^2(\Omega)}, \|f\|_{L^2(\Omega)} \right)$$

and

$$\int_0^T \|\nabla u\|_{L^{1+\gamma}(\Omega)} \leq C \left(\|u_0\|_{L^1(0,T,L^2(\Omega))}, \|f\|_{L^1(0,T,L^2(\Omega))} \right)$$

55

In our error analysis, we obtain stability and *a priori* error estimates for the approximations of u, and ∇u, namely U and ∇U, in the L^2-norm, by assuming the appropriate regularity on u, nondegeneracy and the physical-consistency assumption on the uniform boundedness of ∇u_ϵ. Working with a norm that is not naturally induced by the problem has advantages and disadvantages. The advantages are that the arguments used to prove estimates in our study extend naturally from the classical arguments developed by Wheeler [67], and Douglas and Dupont [32] for nonlinear parabolic problems. The disadvantage is that the error bounds may be too conservative. The previous statement is supported by the numerical findings on the performance of our method presented in Section 4.4 which show higher convergence rates than those ensured by our analysis.

4.2   Regularized problem

In this section we present the nondegenerate problem we will approximate numerically in the rest of the Chapter. We begin by introducing the nondegenerate version of the IBVP (2.1) obtained by replacing the function $(s - z)^\alpha$ with a sequence of bounded Lipschitz functions $\{\beta_\epsilon(s)\}$, with the properties that *(i)* $\{\beta_\epsilon(s)\}$ converges uniformly to $(s - z)^\alpha$ as $\epsilon \to 0$, and *(ii)* for small $\epsilon > 0$ the following holds $\beta_\epsilon(s) \geq \epsilon$ for all $t \in [0, T]$. To this end, the bathymetry $z(x)$ will be assumed to be a smooth and bounded time independent function defined in Ω. The nondegenerate IBVP is given by

$$\begin{cases} \dfrac{\partial u}{\partial t} - \nabla \cdot \left(\beta_\epsilon(u) \dfrac{\nabla u}{|\nabla u|^{1-\gamma}} \right) = f & \text{on } \Omega \times (0, T] \\[2mm] \qquad\qquad\qquad u = u_0 & \text{on } \Omega \times \{t = 0\} \\[2mm] \left(\dfrac{\beta_\epsilon(u)}{|\nabla u|^{1-\gamma}} \nabla u \right) \cdot n = B_N & \text{on } \partial\Omega \cap \Gamma_N \times (0, T] \\[2mm] \qquad\qquad\qquad u = B_D & \text{on } \partial\Omega \cap \Gamma_D \times (0, T]. \end{cases} \tag{4.4}$$

In the next section we develop a numerical scheme to approximate this nondegenerate problem as explained in section 4.1. The fact that solutions to the nondegenerate problem (4.4) are close to the original solution to problem (2.1) as $\epsilon \to 0$ will be understood as in Chapter 3 for $z = 0$ and will be assumed for the general case $z \neq 0$.

Remark 4.2.1. In (4.4), for each ϵ and thus for each $\beta_\epsilon(u)$, one has a solution u_ϵ.

Remark 4.2.2. Frequently, in the actual computational code one does not need to implement the sequence of $\{\beta_\epsilon(u)\}$, however it becomes crucial to use this sequence if one wants to find error estimates. For intuition purposes one could choose for example the following sequence $\beta_\epsilon(u) = (u - z)^\alpha + \epsilon$.

4.2.1 Auxiliary results

We will present some results that will be useful in the subsequent sections.

Lemma 4.2.1. *Let u_1 and u_2 be non negative $L^\infty(\Omega)$ functions, then for $\alpha \geq 1$*

$$|u_1^\alpha - u_2^\alpha| \leq \alpha \left(max(\|u_1\|_{L^\infty(\Omega)}, \|u_2\|_{L^\infty(\Omega)}) \right)^{\alpha-1} |u_1 - u_2| \tag{4.5}$$

Proof. We can express

$$
\begin{aligned}
|u_1^\alpha - u_2^\alpha| &= \left| \int_0^1 \frac{d}{d\tau}(\tau u_1 + (1-\tau)u_2)^\alpha d\tau \right| \\
&\leq \alpha |u_1 - u_2| \int_0^1 (\tau u_1 + (1-\tau)u_2)^{\alpha-1} d\tau \\
&\leq \alpha |u_1 - u_2| \int_0^1 (\tau \|u_1\|_{L^\infty(\Omega)} + (1-\tau)\|u_2\|_{L^\infty(\Omega)})^{\alpha-1} d\tau \\
&\leq \alpha |u_1 - u_2| \left(max(\|u_1\|_{L^\infty(\Omega)}, \|u_2\|_{L^\infty(\Omega)}) \right)^{\alpha-1} \qquad \square
\end{aligned}
$$

Lemma 4.2.2 (Coercivity and continuity)**.** *Let u_1 and u_2 be $L^\infty(\Omega)$ positive functions with the property that $\nabla u_1, \nabla u_2 \in L^\infty(\Omega)$ then the following estimates hold true*

$$\gamma \mathcal{A}_0 |\nabla u_1 - \nabla u_2|^2 \leq \left(\frac{\nabla u_1}{|\nabla u_1|^{1-\gamma}} - \frac{\nabla u_2}{|\nabla u_2|^{1-\gamma}} \right) (\nabla u_1 - \nabla u_2) \tag{4.6}$$

and

$$\left| \frac{\nabla u_1}{|\nabla u_1|^{1-\gamma}} - \frac{\nabla u_2}{|\nabla u_2|^{1-\gamma}} \right| \leq \mathcal{A}_0 |\nabla u_1 - \nabla u_2| \leq \frac{2}{\gamma} |\nabla u_1 - \nabla u_2|^\gamma \tag{4.7}$$

where

$$\mathcal{A}_0 := \int_0^1 |\lambda \nabla u_1 + (1-\lambda) \nabla u_2|^{\gamma-1} d\lambda$$

Proof. See [31], pp. 348-350. $\qquad \square$

See also [9] and the references therein for a more general result.

4.3 The Continuous Galerkin Method

In this section we will use the Galerkin method in order to numerically approximate the solution of the initial/boundary value problem (4.4). We will provide *a priori* error estimates between the true solution of (4.4), u_ϵ, and the Galerkin approximate solutions U_ϵ^h, both in the semidiscrete and fully discrete cases for the zero Dirichlet and Newmann boundary conditions. The analysis will be an extension of the techniques presented in [67] and [32]

and holds true for any sequence of Lipschitz functions, $\beta_\epsilon(u)$, with the properties *(i)* and *(ii)* described in section 4.2. The (unique) solution u_ϵ to the nondegenerate problem (4.4) and the Galerkin approximate solution U_ϵ^h will be denoted with u and U, respectively, in the following paragraphs. As discussed in section 4.1, and based on Corollaries 3.1.1 and 3.2.2, the following assumptions will be made for the general case when $z \neq 0$, :

- Solutions to the nondegenerate problem (4.4) are close to the original solution of problem in some sense as $\epsilon \to 0$. See *Step 5* of the proof of Theorem 3.1.1.

- $u \in L^\infty(0, T; L^\infty(\Omega))$. See Corollary 3.2.2.

- $\nabla u \in L^\infty(0, T; L^\infty(\Omega))$

The latter assumption restricts our analysis to physically consistent solutions in the context of shallow water modeling. Our numerical analysis is carried out for piecewise polynomial basis functions of order k. However, the limited regularity of solutions of the DSW calls in general for lower order approximation spaces.

4.3.1 The semi-discrete case

In the Galerkin method, we seek a differentiable function $U(\cdot, t) \in \mathcal{M}$, a finite dimensional subspace of $H^1(\Omega)$ if the boundary conditions in problem (4.4) are of Newman-type, or $H_0^1(\Omega)$) if they are of Dirichlet-type, such that it satisfies the following weak form

$$
\begin{cases}
\left(\dfrac{\partial U}{\partial t}, v \right) + \left(\beta_\epsilon(U) \dfrac{\nabla U}{|\nabla U|^{1-\gamma}}, \nabla v \right) = (f, v) & t > 0, \ \forall v \in \mathcal{M}, \\
\text{and} \hspace{5.5cm} (U(\cdot, 0), v) = (u_0, v) & t = 0, \ \forall v \in \mathcal{M},
\end{cases}
\tag{4.8}
$$

where $\mathcal{M}$ denotes the span $\{v_i\}_{i=1}^M$, and $v_1, ..., v_M$ are linearly independent functions in $H^1(\Omega)$. By construction, we can represent any function in $\mathcal{M}$ as a linear combination of the family $\{v_i\}$, thus, in particular

$$
U(x, t) = \sum_{i=1}^M \zeta_i(t) v_i(x).
\tag{4.9}
$$

Substituting (4.9) in (4.8) we observe that the semidiscrete problem can be stated: Find coefficients $\zeta_i(t)$ in (4.9) such that

$$
\sum_{i=1}^M \zeta_i'(t)(v_i, v_j) + \sum_{i=1}^M \zeta_i(t) \left(\beta_\epsilon^*(\zeta) \nabla v_i, \nabla v_j \right) = (f, v_j) \quad \text{for } j = 1, ..., M.
\tag{4.10}
$$

with $\sum_{i=1}^{M} \zeta_i(0)(v_i, v_j) = (u_0, v_j)$, and

$$\beta_\epsilon^*(\zeta) := \beta_\epsilon \left(\sum_{i=1}^{M} \zeta_i(t)v_i(x) \right) \left(\sum_{i=1}^{M} [\zeta_i(t)\nabla v_i(x)]^2 \right)^{\frac{\gamma-1}{2}}. \tag{4.11}$$

Equivalently, we can express the previous problem as the initial value problem for the system of nonlinear ordinary differential equations given by

$$\begin{cases} G\zeta'(t) & = -B(\zeta)\zeta + F, \\ G\zeta(0) & = b, \end{cases} \tag{4.12}$$

where the entries of the matrix $G = (G_{ij})$ are given by $G_{ij} = (v_i, v_j)$, the entries of the matrix $B(\zeta) = (B_{ij}(\zeta))$ are given by

$$B_{ij}(\zeta) = (\beta_\epsilon^*(\zeta)\nabla v_i, \nabla v_j), \tag{4.13}$$

the components of the vectors $b = (b_j)$ and $F = (F_j)$ are given by $b_j = (u_0, v_j)$ and $F_j = (f, v_j)$ respectively, and the vector of unknowns is $\zeta(t) = (\zeta_j(t))$.

Whenever $U(x, t)$ given by (4.9) exists, it is called the *continuous in time Galerkin approximation* or *semi-discrete approximation* to the weak solution of problem (4.4). Though this approximation is never computed in practice, it is easy to understand and gives us insight in our method development. For computational purposes, the variable t is discretized and a fully algebraic nonlinear system of equations needs to be solved at each time step, in order to obtain a fully discrete approximation to the solution. This is studied in section 4.3.4.

4.3.2   Existence of the continuous in time Galerkin approximation

Theorem 4.3.1. *There exists at least one solution to problem (4.12).*

Proof. Since the family $\{v_i\}$ is linearly independent, the mass matrix G is a Gram matrix, and thus in particular it is a positive definite and invertible matrix. Hence, problem (4.12) can be written as:

$$\begin{cases} \zeta'(t) & = -G^{-1}B(\zeta)\zeta + G^{-1}F, \\ \zeta(0) & = G^{-1}b. \end{cases} \tag{4.14}$$

Note that the mapping $\Theta(\zeta) : \mathbb{R}^M \to \mathbb{R}^M$ defined by $\Theta(\zeta) = -G^{-1}B(\zeta)\zeta + G^{-1}F$ is γ-Hölder continuous and thus, by *Peano's* existence theorem for ordinary differential equations (See

59

[43] pp. 10), there exists at least one solution to problem (4.14). Which immediately implies the statement of the Theorem. $\square$

Remark 4.3.1. The fact that $\Theta(\zeta)$ is γ-Hölder continuous is a consequence of $B(\zeta)\zeta$ being γ-Hölder continuous itself. This can be easily verified recalling definition (4.13). Indeed, if we let $Y = \nabla U$, the diffusion operator $\beta_\epsilon(U)Y|Y|^{\gamma-1}$ is Lipschitz continuous with respect to U and γ-Hölder continuous w.r.t. Y. Since U and Y depend linearly on ζ, the diffusion operator will be γ-Hölder continuous w.r.t. ζ, and thus the discrete diffusion operator defined by (4.13) will inherit this property. See Lemma A.0.2 in Appendix A.

4.3.3 Continuous in time *a priori* error estimate

In this section we will study how close (possibly nonunique) solutions to the continuous in time Galerkin approximation problem (4.12), U, are to the true weak solution u of problem (4.4). We focus the analysis to the case when $\nabla u \in L^\infty(0, T, L^\infty(\Omega))$. In the following paragraphs we will assume that there exists a function $\hat{u}$ called the *interpolant* of u, provided u belongs to some Banach space with certain regularity. The interpolant could be for example the L^2 projection as defined in Lemma 2.4.1. Further assumptions include that $\beta_\epsilon(\hat{u}), \beta_\epsilon(U), \nabla \hat{u}, \nabla U \in L^\infty(0, T, L^\infty(\Omega))$. See Remark 4.3.2 below. The error between solutions of problem (4.12) and the solution of problem (4.4), $\|u - U\|_{L^2}$, will be shown to be bounded by terms that only involve approximation errors between the interpolant and the true solution of problem (4.4), $\|u - \hat{u}\|_{L^2}$. Thus, reducing the global problem to well known results in interpolation theory in Hilbert spaces, such as those presented in section 2.4. To simplify the notation in the analysis we will denote with $\beta(x)$ any element of the sequence $\{\beta_\epsilon(x)\}$ as described in section 4.2.

Remark 4.3.2. The fact that $\beta_\epsilon(\hat{u}), \nabla \hat{u} \in L^\infty(0, T, L^\infty(\Omega))$, for particular finite element approximation spaces, is a direct consequence of the assumptions that $\nabla u \in L^\infty(0, T, L^\infty(\Omega))$, $(u - z) \in L^\infty(0, T, L^\infty(\Omega))$, and Theorem 4.8.7 in [18]. We will further show that, provided $(u - z) \le K_1$ and $\nabla u \le K_2$ for some (possibly large) constants $K_1, K_2 > 0$, then $(U - z) \le K_1(1 + \epsilon_{k_1})$ and $\nabla U \le K_2(1 + \epsilon_{k_2})$, for small parameters ϵ_{k_1} and ϵ_{k_2}, and for particular finite element approximation spaces. See Lemmas 4.3.1 and 4.3.2 below.

Theorem 4.3.2. *Let $u \in W^{1,\infty}(\Omega)$ for all $t \in [0, T]$ be the solution of problem (4.4) and let U be a solution of problem (4.12). Let $\chi = u - \hat{u}$ be the approximation error between the interpolant and the true solution of problem (2.1). Further, assume that $\nabla \hat{u}, \nabla U \in$*

$L^\infty(0, T, L^\infty(\Omega))$. *Then for all $t \in [0, T]$*

$$\|u - U\|^2_{L^2(\Omega)} + \|\nabla u - \nabla U\|^2_{L^2(0,T,L^2(\Omega))} \leq \|\chi(t)\|^2_{H^1(\Omega)} + C \left(\|b - u_0\|^2_{L^2(\Omega)} + \|\chi(0)\|^2_{L^2(\Omega)} + \right.$$
$$\left. + \int_0^T \|\chi_t\|^2_{L^2(\Omega)} + \int_0^T \|\chi\|^2_{L^2(\Omega)} + \int_0^T \int_\Omega |\nabla\chi|^{2\gamma} \right) \quad (4.15)$$

Proof. Note that a weak solution u of problem (4.4) satisfies the weak form given by (4.8). Thus, in particular, the following holds

$$\left(\frac{\partial \hat{u}}{\partial t}, v \right) + \left(\beta(U) \frac{\nabla \hat{u}}{|\nabla \hat{u}|^{1-\gamma}}, \nabla v \right) = (f, v) - \left(\frac{\partial(u - \hat{u})}{\partial t}, v \right) +$$
$$+ \left(\beta(u) \left(\frac{\nabla \hat{u}}{|\nabla \hat{u}|^{1-\gamma}} - \frac{\nabla u}{|\nabla u|^{1-\gamma}} \right), \nabla v \right) + \left((\beta(U) - \beta(u)) \frac{\nabla \hat{u}}{|\nabla \hat{u}|^{1-\gamma}}, \nabla v \right) \quad (4.16)$$

Subtracting (4.16) from (4.8) we obtain that

$$\left(\frac{\partial(U - \hat{u})}{\partial t}, v \right) + \left(\beta(U) \left(\frac{\nabla U}{|\nabla U|^{1-\gamma}} - \frac{\nabla \hat{u}}{|\nabla \hat{u}|^{1-\gamma}} \right), \nabla v \right) = \left(\frac{\partial(u - \hat{u})}{\partial t}, v \right) +$$
$$+ \left(\beta(u) \left(\frac{\nabla u}{|\nabla u|^{1-\gamma}} - \frac{\nabla \hat{u}}{|\nabla \hat{u}|^{1-\gamma}} \right), \nabla v \right) + \left((\beta(u) - \beta(U)) \frac{\nabla \hat{u}}{|\nabla \hat{u}|^{1-\gamma}}, \nabla v \right) \quad (4.17)$$

Let $\xi = U - \hat{u}$ and $\chi = u - \hat{u}$. Set $v = \xi$ in the previous expression to find

$$\left(\frac{\partial \xi}{\partial t}, \xi \right) + \left(\beta(U) \left(\frac{\nabla U}{|\nabla U|^{1-\gamma}} - \frac{\nabla \hat{u}}{|\nabla \hat{u}|^{1-\gamma}} \right), \nabla \xi \right) = \left(\frac{\partial \chi}{\partial t}, \xi \right) +$$
$$+ \left(\beta(u) \left(\frac{\nabla u}{|\nabla u|^{1-\gamma}} - \frac{\nabla \hat{u}}{|\nabla \hat{u}|^{1-\gamma}} \right), \nabla \xi \right) + \left((\beta(u) - \beta(U)) \frac{\nabla \hat{u}}{|\nabla \hat{u}|^{1-\gamma}}, \nabla \xi \right) \quad (4.18)$$

The above expression and estimate (4.6) lead to the following inequality

$$\frac{1}{2} \frac{\partial}{\partial t} \|\xi\|^2_{L^2(\Omega)} + \gamma \epsilon \, \mathcal{A} \|\nabla \xi\|^2_{L^2(\Omega)} \leq \left(\frac{\partial \chi}{\partial t}, \xi \right) + \left(\beta(u) \left(\frac{\nabla u}{|\nabla u|^{1-\gamma}} - \frac{\nabla \hat{u}}{|\nabla \hat{u}|^{1-\gamma}} \right), \nabla \xi \right)$$
$$+ \left((\beta(u) - \beta(U)) \frac{\nabla \hat{u}}{|\nabla \hat{u}|^{1-\gamma}}, \nabla \xi \right), \quad (4.19)$$

where[1] $\mathcal{A} := \inf_{(0,T)\times\Omega} (\mathcal{A}_0) = \inf_{(0,T)\times\Omega} 1 / \left(\|\nabla U\|_{L^\infty(0,T,L^\infty(\Omega))} + \|\nabla \hat{u}\|_{L^\infty(0,T,L^\infty(\Omega))} \right)^{1-\gamma}$.

The terms on the right hand side can be bounded in the following way. The first one, using

---

[1]See remark 4.3.3

Young's inequality:

$$\left| \int_\Omega \frac{\partial \chi}{\partial t} \xi \right| \le \frac{1}{2} \left(\|\chi_t\|_{L^2(\Omega)}^2 + \|\xi\|_{L^2(\Omega)}^2 \right). \tag{4.20}$$

The second one, using estimate (4.7) and Young's inequality with ϵ_1:

$$\left| \int_\Omega \beta(u) \left(\frac{\nabla u}{|\nabla u|^{1-\gamma}} - \frac{\nabla \hat{u}}{|\nabla \hat{u}|^{1-\gamma}} \right) \nabla \xi \right| \le \frac{2}{\gamma} M \int_\Omega |\nabla \chi|^\gamma |\nabla \xi| \tag{4.21}$$

$$\le \frac{2}{\gamma} M \left(\frac{1}{2\epsilon_1} \int_\Omega |\nabla \chi|^{2\gamma} + \frac{\epsilon_1}{2} \|\nabla \xi\|_{L^2(\Omega)}^2 \right) \tag{4.22}$$

where $M = \|\beta(u)\|_{L^\infty(\Omega)}$. The third one, using estimate (4.5), Cauchy-Schwarz, and Young's inequalities with ϵ_2

$$\left| \int_\Omega (\beta(u) - \beta(U)) \frac{\nabla \hat{u}}{|\nabla \hat{u}|^{1-\gamma}} \nabla \xi \right| \le M^* \int_\Omega |u - U||\nabla \hat{u}|^\gamma |\nabla \xi| \tag{4.23}$$

$$\le M^* \|\nabla \hat{u}\|_{L^\infty(\Omega)}^\gamma \|u - U\|_{L^2(\Omega)} \|\nabla \xi\|_{L^2(\Omega)}$$

$$\le M^* \|\nabla \hat{u}\|_{L^\infty(\Omega)}^\gamma \left(\frac{1}{2\epsilon_2} (\|\chi\|_{L^2(\Omega)}^2 + \|\xi\|_{L^2(\Omega)}^2) + \frac{\epsilon_2}{2} \|\nabla \xi\|_{L^2(\Omega)}^2 \right),$$

where $M^* = \alpha \max(\|\beta(u)\|_{L^\infty(\Omega)}, \|\beta(U)\|_{L^\infty(\Omega)})^{\alpha-1}$.

From estimate (4.6), provided $\nabla \hat{u}, \nabla U \in L^\infty(0,T,L^\infty(\Omega))$, there exists a constant $\epsilon_3 > 0$ such that

$$\gamma \epsilon \, \mathcal{A} \ge \epsilon_3 \qquad \text{for all } t \in [0,T] \tag{4.24}$$

we can combine expressions (4.19)-(4.23) and choose ϵ_1 and ϵ_2 small enough to obtain that for some $\bar{\epsilon} > 0$ and some constants $C_i > 0$ $(1 \le i \le 4)$

$$\frac{1}{2} \frac{\partial}{\partial t} \|\xi\|_{L^2(\Omega)}^2 + \bar{\epsilon} \|\nabla \xi\|_{L^2(\Omega)}^2 \le C_1 \|\xi\|_{L^2(\Omega)}^2 + C_2 \|\chi_t\|_{L^2(\Omega)}^2 + C_3 \|\chi\|_{L^2(\Omega)}^2 + C_4 \int_\Omega |\nabla \chi|^{2\gamma} \tag{4.25}$$

Since the second term on the left hand side is nonnegative, we can use Gronwall's Lemma in the previous expression to obtain that for all $t \in [0,T]$

$$\|\xi(t)\|_{L^2(\Omega)}^2 \le C_5(T) \left(\|\xi(0)\|_{L^2(\Omega)}^2 + C_2 \int_0^T \|\chi_t\|_{L^2(\Omega)}^2 + C_3 \int_0^T \|\chi\|_{L^2(\Omega)}^2 + C_4 \int_0^T \int_\Omega |\nabla \chi|^{2\gamma} \right). \tag{4.26}$$

Observe that

$$\|u - U\|_{L^2(\Omega)}^2 \le \|U - \hat{u}\|_{L^2(\Omega)}^2 + \|u - \hat{u}\|_{L^2(\Omega)}^2 = \|\xi(t)\|_{L^2(\Omega)}^2 + \|\chi(t)\|_{L^2(\Omega)}^2. \tag{4.27}$$

and

$$\|\xi(0)\|_{L^2(\Omega)}^2 = \|b - \hat{u}_0\|_{L^2(\Omega)}^2 \quad \leq \quad \|b - u_0\|_{L^2(\Omega)}^2 + \|u_0 - \hat{u}_0\|_{L^2(\Omega)}^2$$

$$\leq \quad \|b - u_0\|_{L^2(\Omega)}^2 + \|\chi(0)\|_{L^2(\Omega)}^2 \qquad (4.28)$$

Thus, combining result (4.26), the two previous expressions and letting $C = \max(C_5 C_i)$ for $i = 2, 3, 4.$ we can establish the first portion of the statement of the Theorem (4.15).

In order to complete the proof we need to find a bound for the gradients. This is done integrating expression (4.25) in time. Note that on the left hand side we have

$$\frac{1}{2}\|\xi(T)\|_{L^2(\Omega)}^2 - \frac{1}{2}\|\xi(0)\|_{L^2(\Omega)}^2 + \epsilon\|\nabla u - \nabla U\|_{L^2(0,T,L^2(\Omega))}^2. \qquad (4.29)$$

Since the first term is nonnegative, the following holds from the time integration of (4.25)

$$\|\nabla\xi\|_{L^2(0,T,L^2(\Omega))}^2 \leq C\left(\frac{1}{2}\|\xi(0)\|_{L^2(\Omega)}^2 + \int_0^T \|\xi\|_{L^2(\Omega)}^2 + \int_0^T \|\chi_t\|_{L^2(\Omega)}^2 + \int_0^T \|\chi\|_{L^2(\Omega)}^2 + \right.$$

$$\left. + \int_0^T \int_\Omega |\nabla\chi|^{2\gamma}\right). \qquad (4.30)$$

Estimate (4.15) follows from the observation that

$$\|\nabla u - \nabla U\|_{L^2(0,T,L^2(\Omega))}^2 \quad \leq \quad \|\nabla U - \nabla\hat{u}\|_{L^2(0,T,L^2(\Omega))}^2 + \|\nabla u - \nabla\hat{u}\|_{L^2(0,T,L^2(\Omega))}^2$$

$$\leq \quad \|\nabla\xi(t)\|_{L^2(0,T,L^2(\Omega))}^2 + \|\nabla\chi(t)\|_{L^2(0,T,L^2(\Omega))}^2.$$

and the combination of (4.26) and (4.30). This concludes the proof. $\qquad\square$

Remark 4.3.3. Note that the error estimate collapses if condition (4.24) is not satisfied. This is the reason why we need to use both: the family of nondegenerate $\{\beta_\epsilon(x)\} > \epsilon$, and the assumptions that $\nabla u \in L^\infty(0, T, L^\infty(\Omega))$ and $\nabla U \in L^\infty(0, T, L^\infty(\Omega))$. The latter assumptions ensure that $\mathcal{A} > 0$ since $\nabla u \in L^\infty(0, T, L^\infty(\Omega))$ implies that $\nabla\hat{u} \in L^\infty(0, T, L^\infty(\Omega))$ for an appropriate finite element space.

Corollary 4.3.1 (Stability). *Under the conditions of Theorem 4.3.2, the method is stable.*

Proof. Write $U(t) = U(t) - u(t) + u(t)$ and use the triangle inequality to find

$$\|U(t)\|_{L^2(\Omega)} \leq \|U(t) - u(t)\|_{L^2(\Omega)} + \|u(t)\|_{L^2(\Omega)}$$

Using the previous theorem and assuming that $u \in L^2(\Omega)$ the result is immediate. $\qquad\square$

Corollary 4.3.2. *Under the assumptions of Theorem 4.3.2, if $u \in W^{1,\infty}(\Omega) \cap H^{k+1}(\Omega)$ is the solution of problem (4.4) and U is a solution of problem (4.8), constructed with piecewise polynomials of degree at most k. Then for all $t \in [0,T]$*

$$\|u - U\|_{L^2(\Omega)} + \|\nabla u - \nabla U\|_{L^2(0,T,L^2(\Omega))} \leq C\,h^{k\gamma} \left(\int_0^T \|u\|_{H^{k+1}(\Omega)}^{2\gamma} \right)^{\frac{1}{2}}. \tag{4.31}$$

Proof. Assuming such regularity on u, the estimates given by Lemma 2.4.1 hold. Thus, by applying the result of Theorem 4.3.2, and using Hölder's inequality with $p = 1/\gamma \geq 1$ and $p^* = 1/(1-\gamma) \geq 1$ in the following expression,

$$\int_\Omega |\nabla \chi|^{2\gamma} \leq \left(\int_\Omega |\nabla \chi|^2 \right)^{\frac{2\gamma}{2}} |\Omega|^{1-\gamma} = \|\nabla \chi\|_{L^2(\Omega)}^{2\gamma} |\Omega|^{1-\gamma} \tag{4.32}$$

we obtain for small h

$$
\begin{aligned}
\|u - U\|_{L^2(\Omega)}^2 + \|\nabla u - \nabla U\|_{L^2(0,T,L^2(\Omega))}^2 \;\leq\;& h^{2k} \|u\|_{H^{k+1}(\Omega)}^2 + C \left(h^{2(k+1)} \|u_0\|_{H^{k+1}(\Omega)}^2 \right. \\
& \left. + h^{2k\gamma} \int_0^T \|u\|_{H^{k+1}(\Omega)}^{2\gamma} \right) \\
\leq\;& C\,h^{2k\gamma} \int_0^T \|u\|_{H^{k+1}(\Omega)}^{2\gamma}
\end{aligned}
$$

since $\|b - u_0\|_{L^2(\Omega)}^2 \leq h^{2(k+1)} \|u_0\|_{H^{k+1}(\Omega)}^2$. The result of the Corollary comes from the observation that for p, q, s positive numbers, $p^2 + q^2 \leq s^2$ implies that $p + q \leq \sqrt{2}\,s$. $\square$

Remark 4.3.4. Note that the error estimate obtained in Corollary 4.3.2 is constrained by the value of $\gamma \in (0,1)$. In the hydraulic context, $\gamma = 1/2$ for both Manning's and Chézy's formulas. According to Corollary 4.3.2 we can ensure $\mathcal{O}(h)$ convergence, for $\gamma = 1/2$, by using quadratic basis functions ($k = 2$) to approximate a very regular solution $u \in W^{1,\infty}(\Omega) \cap H^3(\Omega)$ of problem (4.4). In Section 4.4 we present numerical experiments that show that our analysis is conservative. In fact, we show that our method, implemented with piecewise linear basis functions, can approximate the true solution as $\mathcal{O}(h^2)$ in nondegenerate scenarios, and as $\mathcal{O}(h)$ even when degeneracy happens in subregions of the domain Ω.

Lemma 4.3.1 (Boundedness of approximation). *Under the assumptions of Theorem 4.3.2, choosing $\gamma \geq 1/2$, and provided h is sufficiently small, if $(u - z) \in L^\infty(0,T,L^\infty(\Omega)) \cap H^4(\Omega)$ and $\|u - z\|_{L^\infty(0,T,L^\infty(\Omega))} \leq K_1$, then $\|U - z\|_{L^\infty(0,T,L^\infty(\Omega))} \leq K_1(1 + \epsilon_{k_1})$ for a small parameter ϵ_{k_1}, and for a finite element approximation space consisting of at least piecewise*

cubic basis functions.

Proof. Write $(U - z)$ as $(U - \hat{u} + \hat{u} - z)$ to find

$$\|U - z\|_{L^\infty(0,T,L^\infty(\Omega))} \leq \|U - \hat{u}\|_{L^\infty(0,T,L^\infty(\Omega))} + \|\hat{u} - z\|_{L^\infty(0,T,L^\infty(\Omega))}. \tag{4.33}$$

Choose $\hat{u}$ as in the definition (2.8), and choose at least $\mathcal{M} = \mathcal{P}^3$ (piecewise cubic basis functions, $k = 3$). The fact that the interpolant $(\hat{u} - z) \in L^\infty(0,T,L^\infty(\Omega))$ provided $(u - z) \in L^\infty(0,T,L^\infty(\Omega)) \cap H^4(\Omega)$, for a bounded and smooth time independent function z, is a consequence of Theorem 4.8.7 in [18]. Thus, $\|\hat{u} - z\|_{L^\infty(0,T,L^\infty(\Omega))} \leq K_1$ for sufficiently small h. Now, from Corollary 4.3.2 and Lemma 2.4.2 we obtain

$$\begin{aligned} \|U - z\|_{L^\infty(0,T,L^\infty(\Omega))} &\leq \|U - \hat{u}\|_{L^\infty(0,T,L^\infty(\Omega))} + K_1 \\ &\leq K_0\, h^{-1}\|U - \hat{u}\|_{L^\infty(0,T,L^2(\Omega))} + K_1 \\ &\leq K_0\, h^{3\gamma-1}\left(\int_0^T \|u\|_{H^4(\Omega)}^{2\gamma}\right)^{\frac{1}{2}} + K_1 \end{aligned}$$

Thus, we can choose a sufficiently small h so that $K_0\, h^{1/2}\left(\int_0^T \|u\|_{H^4(\Omega)}^{2\gamma}\right)^{\frac{1}{2}} \leq \epsilon_{k_1} K_1$, which implies

$$\|U - z\|_{L^\infty(0,T,L^\infty(\Omega))} \leq K_1(1 + \epsilon_{k_1})$$

This establishes the result of the Lemma. $\qquad\square$

Lemma 4.3.2 (Boundedness of the gradient of the approximation). *Under the assumptions of Theorem 4.3.2, choosing $\gamma \geq 1/2$, and provided h is sufficiently small, if $u \in L^\infty(0,T,L^\infty(\Omega)) \cap H^5(\Omega)$, $\nabla u \in L^\infty(0,T,L^\infty(\Omega))$ and $\|\nabla u\|_{L^\infty(0,T,L^\infty(\Omega))} \leq K_2$, then $\|\nabla U\|_{L^\infty(0,T,L^\infty(\Omega))} \leq K_2(1 + \epsilon_{k_2})$ for a finite element approximation space consisting of at least fourth degree piecewise polynomial basis functions.*

Proof. Similar to the one in Lemma (4.3.1). $\qquad\square$

4.3.4 Fully discrete approximation

In this section we will further proceed to consider discretization with respect to time. We will denote by dt the time step and with U^n the approximation of $u(t)$ at time $t = t_n = n\, dt$. In order to write down the method we will replace the time derivative in (4.8) by the quotient

$$\delta U^n = \frac{U^n - U^{n-1}}{dt}, \tag{4.34}$$

to obtain the following Backward Euler scheme

$$
\begin{cases}
(\delta U^n, v) & + \left(\dfrac{\beta_\epsilon(U^n)}{|\nabla U^n|^{1-\gamma}} \, \nabla U^n, \nabla v \right) & = (f^n, v) & n\, dt \in [0, T], \quad \forall v \in \mathcal{M}, \\
\text{and} & (U^0, v) & = (u_0, v) & \forall v \in \mathcal{M}.
\end{cases}
\tag{4.35}
$$

The previous expression defines U^n implicitly for U^{n-1} given and it can be written as follows

$$
(U^n, v) + dt \, \left(\frac{\beta_\epsilon(U^n)}{|\nabla U^n|^{1-\gamma}} \, \nabla U^n, \nabla v \right) = (U^{n-1} + dt\, f^n, v) \quad \forall v \in \mathcal{M},
\tag{4.36}
$$

or in matrix form, using the definitions explained in section 4.3.1,

$$
(G + dt\, B(\zeta^n))\zeta^n = G\zeta^{n-1} + dt\, F(t_n),
\tag{4.37}
$$

where f^n (alt. $F(t_n)$) is a known function (alt. matrix)

4.3.4.1 Fully discrete A *priori* error estimate

In this section we will study how close solutions to the fully discrete Galerkin approximation problem (4.35) are to the true weak solution of problem (2.1). The proof follows immediately from the semi-discrete estimates and it is presented for completeness. Once more, in the analysis we will denote with $\beta(x)$ any element of the sequence $\{\beta_\epsilon(x)\}$ as described in section 4.2.

Theorem 4.3.3. *Let $u \in W^{1,\infty}(\Omega)$ for all $t \in [0, T]$ be the solution of problem (4.4) and let U^n be a solution of problem (4.35) at $t_n = n\, dt$. Let also $\chi^n = u(t_n) - \hat{u}(t_n)$ be the approximation error between the interpolant and the true solution of problem (2.1). Then for all $t_n \in [0, T]$*

$$
\|u(t_n) - U^n\|^2_{L^2(\Omega)} + dt \sum_{j=1}^{n} \|\nabla u(t_j) - \nabla U^j\|^2_{L^2(\Omega)} \leq C\|b - u_0\|^2_{L^2(\Omega)} + \|\chi^0\|^2_{L^2(\Omega)} +
$$

$$
+ dt\, C \sum_{j=1}^{n} \left(\|u_t(t_j) - \delta \hat{u}(t_j)\|^2_{L^2(\Omega)} + \|\chi^j\|^2_{L^2(\Omega)} + \int_\Omega |\nabla \chi^j|^{2\gamma} \right)
\tag{4.38}
$$

Proof. Observe that identifying $u^n = u(t_n)$ and $\hat{u}^n = \hat{u}(t_n)$ a similar calculation to (4.17)

66

yields

$$\left(\delta(U^n - \hat{u}^n), v\right) + \left(\beta(U^n)\left(\frac{\nabla U^n}{|\nabla U^n|^{1-\gamma}} - \frac{\nabla \hat{u}^n}{|\nabla \hat{u}^n|^{1-\gamma}}\right), \nabla v\right) = (u_t^n - \delta \hat{u}^n, v) +$$

$$+ \left(\beta(u^n)\left(\frac{\nabla u^n}{|\nabla u^n|^{1-\gamma}} - \frac{\nabla \hat{u}^n}{|\nabla \hat{u}^n|^{1-\gamma}}\right), \nabla v\right) + \left((\beta(u^n) - \beta(U^n))\frac{\nabla \hat{u}^n}{|\nabla \hat{u}^n|^{1-\gamma}}, \nabla v\right) \quad (4.39)$$

Now, let $\xi^n = U^n - \hat{u}(t_n)$ and $\chi^n = u(t_n) - \hat{u}(t_n)$ and choose $v = \xi^n$. Using estimates (4.21) and (4.23) we obtain

$$(\delta \xi^n, \xi^n) + \gamma(\beta(U^n))\mathcal{A}\|\nabla \xi^n\|^2_{L^2(\Omega)} \leq (u_t^n - \delta \hat{u}^n, \xi^n) + \frac{1}{\epsilon_2}C\|\xi^n\|^2_{L^2(\Omega)} + C(\epsilon_2 + \epsilon_1)\|\nabla \xi^n\|^2_{L^2(\Omega)} +$$

$$+ \frac{1}{2\epsilon_2}\|\chi^n\|^2_{L^2(\Omega)} + \frac{1}{2\epsilon_1}C\int_\Omega |\nabla \chi^n|^{2\gamma}. \quad (4.40)$$

If condition (4.24) is satisfied, then for ϵ_1 and ϵ_2 small enough, the previous expression implies

$$\|\xi^n\|^2_{L^2(\Omega)} \leq (\xi^{n-1}, \xi^n) + dt\,(u_t^n - \delta \hat{u}^n, \xi^n) + dt\,C\|\xi^n\|^2_{L^2(\Omega)} + dt\,C\|\chi^n\|^2_{L^2(\Omega)} + dt\,C\int_\Omega |\nabla \chi^n|^{2\gamma},$$
$$(4.41)$$

which yields

$$(1 - dt\,C)\|\xi^n\|^2_{L^2(\Omega)} \leq \|\xi^{n-1}\|^2_{L^2(\Omega)} + dt\,\|u_t^n - \delta \hat{u}^n\|^2_{L^2(\Omega)} + dt\,C\|\chi^n\|^2_{L^2(\Omega)} + dt\,C\int_\Omega |\nabla \chi^n|^{2\gamma}.$$
$$(4.42)$$

Using the Taylor expansion for $(1 - dt\,C)^{-1}$ around zero we can rewrite the previous expression for small dt:

$$\|\xi^n\|^2_{L^2(\Omega)} \leq (1 + dt\,C)\|\xi^{n-1}\|^2_{L^2(\Omega)} + dt\,C\,R_n \quad (4.43)$$

where
$$R_n = \|u_t^n - \delta \hat{u}^n\|^2_{L^2(\Omega)} + \|\chi^n\|^2_{L^2(\Omega)} + \int_\Omega |\nabla \chi^n|^{2\gamma}$$

Making use of inequality (4.43) repeatedly we find that

$$\|\xi^n\|^2_{L^2(\Omega)} \leq (1 + dt\,C)^n\|\xi^0\|^2_{L^2(\Omega)} + dt\,C\sum_{j=1}^n (1 + dt\,C)^{n-j} R_j$$

$$\leq C\|\xi^0\|^2_{L^2(\Omega)} + dt\,C\sum_{j=1}^n R_j \quad for \quad t_n = dt\,n \in [0, T] \quad (4.44)$$

which together with

$$\|\xi^0\|_{L^2(\Omega)}^2 = \|b - \hat{u}^0\|_{L^2(\Omega)}^2 \leq \|b - u_0\|_{L^2(\Omega)}^2 + \|u_0 - \hat{u}^0\|_{L^2(\Omega)}^2$$
$$\leq \|b - u_0\|_{L^2(\Omega)}^2 + \|\chi^0\|_{L^2(\Omega)}^2$$

establish the first portion of the statement of the Theorem. In order to find the appropriate estimate on the gradients a similar argument to the one in the proof of the continuous in time *a priori* error estimate has to be used. Both estimates lead to the result of the Theorem. $\square$

Corollary 4.3.3. *If $u \in W^{1,\infty}(\Omega) \cap H^{k+1}(\Omega)$ is the solution of problem (4.4) and U^n is a solution of problem (4.35) at $t_n = n\,dt$, constructed with piecewise polynomials of degree at most k. Then for all $t_n \in [0, T]$*

$$\|u(t_n) - U^n\|_{L^2(\Omega)} + dt \sum_n \|\nabla u(t_n) - \nabla U^n\|_{L^2(\Omega)} \leq C(u, t_n)\,(dt + h^{k\gamma}) \qquad (4.45)$$

Proof. Given the regularity of u, the quantity R_j can be bounded using the estimates given by Lemma 2.4.1, the fact that $\|u_t^n - \delta\hat{u}^n\|_{L^2(\Omega)}^2 \leq C^*(u)dt^2$ (see [60]), and applying the results of Theorem 4.3.3. Thus,

$$\|u(t_n) - U^n\|_{L^2(\Omega)}^2 + dt \sum_n \|\nabla u(t_n) - \nabla U^n\|_{L^2(\Omega)}^2 \leq C\|b - u_0\|_{L^2(\Omega)}^2 + h^{2(k+1)}\|u_0\|_{H^{k+1}(\Omega)}^2 +$$
$$dt\,C\sum_{j=1}^{n}\left(C^*(u)dt^2 + h^{2k}\|u(t_n)\|_{H^{k+1}(\Omega)}^2 + \right.$$
$$\left. + h^{2k\gamma}\|u(t_n)\|_{H^{k+1}(\Omega)}^{2\gamma}\right)$$
$$\leq C(u, t_n)\,(dt + h^{k\gamma})^2. \qquad (4.46)$$

Recall that for p, q, s positive numbers, $p^2 + q^2 \leq s^2$ implies that $p + q \leq \sqrt{2}\,s$. Thus, the result of the Corollary follows immediately. $\square$

4.4 Numerical Experiments. 1D

A lumped mass continuous Galerkin code with piecewise linear basis functions was implemented in order to perform numerical experiments aimed at investigating: *(i)* The accuracy and validity of the numerical method previously described to solve the DSW equation for the case when $z = 0$, see section 4.4.1, *(ii)* the applicability of the DSW equation as a

model to simulate observed quantities (such as water discharge and depth profile) in an experimental setting for a prescribed bathymetry $z = z(x) \neq 0$, see section 4.4.2, and *(iii)* the qualitative properties of solutions to the DSW in its general form (2.1).

Even though Lemmas 4.3.1 and 4.3.2 suggest that one should use at least fourth order polynomial basis functions in order to ensure the boundedness of both $U - z$ and ∇U and thus convergence of the numerical scheme, in practice, we found that the use of piecewise linear basis functions was appropriate. Furthermore, our numerical experiments showed that the implementation of the regularizing Lipschitz functional $\beta_\epsilon(u) > \epsilon$ (instead of $u - z$), as described in section 4.2, was not necessary. In fact, we found that the stability of the code behaved similarly with or without the implementation of the $\beta_\epsilon(u) > 0$, and generally $\|U_\epsilon - U\|_{L^2(\Omega)} \backsim \mathcal{O}(\epsilon)$, with U_ϵ the numerical solution of the nondegenerate problem (4.4), and U the numerical solution of the possibly degenerate problem (2.1).

A lumped mass approach was chosen since for *well behaved* nonlinear parabolic equations such schemes satisfy a maximum principle and thus provide a monotone and physically consistent way to approach the solution, see [60]. For computational purposes, problem (4.37) was approximated by the semi-implicit scheme

$$(G + dt\, B(\zeta_{(l)}))\zeta^n = G\zeta^{n-1} + dt\, F(t_n). \tag{4.47}$$

A Picard Iteration approach

$$\zeta_{(l)} = (G + dt\, B(\zeta_{(l-1)}))^{-1}(G\zeta^{n-1} + dt\, F(t_n)),$$

with an initial guess $\zeta_{(0)} = \zeta^{n-1}$ was used in order to resolve the nonlinearity, with the assumption that

$$\lim_{l \to \infty} (G + dt\, B(\zeta_{(l)})) = (G + dt\, B(\zeta^n)). \tag{4.48}$$

In practice, the iteration process was stopped when $\|\zeta_{(l)} - \zeta_{(l-1)}\|_{l^2(\Omega)}^2 \leq \tau$ for a prescribed tolerance τ, and ζ^n was set equal to the value of $\zeta_{(l)}$ in the last iteration. In all our experiments we chose $\alpha = 5/3$ and $\gamma = 1/2$ (as in [73] and [38]). These values correspond to Manning's formula. Numerical studies addressing the case when $z = 0$, $\alpha \geq 1$ and $\gamma = 1$, which corresponds to the Porous Medium Equation, include for example: [53] and [72].

4.4.1 Convergence analysis

As suggested in the previous sections, convergence of the numerical method proposed to approximate the DSW equation may fail if the depth $u - z$ is zero or if its gradient ∇u
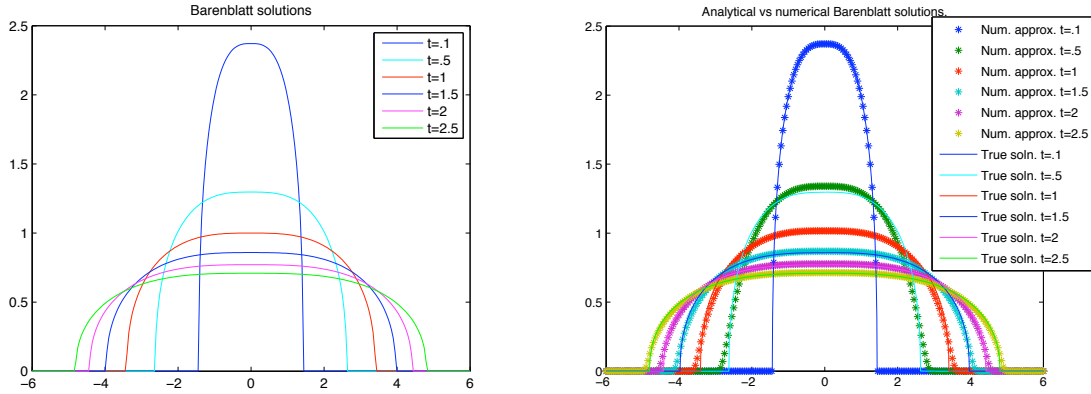
FIGURE 4.1: *Comparison of numerically simulated compactly supported solutions for $\alpha = 5/3$ and $\gamma = 1/2$. **(Left)** Analytical Barenblatt solutions ($z = 0$), **(Right)** Numerical vs analytical solutions ($z = 0$).*

is unbounded. In order to investigate the accuracy and validity of the numerical method in different circumstances we chose two approaches: the first one consisted of reproducing a Lipschitz continuous compactly supported solution of (2.1) presented in [35] for the 1-dimensional case, for $z = 0$, $f = 0$, and $\Omega = \mathbb{R}$. The second one consisted of choosing a simple function $u(x, t) \geq 0$ with unbounded gradient at $x = 0$ that we used to create a right hand side $f(x, t)$. This was done by applying the differential operator defined by the left hand side of (2.1). We then used our method to approximate this u. For both cases we obtained convergence rates for a variety of scenarios and present them in the next paragraphs.

Remark 4.4.1. Despite the fact that Corollary 4.3.3 would only ensure convergence results of the type, $\|u(t_n) - U^n\|_{L^2(\Omega)} \leq C(u, t_n) (dt + h^{1/2})$ for $\gamma = 1/2$, and nondegenerate solutions $u \in H^2(\Omega)$ using piecewise linear basis functions ($k = 1$), we chose the time step comparable to or smaller than the square of the grid diameter, $dt \lesssim h^2 = dx^2$, in our convergence experiments. This was done in order to investigate if optimality in the convergence rates (*i.e.* $\|u(t_n) - U^n\|_{L^2(\Omega)} \leq C(u, t_n) (dt + h^2)$) could be achieved for say, nondegenerate solutions with bounded gradients. Intuitively, under these conditions, (2.1) should resemble a *well behaved* nonlinear parabolic problem.

4.4.1.1 Compactly supported solution

Most of the Barenblatt solutions presented in this subsection do not exhibit the properties of solutions assumed in order to derive the DSW equation from the SWE (such as the fact

that the gradient of the solution be comparable to the gradient of the bathymetry) and thus they would be incorrect models for shallow water flows. However, from the numerical point of view, having an analytic expression of the solution makes the convergence analysis much simpler. The explicit expression of a Barenblatt solution for (2.1) is presented in [35] and given by

$$u(x,t) = t^{-\frac{1}{\gamma(m+1)}} \left[C - k(m,\gamma) \, |\Phi|^{\frac{\gamma+1}{\gamma}} \right]_+^{\frac{\gamma}{m\gamma-1}} \tag{4.49}$$

where $[s(x)]_+$ denotes the positive part of $s(x)$, $m = 1 + \alpha/\gamma$, C is a positive function related to the initial mass M, given by

$$M = \int_{-\infty}^{\infty} u(x,t)dx,$$

and

$$k(m,\gamma) = \frac{m\gamma - 1}{m(\gamma + 1)} \left(\frac{1}{\gamma(m+1)} \right)^{\frac{1}{\gamma}}, \qquad \text{and} \qquad \Phi = x \, t^{-\frac{1}{\gamma(m+1)}}$$

The function given by (4.49) is Lipschitz continuous and compactly supported, thus it is almost everywhere differentiable and its gradient is bounded wherever it exists. By changing the numerical domain Ω we managed to verify how the numerical method approximates the solution both when it is globally not degenerate, $u > 0$, and when it degenerates in some subsets of Ω. For our calculations in the degenerate case, we chose a numerical interval $[-L, L]$ big enough so that the free boundary would always be inside the domain Ω for $t \in [t_0, t_f]$. The results are shown in Table 4.1(a) for the nondegenerate case, and in Table 4.1(b) for the degenerate case. In all cases $t_0 = 2$ and $t_f = 2.1$ and the Picard Iteration scheme was run until the tolerance value met the condition $\tau \leq 10^{-10}$. We computed the numerical solutions and compared them to the true solution using $dt = 1/10$ and $dx = 1/20$ to produce the plots in Figure 4.1. The computed mass M of the numerical solution was calculated as a function of time t and it was observed to be close to the constant $M = 5.8465$ which corresponds to the value of the mass of the true solution (4.49) in the time interval $t \in [.1, 2.5]$.

4.4.1.2   Artificial right hand side

The function $u(x,t) \geq 0$ with unbounded gradient at $x = 0$ that we used to create a right hand side $f(x,t)$ by applying the differential operator defined by the left hand side of (2.1)

71

TABLE 4.1: *Convergence rates. Barenblatt solutions for $\alpha = 5/3$ and $\gamma = 1/2$.*

(a) Nondegenerate case

dt	dx	$\|U-u\|_{L^2(\Omega)}$	Conv. rate
1/50	1	$6.34\text{x}10^{-3}$	.
1/50	1/2	$1.79\text{x}10^{-3}$	1.82
1/100	1/4	$4.87\text{x}10^{-4}$	1.88
1/200	1/8	$1.21\text{x}10^{-4}$	2.00
1/400	1/16	$2.88\text{x}10^{-5}$	2.08
1/1000	1/32	$7.37\text{x}10^{-6}$	1.97
1/4000	1/64	$1.87\text{x}10^{-6}$	1.98

(b) Degenerate case

dt	dx	$\|U-u\|_{L^2(\Omega)}$	Conv. rate
1/50	1	$1.32\text{x}10^{-1}$	.
1/50	1/2	$8.39\text{x}10^{-2}$	0.65
1/100	1/4	$3.97\text{x}10^{-2}$	1.08
1/200	1/8	$2.47\text{x}10^{-2}$	0.69
1/400	1/16	$1.36\text{x}10^{-2}$	0.85
1/1000	1/32	$7.83\text{x}10^{-3}$	0.80
1/4000	1/64	$4.74\text{x}10^{-3}$	0.72

was:

$$u = \begin{cases} (100 - t^2)\sqrt{x} & \text{if } x \geq 0, \\ 0 & \text{if } x < 0. \end{cases} \tag{4.50}$$

with $\alpha = \frac{5}{3}$ and $\gamma = \frac{1}{2}$, the right hand side f is given by

$$f = \frac{\partial u}{\partial t} - \frac{\partial}{\partial x}\left(\frac{u^{\frac{5}{3}}}{|\frac{\partial u}{\partial x}|^{\frac{1}{2}}} \frac{\partial u}{\partial x} \right) = \begin{cases} -2t\sqrt{x} - \dfrac{7}{12\sqrt{2}}(100-t^2)^{\frac{13}{6}} x^{\frac{-5}{12}} & \text{if } x \geq 0, \\ 0 & \text{if } x < 0, \end{cases}$$

Examples of convergence rates between the numerical solution and the true solution (4.50) in three different domains Ω can be found in Tables 4.2(a), 4.2(b), and 4.2(c). All results were calculated with $t_0 = 9$ and $t_f = 9.1$ and the Picard Iteration scheme was run until the tolerance value met the condition $\tau \leq 10^{-10}$. Table 4.2(a) shows results for $\Omega = [5, 10]$, where the solution u is not degenerate and its gradient is bounded. 4.2(b) shows results corresponding to $\Omega = [0, 5]$. In this case, the gradient of the solution u is unbounded and $u = 0$ at $x = 0$. Finally, for $\Omega = [-0.5, 4.5]$ the results are shown in Table 4.2(c). Here the solution u degenerates in the interval $[-0.5, 0]$ and the gradient of u is unbounded at $x = 0$.

4.4.1.3 Convergence results discussion

Corollary 4.3.3 establishes that the numerical scheme will approximate the true solution as $\mathcal{O}(h^{k\gamma})$, whenever piecewise polynomials of degree at most k are used and the true solution $u \in H^{k+1}(\Omega)$ is nondegenerate and its gradient is bounded. In Tables 4.1(a) and 4.2(a) we show the convergence rates of the numerical method in cases when the true solution is nondegenerate and its gradient is bounded. Note that the numerical solution

TABLE 4.2: *Convergence rates. Artificial right hand side with $t_0 = 9$ and $t_f = 9.1$*

(a) Nondeg. & bdd gradient, $\Omega=(5,10)$

dt	dx	$\|U - u\|_{L^2(\Omega)}$	Conv. rate
1/50	1	$1.28\mathrm{x}10^{-1}$	.
1/50	1/2	$3.52\mathrm{x}10^{-2}$	1.86
1/100	1/4	$9.57\mathrm{x}10^{-3}$	1.88
1/200	1/8	$2.62\mathrm{x}10^{-3}$	1.87
1/400	1/16	$7.53\mathrm{x}10^{-4}$	1.80
1/1000	1/32	$2.16\mathrm{x}10^{-4}$	1.80
1/4000	1/64	$5.44\mathrm{x}10^{-5}$	1.99

(b) Unbdd gradient, $\Omega=(0,5)$

dt	dx	$\|U - u\|_{L^2(\Omega)}$	Conv. rate
1/50	1	2.83	.
1/50	1/2	1.48	0.93
1/100	1/4	$7.85\mathrm{x}10^{-1}$	0.92
1/200	1/8	$4.05\mathrm{x}10^{-1}$	0.95
1/400	1/16	$2.04\mathrm{x}10^{-1}$	0.98
1/1000	1/32	$1.03\mathrm{x}10^{-1}$	1.00
1/4000	1/64	$5.14\mathrm{x}10^{-2}$	1.00

(c) Deg. & unbdd gradient, $\Omega=(-0.5,4.5)$

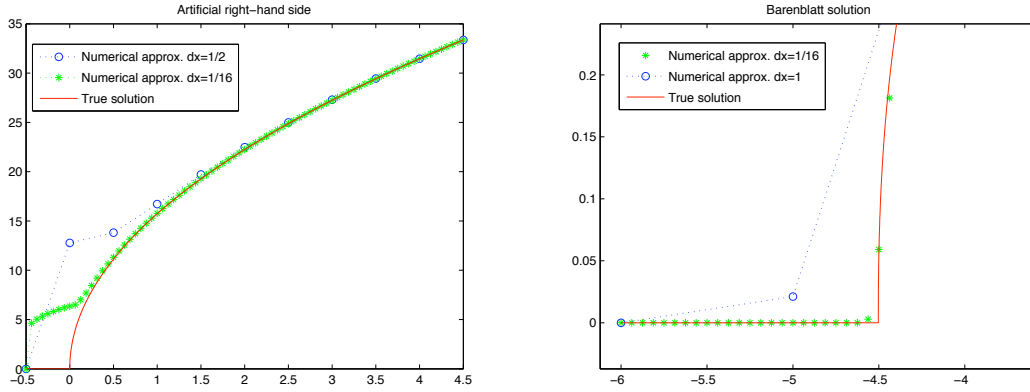| dt	dx	$\|U - u\|_{L^2(\Omega)}$	Conv. rate
1/50	1/2	4.99	.
1/100	1/4	3.61	0.47
1/200	1/8	2.64	0.45
1/400	1/16	1.96	0.43
1/1000	1/32	1.48	0.41
1/4000	1/64	1.11	0.41

FIGURE 4.2: *Comparison of numerically computed solutions **(Left)** (with initial condition $u_0 = 19\sqrt{x}$ at $t = 9$) vs true solutions $u = (100 - t^2)\sqrt{x}$ at time $t = 9.2$ using an artificial right hand side $\Omega = [-0.5, 4.5]$ **(Right)** Numerical approx. of Barenblatt $\Omega = [-6, 6]$.*

converges quadratically to the true solution ($\mathcal{O}(h^2)$) for piecewise linear basis functions ($k = 1$), suggesting that, under these conditions, (2.1) becomes a *well behaved* nonlinear parabolic problem and the expected convergence rates from Corollary 4.3.3 appear to be too conservative. On the other hand, when the solution degenerates at one point and the gradient is unbounded (thus, the conditions to ensure convergence according to Corollary 4.3.3 are not satisfied), optimality is lost but the numerical scheme still converges linearly to the true solution ($\mathcal{O}(h)$), see Table 4.2(b). Close to linear convergence is observed when degeneracy happens and the gradient of the solution is discontinuous, see Table 4.1(b). The worst case scenario takes place when degeneracy happens in an interval and the gradient of the solution is unbounded, see Table 4.2(c). In this case the order of convergence seems to behave close to $\mathcal{O}(h^{2/5})$. It is important to mention at this point that the regions of the domain where the numerical approximation differs mostly from the true solution correspond to regions where the gradient is discontinuous, which generally takes place near the free boundary. We show this in Figure 4.2.

4.4.2 Validation: Iwagaki's experiment

In order to validate our code in the case when $z \neq 0$, we chose to reproduce numerically a set of laboratory experiments conducted by Iwagaki [46]. This approach was followed by Zhang and Cundy [73], and by Feng and Molz [38]. We used the friction parameters reported by Iwagaki and no calibration was pursued. The numerical simulations were obtained using a regular mesh with element diameter $\Delta x = 1/2$ meters and a time step $\Delta t = 1/2$ seconds. Iwagaki's experiments were designed to produce unsteady flows in a channel 24 m long with
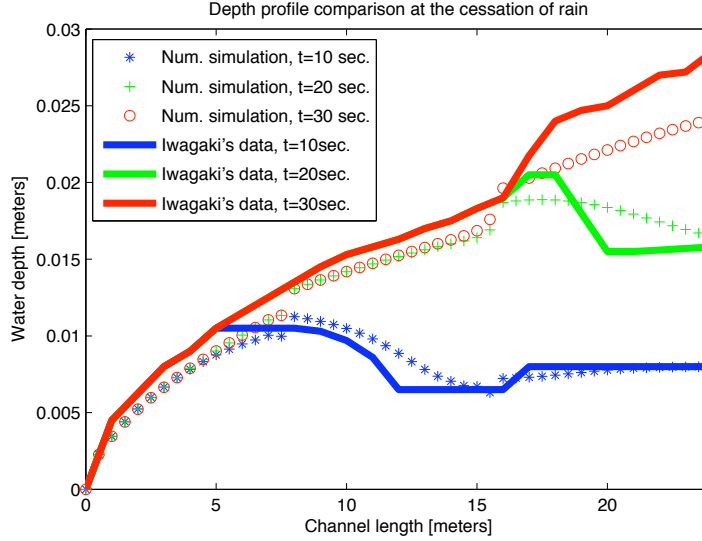
FIGURE 4.3: *Comparison of the numerically calculated depth profiles at the cessation of rainfall with Iwagaki's experimental results on a three plane cascade. Solid lines are experimental results and dashed points are numerical results.*

a cross section of 19.6 cm. The channel was divided into three sections of equal length (8m) and different slopes ($\theta = 0.02, 0.015, 0.01\%$) each. During experiments, three different rainfall intensities ($f = 0.108, 0.064$, and 0.80 cm/s) were simultaneously applied to each section for different time periods ($t = 10, 20$ and 30 sec). Figure 4.3 shows snapshots of the depth profiles along the domain, both measured and numerically simulated, at the cessation of three different rain events lasting $t = 10, 20$ and 30 seconds respectively. We can see that overall, the relevant qualitative nature of the phenomena is captured in the numerical simulations. In Figure 4.4 the simulated water discharge $q = Vh$ as a function of time t, at the lowest end of the domain, $x = 24$m, was plotted and compared to the experimental data for three different rain events[2]. In Figure 4.4, top left, the agreement of the hydrograph with the experimental data is very good. The inability to get the full maximum of the curve as well as the extra spread in the curve by the numerical results is clearly one of the limitations of the diffusive wave approximation. Nevertheless, the time when a maximum discharge is achieved is matched well by the simulation. In Figure 4.4, top right, the agreement of the hydrograph with the experimental data is good. In fact, we can observe that the breakthrough time in the simulation is smaller, however, the area under the curve corresponding to the water discharge volume seems to be in good agreement with the

---

[2]It is important to note that the discharge q is calculated using the expression inside the divergence term in the first equation of the IBVP (2.1).
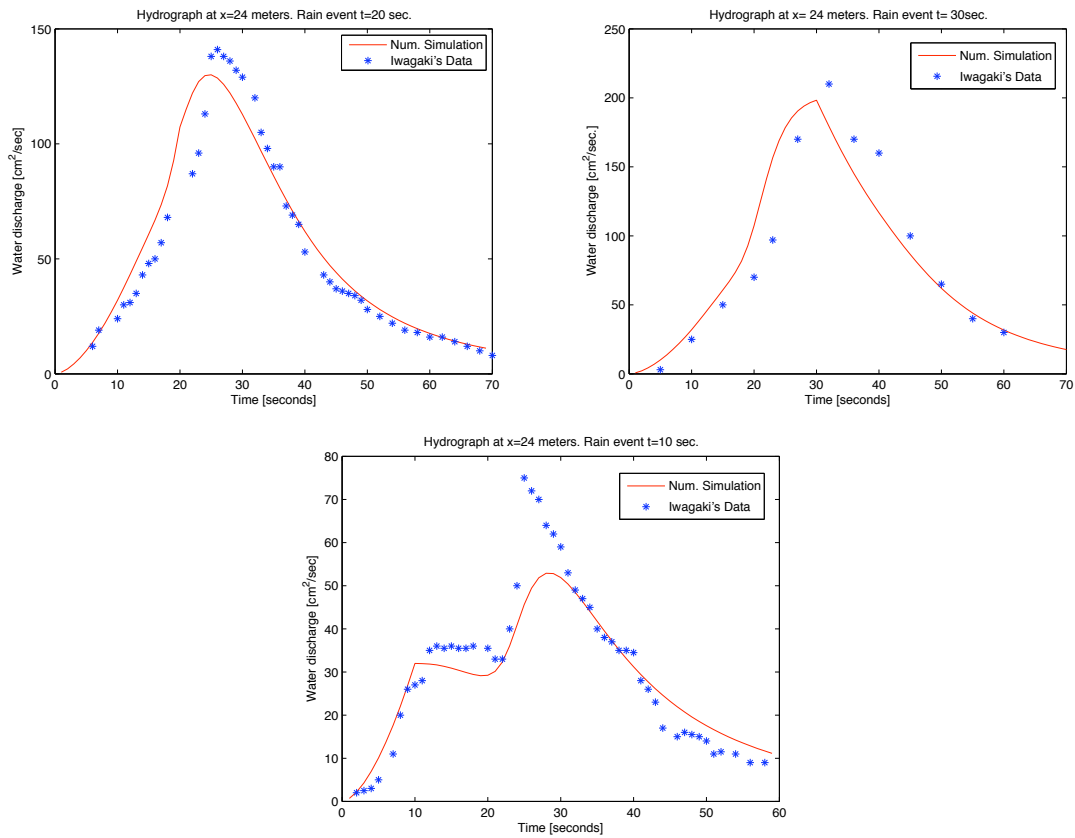
FIGURE 4.4: *Comparison of the numerically calculated hydrographs at x=24 m with Iwagaki's experimental results on a three plane cascade, during a rain event of (**Top left**) 20 seconds, (**Top right**) 30 seconds, and (**Bottom**) 10 seconds. Solid lines are numerical results and dashed points are experimental results.*
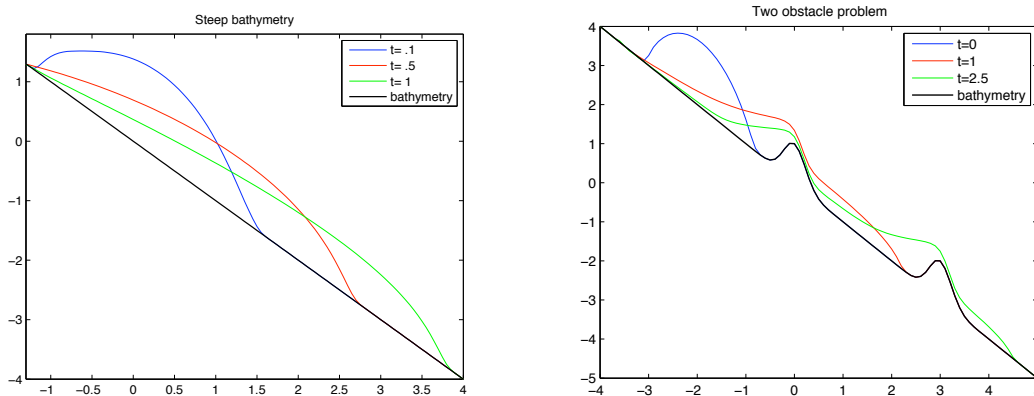
FIGURE 4.5: *Evolution of a compactly supported initial condition simulating a well localized rain event.* ***(Left)*** *Initial condition $u_0 = z + [-2(x-1)(x+1)]_+$ on a steep bathymetry,* ***(Right)*** *Initial condition $u_0 = z + [-2(x+3)(x+1)]_+$ on an inclined bathymetry with two obstacles.*

experimental data. Figure 4.4, bottom, shows the hydrograph for a 10 second rain event. Breakthrough times as well as the main qualitative behavior are captured by the simulation even though the maximum values are not accurately reproduced due to the diffusive nature of the approximation. It must be emphasized that in a flooding event, the most relevant pieces of information obtained from hydrographs are the breakthrough time and the overall discharged water volume, which are nicely modeled by the numerical simulation. Moreover, it is interesting to observe that despite the fact that in the derivation of the DSW equation one assumes uniform flow conditions, the numerical results match reasonably well with the unsteady flow experimental measurements.

4.4.3 Qualitative properties of solutions

In this section we present qualitative properties of solutions to the DSW when the topographic effects are not neglected ($z \neq 0$). Our findings are based on numerically simulated solutions obtained with our code. Our aim was to investigate if the properties of solutions of the DSW equation found in the 1-D case when $z = 0$ in [35], [6], and [2] persist in a more general setting when the bathymetry z is a piece-wise smooth and bounded time independent function. Properties such as boundedness and existence of compactly supported solutions, finite speed of propagation of disturbances, extinction in finite time were found to persist for piece-wise regular $z \neq 0$. We did not pursue any convergence analysis for this case due to the lack of an analytic expression for true solutions of the IBVP (2.1). We present snapshots of solutions at different times for different bathymetries $z(x) \neq 0$ in
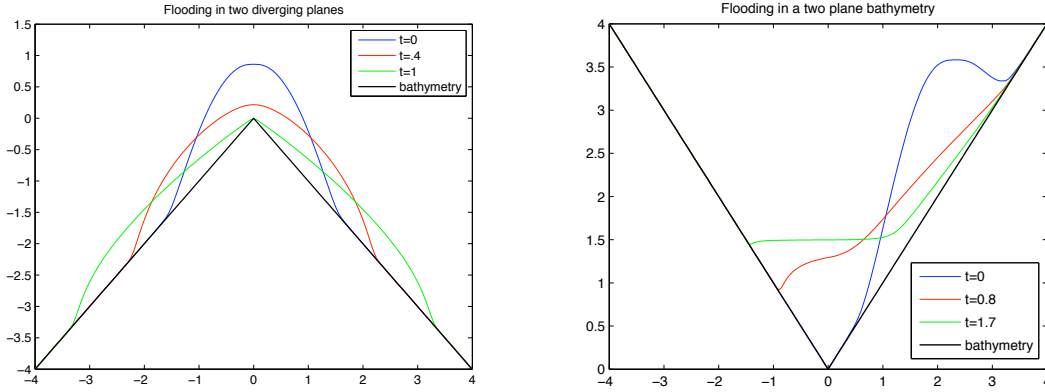
FIGURE 4.6: *Evolution of a compactly supported initial condition simulating a well localized rain event. (**Left**) Initial condition $u_0 = z + 2\exp(-2x^2)$ on two diverging planes, (**Right**) Initial condition $u_0 = z + 2\exp(-2(x-2)^2)$ on a two plane bathymetry.*

Figures 4.5 and 4.6. In all our numerical experiments we used $\Delta x = 1/20$ and $\Delta t = 1/20$. Observe in particular that in Figure 4.6 (Right), the water depth reaches equilibrium without the appearance of sloshing. This is a consequence of the diffusive wave approximation and would model the flow of a very viscous fluid flow or a flow dominated by bottom friction. In Figures 4.5 and 4.6 we observe that the DSW equation has some regularization effect on the solution in regions where the problem is not degenerate. However, the regularity of the solution $u(x)$ corresponds to that of $z(x)$ whenever the problem becomes degenerate (*i.e.* when $u - z = 0$).

Chapter 5

A Discontinuous Galerkin Method

In this Chapter we will study the approximation properties of numerical solutions to the DSW equation, through the regularized initial/boundary-value problem (4.4), obtained using the LDG method. The overview of this Chapter is the following. For consistency we present the well known notation in DG methods in section 5.1. Then we describe the details of the LDG method in section 5.2. In section 5.2.2, we prove properties such as stability and *a priori* error estimates for the numerical formulation of the LDG method. In section 5.3, we show numerical simulations of the evolution of water depth profiles in 2D that capture the salient features of: *(i)* an ideal dam break problem and *(ii)* water flow in a channel containing vegetation.

5.1   Preliminaries

In this Chapter, we will consider a general triangulation τ_h of $\Omega \subset \mathbb{R}^d$ into elements Ω_e of maximum diameter $h > 0$. We will denote with ε_i the set of all interior element faces, with ε_D the set of all element faces along the Dirichlet boundary Γ_D, and ε_N the set of all element faces along the Newmann boundary Γ_N. Now, note that if ε is an interior face in the finite element mesh, then ε has two elements adjacent to it, we will denote them by Ω_e^- and Ω_e^+. Also, if v and $\boldsymbol{w}$ are smooth real valued and vector valued functions, respectively, defined on these elements, we will denote their traces on ε, from the interior of the element Ω_e^-, as v^- and $\boldsymbol{w}^-$; and from the exterior of the element Ω_e^+, as v^+ and $\boldsymbol{w}^+$. We will denote by $\boldsymbol{n}^-$ the outward normal vector to the element Ω_e^- at ε and by $\boldsymbol{n}^+$ the outward normal vector to the element Ω_e^+ at ε. The previous definition implies naturally that $\boldsymbol{n}^+ = -\boldsymbol{n}^-$.

We will define the average $\{\cdot\}$ and the jump $[\![\cdot]\!]$ on the face ε as:

$$\{v\} = \frac{(v^- + v^+)}{2}, \qquad \{\boldsymbol{w}\} = \frac{(\boldsymbol{w}^- + \boldsymbol{w}^+)}{2}, \tag{5.1}$$

$$[\![v]\!] = v^- \boldsymbol{n}^- + v^+ \boldsymbol{n}^+, \qquad [\![\boldsymbol{w}]\!] = \boldsymbol{w}^- \cdot \boldsymbol{n}^- + \boldsymbol{w}^+ \cdot \boldsymbol{n}^+ \tag{5.2}$$

In this Chapter, we will also denote by $(\cdot, \cdot)_E$, the usual L^2 inner product over an d-dimensional domain E (as defined in section 2.3), and by $\langle \cdot, \cdot \rangle_{\partial E}$, the $(d-1)$-dimensional integral over the surface ∂E. To simplify notation, we will omit the dependence on the domain and denote with $(\cdot, \cdot)$ the integrals over the whole domain $E = \Omega$, $(\cdot, \cdot)_\Omega$.

5.2 The Local Discontinuous Galerkin Method

In order to motivate the LDG method it is appropriate to rewrite the nonlinear degenerate parabolic IBVP (2.1) as a degenerate first order system of equations where u, ∇u, and $a(u, \nabla u)$, are now considered as *independent unknowns*:

$$\begin{cases} u_t - \nabla \cdot \boldsymbol{q} = f & \text{on } \Omega \times (0, T] \\ \tilde{\boldsymbol{q}} = \nabla u & \text{on } \Omega \times (0, T] \\ \boldsymbol{q} = a(u, \tilde{\boldsymbol{q}})\tilde{\boldsymbol{q}} & \text{on } \Omega \times (0, T] \end{cases} \tag{5.3}$$

where

$$a(u, \tilde{\boldsymbol{q}}) = \frac{(u - z)^\alpha}{|\tilde{\boldsymbol{q}}|^{1-\gamma}}, \tag{5.4}$$

with initial and boundary conditions given as before by

$$\begin{cases} u = u_0 & \text{on } \Omega \times \{0\} \\ \boldsymbol{q} \cdot n = B_N & \text{on } \partial\Omega \cap \Gamma_N \times (0, T] \\ u = B_D & \text{on } \partial\Omega \cap \Gamma_D \times (0, T] \end{cases} \tag{5.5}$$

where $\partial\Omega = \Gamma = \Gamma_N + \Gamma_D$. Moreover, assuming u, $\boldsymbol{q}$ and $\tilde{\boldsymbol{q}}$ are smooth enough, we multiply each equation in problem (5.3) by a local test functions $w \in W_{h,e}$, $\boldsymbol{v} \in (V_{h,e})^d$ and $\tilde{\boldsymbol{v}} \in (V_{h,e})^d$ respectively (where $d = 1, 2$, is the spatial dimension), and integrate by parts over an element Ω_e to obtain the local weak form of (5.3):

$$\begin{cases} (u_t, w)_{\Omega_e} + (\boldsymbol{q}, \nabla w)_{\Omega_e} - \langle \boldsymbol{q} \cdot \boldsymbol{n}_e, w \rangle_{\partial\Omega_e} = (f, w)_{\Omega_e} & \forall\, w \in W_{h,e} \\ (\tilde{\boldsymbol{q}}, \tilde{\boldsymbol{v}})_{\Omega_e} + (u, \nabla \cdot \tilde{\boldsymbol{v}})_{\Omega_e} - \langle u, \boldsymbol{v} \cdot \boldsymbol{n}_e \rangle_{\partial\Omega_e} = 0 & \forall\, \tilde{\boldsymbol{v}} \in (V_{h,e})^d \\ (\boldsymbol{q}, \boldsymbol{v})_{\Omega_e} - (a(u, \tilde{\boldsymbol{q}})\tilde{\boldsymbol{q}}, \boldsymbol{v})_{\Omega_e} = 0 & \forall\, \boldsymbol{v} \in (V_{h,e})^d \end{cases} \tag{5.6}$$

where $\boldsymbol{n}_e$ represents the outward normal vector in the faces of the element Ω_e. The discontinuous Galerkin method consists of finding an approximation $(U, \boldsymbol{Q}, \tilde{\boldsymbol{Q}})$ to the solution $(u, \boldsymbol{q}, \tilde{\boldsymbol{q}})$ of (5.6), of the form

$$U(x,t) = \sum_{i=1}^{M} \zeta_i(t) w_i(x), \quad \boldsymbol{Q}(x,t) = \sum_{i=1}^{N} \mu_i(t) \boldsymbol{v}_i(x), \quad \text{and} \quad \tilde{\boldsymbol{Q}}(x,t) = \sum_{i=1}^{N} \tilde{\mu}_i(t) \boldsymbol{v}_i(x). \quad (5.7)$$

satisfying for all $t \in [0, T]$

$$\begin{cases} (U_t, w)_{\Omega_e} + (\boldsymbol{Q}, \nabla w)_{\Omega_e} - \langle \widehat{\boldsymbol{Q}} \cdot \boldsymbol{n}_e, w \rangle_{\partial \Omega_e} = (f, w)_{\Omega_e} & \forall\, w \in W_{h,e} \\ (\tilde{\boldsymbol{Q}}, \tilde{\boldsymbol{v}})_{\Omega_e} + (U, \nabla \cdot \tilde{\boldsymbol{v}})_{\Omega_e} - \langle \widehat{U}, \boldsymbol{v} \cdot \boldsymbol{n}_e \rangle_{\partial \Omega_e} = 0 & \forall\, \tilde{\boldsymbol{v}} \in (V_{h,e})^d \\ (\boldsymbol{Q}, \boldsymbol{v})_{\Omega_e} - (a(U, \tilde{\boldsymbol{Q}}) \tilde{\boldsymbol{Q}}, \boldsymbol{v})_{\Omega_e} = 0 & \forall\, \boldsymbol{v} \in (V_{h,e})^d \end{cases} \quad (5.8)$$

for every element Ω_e in the domain Ω. In (5.7), $w_1, ..., w_M$ are linearly independent piecewise polynomial functions in $W_h = \{\cup_{\Omega_e \in \Omega}\, W_{h,e}\}$ and $\boldsymbol{v}_1, ..., \boldsymbol{v}_N$ are vector valued linearly independent piecewise polynomial functions in $(V_h)^d = \{\cup_{\Omega_e \in \Omega}\, (V_{h,e})^d\}$. Generally speaking, $W_{h,e}$ and $V_{h,e}$ are subspaces of $H^1(\Omega_e)$, and in practice they are chosen to be the set of all polynomials of degree at most k inside every element Ω_e. By construction, the approximants $(U, \boldsymbol{Q}, \tilde{\boldsymbol{Q}})$ may be discontinuous across element boundaries since no continuity across elements is assumed in the basis functions spaces $(W_{h,e})$ and $(V_{h,e})^d$. As a consequence, at a given face ε the functions $(U, \boldsymbol{Q}, \tilde{\boldsymbol{Q}})$ may be *multi-valued*. This is why the *numerical fluxes*: $\widehat{\boldsymbol{Q}}$ and $\widehat{U}$ are introduced in (5.8). Furthermore, in order to produce a meaningful DG method and globally determine the coefficients $\zeta_i(t)$, $\mu_i(t)$, and $\tilde{\mu}_i(t)$ appearing in (5.7) the numerical fluxes $\widehat{\boldsymbol{Q}}$ and $\widehat{U}$ need to be carefully defined. This issue is clearly explained in the context of elliptic problems in [22] and in [5], and in the context of nonlinear diffusion problems in [19].

For the LDG method that we will analyse and implement, the numerical fluxes are chosen in the following way:

$$\widehat{U} = \begin{cases} \{U\} & \text{if } \varepsilon \in \varepsilon_i, \\ B_D & \text{if } \varepsilon \in \Gamma_D, \\ U & \text{if } \varepsilon \in \Gamma_N, \end{cases} \quad (5.9)$$

and

$$\widehat{\boldsymbol{Q}} = \begin{cases} \{\boldsymbol{Q}\} - \sigma [\![U]\!] & \text{if } \varepsilon \in \varepsilon_i, \\ \boldsymbol{Q} - \sigma(U\boldsymbol{n} - B_D\boldsymbol{n}) & \text{if } \varepsilon \in \Gamma_D, \\ B_N & \text{if } \varepsilon \in \Gamma_N, \end{cases} \quad (5.10)$$

81

Furthermore, the basis functions spaces $W_{h,e}$ and $V_{h,e}$ will be chosen to be the same, thus simplifying the coding as compared to standard *mixed* methods. Note that for the particular choice of numerical fluxes introduced in (5.9) and (5.10), the numerical flux $\widehat{U}$ does not depend on $\boldsymbol{Q}$. This makes it possible for the local variable $\boldsymbol{Q}$ to be solved in terms of U by using the second equation of (5.8). This is a particular property that distinguishes the LDG method from other DG schemes. The parameter σ appearing in the definition of the numerical fluxes will be chosen carefully in order to enhance the stability and thus, the accuracy of the method.

Remark 5.2.1. The fluxes defined in (5.9) and (5.10) are both *consistent* and *conservative* as defined in [5] and [22].

The resulting LDG formulation is obtained in two steps. First, by summing over all elements Ω_e to find

$$
\begin{cases}
(U_t, w) + (\boldsymbol{Q}, \nabla w) - \langle \widehat{\boldsymbol{Q}}, [\![w]\!] \rangle_{\varepsilon_i} - \langle \widehat{\boldsymbol{Q}} \cdot \boldsymbol{n}, w \rangle_{\partial\Omega} = (f, w) & \forall\, w \in W_h \\
(\tilde{\boldsymbol{Q}}, \tilde{\boldsymbol{v}}) + (U, \nabla \cdot \tilde{\boldsymbol{v}}) - \langle \widehat{U}, [\![\boldsymbol{v}]\!] \rangle_{\varepsilon_i} - \langle \widehat{U}, \boldsymbol{v} \cdot \boldsymbol{n} \rangle_{\partial\Omega} = 0 & \forall\, \tilde{\boldsymbol{v}} \in (W_h)^d \\
(a(U, \tilde{\boldsymbol{Q}})\tilde{\boldsymbol{Q}}, \boldsymbol{v}) - (\boldsymbol{Q}, \boldsymbol{v}) = 0 & \forall\, \boldsymbol{v} \in (W_h)^d
\end{cases}
\tag{5.11}
$$

where we have denoted with $(\cdot, \cdot) = \sum_e (\cdot, \cdot)_{\Omega_e}$ the sum of all element integrals. And second, by substituting the values of the *numerical fluxes* (5.9) and (5.10) in (5.11)

$$
\begin{cases}
(U_t, w) + (\boldsymbol{Q}, \nabla w) - \langle \{\boldsymbol{Q}\}, [\![w]\!] \rangle_{\varepsilon_i} + \langle \sigma [\![U]\!], [\![w]\!] \rangle_{\varepsilon_i} + \\
\quad - \langle B_N, w \rangle_{\Gamma_N} - \langle \boldsymbol{Q} \cdot \boldsymbol{n}, w \rangle_{\Gamma_D} + \langle \sigma(U - B_D), w \rangle_{\Gamma_D} = (f, w) & \forall\, w \in W_h \\
(\tilde{\boldsymbol{Q}}, \tilde{\boldsymbol{v}}) + (U, \nabla \cdot \tilde{\boldsymbol{v}}) - \langle \{U\}, [\![\tilde{\boldsymbol{v}}]\!] \rangle_{\varepsilon_i} - \langle U, \tilde{\boldsymbol{v}} \cdot \boldsymbol{n} \rangle_{\Gamma_N} = \langle B_D, \tilde{\boldsymbol{v}} \cdot \boldsymbol{n} \rangle_{\Gamma_D} & \forall\, \tilde{\boldsymbol{v}} \in (W_h)^d \\
(a(U, \tilde{\boldsymbol{Q}})\tilde{\boldsymbol{Q}}, \boldsymbol{v}) - (\boldsymbol{Q}, \boldsymbol{v}) = 0 & \forall\, \boldsymbol{v} \in (W_h)^d
\end{cases}
\tag{5.12}
$$

where, for simplicity, we have denoted with

$$
\langle \cdot, \cdot \rangle_{\varepsilon_i} := \sum_e \langle \cdot, \cdot \rangle_{\partial\Omega_e \setminus \Gamma}, \quad \langle \cdot, \cdot \rangle_{\Gamma_N} = \sum_e \langle \cdot, \cdot \rangle_{\partial\Omega_e \cap \Gamma_N}, \quad \text{and} \quad \langle \cdot, \cdot \rangle_{\Gamma_D} = \sum_e \langle \cdot, \cdot \rangle_{\partial\Omega_e \cap \Gamma_D}
$$

the sum of the boundary integrals in all interior element boundaries ε_i, in all element boundaries along the Newman boundary Γ_N, and in all element boundaries along the Dirichlet boundary Γ_D, respectively. In order to enforce the initial condition one must demand that

$$
(U_0, w) = (u_0, w) \qquad \forall\, w \in W_h, \qquad t = 0.
\tag{5.13}
$$

Note that using integration by parts for some of the terms in the second equation of (5.12),

the following expression holds,

$$(U, \nabla \cdot \tilde{\boldsymbol{v}}) - \langle \{U\}, [\![\tilde{\boldsymbol{v}}]\!] \rangle_{\varepsilon_i} - \langle U, \tilde{\boldsymbol{v}} \cdot \boldsymbol{n} \rangle_{\Gamma_N} = -(\nabla U, \boldsymbol{v}) + \langle [\![U]\!], \{\tilde{\boldsymbol{v}}\} \rangle_{\varepsilon_i} + \langle U, \tilde{\boldsymbol{v}} \cdot \boldsymbol{n} \rangle_{\Gamma_D} \quad (5.14)$$

Based on the previous observation we will rewrite the LDG formulation for the IBVP (4.4) as,

$$\begin{cases} (U_t, w) + (\boldsymbol{Q}, \nabla w) - \langle \{\boldsymbol{Q}\}, [\![w]\!] \rangle_{\varepsilon_i} + \langle \sigma[\![U]\!], [\![w]\!] \rangle_{\varepsilon_i} + \\ \quad - \langle B_N, w \rangle_{\Gamma_N} - \langle \boldsymbol{Q} \cdot \boldsymbol{n}, w \rangle_{\Gamma_D} + \langle \sigma(U - B_D), w \rangle_{\Gamma_D} = (f, w) \qquad \forall\, w \in W_h \\ \quad (\tilde{\boldsymbol{Q}}, \tilde{\boldsymbol{v}}) - (\nabla U, \tilde{\boldsymbol{v}}) + \langle [\![U]\!], \{\tilde{\boldsymbol{v}}\} \rangle_{\varepsilon_i} + \langle U, \tilde{\boldsymbol{v}} \cdot \boldsymbol{n} \rangle_{\Gamma_D} = \langle B_D, \tilde{\boldsymbol{v}} \cdot \boldsymbol{n} \rangle_{\Gamma_D} \quad \forall\, \tilde{\boldsymbol{v}} \in (W_h)^d \\ \qquad\qquad\qquad (a(U, \tilde{\boldsymbol{Q}})\tilde{\boldsymbol{Q}}, \boldsymbol{v}) - (\boldsymbol{Q}, \boldsymbol{v}) = 0 \qquad\qquad\qquad \forall\, \boldsymbol{v} \in (W_h)^d \end{cases}$$

$$(5.15)$$

Remark 5.2.2. It is clear that any continuous classical solution of problem (5.3)-(5.5) will satisfy problem (5.15) since all terms involving jumps across elements $[\![\cdot]\!]$, will be zero and all boundary terms will satisfy strongly the boundary conditions.

Remark 5.2.3. As in the numerical method presented in Chapter 4, the diffusion coefficient $a(u, \tilde{\boldsymbol{q}})$ in (5.4) will be approximated by the family of Lipschitz nondegenerate diffusion coefficients of the form

$$a_\epsilon(u, \tilde{\boldsymbol{q}}) = \frac{\beta_\epsilon(u)}{|\tilde{\boldsymbol{q}}|^{1-\gamma}}, \tag{5.16}$$

and we will denote with $\beta(\cdot)$ any member of the family $\{\beta_\epsilon(\cdot)\}$ in the subsequent analysis to simplify the notation. Furthermore, note that any solution of the IBVP (4.4), as introduced in Chapter 4, will also be a solution (5.3)-(5.5) with the regularized diffusion coefficient (5.16).

Remark 5.2.4. The system of nonlinear ordinary differential equations arising from substituting the expressions of $(U, \boldsymbol{Q}, \tilde{\boldsymbol{Q}})$ given by (5.7) in (5.15) will be assumed to have at least one solution. This assumption is based on the proof of existence of solutions of the related problem (4.12) provided in Theorem 4.3.1.

5.2.1 Stability analysis

Even though the proof of Theorem 5.2.1 can be established as a Corollary of Theorem 5.2.2, we present it here for clarity. Indeed, many of the mathematical manipulations presented in the proof of Theorem 5.2.1 can be easily followed and will be used in the more elaborate setting of the proof of Theorem 5.2.2.

Theorem 5.2.1 (Stability). *Let U and $\tilde{\boldsymbol{Q}}$ be solutions of (5.15) and (5.13) with $B_N = 0$, and $B_D = 0$. Then*

$$\|U(t)\|_{L^2(\Omega)}^2 + \|\sigma^{\frac{1}{2}}[\![U]\!]\|_{L^2(\varepsilon_i)}^2 + \|\sigma^{\frac{1}{2}}U\|_{L^2(\Gamma_D)}^2 + \|\tilde{\boldsymbol{Q}}\|_{L^{1+\gamma}(0,T,L^{1+\gamma}(\Omega))}^{1+\gamma} \le$$
$$C\left(\|u_0\|_{L^2(\Omega)}^2, \|f\|_{L^2(0,T;L^2(\Omega))}^2\right). \quad (5.17)$$

Proof. Note that choosing $w = U$, $\tilde{\boldsymbol{v}} = \boldsymbol{Q}$, and $\boldsymbol{v} = \tilde{\boldsymbol{Q}}$ and adding all terms on the left hand side in (5.15) we obtain, after several cancellations

$$\frac{1}{2}\frac{\partial}{\partial t}\|U(t)\|_{L^2(\Omega)}^2 + \langle\sigma[\![U]\!], [\![U]\!]\rangle_{\varepsilon_i} + \langle\sigma U, U\rangle_{\Gamma_D} + \left(a(U, \tilde{\boldsymbol{Q}})\tilde{\boldsymbol{Q}}, \tilde{\boldsymbol{Q}}\right) = (f, U) \quad (5.18)$$

From the observation that

$$\epsilon\|\tilde{\boldsymbol{Q}}\|_{L^{1+\gamma}(\Omega)}^{1+\gamma} \le \int_\Omega \beta(U)|\tilde{\boldsymbol{Q}}|^{1+\gamma} = \left(a(U, \tilde{\boldsymbol{Q}})\tilde{\boldsymbol{Q}}, \tilde{\boldsymbol{Q}}\right).$$

equation (5.18) leads to

$$\frac{1}{2}\frac{\partial}{\partial t}\|U(t)\|_{L^2(\Omega)}^2 + \|\sigma^{\frac{1}{2}}[\![U]\!]\|_{L^2(\varepsilon_i)}^2 + \|\sigma^{\frac{1}{2}}U\|_{L^2(\Gamma_D)}^2 + \epsilon\|\tilde{\boldsymbol{Q}}\|_{L^{1+\gamma}(\Omega)}^{1+\gamma} \le (f, U). \quad (5.19)$$

Furthermore, since

$$(f, U) \le \frac{1}{2}\|U(t)\|_{L^2(\Omega)}^2 + \frac{1}{2}\|f\|_{L^2(\Omega)}^2$$

equation (5.19) implies

$$\frac{1}{2}\frac{\partial}{\partial t}\|U(t)\|_{L^2(\Omega)}^2 + \|\sigma^{\frac{1}{2}}[\![U]\!]\|_{L^2(\varepsilon_i)}^2 + \frac{1}{2}\|\sigma^{\frac{1}{2}}U\|_{L^2(\Gamma_D)}^2 + \epsilon\|\tilde{\boldsymbol{Q}}\|_{L^{1+\gamma}(\Omega)}^{1+\gamma} \le \frac{1}{2}\|U(t)\|_{L^2(\Omega)}^2 + \frac{1}{2}\|f\|_{L^2(\Omega)}^2.$$
$$(5.20)$$

Since the second, third, and fourth terms of the left hand side of the previous equation are nonnegative we obtain

$$\frac{1}{2}\frac{\partial}{\partial t}\|U(t)\|_{L^2(\Omega)}^2 \le \frac{1}{2}\|U(t)\|_{L^2(\Omega)}^2 + \frac{1}{2}\|f\|_{L^2(\Omega)}^2$$

which, by Gronwall's Lemma, leads to

$$\|U(t)\|_{L^2(\Omega)}^2 \le C\left(\|U_0\|_{L^2(\Omega)}^2, \|f\|_{L^2(0,T;L^2(\Omega))}^2\right) \quad \text{for all } t \in [0, T]$$

Integrating (5.20) in time from 0 to T, the following must also hold:

$$\|\tilde{\boldsymbol{Q}}\|_{L^{1+\gamma}(0,T,L^{1+\gamma}(\Omega))}^{1+\gamma} \le C\left(\|U_0\|_{L^2(\Omega)}^2, \|f\|_{L^2(0,T;L^2(\Omega))}^2\right) \quad (5.21)$$

Likewise for the second and third terms of (5.20). Finally, by choosing $w = U_0$ in the first equation of (5.13) we obtain

$$(U_0, U_0) = (u_0, U_0) \leq \frac{1}{2}\|U_0\|_{L^2(\Omega)}^2 + \frac{1}{2}\|u_0\|_{L^2(\Omega)}^2$$

which implies

$$\|U_0\|_{L^2(\Omega)}^2 \leq \|u_0\|_{L^2(\Omega)}^2$$

Thus, the result of the Theorem follows at once. $\qquad\square$

5.2.2   Continuous in time *a priori* error analysis

In this section we will study how close (possibly nonunique) solutions to the LDG approximation problem (5.15), U, are to the true weak solution u of problem (5.3)-(5.5) with the regularized diffusion coefficient (5.16). We will focus the present analysis to the case when $\nabla u \in L^\infty(0, T, L^\infty(\Omega))$. In the following paragraphs we will assume that u is regular enough so that:

(i) the continuous interpolant of u in $\mathcal{P}^k$, $\hat{u}$, is well defined and satisfies

$$\int_\Omega (u - \hat{u})w = 0 \quad \forall\, w \in \mathcal{P}^k \tag{5.22}$$

(ii) the L^2-projection of $\boldsymbol{q}$ in $(\mathcal{P}^k)^d$, $\hat{\boldsymbol{q}}$, is well defined and given by

$$\int_\Omega (\boldsymbol{q} - \hat{\boldsymbol{q}})\boldsymbol{v} = 0 \quad \forall\, \boldsymbol{v} \in (\mathcal{P}^k)^d \tag{5.23}$$

(iii) inside every element Ω_e,

$$\hat{\tilde{\boldsymbol{q}}} = \nabla \hat{u}. \tag{5.24}$$

Further assumptions include that $\beta_\epsilon(\hat{u}), \beta_\epsilon(U), \nabla \hat{u}, \nabla U \in L^\infty(0, T, L^\infty(\Omega))$.

Theorem 5.2.2. *Let $(u, \tilde{\boldsymbol{q}}, \boldsymbol{q}) \in L^\infty(\Omega) \times (L^\infty(\Omega))^d \times (L^\infty(\Omega))^d$ be the solution of problem (5.3)-(5.5) and let $U, \tilde{\boldsymbol{Q}}, \boldsymbol{Q}$ be a solution of problem (5.15). Let $\chi_u = u - \hat{u}$, $\chi_{\boldsymbol{q}} = \tilde{\boldsymbol{q}} - \hat{\tilde{\boldsymbol{q}}}$*

and $\chi_{\boldsymbol{q}} = \boldsymbol{q} - \hat{\boldsymbol{q}}$. *Further, assume that* $\nabla \hat{u}, \tilde{\boldsymbol{Q}} \in L^\infty(0, T, L^\infty(\Omega))$. *Then for all* $t \in [0, T]$

$$\|U(t) - u(t)\|^2_{L^2(\Omega)} + \|\tilde{\boldsymbol{Q}} - \tilde{\boldsymbol{q}}\|^2_{L^2(0,T,L^2(\Omega))} +$$
$$+ \int_0^T \left(\|\sigma^{\frac{1}{2}} [\![U(t) - u(t)]\!] \|^2_{L^2(\varepsilon_i)} + \|\sigma^{\frac{1}{2}} (U(t) - u(t)) \|^2_{L^2(\Gamma_D)} \right) \leq$$
$$\|\chi_u(t)\|^2_{L^2(\Omega)} + \|\sigma^{\frac{1}{2}} \chi_u\|^2_{L^2(\Gamma_D)} + \|\tilde{\chi}_{\boldsymbol{q}}\|^2_{L^2(\Omega)} + C \left(\|\xi_u(0)\|^2_{L^2(\Omega)} + \|\chi_u\|^2_{L^2(0,T,L^2(\Omega))} + \right.$$
$$\left. + \int_0^T \left(\|\sigma^{-\frac{1}{2}} \{ \chi_{\boldsymbol{q}} \} \|^2_{L^2(\varepsilon_i)} + \|\sigma^{\frac{1}{2}} \chi_u\|^2_{L^2(\Gamma_D)} \right) + \int_0^T \int_\Omega |\tilde{\chi}_{\boldsymbol{q}}|^{2\gamma} \right). \quad (5.25)$$

Proof. Since any solution u of (5.3)-(5.5) satisfies the weak form (5.15), the following three equations must hold:

$$(U_t - \hat{u}_t, w) + (\boldsymbol{Q} - \hat{\boldsymbol{q}}, \nabla w) - \langle \{ \boldsymbol{Q} - \hat{\boldsymbol{q}} \}, [\![w]\!] \rangle_{\varepsilon_i} + \langle \sigma [\![U - \hat{u}]\!], [\![w]\!] \rangle_{\varepsilon_i} +$$
$$- \langle (\boldsymbol{Q} - \hat{\boldsymbol{q}}) \cdot \boldsymbol{n}, w \rangle_{\Gamma_D} + \langle \sigma(U - \hat{u}), w \rangle_{\Gamma_D}$$
$$= (u_t - \hat{u}_t, w) + (\boldsymbol{q} - \hat{\boldsymbol{q}}, \nabla w) - \langle \{ \boldsymbol{q} - \hat{\boldsymbol{q}} \}, [\![w]\!] \rangle_{\varepsilon_i} + \langle \sigma [\![u - \hat{u}]\!], [\![w]\!] \rangle_{\varepsilon_i} +$$
$$- \langle (\boldsymbol{q} - \hat{\boldsymbol{q}}) \cdot \boldsymbol{n}, w \rangle_{\Gamma_D} + \langle \sigma(u - \hat{u}), w \rangle_{\Gamma_D} \ , \quad (5.26)$$

$$(\tilde{\boldsymbol{Q}} - \hat{\tilde{\boldsymbol{q}}}, \tilde{\boldsymbol{v}}) - (\nabla(U - \hat{u}), \tilde{\boldsymbol{v}}) + \langle [\![U - \hat{u}]\!], \{ \tilde{\boldsymbol{v}} \} \rangle_{\varepsilon_i} + \langle U - \hat{u}, \tilde{\boldsymbol{v}} \cdot \boldsymbol{n} \rangle_{\Gamma_D}$$
$$= (\tilde{\boldsymbol{q}} - \hat{\tilde{\boldsymbol{q}}}, \tilde{\boldsymbol{v}}) - (\nabla(u - \hat{u}), \tilde{\boldsymbol{v}}) + \langle [\![u - \hat{u}]\!], \{ \tilde{\boldsymbol{v}} \} \rangle_{\varepsilon_i} + \langle u - \hat{u}, \tilde{\boldsymbol{v}} \cdot \boldsymbol{n} \rangle_{\Gamma_D} \ , \quad (5.27)$$

and

$$\left(\beta(U) \left(\frac{\tilde{\boldsymbol{Q}}}{|\tilde{\boldsymbol{Q}}|^{1-\gamma}} - \frac{\hat{\tilde{\boldsymbol{q}}}}{|\hat{\tilde{\boldsymbol{q}}}|^{1-\gamma}} \right), \boldsymbol{v} \right) - (\boldsymbol{Q} - \hat{\boldsymbol{q}}, \boldsymbol{v}) = \left(\beta(u) \left(\frac{\tilde{\boldsymbol{q}}}{|\tilde{\boldsymbol{q}}|^{1-\gamma}} - \frac{\hat{\tilde{\boldsymbol{q}}}}{|\hat{\tilde{\boldsymbol{q}}}|^{1-\gamma}} \right), \boldsymbol{v} \right) +$$
$$- (\boldsymbol{q} - \hat{\boldsymbol{q}}, \boldsymbol{v}) - \left((\beta(U) - \beta(u)) \frac{\hat{\tilde{\boldsymbol{q}}}}{|\hat{\tilde{\boldsymbol{q}}}|^{1-\gamma}}, \boldsymbol{v} \right). \quad (5.28)$$

Note that the second term of the right hand side of (5.26) and the second term of the right hand side of (5.28) are zero since $\hat{\boldsymbol{q}}$ satisfies (5.23). Thus,

$$(\boldsymbol{q} - \hat{\boldsymbol{q}}, \nabla w) = 0 \qquad \text{and} \qquad (\boldsymbol{q} - \hat{\boldsymbol{q}}, \boldsymbol{v}) = 0$$

To simplify notation, let $\xi_u = U - \hat{u}$, $\tilde{\boldsymbol{\xi}}_{\boldsymbol{q}} = \tilde{\boldsymbol{Q}} - \hat{\tilde{\boldsymbol{q}}}$ and $\boldsymbol{\xi}_{\boldsymbol{q}} = \boldsymbol{Q} - \hat{\boldsymbol{q}}$. Now, choosing $w = \xi_u$, $\tilde{\boldsymbol{v}} = \boldsymbol{\xi}_{\boldsymbol{q}}$, and $\boldsymbol{v} = \tilde{\boldsymbol{\xi}}_{\boldsymbol{q}}$, and adding equations (5.26),(5.27), and (5.28) we obtain, after multiple cancellations:

86

$$\frac{1}{2}\frac{\partial}{\partial t}\|\xi_u(t)\|^2_{L^2(\Omega)} \quad + \quad \langle\sigma[\![\xi_u]\!],[\![\xi_u]\!]\rangle_{\varepsilon_i} + \langle\sigma\xi_u,\xi_u\rangle_{\Gamma_D} + \left(\beta(U)\left(\frac{\tilde{Q}}{|\tilde{Q}|^{1-\gamma}} - \frac{\hat{\tilde{q}}}{|\hat{\tilde{q}}|^{1-\gamma}}\right),\tilde{\xi}_q\right) =$$

$$((\chi_u)_t,\xi_u) - \langle\{\chi_q\},[\![\xi_u]\!]\rangle_{\varepsilon_i} + \langle\sigma[\![\chi_u]\!],[\![\xi_u]\!]\rangle_{\varepsilon_i} +$$

$$+ \quad (\tilde{\chi}_q,\xi_q) - (\nabla\chi_u,\xi_q) + \langle[\![\chi_u]\!],\{\xi_q\}\rangle_{\varepsilon_i} + \langle\sigma\chi_u,\xi_u\rangle_{\Gamma_D}$$

$$+ \quad \left(\beta(u)\left(\frac{\tilde{q}}{|\tilde{q}|^{1-\gamma}} - \frac{\hat{\tilde{q}}}{|\hat{\tilde{q}}|^{1-\gamma}}\right),\tilde{\xi}_q\right) - \left((\beta(U)-\beta(u))\frac{\hat{\tilde{q}}}{|\hat{\tilde{q}}|^{1-\gamma}},\tilde{\xi}_q\right)\hspace{1em}(5.29)$$

Furthermore, from the result of Lemma 4.2.2 and provided $\beta(U) \geq \epsilon > 0$ the following must hold

$$\gamma\epsilon\,\mathcal{A}\|\tilde{\xi}_q\|^2_{L^2(\Omega)} \leq \left(\beta(U)\left(\frac{\tilde{Q}}{|\tilde{Q}|^{1-\gamma}} - \frac{\hat{\tilde{q}}}{|\hat{\tilde{q}}|^{1-\gamma}}\right),\tilde{Q}-\hat{\tilde{q}}\right) \hspace{2em}(5.30)$$

where $\mathcal{A} := \inf\limits_{(0,T)\times\Omega}(\mathcal{A}_0) = \inf\limits_{(0,T)\times\Omega} 1/\left(\|\tilde{Q}\|_{L^\infty(0,T,L^\infty(\Omega))} + \|\hat{\tilde{q}}\|_{L^\infty(0,T,L^\infty(\Omega))}\right)^{1-\gamma}$

Using the triangle inequality in (5.29) and using the previous inequality, we can establish that

$$\frac{1}{2}\frac{\partial}{\partial t}\|\xi_u(t)\|^2_{L^2(\Omega)} + \|\sigma^{\frac{1}{2}}[\![\xi_u]\!]\|^2_{L^2(\varepsilon_i)} + \|\sigma^{\frac{1}{2}}\xi_u\|^2_{L^2(\Gamma_D)} + \gamma\epsilon\,\mathcal{A}\|\tilde{\xi}_q\|^2_{L^2(\Omega)} \leq \sum_{i=1}^{10} T_i \hspace{1em}(5.31)$$

where T_i, $i = 1,...,9$, are the terms arising from the right hand side of (5.29). We now proceed to bound the terms $|T_i|$ for $i = 1,...,9$.

Note that T_1 is zero by (5.22). For T_2, we multiply and divide by $\sigma^{\frac{1}{2}}$ to get

$$T_2 = \langle\sigma^{-\frac{1}{2}}\{\chi_q\},\sigma^{\frac{1}{2}}[\![\xi_u]\!]\rangle_{\varepsilon_i} \leq \frac{\epsilon_1}{2}\|\sigma^{\frac{1}{2}}[\![\xi_u]\!]\|^2_{L^2(\varepsilon_i)} + \frac{1}{2\epsilon_1}\|\sigma^{-\frac{1}{2}}\{\chi_q\}\|^2_{L^2(\varepsilon_i)} \hspace{1em}(5.32)$$

For T_3

$$T_3 = \langle\sigma^{\frac{1}{2}}[\![\chi_u]\!],\sigma^{\frac{1}{2}}[\![\xi_u]\!]\rangle_{\varepsilon_i} \leq \frac{1}{2}\|\sigma^{\frac{1}{2}}[\![\xi_u]\!]\|^2_{L^2(\varepsilon_i)} + \frac{1}{2}\|\sigma^{\frac{1}{2}}[\![\chi_u]\!]\|^2_{L^2(\varepsilon_i)} \hspace{1em}(5.33)$$

Note that the choice of $\hat{\tilde{q}}$ and $\hat{u}$ in (5.24) and (5.22), respectively, and the fact that $\xi_q \in \mathcal{P}^k$ ensures the three following equalities:

$$\|\sigma^{\frac{1}{2}}[\![\chi_u]\!]\|^2_{L^2(\varepsilon_i)} = 0, \qquad T_6 = \langle[\![\chi_u]\!],\{\xi_q\}\rangle_{\varepsilon_i} = 0, \hspace{2em}(5.34)$$

and

$$T_4 = (\tilde{\chi}_q,\xi_q) = (\nabla\chi_u,\xi_q) = T_5. \hspace{3em}(5.35)$$

Thus, $T_4 - T_5 = 0$ and $T_6 = 0$.

For term T_7 we have the following inequality

$$T_7 = \langle \sigma \chi_u, \xi_u \rangle_{\Gamma_D} \leq \frac{1}{2} \|\sigma^{\frac{1}{2}} \chi_u\|^2_{L^2(\Gamma_D)} + \frac{1}{2}\|\sigma^{\frac{1}{2}}\xi_u\|^2_{L^2(\Gamma_D)} \tag{5.36}$$

As for terms T_8 and T_9, note that

$$
\begin{aligned}
T_8 = \int_\Omega \beta(u) \left(\frac{\tilde{q}}{|\tilde{q}|^{1-\gamma}} - \frac{\hat{\tilde{q}}}{|\hat{\tilde{q}}|^{1-\gamma}} \right) \tilde{\xi}_q
&\leq \frac{2}{\gamma} M \int_\Omega |\tilde{\chi}_q|^\gamma |\tilde{\xi}_q| \\
&\leq \frac{2}{\gamma} M \left(\frac{1}{2\epsilon_2} \int_\Omega |\tilde{\chi}_q|^{2\gamma} + \frac{\epsilon_2}{2}\|\tilde{\xi}_q\|^2_{L^2(\Omega)} \right)
\end{aligned}
\tag{5.37}
$$

and for $T_9 = \left((\beta(U) - \beta(u)) \dfrac{\hat{\tilde{q}}}{|\hat{\tilde{q}}|^{1-\gamma}}, \tilde{\xi}_q \right)$ we have

$$
\begin{aligned}
T_9 &\leq M^* \int_\Omega |u - U||\hat{\tilde{q}}|^\gamma |\tilde{\xi}_q| \\
&\leq M^* \|\hat{\tilde{q}}\|^\gamma_{L^\infty(\Omega)} \|u - U\|_{L^2(\Omega)} \|\tilde{\xi}_q\|_{L^2(\Omega)} \\
&\leq M^* \|\hat{\tilde{q}}\|^\gamma_{L^\infty(\Omega)} \left(\frac{1}{2\epsilon_3}(\|\chi_u\|^2_{L^2(\Omega)} + \|\xi_u\|^2_{L^2(\Omega)}) + \frac{\epsilon_3}{2}\|\tilde{\xi}_q\|^2_{L^2(\Omega)} \right),
\end{aligned}
\tag{5.38}
$$

where $M = \|\beta(u)\|_{L^\infty(\Omega)}$ and $M^* = \alpha \max(\|\beta(u)\|_{L^\infty(\Omega)}, \|\beta(U)\|_{L^\infty(\Omega)})^{\alpha-1}$.

From (5.31) and (5.32), (5.33), (5.35), (5.34), (5.36), (5.37), and (5.38) and choosing ϵ_1, ϵ_2 and ϵ_3 small enough so that for $\bar{\epsilon}$ and ϵ^* small positive numbers, $0 < \bar{\epsilon} \leq \gamma \epsilon \, \mathcal{A} - \left(\frac{1}{\gamma} M \epsilon_2 + \frac{1}{2} M^* \|\hat{\tilde{q}}\|^\gamma_{L^\infty} \epsilon_3 \right)$ and $0 < \epsilon^* \leq \frac{1}{2}(1 - \epsilon_1)$, we obtain

$$
\begin{aligned}
\frac{1}{2}\frac{\partial}{\partial t}\|\xi_u(t)\|^2_{L^2(\Omega)} + \epsilon^*\|\sigma^{\frac{1}{2}}[\![\xi_u]\!]\|^2_{L^2(\varepsilon_i)} + \frac{1}{2}\|\sigma^{\frac{1}{2}}\xi_u\|^2_{L^2(\Gamma_D)} + \bar{\epsilon}\,\|\tilde{\xi}_q\|^2_{L^2(\Omega)} &\leq \\
C \Big(\|\xi_u\|^2_{L^2(\Omega)} + \|\chi_u\|^2_{L^2(\Omega)} + & \\
+ \|\sigma^{-\frac{1}{2}}\{\tilde{\chi}_q\}\|^2_{L^2(\varepsilon_i)} + \|\sigma^{\frac{1}{2}}\chi_u\|^2_{L^2(\Gamma_D)} + \int_\Omega |\tilde{\chi}_q|^{2\gamma} \Big)
\end{aligned}
\tag{5.39}
$$

Since the second, third, and fourth terms of the left hand side in the previous inequality are nonnegative, we can use Gronwall's Lemma to find that for all $t \in [0, T]$,

$$
\begin{aligned}
\|\xi_u(t)\|^2_{L^2(\Omega)} \leq C \Big(\|\xi_u(0)\|^2_{L^2(\Omega)} + \|\chi_u\|^2_{L^2(0,T,L^2(\Omega))} + & \\
+ \int_0^T \left(\|\sigma^{-\frac{1}{2}}\{\tilde{\chi}_q\}\|^2_{L^2(\varepsilon_i)} + \|\sigma^{\frac{1}{2}}\chi_u\|^2_{L^2(\Gamma_D)} \right) + \int_0^T \int_\Omega |\tilde{\chi}_q|^{2\gamma} \Big)
\end{aligned}
\tag{5.40}
$$

Likewise, integrating (5.39) in time from 0 to T, we can establish the boundedness of the three remaining terms of the left hand side of (5.39):

$$\int_0^T \left(\|\sigma^{\frac{1}{2}} [\![\xi_u]\!]\|_{L^2(\varepsilon_i)}^2 + \|\sigma^{\frac{1}{2}}\xi_u\|_{L^2(\Gamma_D)}^2 \right) \quad \text{and} \quad \|\tilde{\boldsymbol{\xi}}_{\boldsymbol{q}}\|_{L^2(0,T,L^2(\Omega))}^2. \tag{5.41}$$

Finally, observe that

$$
\begin{aligned}
\|U(t) - u(t)\|_{L^2(\Omega)}^2 &\leq \|\xi_u(t)\|_{L^2(\Omega)}^2 + \|\chi_u(t)\|_{L^2(\Omega)}^2, \\
\|\tilde{\boldsymbol{Q}} - \tilde{\boldsymbol{q}}\|_{L^2(\Omega)}^2 &\leq \|\tilde{\boldsymbol{\xi}}_{\boldsymbol{q}}\|_{L^2(\Omega)}^2 + \|\tilde{\boldsymbol{\chi}}_{\boldsymbol{q}}\|_{L^2(\Omega)}^2, \\
\|\sigma^{\frac{1}{2}} \left(U(t) - u(t)\right)\|_{L^2(\Gamma_D)}^2 &\leq \|\sigma^{\frac{1}{2}}\xi_u\|_{L^2(\Gamma_D)}^2 + \|\sigma^{\frac{1}{2}}\chi_u\|_{L^2(\Gamma_D)}^2
\end{aligned}
$$

and
$$\|\sigma^{\frac{1}{2}}[\![U(t) - u(t)]\!]\|_{L^2(\varepsilon_i)}^2 = \|\sigma^{\frac{1}{2}}[\![\xi_u]\!]\|_{L^2(\varepsilon_i)}^2, \quad \text{since} \quad \|\sigma^{\frac{1}{2}}[\![u(t)]\!]\|_{L^2(\varepsilon_i)}^2 = 0.$$

The result of the Theorem follows immediately. $\qquad\square$

Corollary 5.2.1. *If $u \in W^{1,\infty}(\Omega) \cap H^{k+1}(\Omega)$, $\tilde{\boldsymbol{q}}, \boldsymbol{q} \in L^\infty(\Omega)$ are the solution of problem (5.4) and $U, \tilde{\boldsymbol{Q}}, \boldsymbol{Q}$ are a solution of problem (5.15), constructed with piecewise polynomials of degree at most k. Then for all $t \in [0,T]$ and $\sigma = \frac{1}{h}$*

$$
\begin{aligned}
&\|U(t) - u(t)\|_{L^2(\Omega)} + \|\tilde{\boldsymbol{Q}}(t) - \tilde{\boldsymbol{q}}(t)\|_{L^2(0,T,L^2(\Omega))} + \\
&+ \int_0^T \frac{1}{h} \left(\|[\![U(t) - u(t)]\!]\|_{L^2(\varepsilon_i)} + \|\left(U(t) - u(t)\right)\|_{L^2(\Gamma_D)} \right) \leq C(T) h^{k\gamma} \left(\int_0^T \|u(t)\|_{H^{k+1}(\Omega)}^{2\gamma} \right)^{\frac{1}{2}}
\end{aligned}
\tag{5.42}
$$

Proof. Based on the result of Theorem 5.2.2, we need to show how all the terms of the right hand side of (5.25) are bounded. For such regularity on u, the approximation estimates shown in Lemma 2.4.1 must hold. Thus,

$$
\begin{aligned}
\|\chi_u(t)\|_{L^2(\Omega)}^2 &\leq C\, h^{2(k+1)} \|u\|_{H^{k+1}(\Omega)}^2, &\tag{5.43} \\
\|\tilde{\boldsymbol{\chi}}_{\boldsymbol{q}}\|_{[L^2(\Omega)]^2}^2 &\leq C\, h^{2k} \|u\|_{[H^{k+1}]^2(\Omega)}^2, &\tag{5.44} \\
\|(\chi_u)_t\|_{L^2(0,T,L^2(\Omega))}^2 &\leq C(T)\, h^{2(k+1)} \int_0^T \|u\|_{H^{k+1}(\Omega)}^2, &\tag{5.45}
\end{aligned}
$$

Using the approximation estimates from Lemma 2.5.1 the following must also hold:

$$\|\sigma^{\frac{1}{2}}\chi_u\|_{L^2(\Gamma_D)}^2 \leq \|\sigma^{\frac{1}{2}}\chi_u\|_{L^2(\partial\Omega)}^2 \leq \frac{1}{h} C\, h^{2k+1} \|u\|_{H^{k+1}(\Omega)}^2 \leq C\, h^{2k} \|u\|_{H^{k+1}(\Omega)}^2 \tag{5.46}$$

Note also that

$$\int_0^T \int_\Omega |\tilde{\chi}_q|^{2\gamma} \leq \int_0^T \left(\int_\Omega |\tilde{\chi}_q|^2 \right)^{\frac{2\gamma}{2}} |\Omega|^{1-\gamma} = \int_0^T \|\tilde{\chi}_q\|_{L^2(\Omega)}^{2\gamma} |\Omega|^{1-\gamma}$$
$$\leq C(T) h^{2k\gamma} \|u\|_{H^{k+1}(\Omega)}^{2\gamma} \qquad (5.47)$$

and using Lemma 2.5.1 for the union of elements,

$$\int_0^T \|\sigma^{-\frac{1}{2}}\{\tilde{\chi}_q\}\|_{L^2(\varepsilon_i)}^2 \leq \int_0^T h \left\| \left\{ \beta(u) \frac{\tilde{q}}{|\tilde{q}|^{1-\gamma}} - \hat{q} \right\} \right\|_{L^2(\varepsilon_i)}^2$$
$$\leq \int_0^T h \, h^{2k+1} \left\| \beta(u) \frac{\tilde{q}}{|\tilde{q}|^{1-\gamma}} \right\|_{H^{k+1}(\Omega)}^2$$
$$\leq \int_0^T h^{2(k+1)} \|\beta(u)\|_{L^\infty(\Omega)}^2 \||\tilde{q}|^\gamma\|_{H^{k+1}(\Omega)}^2 \qquad (5.48)$$

From the result of Theorem 5.2.2, and noting that the leading term of estimates (5.43)-(5.48) for small h is the one given by (5.47), the following must hold

$$\|U(t) - u(t)\|_{L^2(\Omega)}^2 + \|\tilde{Q}(t) - \tilde{q}(t)\|_{L^2(0,T,L^2(\Omega))}^2 +$$
$$+ \int_0^T \frac{1}{h} \left(\|[[U(t) - u(t)]]\|_{L^2(\varepsilon_i)}^2 + \|(U(t) - u(t))\|_{L^2(\Gamma_D)}^2 \right) \leq C(T) h^{2k\gamma} \int_0^T \|u(t)\|_{H^{k+1}(\Omega)}^{2\gamma}$$
$$(5.49)$$

The result of the Theorem follows from the observation that for nonnegative numbers p, q, r, s, f, the fact that $p^2 + q^2 + r^2 + s^2 \leq f^2$ implies that there exists a positive constant C such that $p + q + r + s \leq C f$. This concludes the proof. $\square$

5.3 Numerical Experiments. 2D

In this section, we present the results of some numerical experiments aimed at solving two ideal 2D problems: a dam break event, and flow in a channel with vegetation resulting from a dam break event. The main motivation to show these results is to provide the reader with convincing evidence that the DSW equation captures the physics of the aforementioned ideal problems. In fact, the setting of the simulated flow in a channel with vegetation was inspired by an actual experiment shown in [13]. In future work we intend to simulate the results of this real life experiment in order to verify that the computational results match the measurements presented in [13].
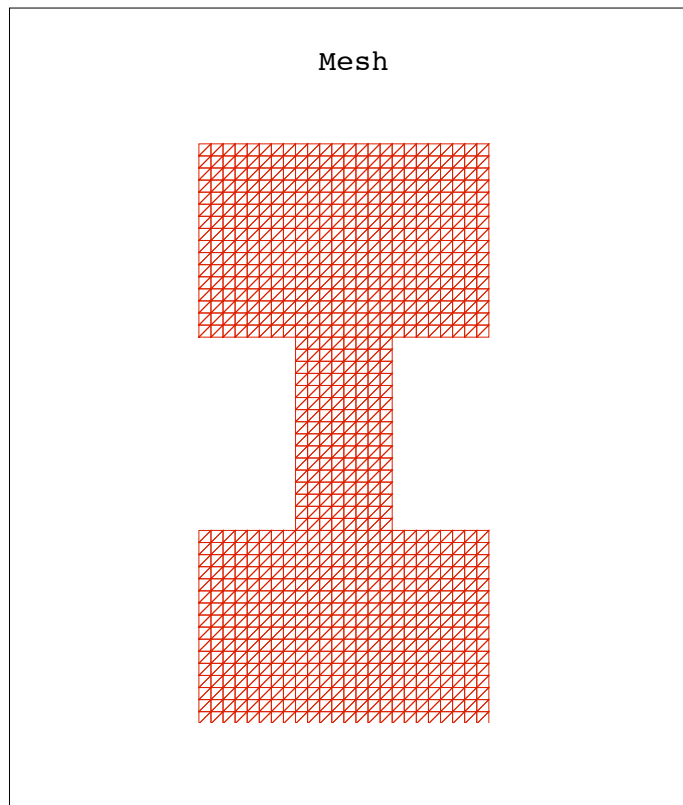
Mesh

FIGURE 5.1: *Dam break simulation. Mesh*

An LDG finite element formulation was coded in order to carry out the numerical experiments. This code was implemented in FORTRAN, based on a code originally designed by C. Dawson, to solve Richard's equation using an LDG formulation in 2D. A Picard iteration approach was used to solve the nonlinear problem (as discussed in section 4.4) and the conjugate gradient method was used to solve the resulting linear systems.

5.3.1 The Dam Break problem

In this section we present the results of 2-D simulations of the evolution of water depth profiles in an ideal dam break problem. This problem consists of simulating the water flow resulting from removing an ideal dam that keeps water on a confined area of the domain. The set up is as follows, a channel was designed to connect two reservoirs, one completely filled with water (up hill) and the other completely empty (down hill). The channel is considered to be dry at the beginning as well. When the ideal dam is removed from the upper reservoir, water is expected to flow downhill, flooding first the channel with a well defined front, and later flooding the lower reservoir; first with a well defined and radially symmetric front, and later filling it gradually. This process is expected to continue until all the water is transferred fully to the lower reservoir.

The units used in this ideal setting were meters for the water depth and height, and seconds for the time. This numerical experiment was computed in a domain with a uniform friction coefficient $c_f = 1$ (this value was chosen for simplicity and without any physical meaning) and with *zero* Newmann boundary conditions on $\partial\Omega$. The mesh of the computational domain is shown in Figures 5.1, the initial condition and water bed of this problem are presented at the top of Figure 5.2. The mesh radius is of the order $h \sim 0.125$ meters (in a domain with characteristic lengths of order $L \sim 6$ meters and $W \sim 3$ meters, respectively. See Figure 5.1 for a feel of the ratios) and the time step was comparable in size, *i.e.* $dt = 0.125$ seconds. The experiment was run from $t = 0.0$ seconds to $t = 70.0$ seconds. 3D and 2D views of the numerically simulated evolution of the water depth are presented in Figures 5.2, 5.3, 5.4. Side views of the evolution of the water depth profiles are shown in Figure 5.5. As discussed before, the main features of the phenomenon are captured, these include:

- The down-hill flow of water,

- The appearance of a flooding wave with a well defined front propagating in the direction of lowest potential energy points (lowest points in space), see Figures 5.2, 5.3, and 5.5,
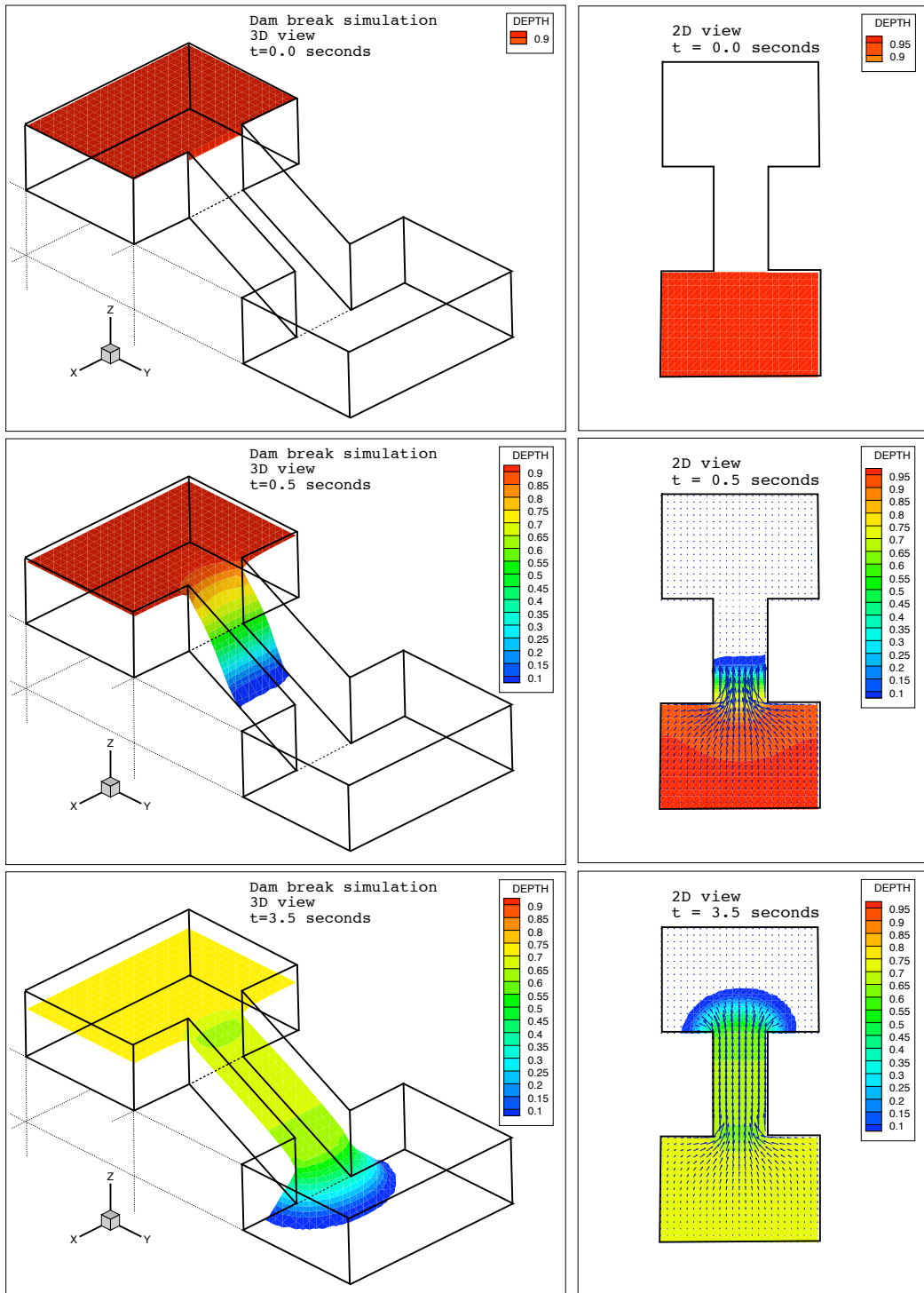
FIGURE 5.2: *Dam break simulation. Figures showing evolution of water depth (meters) at times=0, 0.5, and 3.5 seconds.* **(Left)** *3D views,* **(Right)** *2D views*
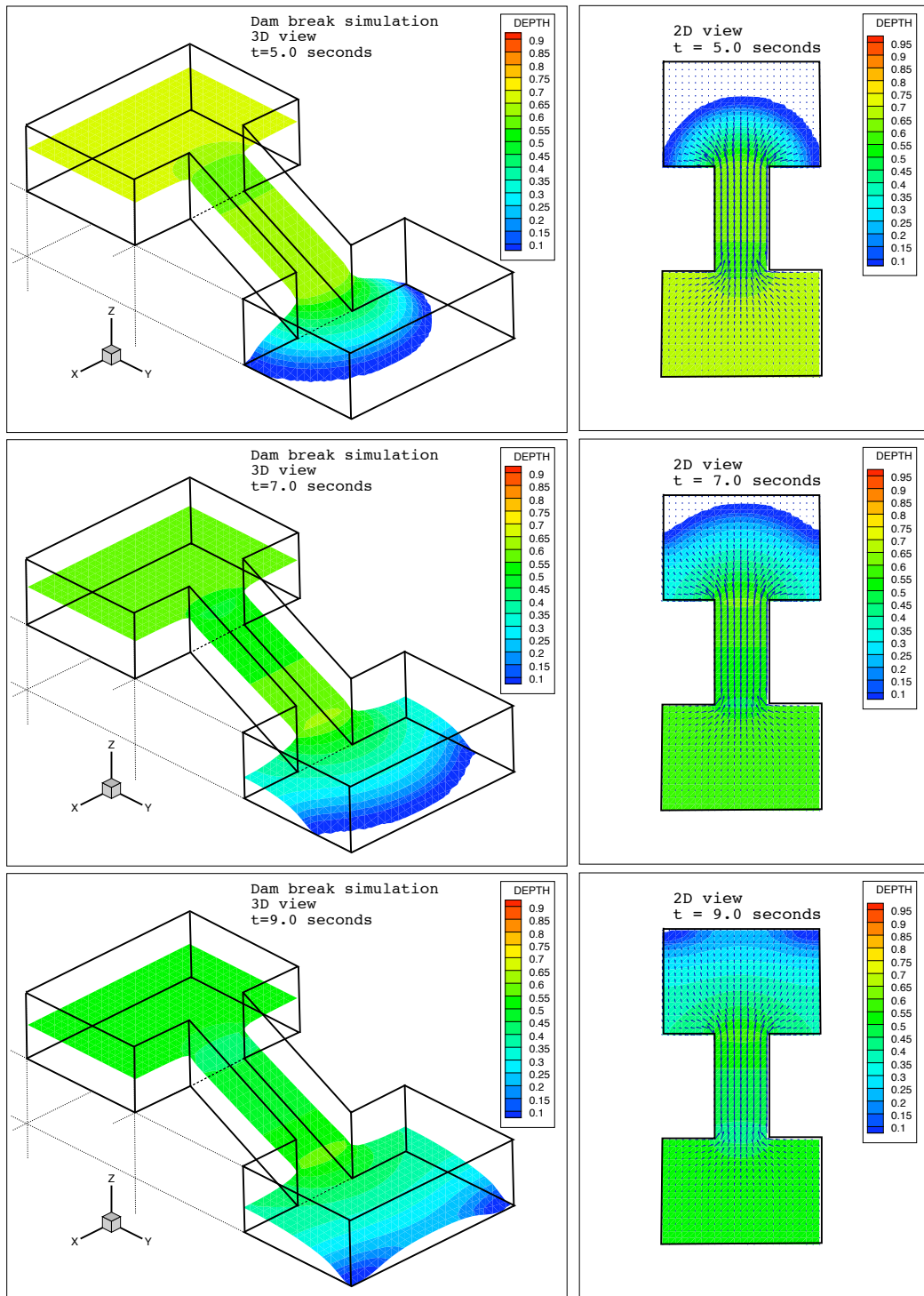
FIGURE 5.3: *Dam break simulation. Figures showing evolution of water depth (meters) at times=5.0, 7.0, and 9.0 seconds.* **(Left)** *3D views,* **(Right)** *2D views*
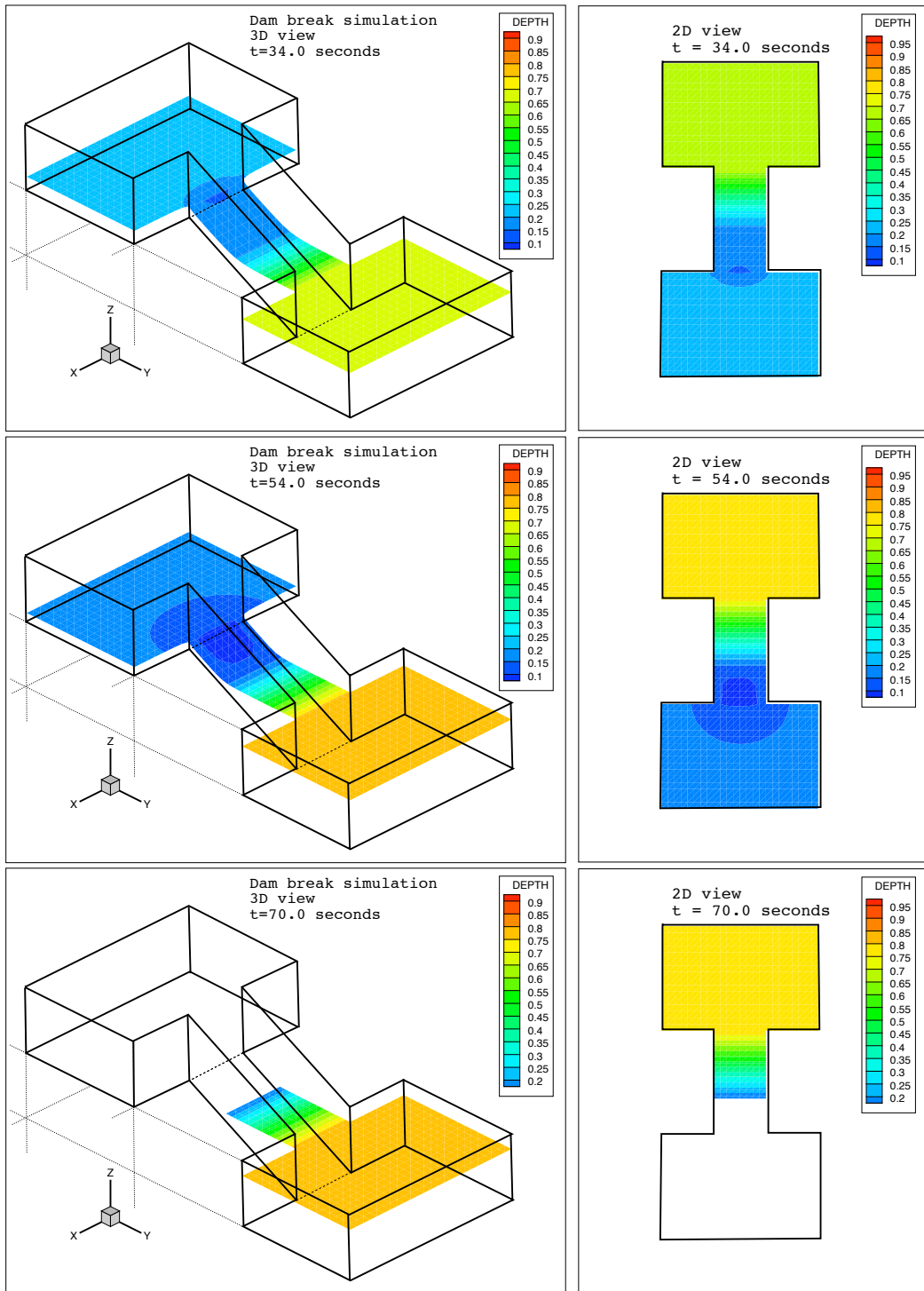
FIGURE 5.4: *Dam break simulation. Figures showing evolution of water depth (meters) at times=34.0, 54.0, and 70.0 seconds.* **(Left)** *3D views,* **(Right)** *2D views*
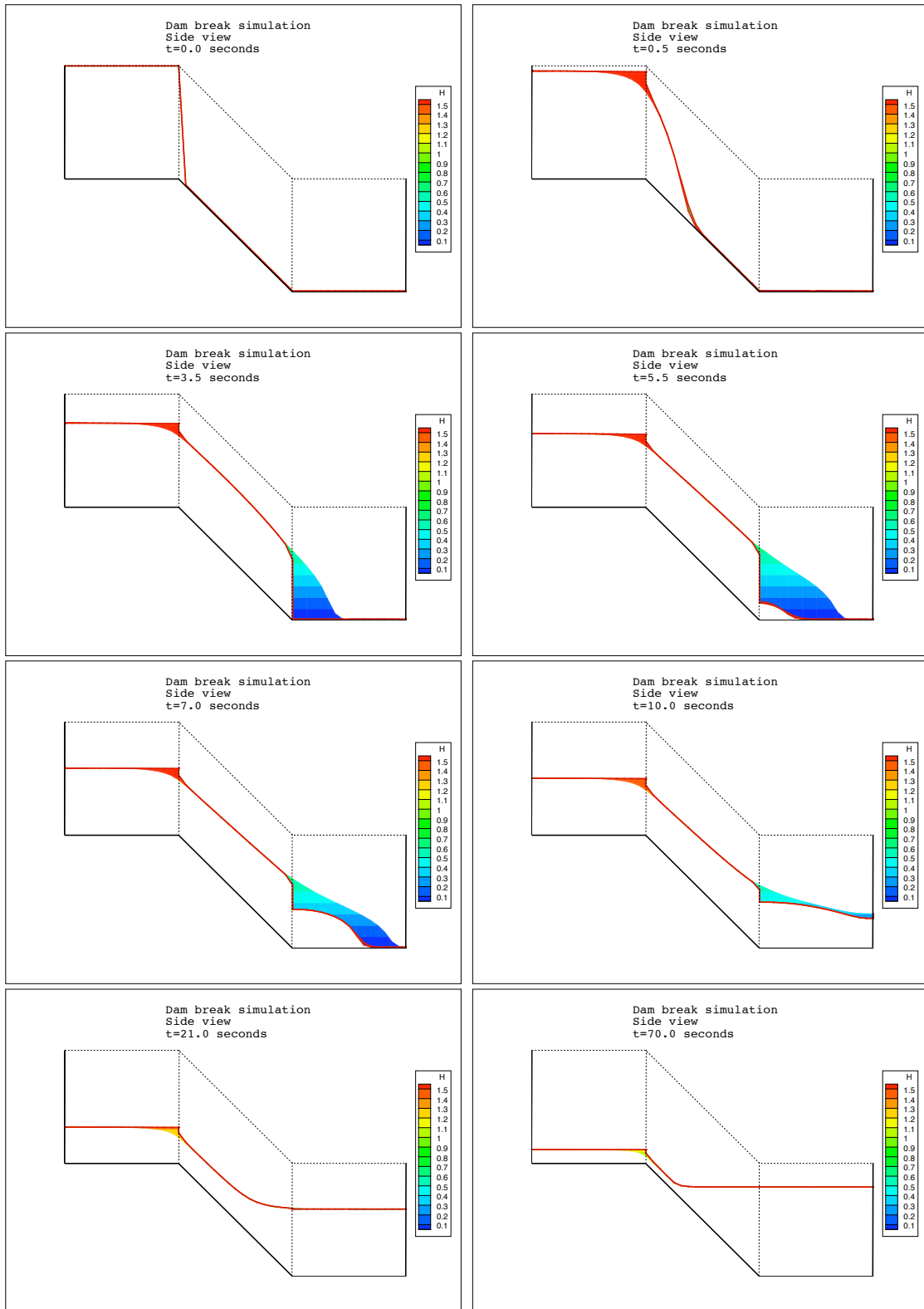
FIGURE 5.5: *Dam break simulation. Side views showing evolution of water height (meters) at different times.*

- The radial symmetry of the water flow both, at the entrance of the channel (uphill) as well as at the exit of the channel (downhill) throughout the event,

- The radial symmetry in the flooding front when reaching the lowest reservoir, see upper views of Figure 5.3, (this is a consequence of the previous observation),

- The eventually gradual transfer of water from the upper part to the lowest one.

Some of the characteristics of the phenomenon that are *not* captured are mostly related to two factors: the diffusive nature of the DSW equation, and the vertical integration utilized to derive it. Related to the first factor, the physical interaction of the water flow with the walls is not captured. For example, when water flows in a confined channel, ripples form as a consequence of momentum transfers between the water and the walls (as well as friction). Also, when water frontally hits a wall (as it happens in the lower views of Figure 5.3) water slushes and forms reflecting waves. These features are not present in the experiments we show. Another obvious characteristic *not* captured with the DSW equation as a model, and related to the second factor, is the vertical velocity profile of the water flow.

5.3.2 The Dam Break problem with vegetation

In this section we present the results of 2-D simulations of the evolution of water depth profiles in an ideal dam break problem with vegetation in some regions of the domain. This problem was inspired by the experimental setting shown in [13]. The numerical implementation was set up similarly to the one presented in section 5.3.1. The main difference consists of including three islands of vegetation in different locations of the domain. These vegetated regions, considered to have the same vegetation density, modify the water flow lines in the experimental setting of [13]. It is observed, as intuition would suggest, that water flows more rapidly away from them. Their inclusion in the numerical simulations is done only by assigning a higher value of the friction coefficient c_f inside these areas. Throughout the domain $c_f = 1$ and in the vegetated regions $c_f = 5$. The bathymetry remained the same as well as all the remaining computational variables presented in the dam break problem in section 5.3.1. Again, we chose to simulate the water flow resulting from removing an ideal dam that keeps water on a confined (uphill) area of the domain.

The mesh for this problem and the location of the islands of vegetation are shown in Figure 5.6. This experiment was run from $t = 0.0$ to $t = 70.0$ as well. However, since the most relevant features of this event take place before $t = 20.0$, only views for $t \in [0, 20]$ are presented. 3D and 2D views of the numerically simulated evolution of the water depth are presented in Figures 5.7, 5.8, 5.9.
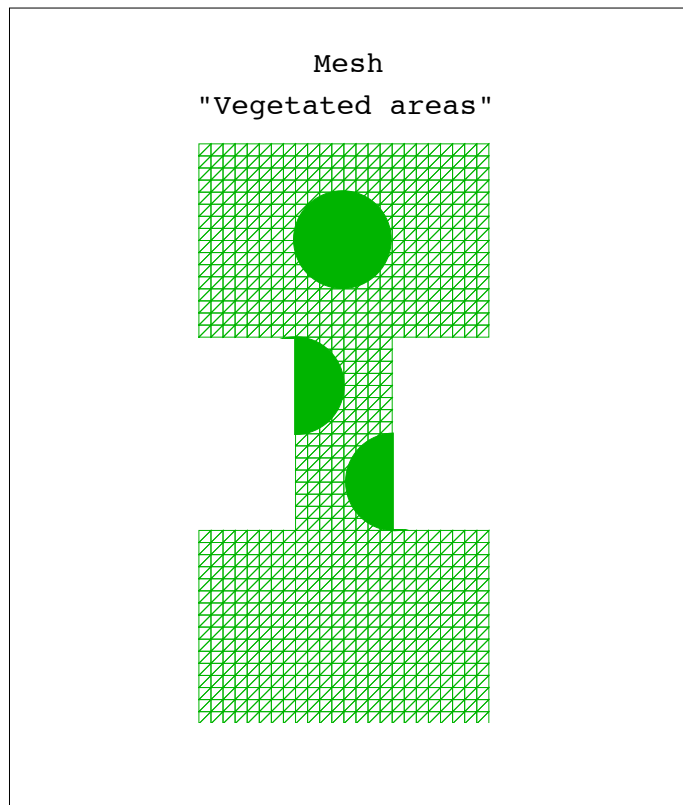
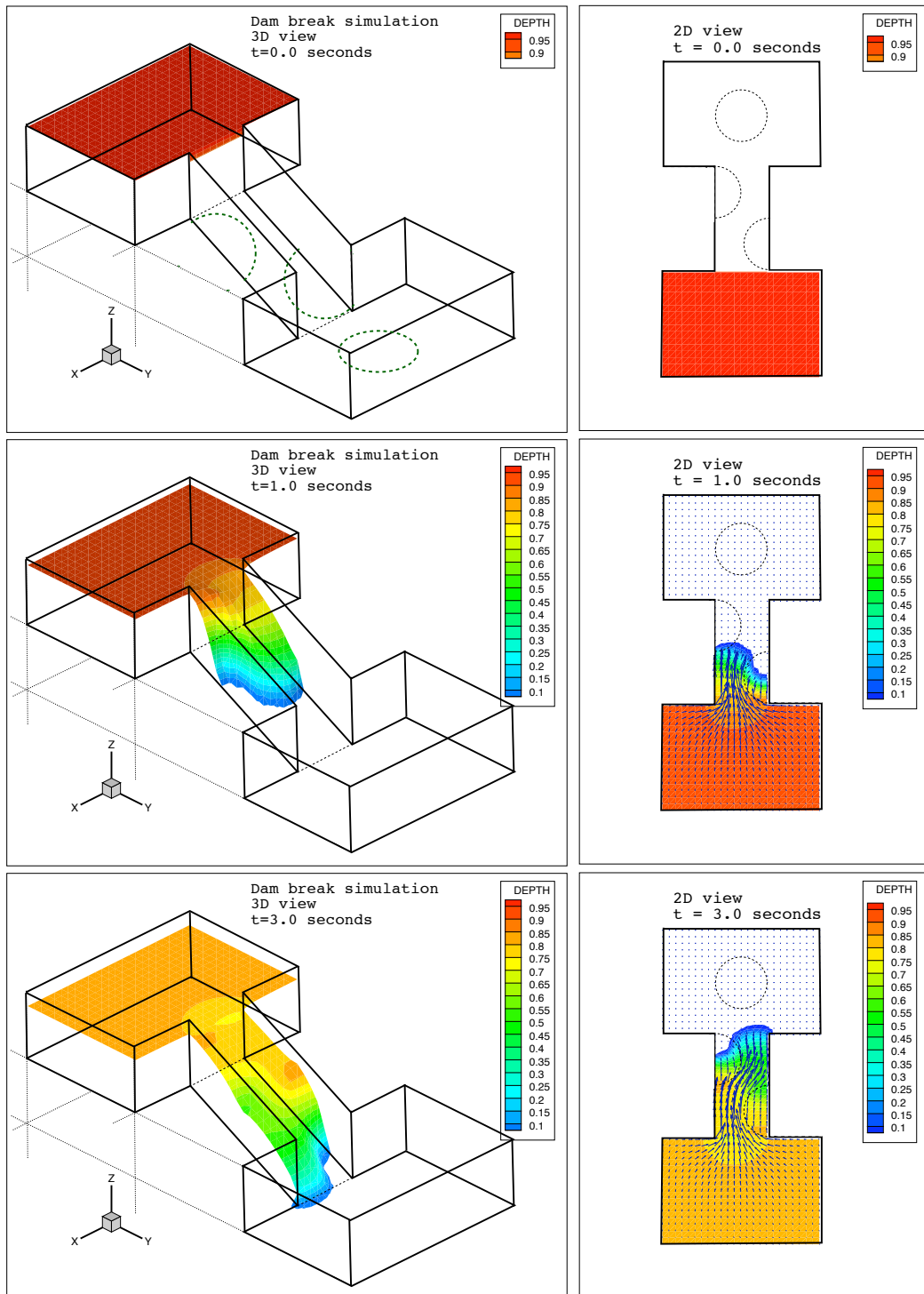Figure 5.6: *Dam break simulation with vegetation. Mesh*

FIGURE 5.7: *Dam break simulation with vegetation. Figures showing evolution of water depth (meters) at times=0, 1.0, and 3.0 seconds.* **(Left)** *3D views,* **(Right)** *2D views*
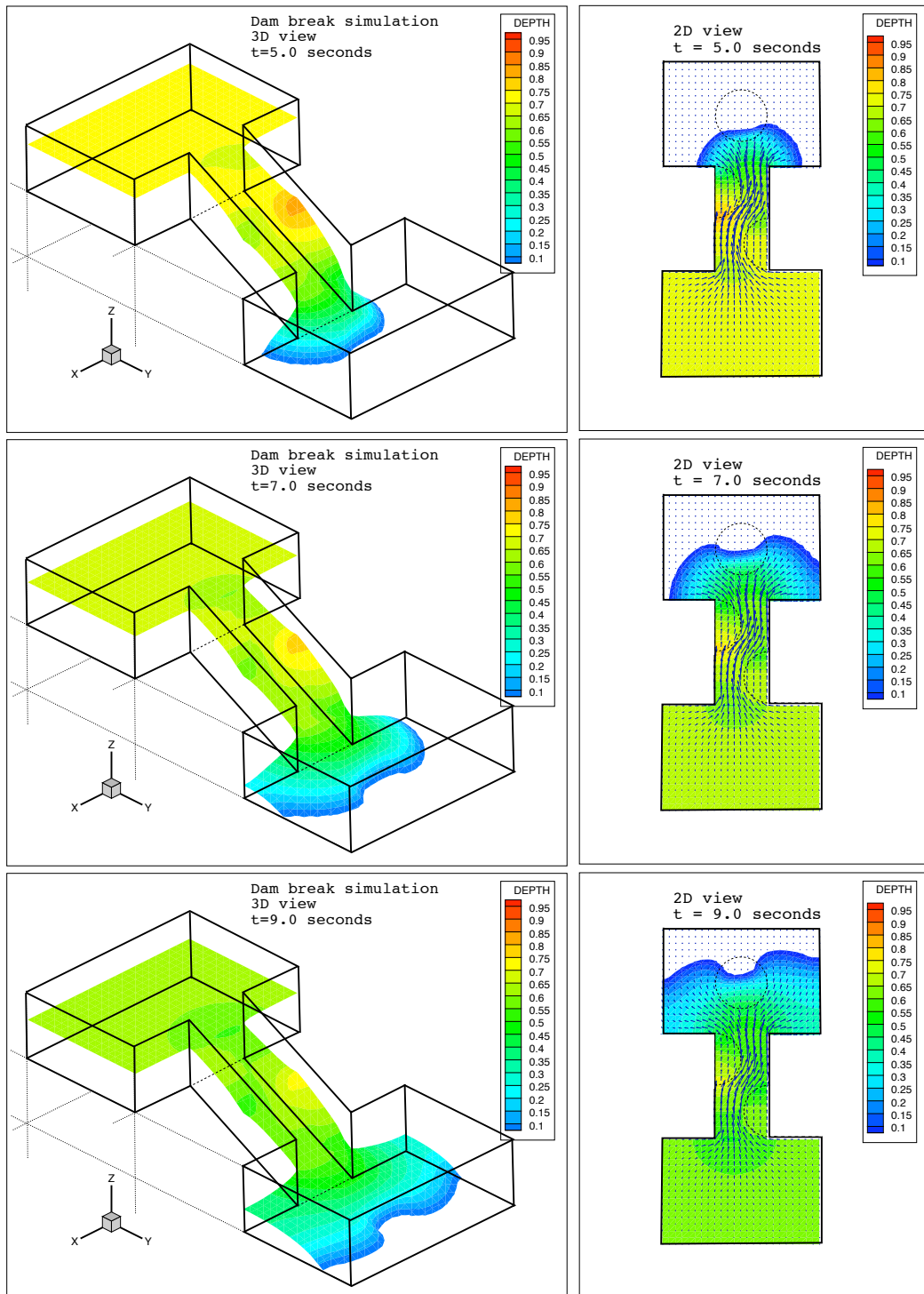
FIGURE 5.8: *Dam break simulation with vegetation. Figures showing evolution of water depth (meters) at times=5.0, 7.0, and 9.0 seconds. (**Left**) 3D views, (**Right**) 2D views*
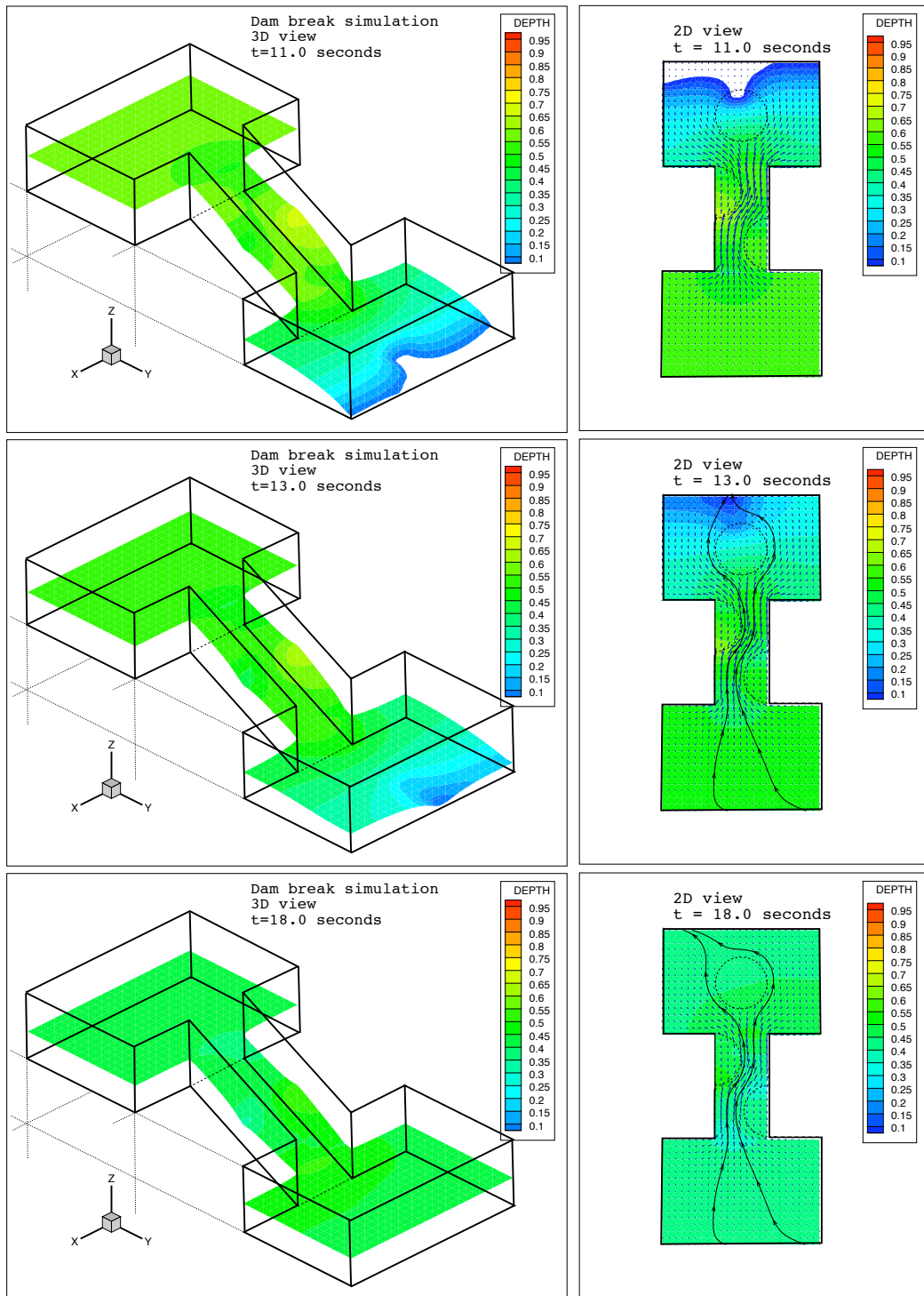
FIGURE 5.9: *Dam break simulation with vegetation. Figures showing evolution of water depth (meters) at times=11.0, 13.0, and 17.0 seconds.* **(Left)** *3D views,* **(Right)** *2D views*

Figures 5.7-5.9 show very good agreement with the expected features of the phenomenon. In particular, they clearly display the fact that, as expected, water flows more rapidly away from the vegetated areas. Also, the flooding front propagates throughout the domain in a way that qualitatively captures the expected dynamics. Again, the limitations of the DSW equation as a model appear as described in the previous section.

Chapter 6

Concluding Remarks and Future Work

In this dissertation, a mathematical model aimed at describing the dynamics of shallow water flow in environments such as wetlands is studied in depth. A concise derivation of this mathematical model is proposed in this work as a combination of, scaling arguments that characterize the *diffusive wave approximation of the shallow water equations* (DSW), and experimental studies aimed at understanding the dynamics of shallow water flow in vegetated areas. The emerging *initial boundary value problem* (IBVP), associated to this model, is stated in Chapter 2 as the IBVP (2.1).

The mathematical properties of the IBVP (2.1) are initially explored in Chapter 1, in the context of the theory of doubly nonlinear parabolic equations, by carefully identifying an alternative formulation of the IBVP (2.1) for flat topographies, namely the IBVP (1.7). In Chapter 3, a study of basic properties of nonnegative solutions for the DSW equation is presented in a hydrological context. Furthermore, an original proof of existence of weak solutions using constructive techniques that directly lead to the implementation of numerical algorithms to obtain approximate solutions is proposed. Also, alternative proofs of the most relevant results existing in the literature are presented. The open problem that arises, when topographic effects are considered (obstacle problem) in the DSW, is discussed and introduced as a new avenue of research in the area of theoretical PDEs.

In Chapter 4 the results of a numerical approach to study the properties of solutions of the IBVP (2.1) are shown. The emphasis is placed in analysing the mathematical properties of the partial differential equation appearing in the IBVP (2.1), using a regu-

larization technique introduced in Chapter 3. The analysis is carried out in order to find estimates for the error between numerical solutions, constructed using the Galerkin method, and true solutions of this equation. A proof that the numerical solutions converge to the true solution of the *regularized* DSW equation (4.4) under certain physically consistent conditions is presented. These conditions emerge from physical consistency and are given by $u \in L^\infty(0, T; L^\infty(\Omega))$ and $\nabla u \in L^\infty(0, T; L^\infty(\Omega))$. The absence of appropriate conditions leading to a proof to ensure uniqueness of solutions of the DSW equation, in its general form (2.1), imposes restrictions in the analysis shown, and as a consequence the analytical *a priori* error estimates presented in sections 4.3.3 and 4.3.4 are not optimal. In section 4.4, numerical evidence that shows that the proposed numerical method converges to the true weak solution of the IBVP (2.1) is presented even when the conditions for Theorem 4.3.2 to hold are not met. In fact, it is found that in regions where the solution does not degenerate $(u - z > 0)$ the numerical method reaches optimal convergence rates. Despite the fact that the IBVP (2.1) has not been fully studied analytically when $z \neq 0$, properties such as boundedness and existence of compactly supported solutions, finite speed of propagation of disturbances, and extinction in finite time found in the 1-D case when $z = 0$ in [35], [6], and [2], were found to persist for a bounded and smooth bathymetry $z \neq 0$, based on numerical evidence.

In Chapter 4, solutions of the DSW equation that locally violate some of the essential assumptions used to derive the DSW equation from the Navier-Stokes equations are introduced. Such is the case for the family of compactly supported Barenblatt solutions exhibited when $z = 0$ in section 4.4.1. The gradient of these solutions (water surface slope) is not comparable to the gradient of the bathymetry ∇z close to the free boundary. A more extreme case of solutions that violate the uniform flow conditions happen even when the DSW equation becomes the PME ($z = 0$ and $\gamma = 1$) in two or higher dimensions. In Chapter 19 of [62], Vázquez shows that there exists a class of solutions called *focusing solutions* that exhibit no local regularity on the gradient in subsets of Ω. The existence of these kinds of solutions serves as a reminder of the limitations of using the DSW equation as a hydrodynamical model.

One dimensional experiments presented in section 4.4.2, show that despite the limitations of the diffusive nature of the DSW equation, the main qualitative behavior of water flow in an experimental setting designed to produce unsteady flows [46], such as breakthrough time and discharge, were captured by the simulation using the DSW as a model.

In Chapter 5, the approximation properties of numerical solutions to the DSW equation obtained using the *local discontinuous Galerkin* (LDG) method are studied. In section 5.2.2, properties such as stability and *a priori* error estimates for the numerical formulation of the LDG method are established. In section 5.3, numerical simulations of the evolution of water depth profiles in 2D are presented, as qualitative evidence, to support the fact that the DSW as a hydrologycal model, captures the salient features of: *(i)* an ideal dam break problem and *(ii)* water flow in a channel containing vegetation.

6.1 Future work

Important issues to be addressed in future works in the realm of PDE theory should include:

- An appropriate study of existence and uniqueness of weak solutions of problem (2.1) when topographic effects are considered ($z \neq 0$).

- Regularity of the free boundary for the two dimensional case both when $z = 0$ and $z \neq 0$.

- The connection between the regularity of the bathymetry z and the resulting weak solution of the IBVP (2.1).

- Conditions for which the regularity in the time derivative can be improved as well as conditions for which the pointwise gradient can be bounded (for $z = 0$).

In terms of numerical analysis,

- From the theoretical point of view, in order to improve the error estimates and thus, convergence rates, the *Quasi-norm* error estimates techniques recently introduced in [33] and [51] for elliptic p-Laplacian problems should be explored.

From both the implementation and numerical analysis point of view,

- The coupling of the DSW model with other surface water models (such as the SWE) and other subsurface flow models (such as Darcy's flow or other infiltration models) should be pursued.

- The LDG formulation provides a natural environment to easily handle these couplings

105

Appendix A

Additional results

Lemma A.0.1. *The operator $\mathcal{A}(x) : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ defined by*

$$\mathcal{A}(x) = \frac{x}{|x|^{1-\gamma}} \tag{A.1}$$

is monotone, i.e., for any $x, y \in \mathbb{R}^n$

$$(\mathcal{A}(x) - \mathcal{A}(y)) \cdot (x - y) \geq 0.$$

Proof. Define the function $\mathcal{B}(x) : \mathbb{R}^n \longrightarrow \mathbb{R}$ as

$$\mathcal{B}(x) = |x|^{\gamma+1} \qquad \text{where} \qquad |x| = \left(\sum_{j=1}^{n} x_j^2 \right)^{\frac{1}{2}}$$

and note that

$$\frac{\partial}{\partial x_i} |x|^{\gamma+1} = (\gamma+1)|x|^{\gamma-1} x_i \qquad \Longrightarrow \qquad \frac{1}{\gamma+1} \nabla \mathcal{B}(x) = \mathcal{A}(x).$$

Since $\gamma + 1 > 1$, the function $\mathcal{B}(x)$ is strictly convex. The gradient of a convex function is strictly increasing in each and all of its components, thus the result of the lemma holds true. $\square$

Theorem A.0.1. *(Calculus in abstract space) Let X a Banach space and let $u \in W^{1,p}(0,T;X)$ for some $1 \leq p \leq \infty$. Then*

(i) $u \in C([0,T];X)$ (after possibly being redefined on a set of measure zero), and

(ii) $u(t_1) = u(t_0) + \int_{t_0}^{t_1} u_t(\tau)d\tau$ for all $0 \leq t_0 \leq t_1 \leq T$.

Proof. See [36]. □

Assume that Ω is an open, bounded set, with smooth boundary, and $T > 0$. We have

Theorem A.0.2. *Let ψ be a real valued, absolutely continuous and monotone function, and let $0 < \eta \leq \gamma \leq 1$. Assume that*

(i) ψ is an η-Hölder continuous function with $\psi(0) = 0$.

(ii) $u_\mu \rightharpoonup u$ in $L^{1+\gamma}(0, T; W^{1,1+\gamma}(\Omega))$.

(iii) $\psi(u_\mu)_t \rightharpoonup v$ in $L^{(1+\gamma)^*}(0, T; W^{-1,(1+\gamma)^*}(\Omega))$.

Then, $v = \psi(u)_t$.

Proof. During the proof every subsequence obtained by a compact argument will be relabeled with the index μ for clarity. Set $p = \frac{1+\gamma}{\eta}$ and note that by (i) and (ii) we have

$$\|\psi(u_\mu)\|^p_{L^p(0,T;L^p(\Omega))} \leq \|u_\mu\|^{1+\gamma}_{L^{(1+\gamma)}(0,T;L^{(1+\gamma)}(\Omega))} \leq C. \tag{A.2}$$

Since $L^p(0, T; L^p(\Omega))$ is a separable and reflexive Banach space, inequality (A.2) implies that

$$\psi(u_\mu) \rightharpoonup \xi \quad \text{weakly in} \quad L^p(0, T; L^p(\Omega)). \tag{A.3}$$

Since

$$\frac{1+\gamma}{\eta} \geq \frac{1+\gamma}{\gamma} = (1+\gamma)^*,$$

it follows that

$$L^p(0, T; L^p(\Omega)) \subset L^{(1+\gamma)^*}(0, T; L^{(1+\gamma)^*}(\Omega)) \subset L^{(1+\gamma)^*}(0, T; W^{-1,(1+\gamma)^*}(\Omega)).$$

Then, for any $\varphi \in C^1_c(0, T)$ and $\omega \in W^{1,1+\gamma}(\Omega)$ we obtain

$$\int_0^T \langle \psi(u_\mu)_t, \varphi\, \omega \rangle = -\int_0^T (\psi(u_\mu), \varphi_t\, \omega).$$

Send $\mu \to \infty$ in this equality to obtain that

$$\int_0^T \langle v, \varphi\, \omega \rangle = -\int_0^T (\xi, \varphi_t\, \omega). \tag{A.4}$$

Thus, it remains to prove that $\xi = \psi(u)$.

To this end, we will first prove this for a $\psi' \in L^\infty(\mathbb{R})$. Observe that as a consequence of

107

the chain rule, the sequence

$$\{\psi(u_\mu)\} \subset L^{1+\gamma}(0, T; W^{1,1+\gamma}(\Omega))$$

is uniformly bounded thanks to (A.2) and (ii). Since the sequence

$$\{\psi(u_\mu)_t\} \subset L^{(1+\gamma)^*}(0, T; W^{-1,(1+\gamma)^*}(\Omega)) \subset L^{1+\gamma}(0, T; W^{-1,1+\gamma}(\Omega))$$

is uniformly bounded by assumption (iii), and

$$L^{1+\gamma}(0, T; W^{1,1+\gamma}(\Omega)) \subset L^{1+\gamma}(0, T; L^{1+\gamma}(\Omega)) \subset L^{1+\gamma}(0, T; W^{-1,1+\gamma}(\Omega))$$

with the compact embedding

$$W^{1,1+\gamma}(\Omega) \hookrightarrow L^{1+\gamma}(\Omega).$$

Thus, one concludes by a compactness criterion in the spaces $L^p(0, T; X)$ that

$$\psi(u_\mu) \to \xi \quad \text{strongly in} \quad L^{1+\gamma}(0, T; L^{1+\gamma}(\Omega)), \tag{A.5}$$

see [50]. This convergence is a.e. in $(0, T) \times \Omega$ as well. Since ψ is invertible, u_μ converges a.e. in $(0, T) \times \Omega$ to $\psi^{-1}(\xi)$. Using (ii) we conclude that

$$u = \psi^{-1}(\xi). \tag{A.6}$$

Now we proceed to extend the previous result for a general ψ having the conditions stated in the hypothesis of the theorem. Let $\{\psi_\epsilon\}$ be a family of absolutely continuous and increasing functions with bounded derivative (for fix $\epsilon > 0$) and fulfilling condition (i), such that for some $k > 0$

$$\sup_s |\psi(s) - \psi_\epsilon(s)| \leq k\epsilon^\eta.$$

Hence for any $\varphi \in C_c^1(0, T)$ and $\omega \in W^{1,1+\gamma}(\Omega)$,

$$
\left| \int_0^T \langle \psi(u_\mu)_t - \psi(u)_t, \varphi\,\omega \rangle \right| = \left| \int_0^T (\psi(u_\mu) - \psi(u), \varphi_t\,\omega) \right|
$$

$$
= \left| \int_0^T (\psi(u_\mu) - \psi_\epsilon(u_\mu) + \psi_\epsilon(u_\mu) - \psi(u), \varphi_t\,\omega) \right|
$$

$$
\leq k\epsilon^\eta \int_0^T (1, |\varphi_t\,\omega|) + \left| \int_0^T (\psi_\epsilon(u_\mu) - \psi(u), \varphi_t\,\omega) \right|. \tag{A.7}
$$

Similarly,

$$\left| \int_0^T (\psi_\epsilon(u_\mu) - \psi(u), \varphi_t\, \omega) \right| \le k\epsilon^\eta \int_0^T (1, |\varphi_t\, \omega|) + \left| \int_0^T (\psi_\epsilon(u_\mu) - \psi_\epsilon(u), \varphi_t\, \omega) \right|. \qquad \text{(A.8)}$$

Using (A.7) and (A.8) we send $\mu \to \infty$ to obtain that

$$\limsup_\mu \left| \int_0^T \langle \psi(u_\mu)_t - \psi(u)_t, \varphi\, \omega \rangle \right| \le 2k\epsilon^\eta \int_0^T (1, |\varphi_t\, \omega|)\,.$$

Finally, let $\epsilon \to 0$ to conclude. $\qquad\square$

Theorem A.0.3. *Assume that Ω is measurable and $|\Omega| < \infty$. Assume also that $f \in L^p(\Omega)$ for any $1 \le p < \infty$ and $\|f\|_{L^p(\Omega)} \le M$ for some $M > 0$. Then*

$$f \in L^\infty(\Omega) \qquad and \qquad \|f\|_{L^\infty(\Omega)} \le M. \qquad \text{(A.9)}$$

Proof. See [66, p. 126] for a version of this result. A slight modification of this proof will work for this version. $\qquad\square$

Lemma A.0.2. *Let $x \in \mathbb{R}^M$ and $f(x), g(x)$ be L^∞ functions. If $f(x)$ is Lipschitz continuous and $g(x)$ is γ- Hölder continuous, with $0 < \gamma < 1$, then the product $f(x)g(x)$ is γ- Hölder continuous for x_1 and x_2 in a bounded domain Ω.*

Proof. Observe that

$$\begin{aligned}
|f(x_1)g(x_1) - f(x_2)g(x_2)| &\le& |f(x_1)(g(x_1) - g(x_2))| + |g(x_2)(f(x_1) - f(x_2))| \\
&\le& \|f(x)\|_{L^\infty}|x_1 - x_2|^\gamma + \|g(x)\|_{L^\infty}|x_1 - x_2| \\
&\le& (\|f(x)\|_{L^\infty} + C\|g(x)\|_{L^\infty})\,|x_1 - x_2|^\gamma
\end{aligned}$$

$\qquad\square$

Bibliography

[1] V. Aizinger, C. Dawson, B. Cockburn, and P. Castillo. The local discontinuous galerkin method for contaminant transport. *Adv. Water Resour.*, 24(1):73–87, 2000.

[2] R. Alonso, M. Santillana, and C. Dawson. On the difussive wave approximation of the shallow water equations. *European Journal of Applied Mathematics*, 2007. To appear.

[3] H. Alt and S. Luckhaus. Quasilinear elliptic-parabolic differential equations. *Mathematische Zeitschrift*, 183:311–341, 1983.

[4] T. Arbogast, M.F. Wheeler, and N.Y. Zhang. A nonlinear mixed finite element method for a degenerate parabolic equation arising in flow in porous media. *SIAM Journal on Numerical Analysis*, 33(4):1669–1687, 1996.

[5] Douglas N. Arnold, Franco Brezzi, Bernardo Cockburn, and L. Donatella Marini. Unified analysis of discontinuous Galerkin methods for elliptic problems. *SIAM Journal of Numerical Analysis*, 39(5):1749–1779, 2002.

[6] A. Bamberger. Étude d'une équation doublement non linéaire. *Journal of Functional Analysis*, 24:148–155, 1977.

[7] G.I. Barenblatt. On self-similar motions of compressible fuids in porous media. *Prikl, Math. Mech.*, 16:679–698, 1952. In Russian.

[8] J. W. Barrett and W. B. Liu. Finite element approximation of the parabolic p-laplacian. *SIAM J. Numer. Anal.*, 31(2):413–428, 1994.

[9] John W. Barrett and W. B. Liu. Finite element approximation of the parabolic p-laplacian. *SIAM J. Numer. Anal.*, 31(2):413–428, 1994.

[10] F. Bassi and S. Rebay. A high-order accurate discontinuous finite element method for the numerical solution of the compressible navier-stokes equations. *J. Comput. Phys.*, 131(2):267–279, 1997.

[11] P. Bauer, T. Gumbricht, and W. Kinzelbach. A regional coupled surface water/groundwater model of the Okavango delta, Botzwana. *Water Resources Research*, 42, 2006.

[12] S. Bennett and A. Simon, editors. *Riparian Vegetation and Fluvial Geomorphology.* Water Science and application. American Geophysical Union, 2004.

[13] S.J. Bennett, T. Pirim, and B.D. Barkdoll. Using simulated emergent vegetation to alter stream flow direction within a straight experimental channel. *Geomorphology*, 44:115–126, 2002.

[14] F. Bernis. Existence results for doubly nonlinear higher order parabolic equations on unbounded domains. *Mathematische Annalen*, 279:373–394, 1988.

[15] D. Blanchard and G. Francfort. Study of a nonlinear heat equation with no growth assumptions on the parabolic term. *SIAM Journal of Mathematical Analysis*, 19(5):1032–1056, 1988.

[16] C.H. Bolster and J.E. Saiers. Development and evaluation of a mathematical model for surface-water flow within the shark river slough of the florida everglades. *Journal of Hydrology*, 259:221–235, 2002.

[17] D. S. Bowles and P.E. O'Connell, editors. *Recent Advances in the Modeling of Hydrologic Systems*, chapter 8. Overland Flow: A two dimensional Modeling Approach, pages 153–166. Springer, 1991.

[18] Susanne C. Brenner and Ridgway Scott. *The Mathematical Theory of Finite Element Methods.* Springer-Verlag, New York, 1994.

[19] R. Bustinza and G. N. Gatica. A local discontinuous galerkin method for nonlinear diffusion problems with mixed boundary conditions. *SIAM Journal on Scientific Computing*, 26(1):152–177, 2004.

[20] N. Calvo, J. I. Díaz, J. Durany, E. Schiavi, and C. Vázquez. On a doubly nonlinear parabolic obstacle problem modelling ice sheet dynamics. *SIAM Journal of Applied Mathematics*, 63(2):683–707, 2003.

[21] José Carrillo. Entropy solutions for nonlinear degenerate problems. *Arch. Ration. Mech. Anal.*, 147(4):269–361, 1999.

[22] P. Castillo, B. Cockburn, I. Perugia, and D. Schötzau. An a priori error analysis of the local discontinuous galerkin method for elliptic problems. *SIAM Journal on Numerical Analysis*, 38(5):1676–1706, 2000.

[23] P Castillo, B. Cockburn, D. Schötzau, and C. Schwab. Optimal a priori error estimates for the *hp*-version of the local discontinuous galerkin method for convection-diffusion problems. *Mathematics of Computation*, 2002.

[24] F. Catté, P. L. Lions, J. M. Morel, and T. Coll. Image selective smoothing and edge detection by nonlinear diffusion. *SIAM J. Numer. Anal.*, 29(1):182–193, 1992.

[25] P. G. Ciarlet. *Finite Element Method for Elliptic Problems.* Society for Industrial and Applied Mathematics, Philadelphia, PA, USA, 2002.

[26] B. Cockburn and C. Dawson. Some extensions of the local discontinuous galerkin method for convection-diffussion equations. In *The Proceedings of the Conference on the Mathematics of Finite Elements and Applications*, pages 225–238. Elsevier, 2000.

[27] B. Cockburn, G. E. Karniadakis, and C. W. Shu, editors. *Discontinuous Galerkin Methods*. Springer, 2000.

[28] B. Cockburn and C. W. Shu. The local discontinuous galerkin method for time-dependent convection-diffusion systems. *SIAM J. Numer. Anal.*, 35(6):2440–2463, 1998.

[29] R. Courant, K. Friedrihs, and H. Lewy. Uber die partiellen differenzengleichungen der mathematischen physik. *Math. Ann.*, 100:32–74, 1928.

[30] R. Daugherty, J. Franzini, and J. Finnemore. *Fluid Mechanics with Engineering Applications*. McGraw-Hill, 1985.

[31] Emmanuel DiBenedetto. *Degenerate Parabolic Equations*. Springer-Verlag, New York, 1993.

[32] J. Douglas and T.F. Dupont. Galerkin methods for parabolic equations. *SIAM Journal on Numericla Analysis*, 7:575–626, 1970.

[33] C. Ebmeyer and W. B. Liu. Quasi-norm interpolation error estimates for the piecewise linear finite element approximation of p-laplacian problems. *Numer. Math.*, 100(2):233–258, 2005.

[34] Alexandre Ern and Jean-Luc Guermond. *Theory and practice of finite elements*. Springer-Verlag, New York City, 2004.

[35] J.R. Esteban and J.L. Vázquez. Homogeneous diffusion in $\mathbb{R}$ with power-like nonlinear diffusivity. *Archive Rat. Mech. Analysis*, 103:39–80, 1988.

[36] Lawrence C. Evans. *Partial Differential Equations*. American Mathematical Society, Providence, 2002.

[37] S. Evje and K. Karlsen. Discrete approximations of bv solutions to doubly nonlinear degenerate parabolic equations.

[38] K. Feng and F.J Molz. A 2-d diffusion based, wetland flow model. *Journal of Hydrology*, 196:230–250, 1997.

[39] P Di Giammarco, E. Todini, and P. Lamberti. A conservative finite elements approach to overland flow: the control volume finite element formulation. *Journal of Hydrology*, 175:267–291, 1996.

[40] G. Gioia and F.A. Bombardelli. Scaling and similarity in rough channel flows. *Physical Review Letters*, 88(1), 2002.

[41] O. Grange and F. Mignot. Sur la résolution d'une équation et d'une inéquation paraboliques non linéaires. *Journal of Functional Analysis*, 11:77–92, 1972.

[42] E. Hansen and A. Ostermann. Finite element runge–kutta discretizations of porous medium–type equations. *SIAM Journal on Numerical Analysis*, 46(4):1769–1779, 2008.

[43] Philip Hartman. *Ordinary Differential Equations*. Society for Industrial and Applied Mathematics, Philadelphia, PA, USA, 2002.

[44] T.V. Hromadka, C.E. Berenbrock, J.R. Freckleton, and G.L. Guymon. A two-dimensional dam-break flood plain model. *Adv. Water Resources*, 8, 1985.

[45] K. Ishige. On the existence of sotutions of the cauch problem for a doubly nonlinear parabolic equation. *SIAM Lournal on Mathematical Analysis*, 27(5):1235–1260, 1996.

[46] Y. Iwagaki. Fundamental studies of runoff analysis by characteristics. Bulletin 10,Disaster Prevention Res. Inst., Kyoto University, Kyoto, Japan, 1955. 25pp.

[47] S.C. Jain. *Open-channel flow*. John Wiley & Sons, 2001.

[48] N. Ju. Numerical analysis of parabolic p-laplacian: Approximation of trajectories. *SIAM Journal on Numerical Analysis*, 37(6):1861–1884, 2000.

[49] H. Lia, M.W. Farthinga, C.N. Dawson, and C.T. Miller. Local discontinuous galerkin approximations to richards equation. *Advances in Water Resources*, 30:555–575, 2007.

[50] J.L. Lions. *Quelques Méthodes de Résolution des Problèmes aux Limites Non Linéaires*. Dunod Gauthier-Villars, Paris, 1969.

[51] W. Liu and N. Yan. Quasi-norm local error estimators for p-laplacian. *SIAM J. Numer. Anal.*, 39(1):100–127, 2001.

[52] J. Nitsche. Über ein variationsprinzip zur lösung von dirichlet-problemen bei verwendung von teilräumen, die keinen randbedingungen unterworfen sind. *Abh. Math. Sem. Univ. Hamburg*, 36, 1971.

[53] R. H. Nochetto and C. Verdi. Approximation of degenerate parabolic problems using numerical integration. *SIAM J. Numer. Anal.*, 25:784 – 814, 1988.

[54] V.M. Ponce, R.M. Li, and D.B. Simons. Aplicability of kinematic and diffusion models. *Journal of the Hydraulics Division*, pages 353–360, 1978.

[55] V.M. Ponce and D.B. Simons. Shallow wave propagation in open channel flow. *Journal of the Hydraulics Division*, pages 1461–1476, 1977.

[56] P.A. Raviart. Sur la résolution de certaines equations paraboliques non linéaires. *Journal of Functional Analysis*, 5:299–328, 1970.

[57] J. Rulla and N. J. Walkington. Optimal rates of convergence for degenerate parabolic problems in two dimensions. *SIAM Journal on Numerical Analysis*, 33(1):56–67, 1996.

[58] M. Santillana and C. Dawson. A numerical approach to study the properties of solutions of the diffusive wave approximation of the shallow water equations. *Computational Geosciences*, 2008. Submitted.

[59] J.W. Thomas. *Numerical Partial Differential Equations: Finite Difference Methods*. Number 22 in Texts in Applied Mathematics. Springer, New York, 1995.

[60] Vidar Thomée. *Galerkin Finite Element Methods for Parabolic Problems*. Number 25 in Springer Series in Computational Mathematics. Springer, 1997.

[61] A.K. Turner and N. Chanmeesri. Shallow flow of water through non-submerged vegetation. *Agricultural Water Management*, 8:375–385, 1984.

[62] J. L. Vázquez. *The Porous Medium Equation. Mathematical Theory*. Oxford University Press, USA, 2006.

[63] C.B. Vreugdenhil. *Numerical Methods for Shallow-Water Flow*. Kluwer Academic Publishers, The Netherlands, 1998.

[64] Dongming Wei and Lew Lefton. A priori L^ρ error estimates for galerkin approximations to porous medium and fast diffusion equations. *Math. Comput.*, 68(227):971–989, 1999.

[65] T. Weiyan. *Shallow Water Hydrodynamics: Mathematical Theory and Numerical Solution for a Two-dimensional System of Shallow Water Equations*, volume 55. Elsevier, Amsterdam, 1992.

[66] Richard L. Wheeden and Antoni Zygmund. *Measure and Integral: An Introduction to Real Analysis*. A series of Monographs and Textbooks. Marcel Dekker, Inc., New York, 1977.

[67] M. F. Wheeler. A priori L^2 error estimates for galerkin approximations to parabolic partial differential equations. *SIAM Journal on Numerical Analysis*, 10(4):723–759, 1973.

[68] M. F. Wheeler. An elliptic collocation finite element method with interior penalties. *SIAM Journal on Numerical Analysis*, 15(1):152–161, 1978.

[69] Th. Xanthopoulos and Ch. Koutitas. Numerical simulation of a two dimensional flood wave propagation due to dam failure. *Journal of Hydraulic Research*, 14(4):321–331, 1976.

[70] J. Yan and C. W. Shu. A local discontinuous galerkin method for kdv type equations. *SIAM J. Numer. Anal.*, 40(2):769–791, 2002.

[71] J. Yan and C. W. Shu. Local discontinuous galerkin methods for partial differential equations with higher order derivatives. *J. Sci. Comput.*, 17(1-4):27–47, 2002.

[72] W. Yong and I. Pop. A numerical approach to porous medium equations, 1996.

[73] W. Zhang and T.W. Cundy. Modeling of two-dimensional overland flow. *Water Resources Res.*, 25:2019–2035, 1989.

Vita

Mauricio Santillana, son of Humberto Santillana Loyo and Luz Patricia Guzmán Olivares, was born on July 9th in the year 1976, in Mexico City. In 2001, he received a B.S. in Physics with honors from the Universidad Nacional Autónoma de México in Mexico City. After which, he started graduate school at the University of Texas at Austin, earning a M.S. in Computational and Applied Mathematics in 2003. Since 2003, he has been part of the inter-disciplinary research team at the Research Institute for Geography and Geomatics, CentroGeo, based in Mexico City and part of the National Council for Science and Technology of Mexico. In 2004, Mauricio started his Ph.D. studies at the Institute for Computational Engineering and Sciences at the University of Texas at Austin.

Permanent Address: 3914 Avenue H
                            Austin, TX 78751

This dissertation was typeset with $\text{\LaTeX}\,2_\varepsilon$[1] by the author.

---

[1] $\text{\LaTeX}\,2_\varepsilon$ is an extension of $\text{\LaTeX}$. $\text{\LaTeX}$ is a collection of macros for TeX. TeX is a trademark of the American Mathematical Society. The macros used in formatting this dissertation were written by Dinesh Das, Department of Computer Sciences, The University of Texas at Austin, and extended by Bert Kay, James A. Bednar, and Ayman El-Khashab.