SAPIENZA
UNIVERSITÀ DI ROMA

FACOLTÀ DI SCIENZE MATEMATICHE FISICHE E NATURALI
Dottorato di Ricerca in Matematica XXX ciclo

Adaptive Filtered Schemes for first order Hamilton-Jacobi equations and applications

Candidate:
Giulio Paolucci
matricola 1273737

Anno Accademico 2017-2018
Dipartimento di Matematica 'Guido Castelnuovo'

Acknowledgments

I would like to express my sincere gratitude to my advisor Prof. Maurizio Falcone for giving me the opportunity to work under his precious guidance, showing immense patience in bearing all my continuous doubts and delays. His teachings have been of great importance during my university studies and the writing of my master thesis, his support and encouragement have brought me to see the end of this Phd thesis. The things I learned will help me face the challenges that await me.

A very special thank to my friend and co-advisor Dr. Silvia Tozza for having helped and supervised most of the writing of this work. Her presence and support have been of great help in focusing my research, her knowledge and rigor have opened my eyes on many aspects of the life of a mathematician. Thanks for the patience and the great deal of time invested on me.

I would like to thank Prof. Matteo Semplice for his time and help in clarifying some important issues of my thesis, which led to interesting consequences. Many thanks also to my friend Dr. Smita Sahu for her very positive presence and for the useful discussions on the topic of this work.

I wish to particularly thank "my" coordinators of the Phd program at our department 'Guido Castelnuovo', Prof. Paolo Piazza and Prof. Adriana Garroni, for always being comprehending and supporting, especially in the most difficult period of my work.

A comprehensive thanks to all the people of my department, whose presence and friendship have helped me bear all the difficult moments and enjoy the precious ones, from my first day at university till today.

I would like to give many other thanks, name to name, to all my friends but it would probably take another whole year to find the right words to express what I feel.

Last, but clearly not even close to being the least, I would like to express in this few lines all my gratitude to my parents, Carla and Giancarlo, my brother Matteo, my grandparents, Gilda and Donato, and all my family, for the love, trust, safety and freedom I always felt during my whole life. Thanks for being the reason of who I am today, luckily or not.

Contents

Introduction

Contexts and motivations

This PhD thesis deals with the numerical approximation of the solution of *time-dependent first-order Hamilton-Jacobi equations*, in the form

$$\begin{cases} v_t + H(x, \nabla v) = 0 & (t,x) \in \mathbb{R}^+ \times \mathbb{R}^N \\ v(0,x) = v_0(x) & x \in \mathbb{R}^N, \end{cases} \tag{HJ}$$

where the hamiltonian $H : \mathbb{R}^N \times \mathbb{R}^N \to \mathbb{R}$ and the initial condition v_0 are usually, at least, continuous functions. It is well-known that in general this problem does not admit classical solutions ($v \in C^1$), independently on the regularity of the initial condition, since its solutions may develop discontinuities in the gradient in finite time. Consequently, it was necessary to introduce some appropriate definition of solution in "weak sense". This was done by Crandall and Lions in [CL83], in which they introduce (and deeply study) the definition of *viscosity solution*, which precise statement will be given in Section 1.1, using argument based on the vanishing method. The theory of viscosity solutions allows the proof of existence and, more importantly, uniqueness in a wide range of situations, even for more general hamiltonians, possibly dependent also on v, and for less regular solutions (e.g. discontinuous solutions) or hamiltonians (e.g. with discontinuous data). Although the theory is still undergoing technical developments, especially regarding the above mentioned situations, in the case of continuous solutions, which is the main focus of this thesis, it is in fact a well established matter, as witnessed by the books by Lions [L82] and Barles [B98]. In particular, the latter contains a chapter on discontinuous viscosity solutions. This theory also naturally extends to *second order degenerate elliptic equations*, for which we mention the survey paper [CIL92].

Moreover, it has been found that the viscosity solution is the natural concept to use in many applications involving Hamilton-Jacobi equations, including for example optimal control (the *Bellman equation*), differential games (the *Isaacs equation*), front evolution (via the *Level Set Method*) and image processing (*Shape from Shading* and segmentation problems). This will also be argument of the present work, mainly in Chapter 4, in which we apply our scheme to the problem of the segmentation of an image.

The lack of smoothness of viscosity solutions makes it difficult to develop efficient approximations. Starting from the 80's, monotone finite difference methods have been proposed by Crandall and Lions [CL84] using the fact that the solution of (HJ) is the integral of the entropy solution of a strictly related conservation law, which is valid only in the one-dimensional case (see e.g. [CFN95]). For the possible extension of the previous relation in higher dimensions we mention the work [JX98]. On this basis, monotone finite difference methods conceived for the conservation laws have been adapted to the

approximation of the HJ equations and then extended to more general hamiltonians by Souganidis ([So85a]) and Barles ([BS91]). However, monotone schemes are typically very diffusive, a serious drawback when treating nonsmooth solutions, and are limited to first order accuracy, in the sense of the consistency error, also in the case of very regular solutions. Thus, if we want to achieve high-order consistency, at least in regular regions of the solution, we have to look for non-monotone methods, then the real problem is to prove convergence.

In order to summarize the previous discussion and focus on the analysis ahead, we state the properties that should be satisfied by the "high-order" or "high-resolution" schemes for the approximation of (HJ):

- at least second order accuracy in regions of the domain where the solution is regular;
- high-resolution of the singularities, avoiding the usual "rounding" produced by the first-order monotone schemes;
- the artificial oscillations introduced should rapidly disappear as the grid is refined.

The use of high-order schemes is preferable mainly because of computational cost reasons, especially when working with multidimensional problems, since they allow the use of coarser grids, but on the other hand, proving convergence becomes a really difficult matter. A key tool for proving convergence of numerical approximations would be some stability property of discrete solutions; for example, when working with continuous solutions it is common to require the boundedness of the family of approximate solutions in $W^{1,\infty}$, so that the existence would follow by compactness by the classical Ascoli-Arzelà theorem. Such estimates, may be obtained easily for monotone schemes and with some effort for high-order Semi-Lagrangian schemes (see e.g. [FF14]), but in general are not sufficient to infer convergence. In fact, at least to our knowledge, this theory is still lacking a general stability concept, as the *TV-stability* for conservation laws for example, that is able to secure the convergence of the numerical scheme to the viscosity solution of the problem.

Apparently, a key to obtain some generality is to slightly loosen the monotonicity property to the so-called *ε-monotonicity*, precisely defined in the sequel. Some hints on this generalization are already present in [BS91], while a formal definition has been given, for example, in [CFF10] for second order equations, in [BFFKZ14] for stationary equations and by many other authors for first-order evolutive equations, which will be cited later in this introduction. Moreover, in the convex case an equally useful concept is the *discrete uniform semiconcavity*, introduced in [LT01], which allows to prove convergence and, eventually with some further (not easy) computations, obtain convergence estimates in the L^1 norm.

Despite the difficulties in proving convergence, in the last decades a lot of effort has been directed by many authors on the development of efficient and high-order methods for the solution of (HJ), using different approaches. As a first example, we cite a very efficient class of finite difference schemes, the *Essentially NonOscillatory* (ENO) and

Weighted ENO (WENO) schemes, which ideas have been developded for the treatment of hyperbolic conservation laws, respectively, by Harten et al. in [HEOC87] and by Jiang and Shu in [JS96], and then extended to the case of (HJ), using the recalled relation between the respective solutions, by Osher and Shu in [OSh91] and, later, by Jiang and Peng in [JP00]. ENO and WENO schemes are of high-order, in the sense of the consistency error, and because of their reliability, have been quite successful in many applications. Unfortunately, in the general case although there is numerical evidence that they convergence to the viscosity solution of (HJ), till now and to the best of our knowledge, there is no convergence proof for these classes of schemes. However, convergence may hold for related schemes (see for instance [LS95]), while minor results have been proven by Fjordholm et al. in [FMT12], where they show some stability property of the ENO interpolation although the result is not sufficient to obtain the total variation boundedness (TVB), and in [JP00], where convergence for smooth solutions is proved. Among the other possible approaches, let us mention the semi-discrete central upwind schemes of [KNP01] (see also the references therein) and the discontinuous Galerkin (DG) finite element schemes, introduced by Cockburn and Shu in [CS89] for conservation laws and later extended to HJ equations. These methods, as the ENO and WENO schemes, typically require the use of Runke-Kutta (RK) time discretizations to achieve high-order accuracy in time, which leads to the use of bigger numerical stencils (that are the set of nodes used in the numerical scheme). In particular, the RKDG methods are very flexible, being able to treat complicated geometries and different boundary conditions, and have compact stencils, therefore are particularly indicated for parallel implementation, but they require a stability condition which is usually costly to enforce. In this brief list of contributions to the theory, a special mention is due to the Semi-Lagrangian discretizations, developed by Falcone and Ferretti ([FF94], [FF14]) and Carlini ([CFR05], [CFF10]), which allows the use of very large-time steps, do not require the use of RK procedures, since they are based on some direct representation formula for the solution, and have been proven convergent, in the convex case, for a wide range of possible reconstruction procedures. The proof there relies heavily on the work of Ferretti ([F02]), where a first hint on the ε-monotonicity property can be found in the hypothesis on the high-order reconstruction procedure, which has to be close enough to the linear (monotone) interpolation.

Following on this idea, in recent years many authors faced the problem using the ε-monotone (or similar definitions) stability property in order to prove convergence, mainly devising different procedures to "filter" an high-order scheme, possibly unstable, with a convergent monotone scheme. This idea was already present in [LS95], in the context of implicit schemes, and has been further developed by Abragall in [A09] and Oberman and Salvador in [ObSa15], for first-order stationary equation (in the latter the main focus is on the eikonal equation) and by Froese and Oberman in [FO13], for particular second order equations. Regarding the convergence of ε-monotone scheme for stationary equations we also recall the above mentioned [BFFKZ14]. In this thesis we mainly follow the approach of Bokanowski, Falcone and Sahu in [BFS16] on *filtered schemes* for evolutive first-order Hamilton-Jacobi equations, in which a convergence result with error estimate is given. In the present work we try to deepen the study

of the filtering process in order to better exploit the possibility offered by the relaxed stability property.

Finally, it is worth saying that the development of this theoretical framework (i.e. convergence to the viscosity solution of HJ equations) has been widely driven by the above mentioned applications to control problems, image processing and Fluid Dynamics. Regarding the applications to control problems we refer to the book of Bardi and Capuzzo Dolcetta [BCD97], in which the focus is mainly on Hamilton-Jacobi-Bellman equations, arising via the Dynamic Programming Principle in optimal control problems.

Another typical application of HJ equations, which will also be argument of the present work, is the study and numerical approximation of fronts propagating in their normal direction. This is done via the so-called *Level Set Method*, first introduced by Sethian in [Se85], [Se90] and Osher in [OSe88], and briefly recalled in Section 1.3. It is known that this construction can be applied to model many physical problems, from crystal growth to bubbles moving in a fluid, and so forth. Many efficient schemes have been proposed in recent years, we mention the book of Sethian [Se90] for some examples of finite difference schemes, the works of Bardi and Falcone [BF90] and of Falcone at al. [FGL94] on the Semi-Lagrangian approximation for minimum time-problem, and the paper of Bokanowski et al. [BCS13], where the RKDG method is used. In this thesis we apply the method for the problem of image segmentation, following mainly the approach in [MSV93], but proposing a slight modification in the velocity c, which seems to produce a stable and reliable scheme. For another possible application of HJ equations in the context of image processing, we mention also the widely studied Shape from Shading problem, for which we refer to the book [HB89] and to the more recent work of Falcone and Tozza [FT16] (see also the references therein), where Semi-Lagrangian discretizations are used to approximate the solution of Lambertian and non-Lambertian models. Moreover, they give a unified mathematical formulation of some popular orthographic non-Lambertian models.

Contributions

Briefly speaking, the aim of this thesis is to develop convergent schemes to the solution of (HJ) which are of high-order in regions of regularity, deepening the study of [BFS16] on filtered schemes. The main novelty here is that the defined schemes have the desired properties without the need to tune any parameter manually. This is done by the introduction of a smoothness indicator function and a precise numerical formula for the switching parameter ε. The new Adaptive Filtered Scheme has the form

$$u_j^{n+1} = S^{AF}(u^n)_j := S^M(u^n)_j + \phi_j^n \varepsilon^n \Delta t F\left(\frac{S^A(u^n)_j - S^M(u^n)_j}{\varepsilon^n \Delta t}\right),$$

where ϕ is the introduced smoothness indicator function and ε^n is the new adaptive switching parameter. For the latter, we design a procedure that automatically tunes the value according to the regularity of the solution at any time step t_n, which is measured by the function ϕ.

In order to construct the smoothness indicator function, we first deeply review the theory of the smoothness indicators defined in [JS96] and [JP00], giving a precise proof of their properties in Proposition 2.1. The idea of the proof is inspired by the one of Corollary 2 in [ABM10], but it is more general. Moreover, the formulation of our proof easily generalizes to get similar results for higher degree derivatives. Another aim of this precise analysis is to clarify some confusion we have encountered in literature, since the scaling factor needed in the definition of these indicators, when working with HJ equations which present discontinuities in the first derivative, is usually mistaken.

Then, differently from the usual WENO procedure, we use the information given by the indicators to construct the function ϕ, which basically measures the regularity of the first derivative of a given function f. Moreover, we propose a genuinely multidimensional version of the above one-dimensional definition on a cartesian grid, giving also a very compact explicit formula for the case $r = 2$, which can be used for an easy implementation. The defined 2D-smoothness indicators are able to precisely inspect the regularity of the gradient of a function, as shown through some numerical test and by the successful application in the construction of the 2D-Adaptive Filtered Scheme. To our knowledge, this is a new contribution.

The main focus of our work is then to construct our new Adaptive Filtered Scheme in one and two dimensions, extending the ideas in [BFS16] through a deeper inspection of the ε-monotonicity property. As previously stated, this is done basically using the smoothness indicator function in order to define the numerical *region of regularity* at a certain time t_n, $\mathcal{R}^n$, as the set of points on the grids in which it is "safe" to compute the regularity threshold ε^n, according to the devised formula. The (main) result of this new formulation is the statement of Theorem 3.11, in which we obtain the same good properties of the convergence theorem in [BFS16], but avoiding some hypotheses which shadow the need to manually tune the parameter ε, thanks to Proposition 3.9. The second consequence, which is mainly testified by the numerical tests at the end of the Chapter 3, is that our scheme is able to exploit better the change in the regularity of the solution, thanks to the tuning of the parameter ε^n at every time step according to the detected regularity. This results in a better behavior in the numerical tests in terms of both errors and stability properties, since our scheme is able to reduce the oscillations otherwise allowed by the constant choice of ε and achieves high-order computed orders in all situations, without the need to introduce any limitator.

Particular attention is given to the class of high-order schemes obtained by the discretization of the Lax-Wendroff method, written compactly for our problem in Lemma 3.2. Using this formula, we are able to devise and successfully test an example of *fourth-order Lax-Wendroff scheme*, which requires only a very compact five-points stencil. Notice that WENO and ENO procedure of higher order usually require many more points for their definitions. Moreover, when applying the Runke-Kutta method, the stencil is increased ad every step required by the order of the time discretization.

Finally we apply our $2D$-Adaptive Filtered scheme to the problem of the segmentation of an image, in the setting of the level set method. The main novelty here is the a new definition of the velocity c, in the aim of [MSV93], designed to force all the level sets to evolve according to the "same" law (in some reasonable sense). With this

definition, justified by the method of characteristics, the velocity becomes anisotropic and dependent also on v. More precisely, it is written as

$$\tilde{c}(x, y, v, v_x, v_y) = c\left(x - d(v)\frac{v_x}{|\nabla v|}, y - d(v)\frac{v_y}{|\nabla v|}\right),$$

where $c(x, y)$ is the velocity of the classical model, computed using the information given by the considered image, and $d(v)$ is a function that measures the (signed) distance between the v-level set and the 0-level set. Then, since this new definition could lead to serious problems in the numerical implementation due to the low regularity of the given data, we avoid most of these complications by proposing a simplified procedure, which in truth should be better justified, especially at those points were the solution is singular and the gradient it is not well defined. However, our simplified approach with the modified velocity $\tilde{c}$ has given very promising preliminary responses, as shown by the numerical tests at the end of Chapter 4. Moreover, the Adaptive Filtered Scheme seems to improve evidently the results obtained by the simple monotone scheme, especially if biomedical images are involved.

At the moment, the only work based on the topics of the present thesis, which is about to be published, is the conference proceeding [FPT18], 'Adaptive Filtered schemes for Hamilton-Jacobi equations'.

Nevertheless, as self-testified by this work, we are almost ready to submit a series of papers containing the new contributions presented throughout these pages. For future references, we would like to mention briefly the ongoing works, in which this thesis will be divided as follows:

- [FPTa], 'Convergence of Adaptive Filtered Schemes for first order time-dependent Hamilton-Jacobi equations', is based on the first parts of Chapter 2 and Chapter 3, and is centered on the proof of the convergence result for the one-dimensional Adaptive Filtered scheme;
- [FPTb], 'Smoothness indicators for multidimensional Adaptive Filtered Schemes', is based on the remaining arguments of Chapter 2 and Chapter 3, that lead to the construction of the 2D-Adaptive Filtered schemes, with main focus on the definition of the multidimensional smoothness indicators;
- [FPTc], 'A High-Order Scheme for Image Segmentation via a modified Level-Set method', is devoted to the application of our approach to the problem of segmentation, presented in Chapter 4.

Organization

The thesis is organized as follows.

In Chapter 1 we prepare the setting for the analysis of the following chapters. In Section 1.1 we briefly review the theoretical results involving existence and uniqueness for (HJ) in the continuous case, starting by the precise definition of viscosity solution and the presentation of the classical method of characteristics. In Section 1.2 we recall the

fundamental properties of monotone schemes, following the pioneering work of [CL84]. Then, in Section 1.3 we give a very minimal presentation of the level set method for front propagation, using the notations in [FF14], and finally, in Section 1.4 we discuss on the applicability of space dimensional splitting methods on first-order Hamilton-Jacobi equations.

In Chapter 2 we face the analysis of the smoothness indicators of [JS96] and [JP00] and the construction of the smoothness indicator function ϕ necessary for the definition of the scheme in the following chapter. In Section 2.1 we prove the main theoretical result on the regularity indicators and present various possible constructions for the function ϕ, mainly different w.r.t. the degree of the interpolating polynomials used in the definition. Then, in Section 2.2 we present the extension of the formulas in the 2D case, giving an explicit formula for the case $r = 2$. Finally, Section 2.3 is dedicated to a series of numerical tests on $1D$- and $2D$- functions with varying regularity, in which we show the implications of Proposition 2.1 and the accuracy of the genuinely $2D$ extension w.r.t. a more simple construction, based on the dimensional splitting of one-dimensional indicators.

In Chapter 3 we define our new Adaptive Filtered Scheme in one and two space dimensions, proving a convergence result with error estimates. In Section 3.1 we begin by giving the first definitions and recalling the setting of the problem. Then, in Section 3.2 we proceed with the construction of the scheme, defining in detail all of its components. Once all of the basic components have been defined, in Section 3.3 we prove the stability and convergence property of our scheme, stated by Proposition 3.9 and Theorem 3.11. Subsequently, in Section 3.4 we extend all the constructions in order to apply the adaptive tuning of ε^n also in the $2D$-setting. As a result, we define a scheme which can be applied, for instance, in Chapter 4 for the segmentation problem. Finally, in Section 3.5 we conclude our main chapter by presenting some numerical tests for one and two dimensional problems, comparing the results obtained by the Adaptive Filtered Scheme with those of the simple Filtered Scheme of [BFS16] and of the WENO scheme of second/third order.

In Chapter 4 we conclude the thesis by presenting an interesting application, coming from image processing concerns. In Section 4.1 we begin by recalling the general setting for segmentation problems, giving also some references on other possible approaches. Then, in Section 4.2 we first recall the theory of the level set method for image segmentation, as presented in [MSV93], whence we propose our modification in the definition of the velocity c and finally, we justify its definition by the method of characteristics. Section 4.3 is focused on the discussion on the numerical issues arising with the modified model, for which we propose a simplification in order to avoid most of the mentioned complications, and on the presentation of our numerical implementation, in the form of a sketched pseudo-code. Lastly, in Section 4.4 we present a collection of numerical test, on both synthetic and real images, with main focus on biomedical applications, in which the effectiveness of the modified model and the improvements of the AFS over the monotone scheme are clearly recognizable.

In Appendix A we collect some technical results needed in Chapter 2 for the proof of Proposition 2.1.

1. Background results

In this chapter we collect some background results useful for a complete understanding of the ideas developed in the following chapters. We will begin by reviewing the classical theory of *viscosity solutions* for Hamilton-Jacobi equations, focusing our attention mainly on the equation

$$v_t + H(x, \nabla v) = 0, \qquad (t, x) \in \mathbb{R}^+ \times \mathbb{R}^N, \tag{1.1}$$

where ∇ (if not specified) denotes the gradient w.r.t. the x variable only, which will be of primary interest in the rest of the thesis. Then, in Section 1.2 we collect some fundamental and well-known results on monotone schemes and finally, in the last two sections, we present two closely related problems: the *level set method* in Section 1.3, and the *dimensional splitting* in Section 1.4.

1.1. Overview on first-order Hamilton-Jacobi equations

In this section we briefly review some classical theoretical results, avoiding to present the details of the proofs in most cases. For a complete presentation we refer the reader, for example, to the books [BCD97], [B98], [FF14] and [E10]

Let us consider a first-order evolutive Hamilton-Jacobi equation, which in the most general case reads

$$\begin{cases} v_t + H(t, x, v, \nabla v) = 0 & (t, x) \in \mathbb{R}^+ \times \mathbb{R}^N \\ v(0, x) = v_0(x) & x \in \mathbb{R}^N, \end{cases} \tag{1.2}$$

where $H : \mathbb{R}^+ \times \mathbb{R}^N \times \mathbb{R} \times \mathbb{R}^N \to \mathbb{R}$ is, at least, a Lipschitz continuous function, generally nonlinear, called *hamiltonian* and $v_0 \in BUC(\mathbb{R}^N)$ is the initial condition (usually also Lipschitz continuous in the sequel). We remind that $BUC(\Omega)$ denotes the space of bounded and uniformly continuous functions on the domain Ω.

It is well-known that in general this problem does not admit classical solutions ($v \in C^1$), independently on the regularity of the initial condition, since its solutions may develop discontinuities in the gradient in finite time. Consequently, we have to look for a suitable definition of solution in some "weak sense". Here we only review the definition for Lipschitz continuous functions ($v \in W^{1,\infty}(\mathbb{R}^+ \times \mathbb{R}^N)$), referring the reader to [B98] for the extension to the discontinuous case.

1.1.1. Method of characteristics

Before introducing the right definition of solution for (1.2), let us recall that, under suitable assumptions, classical solutions may be constructed locally by the so-called

method of characteristics. In order to stay close to the majority of the literature on the topic, we show the procedure considering a general first order equation

$$F\left(y, v, \nabla_y v\right) = 0 \qquad \text{in } \Omega \subseteq \mathbb{R}^d, \tag{1.3}$$

then we give an example of classical solution for the simplified problem

$$\begin{cases} v_t + H(\nabla v) = 0 & \text{in } \Omega = \mathbb{R}^+ \times \mathbb{R}^N \\ v(0,x) = v_0(x) & \forall x \in \mathbb{R}^N. \end{cases} \tag{1.4}$$

Fix a point $\overline{y} \in \Omega$, the idea is to find a curve $y(s)$ inside Ω along which the solution $v(y(s))$ can be easily computed (in the time-dependent case, usually $s = t$ and $y(t)$ connects $\overline{y}$ to $y(0)$, then the initial condition v_0 can be used).

Let us consider a curve $y(s)$ to be determined, and define $z(s) := v(y(s))$, $q(s) := \nabla_y v(y(s))$. Next, we derive equations for y, z, and q. Clearly, we have

$$\dot{z}(s) = \dot{y}(s) \cdot q(s) \quad \text{ and } \quad \dot{q}(s) = \frac{d}{ds}\left(\nabla_y v(y(s))\right) = D_y^2 v(y(s)) \cdot \dot{y}(s),$$

where the hessian matrix $D_y^2 v$ is not known and should be "removed". In order to cancel it, we differentiate equation (1.3) w.r.t. y, obtaining

$$\nabla_y F + F_z \nabla_y v + D_y^2 v \cdot \nabla_q F = 0 \iff D_y^2 v \cdot \nabla_q F = -\nabla_y F - F_z \nabla_y v,$$

whence, it is enough to require $\dot{y}(s) = \nabla_q F$. Therefore, the *characteristics* of the equation are obtained solving the system of ordinary differential equation

$$\begin{cases} \dot{y} = \nabla_q F, \\ \dot{z} = \nabla_q F \cdot q, \\ \dot{q} = -F_z q - \nabla_y F. \end{cases} \tag{1.5}$$

In the case of the evolution equation (1.2), we can rewrite this system in a slightly more convenient form. Using the notations $y = (t, x)$ and $q = (p_0, p) = (v_t, \nabla v)$, we have

$$F(y, z, q) = p_0 + H(y, z, p),$$

and the system (1.5) reads

$$\begin{aligned} &\dot{y} = \nabla_q F = \begin{pmatrix} 1 \\ \nabla_p H \end{pmatrix} \iff \dot{t} = 1, \quad \dot{x} = \nabla_p H, \\ &\dot{z} = \nabla_q F \cdot q = p_0 + \nabla_p H \cdot p = \nabla_p H \cdot p - H, \\ &\dot{q} = -F_z q - \nabla_y F \iff \dot{p}_0 = H_z H - H_t, \quad \dot{p} = -H_z p - \nabla_x H, \end{aligned}$$

then, since we can choose $s = t$ as parametrization and the equation for z does not depend explicitly on p_0, we can consider the reduced *characteristic system*

$$\begin{cases} \dot{x} = \nabla_p H, & x(0) = x, \\ \dot{z} = \nabla_p H \cdot p - H, & z(0) = v_0(x) \\ \dot{p} = -H_z p - \nabla_x H, & p(0) = \nabla v_0(x), \end{cases} \tag{1.6}$$

where now we added the initial conditions to close the problem.

Let us conclude this little digression by going back to our example problem (1.4), for which the system (1.6) simply reads

$$\begin{cases} \dot{x}(t) = \nabla_p H & x(0) = x \\ \dot{z}(t) = \nabla_p H \cdot p - H, & z(0) = v_0(x) \\ \dot{p}(t) = 0, & p(0) = \nabla v_0(x), \end{cases} \tag{1.7}$$

which solution for the first and third equation is

$$x(t;x) = x + t\nabla_p H(\nabla v_0(x)), \qquad p(t) \equiv \nabla v_0(x),$$

as can be easily verified. Therefore, the solution given by the method of characteristics reads

$$v(t,x) = v_0\left(x^{-1}(t;x)\right) + t\left(\nabla_p H \cdot \nabla v_0 - H(\nabla v_0)\right)\left(x^{-1}(t;x)\right), \tag{1.8}$$

which domain in time depends on the invertibility of the map $x \mapsto x(t;x)$. In general, this map can be inverted only for relatively small times $t > 0$ and, consequently, the solution v given by (1.8) it is not globally defined. In any case, under suitably strong assumptions on the functions H and g, the solution previously constructed can be globally defined, as stated by the following theorem, proved in [L82].

Theorem 1.1. *Assume H and v_0 in $C^2(\mathbb{R}^N)$ and convex. Then, the function v defined by (1.8) is a classical solution of the problem (1.4).*

1.1.2. Viscosity solutions

Since the equation is not in divergence form, we can not use integration by parts to define weak solutions, as it can be done instead, for example, in the case of conservation laws

$$u_t + \sum_{i=1}^{N} \frac{\partial}{\partial x_i} f_i(u) = 0 \quad (t,x) \in \mathbb{R}^+ \times \mathbb{R}^N,$$

which in one space dimension are closely related to (1.2) (see e.g. [FF14], Section 2.4.). Therefore, we are forced to use a different technique which leads to the definition of *viscosity solution*, presented in the this section for the general case of (1.2).

First, we have to clarify what we mean for "weak solution" in this setting. Let us consider an open set $\Omega \subseteq \mathbb{R}^+ \times \mathbb{R}^N$ and reintroduce the shorter notation $y = (t,x)$.

Definition 1.2. A function $v : \Omega \to \mathbb{R}$ is called

i) *classical solution* for (1.2) in Ω if $v \in C^1(\Omega)$ and satisfies

$$v_t + H(y, v, \nabla v) = 0, \qquad \forall\ y \in \Omega;$$

ii) *generalized solution* for (1.2) in Ω if v is Lipschitz continuous in Ω and satisfies

$$v_t + H(y, v, \nabla v) = 0, \qquad \text{a.e. in } \Omega.$$

Unfortunately, this last "generalized" definition it is not enough to select a unique solution for the Cauchy problem, as can be noted by considering the classical example

$$\begin{cases} v_t + |v_x| = 0, & (t,x) \in \mathbb{R}^+ \times \mathbb{R}, \\ v_0(x) = |x| & x \in \mathbb{R}. \end{cases}$$

In fact, it is straightforward to verify that the functions $v_1(t,x) = |x| - t$ and $v_2(t,x) = (|x| - t)^+$ are both generalized solutions.

In order to select a unique solution we have to rely on the *vanishing viscosity method* which, in particular, lets us characterize the solution according to the following

Definition 1.3. A function $v \in BUC(\Omega)$ is a *viscosity solution* of (1.2) if and only if, for any $\psi \in C^1(\Omega)$, the following conditions hold:

i) for any $y_0 \in \Omega$ local maximum point for $(v - \psi)(y)$,

$$\psi_t(y_0) + H\left(y_0, v(y_0), \nabla\psi(y_0)\right) \leq 0,$$

(i.e. v is a *viscosity subsolution*);

ii) for any $y_0 \in \Omega$ local minimum point for $(v - \psi)(y)$,

$$\psi_t(y_0) + H\left(y_0, v(y_0), \nabla\psi(y_0)\right) \geq 0,$$

(i.e. v is a *viscosity supersolution*).

This notion of solution was first introduced by Crandall and Lions in [CL83]. It is straightforward to show that Definition 1.3 is coherent with the classical sense, that is, if $v \in C^1(\Omega)$, then v is a viscosity solution if and only if it is a classical solution.

***Remark* 1.1.** In Definition 1.3 we can assume, without loss of generality, that y_0 is a *strict* maximum (respectively, minimum) point for $(v - \psi)(y)$; in fact, it is enough to replace ψ with $\psi(y) + |y - y_0|^2$. Moreover, the test functions can be assumed arbitrary regular (but at least C^1) and the maximum (resp. minimum) points can be chosen *global*.

A fundamental property of viscosity solutions is the stability w.r.t. local uniform convergence, differently from generalized solutions, for which it is possible to exhibit examples of sequences that converge to functions which are not solutions of the problem (in any sense!).

Here, in order to highlight a simple and direct implication of this property, we present the result in the more general contest of *degenerate elliptic equation*, which have the form

$$G\left(y, v, \nabla_y v, D_y^2 v\right) = 0 \qquad \text{in } \Omega,$$

where $G : \Omega \times \mathbb{R} \times \mathbb{R}^{N+1} \times \mathcal{S}^{N+1} \to \mathbb{R}$ is a continuous function that satisfies the *(degenerate) ellipticity condition*

$$(y, z, q, M_1) \leq G(y, z, q, M_2) \qquad \text{if } M_1 \geq M_2,$$

for all $y \in \Omega$, $z \in \mathbb{R}$, $q \in \mathbb{R}^{N+1}$ and M_1, $M_2 \in \mathcal{S}^{N+1}$, where $\mathcal{S}^{N+1}$ is the space of $(N+1) \times (N+1)$ symmetric matrices, with the usual ordering

$$M_1 \geq M_2 \iff (M_1\xi, \xi) \geq (M_2\xi, \xi) \ \forall \xi \in \mathbb{R}^{N+1}.$$

The definition of viscosity solution in this contest is analogous to (1.3). Then, we can state the stability result.

Theorem 1.4. *Assume that, for $\varepsilon > 0$, $v_\varepsilon \in C(\Omega)$ is a viscosity subsolution (resp. supersolution) of the equation*

$$G_\varepsilon\left(y, v_\varepsilon, \nabla_y v_\varepsilon, D_y^2 v_\varepsilon\right) = 0 \qquad \text{in } \Omega,$$

where $\{G_\varepsilon\}_{\varepsilon>0}$ is a family of continuous functions satisfying the ellipticity condition. If $v_\varepsilon \to v$ uniformly on compact sets and $G_\varepsilon \to G$ uniformly on compact sets, then v is a subsolution (resp. supersolution) of the equation

$$G\left(y, v, \nabla_y v, D_y^2 v\right) = 0 \qquad \text{in } \Omega.$$

Proof. (Sketch) We give the proof only in the case of subsolutions, with the other being analogous. Under the hypothesis on v_ε and v, it can be shown that, if $y_0 \in \Omega$ is a strict local maximum point for $v - \psi$, then there exists a sequence of local maximum points for $v_\varepsilon - \psi$, denoted with $\{y_\varepsilon\}_\varepsilon$, such that $y_\varepsilon \to y_0$. At those points, by definition, we have

$$G_\varepsilon\left(y_\varepsilon, v_\varepsilon(y_\varepsilon), \nabla_y \psi(y_\varepsilon), D_y^2 \psi(y_\varepsilon)\right) \leq 0,$$

then, exploiting the regularity of the test functions and the convergence hypothesis, we can conclude

$$G_\varepsilon\left(y_\varepsilon, v_\varepsilon(y_\varepsilon), \nabla_y \psi(y_\varepsilon), D_y^2 \psi(y_\varepsilon)\right) \to G\left(y_0, v(y_0), \nabla_y \psi(y_0), D_y^2 \psi(y_0)\right)$$

whence

$$G\left(y_0, v(y_0), \nabla_y \psi(y_0), D_y^2 \psi(y_0)\right) \leq 0.$$

□

One of the most direct application of these result is the vanishing viscosity method for first-order equations, represented by the sequence of equations

$$-\varepsilon \Delta v_\varepsilon + u_t + H(y, v_\varepsilon, \nabla v_\varepsilon) = 0 \qquad \text{in } \Omega,$$

for which the function G_ε is given by

$$G_\varepsilon(y, v, q, M) = -\varepsilon \mathrm{Tr}(M_{[N,N]}) + p_0 + H(y, v, p),$$

where $q = (p_0, p) \in \mathbb{R}^{N+1}$ and $M_{[N,N]}$ is the $N \times N$ matrix obtained by deleting the first row and first column from M (we are considering only spatial derivatives).

Next, we give an equivalent definition of viscosity solutions which was the first introduced in [CL83] and will let us prove easily the connection between generalized and viscosity solutions. Before that, we have to introduce some standard tools of convex analysis.

Definition 1.5. Consider a continuous function $v : \Omega \to \mathbb{R}$ and $y \in \Omega$.
The *subdifferential* of v at y is defined as the set

$$\nabla^- v(y) = \left\{ q \in \mathbb{R}^{N+1} : \liminf_{z \to y} \frac{v(z) - v(y) - q \cdot (z - y)}{|z - y|} \geq 0 \right\}.$$

An element $q \in \nabla^- v(y)$ is called *subgradient*.
Analogously, the *superdifferential* of v at y is defined as the set

$$\nabla^+ v(y) = \left\{ q \in \mathbb{R}^{N+1} : \limsup_{z \to y} \frac{v(z) - v(y) - q \cdot (z - y)}{|z - y|} \leq 0 \right\}.$$

An element $q \in \nabla^+ v(y)$ is called *supergradient*.

It is good to point out that, if v is differentiable at y, then $\nabla^+ v(y) = \nabla^- v(y) = \{\nabla v(y)\}$. Without giving the proof, which can be found in [BCD97] (Lemma 1.7, p. 29), we characterize the previous sets in terms of test functions.

Proposition 1.6. *Let $v : \Omega \to \mathbb{R}$. Then,*

i) $q \in \nabla^+ v(y)$ if and only if $\exists \psi \in C^1(\Omega)$ such that $\nabla \psi(y) = q$ and $v - \psi$ has a local maximum point at y;

ii) $q \in \nabla^- v(y)$ if and only if $\exists \psi \in C^1(\Omega)$ such that $\nabla \psi(y) = q$ and $v - \psi$ has a local minimum point at y.

Now, it is rather natural the following

Definition 1.7. A function $v \in \mathcal{C}(\Omega)$ is a *viscosity subsolution* of (1.2) in Ω if

$$p_0 + H(y, v(y), p) \leq 0 \qquad \forall y \in \Omega,\ \forall q = (p_0, p) \in \nabla^+ v(y).$$

In the same way, $v \in \mathcal{C}(\Omega)$ is a *viscosity supersolution* of (1.2) in Ω if

$$p_0 + H(y, v(y), p) \geq 0 \qquad \forall y \in \Omega,\ \forall q = (p_0, p) \in \nabla^- v(y).$$

Finally, v is a *viscosity solution* if satisfies both of the above conditions.

Thanks to this new characterization we can easily express the connection between viscosity and generalized solutions, in fact we have the following

Proposition 1.8. *If $v : \Omega \to \mathbb{R}$ is a continuous viscosity solution of (1.2), then*

$$v_t + H(y, v(y), \nabla v(y)) = 0$$

at every point where v is differentiable. Moreover, if v is also Lipschitz continuous, then

$$v_t + H(y, v(y), \nabla v(y)) = 0 \qquad \text{a.e. in } \Omega.$$

Proof. Let y be a point of differentiability for u. Then, $\nabla^+ v(y) \cap \nabla^- v(y) \neq \emptyset$ since this set contains $\nabla v(y)$. Moreover, this intersection reduces to a singleton, so that

$$\nabla^+ v(y) = \nabla^- v(y) = \nabla v(y).$$

Hence, by Definition 1.7 we have

$$0 \leq v_t + H(y, v(y), \nabla v(y)) \leq 0,$$

which proves the first assertion. Then, the second statement follows from the previous reasoning and applying Rademacher's theorem (which states that Lipschitz continuous functions are a.e. differentiable, see e.g. [E10], pp. 280-281 for the proof). □

1.1.3. Existence and uniqueness results

In this section we collect some classical results of uniqueness and existence of viscosity solutions, focusing our attention on the evolutive equation

$$v_t + H(x, \nabla v) = 0, \qquad \text{in } \mathbb{R}^+ \times \mathbb{R}^N \tag{1.9}$$

and giving some hints, if possible, on closely related generalizations. For a more detailed presentation we refer the reader to [B98].

In the following results we will require that $H(x,p)$ satisfies the usual set of assumptions:

$$\lim_{|p| \to +\infty} \inf_{x \in \mathbb{R}^N} H(x,p) = +\infty; \tag{H1}$$

$$H \in BUC(\mathbb{R}^N \times B_r(0)) \quad \forall r > 0; \tag{H2}$$

$$\begin{aligned} &\exists \omega(\cdot) : [0, +\infty) \to [0, +\infty) \text{ modulus of continuity such that} \\ &|H(x,p) - H(y,p)| \leq \omega(|x-y|(1+|p|)) \ \forall x, y \in \mathbb{R}^N, \forall p \in R^N. \end{aligned} \tag{H3}$$

The main issue in the theory of viscosity solutions is to prove uniqueness, which also shows the advantage over the concept of generalized *a.e.* solutions. This is done via the so-called *comparison principle*, also termed *maximum principle*, in the sense of the following

Definition 1.9. Let $u, v \in C(\mathbb{R}^+ \times \mathbb{R}^N)$ be, respectively, a sub- and a supersolution for (1.9). Then, we say that a *comparison principle* holds for u and v if

$$u(0,x) \leq v(0,x) \ \forall x \in \mathbb{R}^N \Rightarrow u(t,x) \leq v(t,x) \ \forall (t,x) \in \mathbb{R}^+ \times \mathbb{R}^N.$$

In the case such a relation holds, the proof of uniqueness is straightforward. At this point, we can state a first fundamental result.

Theorem 1.10. *Let $H : \mathbb{R}^N \times \mathbb{R}^N \to \mathbb{R}$ satisfy (H1), (H2) and (H3). Let $u(t,x)$, $v(t,x) \in UC(\mathbb{R}^+ \times \mathbb{R}^N)$ be, respectively, a sub- and a supersolution of (1.9), then a comparison principle holds for u and v. Moreover, the same results holds if (H3) is replaced by the Lipschitz continuity in x, uniformly w.r.t. t, of u or v.*

Proof. (Sketch) In order to give a sketch of the proof we consider the slightly easier case of u and v also bounded and, to lighten the notation, we name $\Omega = \mathbb{R}^+ \times \mathbb{R}^N$. We argue by contradiction assuming that $u \leq v$ does not hold for all $(t,x) \in \Omega$, whence $\exists \lambda > 0$ such that

$$\sup_{\Omega}\{u(t,x) - v(t,x) - 2\lambda t\} =: \sigma > 0.$$

To exploit the definition of viscosity solution we resort to a "doubling of variables" argument, defining

$$\Phi_{\varepsilon,\alpha}(t,x,s,y) = u(t,x) - v(s,y) - \lambda(t+s) - \frac{(|x-y|^2+|t-s|^2)}{\varepsilon^2} - \alpha(|x|^2+|y|^2+|t|^2+|s|^2).$$

The presence of the penalisation terms and the boundedness of the functions u and v ensures the existence of a point $(t_{\varepsilon,\alpha}, x_{\varepsilon,\alpha}, s_{\varepsilon,\alpha}, y_{\varepsilon,\alpha}) \in \Omega \times \Omega$ such that

$$M_{\varepsilon,\alpha} := \Phi_{\varepsilon,\alpha}(t_{\varepsilon,\alpha}, x_{\varepsilon,\alpha}, s_{\varepsilon,\alpha}, y_{\varepsilon,\alpha}) = \max_{\Omega\times\Omega} \Phi_{\varepsilon,\alpha}.$$

Consider the function

$$\phi(t,x) = u(t,x) - \Phi_{\varepsilon,\alpha}(t,x,s_{\varepsilon,\alpha}, y_{\varepsilon,\alpha}),$$

which clearly is a test function such that $u - \phi$ achieves its maximum at the point $(t_{\varepsilon,\alpha}, x_{\varepsilon,\alpha})$. Then, by definition we have that

$$\lambda + \frac{2(t_{\varepsilon,\alpha} - s_{\varepsilon,\alpha})}{\varepsilon^2} + 2\alpha t_{\varepsilon,\alpha} + H\left(x_{\varepsilon,\alpha}, \frac{2(x_{\varepsilon,\alpha} - y_{\varepsilon,\alpha})}{\varepsilon^2} + 2\alpha x_{\varepsilon,\alpha}\right) \leq 0.$$

Analogously, defining the test function

$$\psi(s,y) = \Phi_{\varepsilon,\alpha}(t_{\varepsilon,\alpha}, x_{\varepsilon,\alpha}, s, y) + v(s,y)$$

such that $v - \psi$ achieves its minimum point at $(s_{\varepsilon,\alpha}, y_{\varepsilon,\alpha})$, we get

$$-\lambda + \frac{2(t_{\varepsilon,\alpha} - s_{\varepsilon,\alpha})}{\varepsilon^2} - 2\alpha s_{\varepsilon,\alpha} + H\left(y_{\varepsilon,\alpha}, \frac{2(x_{\varepsilon,\alpha} - y_{\varepsilon,\alpha})}{\varepsilon^2} - 2\alpha y_{\varepsilon,\alpha}\right) \geq 0.$$

Manipulating the previous relations and exploiting the hypothesis on H, eventually, we can arrive to the inequality

$$2\lambda \leq \eta(2\alpha|y_{\varepsilon,\alpha}|) + \eta(2\alpha|x_{\varepsilon,\alpha}|) + \omega\left(|y_{\varepsilon,\alpha} - x_{\varepsilon,\alpha}|\left(1 + \left|\frac{2(x_{\varepsilon,\alpha} - y_{\varepsilon,\alpha})}{\varepsilon^2} + 2\alpha x_{\varepsilon,\alpha}\right|\right)\right),$$

with η modulus of continuity of H on $\mathbb{R}^N \times B_K$, where K depends on $R = \max\{||u||_\infty, ||v||_\infty\}$ ed ε.

At this point, it is enough to prove that the term on the right side of the previous inequality tends to 0 as $(\varepsilon,\alpha) \to (0,0)$ to get the contradiction $\lambda \leq 0$. This can be done proving that (all the limits are for $(\varepsilon,\alpha) \to (0,0)$):

- $M_{\varepsilon,\alpha} \to \sigma$;
- $|x_{\varepsilon,\alpha} - y_{\varepsilon,\alpha}| + |t_{\varepsilon,\alpha} - s_{\varepsilon,\alpha}| \to 0$;
- $\frac{(|x_{\varepsilon,\alpha}-y_{\varepsilon,\alpha}|^2+|t_{\varepsilon,\alpha}-s_{\varepsilon,\alpha}|^2)}{\varepsilon^2} + \alpha(|x_{\varepsilon,\alpha}|^2+|y_{\varepsilon,\alpha}|^2+|t_{\varepsilon,\alpha}|^2+|s_{\varepsilon,\alpha}|^2) \to 0.$

For the details of the computations we refer to [B98]. □

If we consider a bounded domain of the form $\mathcal{Q} = (0,T) \times \Omega$, with Ω bounded set of $\mathbb{R}^N$, then it is possible to state an analogous result for the equation

$$v_t + H(t,x,\nabla v) = 0. \tag{1.10}$$

In the latter situation the comparison principle will refer to the *parabolic boundary* $\partial_p \mathcal{Q} = \partial\Omega \times [0,T] \cup \overline{\Omega} \times \{0\}$ and will require only hypothesis (H3). In particular, it is worth noticing that no specific assumptions are made on the behavior of H w.r.t. p, then we have a uniqueness result for the equation

$$v_t + H(\nabla v) = f(t,x) \qquad \text{in } \mathcal{Q}, \tag{1.11}$$

with f continuous function on $\mathcal{Q}$, assuming simply the continuity of H.

In the more general case in which the hamiltonian depends also on v a further hypothesis is needed in order to avoid the formation of discontinuities. Usually it is assumed that, for all $0 < R < +\infty$, $\exists \gamma_R \in \mathbb{R}$ such that, $\forall (t,x) \in \mathbb{R}^+ \times \mathbb{R}^N$, $-R \le v \le u \le R$ and $p \in \mathbb{R}^N$,

$$H(x,u,p) - H(x,v,p) \ge \gamma_R(u-v). \tag{H4}$$

Let us consider now equation (1.10) in $[0,T] \times \mathbb{R}^N$ and assume

$$|H(t,x,p) - H(t,x,q)| \le C|p-q|, \qquad \forall x \in \mathbb{R}^N,\ t \in [0,T] \text{ e } p,q \in \mathbb{R}^N, \tag{H5}$$

then, the following *finite propagation speed* type result, which we present without proof (see [B98], pg. 72), holds.

Theorem 1.11. *Assume (H2)-(H5) and let u, $v \in C([0,T] \times \mathbb{R}^N)$ be, respectively, a viscosity sub- and supersolution. If $u(0,x) \le v(0,x)$ in $B_r(0)$ for some $r > 0$, then $u(t,x) \le v(t,x)$ for all x in $B_{r-Ct}(0)$, $Ct \le r$.*

Next, we focus on showing the existence of a solution in the viscosity sense. This is usually done by the *Perron method*, which basically constructs a solution taking the supremum of all the subsolutions of the problem. We give an example of application for the Cauchy problem

$$\begin{cases} v_t + H(x,\nabla v) = 0 & \text{in } \mathbb{R}^+ \times \mathbb{R}^N \\ v(0,x) = v_0(x) & \forall x \in \mathbb{R}^N. \end{cases} \tag{1.12}$$

Define the set

$$\mathcal{S} = \{u \in C(\mathbb{R}^+ \times \mathbb{R}^N) :\ u(0,x) = v_0(x)\ \forall x \in \mathbb{R}^N,\ u(t,x) \text{ viscosity subsolution}\}; \tag{1.13}$$

then, the Perron solution is obtained choosing

$$v(t,x) = \sup_{u\in\mathcal{S}} u(t,x).$$

With this simple construction, we can prove the following existence result.

Theorem 1.12. *Let H satisfy (H1), (H2), (H3) and let $v_0 \in UC(\mathbb{R}^N)$. Then, there exists a unique solution $v(t,x) \in UC(\mathbb{R}^+ \times \mathbb{R}^N)$ of the problem (1.12). If moreover $v_0 \in W^{1,\infty}(\mathbb{R}^N)$, then $v(t,x) \in W^{1,\infty}(\mathbb{R}^+ \times \mathbb{R}^N)$. In particular, if $C > 0$ is such that*

$$||\nabla v_0||_\infty < C \quad \textit{and} \quad \inf\{H : |p| > C\} > \sup\{|H| : |p| \leq ||\nabla v_0||_\infty\}$$

then $||\nabla_x v||_\infty \leq C$.

Proof. (Sketch) Let us give an idea of the proof of the existence in the case $v_0 \in W^{1,\infty}$. First, we notice that the set $\mathcal{S}$ defined by (1.13) is not empty, in fact (H1) and (H2) imply that $\exists C_0 > 0$, $C_1 \in \mathbb{R}$, such that

$$\begin{aligned} &H(x,p) \geq -C_0 \quad \forall\ (x,p) \in \mathbb{R}^N \times \mathbb{R}^N \\ &H(x,p) \leq C_1 \quad \forall\ x \in \mathbb{R}^N,\ \forall\ p \in B_R,\ R = ||\nabla v_0||_\infty, \end{aligned}$$

then, defining $M := \max\{C_0, C_1\}$, it follows that $\underline{v}(t,x) = v_0(x) - Mt$ is a subsolution in $C(\mathbb{R}^+ \times \mathbb{R}^N)$ such that $\underline{v}(0,x) = v_0(x)$, whence $\underline{v} \in \mathcal{S}$. Now, let

$$v(t,x) = \sup_{u\in\mathcal{S}} u(t,x).$$

The next step is to prove that v is a subsolution. This is done by considering the definition $v(t,x) = \sup_{\alpha\in A} v_\alpha$ and then proceed by induction on the cardinality of the set A. First thing is to prove that the supremum of a finite set of subsolutions is a subsolution, then consider a numerable set of subsolutions and, finally, an aribrary set A.
The prove that v is a supersolution we proceed by contradiction. Then, if we assume that v is not a supersolution, we can find $(t_0,x_0) \in \mathbb{R}^+ \times \mathbb{R}^N$, $\delta > 0$, $\beta > 0$ and $r > 0$ such that

$$\psi_t(t,x) + H(x, \nabla_x \psi(t,x)) \leq -\beta, \quad \forall (t,x) \in B_r((t_0,x_0)) \text{ t.c. } |\nabla_x \psi(t,x)| \leq \delta,$$

with ψ such that $v - \psi$ has a maximum at (t_0,x_0). At this point, we consider the function

$$v_\varepsilon(t,x) = \max\{v, \psi + \varepsilon - |t-t_0|^2 - |x-x_0|^2\},$$

with $\varepsilon > 0$ such that $\sqrt{\varepsilon} < r$ and for which

$$(0,x) \notin \{(t,x) : v(t,x) < \psi(t,x) + \varepsilon - |t-t_0|^2 - |x-x_0|^2\},\ \forall x \in \mathbb{R}^N.$$

Finally, exploiting the definition of ε and the fact that function which are separately subsolutions on open sets that cover $\mathbb{R}^N$ are subsolutions of $\mathbb{R}^N$, we can prove that v_ε is a subsolution and $v_\varepsilon(0,x) = v_0(x)$, that is $v_\varepsilon \in \mathcal{S}$. Moreover,

$$v_\varepsilon(t_0,x_0) = \psi(t_0,x_0) + \varepsilon = v(t_0,x_0) + \varepsilon > v(t_0,x_0),$$

which contradicts the maximality of v. □

For completeness of presentation, we recall also the property of the viscosity solution in the simpler case (1.4), but in all generality (we need only the continuity of H). More precisely, defining $S(t) : BUC(\mathbb{R}^N) \to BUC(\mathbb{R}^N)$, as the *solution map* $S(t)(v_0(x)) := v(t,x)$, of problem (1.4), then the following result hols.

Proposition 1.13. *Let $H \in \mathcal{C}(\mathbb{R}^N)$, u_0, $v_0 \in BUC(\mathbb{R}^N)$ and $t \geq 0$. Then,*

i) $||(S(t)u_0 - S(t)v_0)^+||_\infty \leq ||(u_0 - v_0)^+||_\infty$;

ii) $||S(t)u_0 - S(t)v_0||_\infty \leq ||u_0 - v_0||_\infty$;

iii) $\inf_{\mathbb{R}^N} v_0 \leq tH(0) + S(t)v_0 \leq \sup_{\mathbb{R}^N} v_0$;

iv) $|S(t)v_0(x+y) - S(t)v_0(x)| \leq \sup_{z\in\mathbb{R}^N} |v_0(z+y) - v_0(z)|$, *per* $y \in \mathbb{R}^N$;

v) If v_0 is Lipschitz continuous with constant L_{v_0}, then $S(t)v_0$ is Lipschitz continuous with the same constant, moreover, it holds

$$||S(t)v_0 - S(\tau)v_0||_\infty \leq |t-\tau| \sup_{\mathbb{R}^N} \{|H(p)|\colon |p| \leq L_{v_0}\}.$$

For the proof of this last result we refer the reader to the paper [CL83] (in particular see pp. 38-39), in which they also prove the existence of the solution through the vanishing viscosity method (see also [CL84], p. 16).

1.2. Convergence result for monotone schemes

After having clarified the theoretical setting, we make a first step into the main topic of this thesis, that is the numerical approximation of the solution of the problem

$$\begin{cases} v_t + H(\nabla v) = 0, & \text{in } \Omega := \mathbb{R}^+ \times \mathbb{R}^N \\ v(0,x) = v_0(x), \end{cases} \tag{1.14}$$

where $H : \mathbb{R}^N \to R$ is, at least, a continuous function and v_0 sufficiently regular. Notice that we are considering the problem in the most simple case, in which the hamiltonian depends only on the gradient. That is in fact a rather usual approach, since the main numerical difficulties come from the nonlinear dependence of H on ∇v. Moreover, more general cases usually follow with minor modifications. Here we basically follow the lines of [CL84], aiming to state a convergence result for first-order monotone schemes.

In the following, mainly to lighten the notation, we will present the result for the case $N = 2$. Let us begin by considering a uniform grid in space $(x_j, y_i) = (j\Delta x, i\Delta y)$, $j, i \in \mathbb{Z}$, and in time $t_n = n\Delta t$, $n \in [0, N_T]$, with $(N_T - 1)\Delta t < T \leq N_T \Delta t$. Then, we define the notation $v^n_{i,j}$ for the values of the exact solution at the point (t_n, x_j, y_i) and $u^n_{i,j}$ for the relative values of the approximate solution. A generic two-level explicit scheme will be written as

$$u^{n+1} = S_\Delta(u^n), \tag{1.15}$$

for some operator S_Δ, where $\Delta := (\Delta t, \Delta x, \Delta y)$ represents the dependence on the discretization parameters. In what follows we use the shorter notation S, since it should not cause any confusion.

***Remark* 1.2.** Notice that we are considering the unrealistic case of an infinite grid in space $(i,j) \in \mathbb{Z}^2$. That is because we want to avoid any discussion on the boundary conditions, which should be treated separately. Alternatively, if we consider a bounded domain, then we assume to possess also the external values of such set, if necessary.

Definition 1.14. A scheme (1.15), dependent on $d = (p+q) \times (r+s)$ points, can be written in *differenced form* if there exists a function $h : \mathbb{R}^d \to \mathbb{R}$, called *numerical hamiltonian*, such that

$$S_\Delta(u_{i-p,j-r}, \dots, u_{i+q+1,j+s+1}) = u_{i,j} - \Delta t h\left(D_x^+ u_{i-p,j-r}, \dots, D_x^+ u_{i+q,j+s+1}; D_y^+ u_{i-p,j-r}, \dots, D_y^+ u_{i+q+1,j+s}\right), \quad (1.16)$$

where $D_x^+ u_{i,j}^n := \frac{u_{i,j\pm1}^n - u_{i,j}^n}{\Delta x}$ and $D_y^+ u_{i,j}^n := \frac{u_{i\pm1,j}^n - u_{i,j}^n}{\Delta y}$.

In the case of schemes in differenced form, the consistency condition can be stated in a rather easy formulation, that is

Definition 1.15. A scheme (1.15) in differenced form is *consistent* with the equation (1.14) if, for any $a, b \in \mathbb{R}$,

$$h(a, \dots, a; b \dots b) = H(a, b).$$

A particular class of schemes, that generally satisfy the previous definition, are the so-called *monotone schemes*, which, following [CL84], we define as

Definition 1.16. A scheme S^M of the form (1.15) is *monotone* on $[-R, R]$ if S^M is a nondecreasing function of each argument, as long as $|D_x^+ u^n|, |D_x^+ u^n| \leq R$, where with $D_z^+ u^n$ we mean all the finite differences in (1.16).

***Remark* 1.3.** Notice that the *a priori* bound R on $|v_x|$ and $|v_y|$ is coherent with the fact that the solutions of (1.14) are generally Lipschitz continuous functions.

Now we are in the position to present the main result of this section, for which we give a sketch of the proof, referring to [CL84] for the complete version.

Theorem 1.17 (Crandall-Lions([CL84])). *Let $H : \mathbb{R}^2 \to \mathbb{R}$ be a continuous function and v_0 a bounded and Lipschitz continuous function on $\mathbb{R}^2$ with constant L. For $\lambda^x,\ \lambda^y > 0$ and fixed, let the scheme (1.15) be consistent with (1.14), have differenced form and be monotone on $[-(L+1),\ L+1]$. Assume, moreover, that the numerical hamiltonian h^M is locally Lipschitz continuous. Then, if v is the viscosity solution of 1.14), there exists a constant C_{CL} dependent only on $\sup|v_0|$, L, h^M and $N_T \Delta t$ such that*

$$|u_{i,j}^n - v(t_n, x_j, y_i)| \leq C_{CL}(\sqrt{\Delta t}), \quad (1.17)$$

for $0 \leq n \leq N_T$ and for all i, j, with $\lambda_x := \frac{\Delta t}{\Delta x}$ and $\lambda_y := \frac{\Delta t}{\Delta y}$.

Proof. (Sketch) The idea of the proof is strongly related to the technique used in the proof of Theorem 1.10.

First, in order to lighten the computations, let us assume the further conditions

$$\begin{cases} v(t,x,y) \to 0 & \text{for } |x|+|y| \to \infty \\ u_{i,j}^n \to 0 & \text{for } |i|+|j| \to \infty, \end{cases}$$

then, define the continuous domains $\mathcal{Q} = [0,\infty)\times\mathbb{R}^2$, $\mathcal{Q}_T = [0,T]\times\mathbb{R}^2$ and their discrete counterparts $\mathcal{Q}^d = \{(t_n, x_i, y_j) : i,j \in \mathbb{Z}, n \in \mathbb{N}\}$ and $\mathcal{Q}^d_{N_T} = \{(t_n,x_i,y_j) \in \mathcal{Q}^d : n \leq N_T\}$.
Fix $T > 0$ and $N_T \in \mathbb{N}$ such that $(N_T - 1)k < T \leq N_T \Delta t$, the aim is to get an estimate for $|v(t_n,x_i,y_j) - u_{i,j}^n|$. Let us assume that

$$\sup_{\substack{i,j\in\mathbb{Z} \\ 0\leq n\leq N_T}} \{v(t_n,x_i,y_j) - u_{i,j}^n\} = \sigma > 0,$$

then, the point is to find an upper bound for σ (the proof of the lower bound is analogous).

Define the function $\psi : \mathcal{Q}\times\mathcal{Q}^d \to \mathbb{R}$, according to the following

$$\psi(t,\xi,s,\eta) = v(t,\xi) - u_{i,j}^n - \frac{\sigma}{4T}(t+s) + \left(5M + \frac{\sigma}{2}\right)\beta\left(\frac{\xi-\eta}{\varepsilon}, \frac{t-s}{\varepsilon}\right),$$

where $M = \sup_{|\xi|\in\mathbb{R}^2}\{|v_0(\xi)|+T|H(0,0)|\}+1$, $\varepsilon > 0$ and β a smooth function on $\mathbb{R}^2\times\mathbb{R}$, such that $0 \leq \beta \leq 1$, $\beta(0,0) = 1$ and $\beta(\xi,t) = 0$ for $|\xi|^2+|t|^2 > 1$. At this point, the first step is to prove that there exists a maximum point $(t_0,\xi_0,s_0,\eta_0) \in \mathcal{Q}_T \times \mathcal{Q}^d_{N_T}$ for ψ, whence, considering separately all the possible cases for $t_0 \geq 0$ and $s_0 \geq 0$, it can be shown that an estimate of the form

$$\sigma \leq C(\Delta t + \Delta x + \Delta y)^{\frac{1}{2}},$$

for some constant $C > 0$, holds in any situation.

Finally, to recover the general case it is enough to exploit the finite speed of propagation of both the exact and the approximated solution. □

Next, in order to highlight the strict relation between monotone schemes in differenced form and the exact solution map $S(t)v_0$, we present also the properties of S^M, referring to [CL84] p.8, for the proof. Let us consider, only to simplify the notations, the one-dimensional case and define the set

$$\mathcal{C} = \{u \in l^\infty(\mathbb{Z}) : |\Delta_+ u_j| \leq R\Delta x, j \in \mathbb{Z}\}.$$

Then we can state the following result.

Proposition 1.18. *Let $R > 0$ and S^M a monotone scheme on $[-R,R]$, in differenced form with numerical hamiltonian h^M bounded on bounded sets, such that $S^M : \mathcal{C} \to l^\infty(\mathbb{Z})$. Then, if $u, v \in \mathcal{C}$ and $K \in \mathbb{R}$,*

i) $S^M(u+K) = S^M(u) + K$;

ii) $||S^M(u) - S^M(v)||_\infty \leq ||u-v||_\infty$;

iii) $||\Delta^+ S^M(u)||_\infty \leq ||\Delta^+ u||_\infty$;

iv) $S^M(\mathcal{C}) \subset \mathcal{C}$;

v) $||\left(S^M\right)^{n+j}(u) - \left(S^M\right)^n(u)||_\infty \leq j\Delta t \sup_{|\xi_j|\leq R}|G(\xi)|$;

vi) $||\left(S^M\right)^n(u)||_\infty \leq ||u||_\infty + n\Delta t|H(0)|$,

where we used the notation $\Delta^+ u_j = u_{j+1} - u_j$.

We would like to mention that, as a mere convergence results, Crandall-Lions theorem ([CL84]) was generalized by Barles and Souganidis in [BS91]. Their theorem gives a more abstract and general framework for convergence schemes, including the possibility of treating second-order equations, although no convergence estimate is obtained. Since it would involve various definitions for second-order equations, that go beyond the scope of this first chapter, we avoid to present the precise statement of the theorem. Roughly speaking, it states that *any monotone, stable and consistent scheme converges to the exact solution, provided there exists a comparison principle for the limiting equation.*
We also mention the version of the convergence theorem for first order equations due to Souganidis ([So85b]), in which an estimate of order $\sqrt{\Delta t}$ is obtained without requiring the differenced form of the scheme, although several other (reasonable) conditions must be met.

As we briefly recalled in this section, monotone schemes are indeed effective schemes for the approximation of the solution of (1.14), being stable and convergent. Unfortunately, by a simple Taylor expansion, it can be shown that monotone schemes are limited to *first order accuracy* even for regular solutions. Thus, if we need higher-order accuracy, we have to look for non-monotone schemes. This would be preferable, for example, for multidimensional problems, where the use of coarser grids is of primary importance.

In Chapter 3 we propose a procedure to couple the convergence property of the monotone schemes and the high-order consistency of some other (possibly unstable) scheme, in order to overcome this limitation and still obtain a convergent scheme. This is done basically relaxing the monotonicity property to a less restrictive ε-monotonocity, which will be defined later in this thesis.

1.3. Fronts evolution via the level set method

In this section we review the basic formulation of the Level Set Method, first introduced in [Se85], showing its reasoning and natural application to front propagation problems. For a complete presentation we refer the reader to the book [Se96], here we mainly follow the lines of [FF14].

The main advantage offered by this method, although it requires the embedding in a higher dimensional space, is the possibility to follow fronts with complex evolution and topology to high generality, thanks to the definition of viscosity solution of Hamilton-Jacobi equations. As we will soon see, the problem of following a front in two space dimensions evolving in its normal direction can be written in terms of a fixed level set of the viscosity solution of a first (or second) order Hamilton-Jacobi equation of the form

$$\begin{cases} v_t + H(x, y, \nabla v) = 0, & (t, x, y) \in (0, T) \times \mathbb{R}^2 \\ v(0, x, y) = v_0(x, y) & (x, y) \in \mathbb{R}^2, \end{cases} \tag{1.18}$$

where v_0 must be a *proper representation* of the front Γ_0. More precisely, if we identify the first configuration of the front as the 0-level set of the viscosity solution, then we require

Definition 1.19. Assume that Γ_0 is a closed curve in $\mathbb{R}^2$, and denote by Ω_0 the region enclosed by Γ_0. A continuous function $v_0 : \mathbb{R}^2 \to \mathbb{R}$ is *a proper representation* of Γ_0 if and only if it satisfies the following conditions:

$$\begin{cases} v_0(x, y) < 0, & (x, y) \in \Omega_0, \\ v_0(x, y) = 0, & (x, y) \in \Gamma_0, \\ v_0(x, y) > 0, & (x, y) \in \mathbb{R}^2 \setminus \overline{\Omega}_0. \end{cases} \tag{1.19}$$

Now, if we assume that the front Γ_0 evolves in the normal direction with velocity $c(x, y)$ and orientation dependent on the type of evolution (outward for an expansion and inward for a shrinking), we can easily derive the equation for the front Γ_t at a generic time $t > 0$. Let us denote a point $P \in \Gamma_t$ as $P = (x(t, s), y(t, s))$, where s is a parametric representation of the front, then, since at every point the velocity V has magnitude $c(P)$ and direction $\eta(P)$, in the case of an expansion we can write

$$V(P) = (\dot{x}(t, s), \dot{y}(t, s)) = c(P)\eta(P). \tag{1.20}$$

Now, if we want to identify the front Γ_t with the 0-level set of the solution v of a Cauchy problem with v_0 as initial condition, that is

$$\Gamma_t := \{(x, y) : v(t, x, y) = 0\},$$

we can derive the equation characterizing the evolution simply by computing the total time derivative of $v\,(t, x(t, s), y(t, s))$ and taking into account that this derivative vanishes on a level set of constant value. Whence we compute,

$$\begin{aligned} 0 &= \frac{d}{dt} v(t, x(t, s), y(t, s)) \\ &= v_t(t, x(s, t), y(s, t)) + \nabla v\,(t, x(t, s), y(t, s)) \cdot (\dot{x}(t, s), \dot{y}(t, s)) \\ &= v_t(t, P) + c(P)\nabla v(t, P) \cdot \eta(P) \\ &= v_t(t, P) + c(P)|\nabla v(t, P)|, \end{aligned} \tag{1.21}$$

where we have written the normal direction $\eta(P)$ to the front Γ_t in terms of the function v as $\eta(P) = \nabla v(t,P)/|\nabla v(t,P)|$. Since in this case the velocity c depends only on (x,y), we obtain a first order evolutive Hamilton-Jacobi equation of eikonal type

$$\begin{cases} v_t + c(x,y)|\nabla v| = 0, & (t,x,y) \in (0,T) \times \mathbb{R}^2 \\ v(0,x,y) = v_0(x,y) & (x,y) \in \mathbb{R}^2, \end{cases} \tag{1.22}$$

which solution remains a function at any time $t > 0$ as long as the velocity c is smooth enough. On the other hand, the front Γ_t can have complex shapes (due to breaking or merging) and also develop sharp corners, that is thanks to the definition of viscosity solution which allows the selection of the correct solution even when singularities appear.

In the previous derivation we assumed for simplicity that the velocity c depended only on the position, but the same approach can be applied to more general velocities, such as

- $c(t,x,y)$, isotropic growth with time varying velocity
- $c(t,x,y,\eta)$, anisotropic growth, dependent on normal direction
- $c(t,x,y,k)$, Mean Curvature Motion, with $k(t,x,y)$ mean curvature to the front at time t,
- $c(t,x,y,v)$, velocity dependent on the level set.

It is important to point out that, when using some of the previous velocities (e.g. the definition for the Mean Curvature Motion), the model problem becomes

$$\begin{cases} v_t + c(t,x,y,\nabla v, D^2 v)|\nabla v| = 0, & (t,x,y) \in (0,T) \times \mathbb{R}^2 \\ v(0,x,y) = v_0(x,y) & (x,y) \in \mathbb{R}^2, \end{cases} \tag{1.23}$$

where $c(t,x,y,\nabla v, D^2 v) = -div\left(\frac{\nabla v(t,x,y)}{|\nabla v(t,x,y)|}\right)$, which is a Hamilton-Jacobi equation of second-order.

It is noteworthy that the choice of the level set is arbitrary, and in fact all the level sets of v are moving according to the same law. This fact will represent a major problem in the application to shape recovery due to the discontinuity of the velocity field c, whence a modification to the classical model will be needed (see Chapter 4).

1.4. Dimensional splitting methods

We close this first preliminary chapter by describing a procedure useful to define numerical schemes for multidimensional problems using simple one-dimensional schemes. Unfortunately, such techniques can not be used in all generality, in fact a strong assumption on the nature of the evolution must be required, that is the separability of the hamiltonian H. Nevertheless, it is worth pointing out that, by the following procedures, the number of points necessary to compute the schemes (termed as the *stencil* of the

scheme) increases linearly w.r.t. the number of spatial dimensions, whereas the cardinality of the stencil of full multidimensional schemes usually increases as the power of N, where N is the number of variables.

As before, we focus on the case $N = 2$ and consider problems of the type

$$\begin{cases} v_t + H(v_x, v_y) = 0 \\ v(0, x, y) = v_0(x, y), \end{cases} \tag{1.24}$$

where the hamiltonian is such that $H(v_x, v_y) = H_1(v_x) + H_2(v_y)$, for a couple of one-dimensional hamiltonians H_1, H_2. The first step is to consider the $1D$-problems associated with the *dimensional splitting*, that are

$$v_t + H_1(v_x) = 0 \qquad \text{and} \qquad v_t + H_2(v_y) = 0, \tag{1.25}$$

and then to denote with S_t, S_t^x and S_t^y, the *solution maps* of, respectively, the $2D$ and $1D$ problems at time t.

The issue of the convergence of dimensional splitting methods has been approached in the contest of *Lie-Trotter* products by Souganidis in [So85b], for general equations, with H dependent on (t, x, u, Du). The results of the paper apply directly to equations of the form (1.24) with H, H_1 and H_2 continuous functions and, in particular, prove the convergence of the *Lie-Trotter splitting*, that is

$$S_{\Delta t}^{LT} v^n = S_{\Delta t}^y S_{\Delta t}^x v^n, \tag{1.26}$$

at each time t_n. More precisely, we have the following

Theorem 1.20. *Let H, H_1 and H_2 continuous functions, with $H = H_1 + H_2$ and $v_0 \in BUC(\mathbb{R}^2)$. Then,*

$$\left|\left| S_T v_0 - [S_{\Delta t}^y S_{\Delta t}^x]^n v_0 \right|\right|_\infty \to 0$$

for $\Delta t \to 0$ and $n \to \infty$, with $n\Delta t = T$.

Proof. For the proof we refer to the formulation of the theorem as in [So85b], p. 218. Conditions (F1)-(F6) follow directly from the properties of the viscosity solution map (Proposition 1.13), applied successively first on $S_{\Delta t}^y$ and then on $S_{\Delta t}^x$, and the composition of continuous functions. As an example, for (F3) we have that, for $v \in W^{1,\infty}$,

$$\left|\left| S_{\Delta t}^y S_{\Delta t}^x (v) \right|\right| = \left|\left| S_{\Delta t}^x v - H_2\left((S_{\Delta t}^x (v))_y \right) + \mathcal{O}(\Delta t^2) \right|\right| \leq ||S_{\Delta t}^x v|| + C_1 \Delta t \leq ||v|| + C\Delta t.$$

The same computations also prove condition (F7), in fact, let $\phi \in \mathcal{C}^2(\mathbb{R}^2)$ be bounded and with bounded derivatives, then

$$S_{\Delta t}^y S_{\Delta t}^x \phi = \phi - \Delta t H_1(\phi_x) - \Delta t H_2(\phi_y) + \mathcal{O}(\Delta t^2),$$

whence

$$\left|\left| \frac{S_{\Delta t}^y S_{\Delta t}^x \phi - \phi}{\Delta t} + H(\phi_x, \phi_y) \right|\right| = \mathcal{O}(\Delta t) \to 0, \tag{1.27}$$

for $\Delta t \to 0$. Finally, the statement of the theorem holds true applying the cited theorem in [So85b]. □

In order to prove the applicability of other splitting methods, e.g. the well-known *Strang-Marchuk splitting*, which at each iteration reads

$$S_{\Delta t}^{SM} v^n = S_{\frac{\Delta t}{2}}^{y} S_{\Delta t}^{x} S_{\frac{\Delta t}{2}}^{y} v^n, \tag{1.28}$$

it is necessary to require the convexity of H_1 and H_2 and then apply the following result, proved in [LR86], which holds for generic splittings methods of the form $H(x) = H_1(x) + H_2(x)$, with $x \in \mathbb{R}^N$. We state the result using the notations in [LR86], avoiding to specify the little differences implied, since they should be easily understood from the contest.

Proposition 1.21 (Lions-Rochet([LR86])). *Let $v_0 \in BUC(R^N)$, H_1, H_2 convex functions, then, for all $t, s > 0$, it holds*

$$S_{H_1}(t) S_{H_2}(s) v_0 = S_{H_2}(s) S_{H_1}(t) v_0 = S_{tH_1 + sH_2}(1) v_0. \tag{1.29}$$

Finally, the convergence of the splitting methods can be obtained applying successively the previous formula. Notice that Proposition 1.21 implies a rather interesting consequence. In fact, under such hypothesis, the splitting method is exact, in the sense that it does not produce error in the time evolution. Consequently, we could take as time step Δt the full length of the interval $[0, T]$ and whence obtain a *global splitting*. For example, using Lie-Trotter splitting, first we should solve

$$\begin{cases} u_t + H_1(u_x) & \text{for } (t, x) \in (0, T) \times \mathbb{R} \\ u(0, x, y) = v_0(x, y) & \text{for } x \in \mathbb{R}, \end{cases}$$

and, as a second step, proceed with the solution of

$$\begin{cases} v_t + H_2(v_y) & \text{for } (t, y) \in (0, T) \times \mathbb{R} \\ v(0, x, y) = u(T, x, y) & \text{for } y \in \mathbb{R}, \end{cases}$$

thus obtaining $v(T, x, y) = S_T v_0$.

***Remark* 1.4.** The same results have been proven by Barles and Tourin in [BT01] for hamiltonians $H(x, p) = H_1(x, p) + H_2(x, p)$, with H_1 and H_2 convex functions in $p \in \mathbb{R}^N$, for all $x \in \mathbb{R}^N$, and such that they satisfy the *Lie bracket condition*

$$[H_1(x, p), H_2(x, p)] = \nabla_x H_1(x, p) \nabla_p H_2(x, p) - \nabla_x H_2(x, p) \nabla_p H_1(x, p). \tag{1.30}$$

***Remark* 1.5.** It is worth to notice that if the hamiltonian H is separable simply as $H(x, y, u_x, u_y) = H_1(x, u_x) + H_2(y, u_y)$, then the Lie bracket condition is trivially satisfied. Consequently, using Taylor expansions it can be shown that, for these kind of equations, the Lie-Trotter splitting is in fact a second order method (in time), in the sense of the consistency error. Moreover, if H_1 and H_2 are convex functions in the second argument, then by Proposition 1.21 the dimensional splitting is exact and does not introduce errors in the time evolution.

Now that we have ensured the reliability of the splitting methods in the contest of Hamilton-Jacobi equations, we can pass to the problem of the numerical application, basically substituting the exact solution operators with numerical schemes for the approximation of the one-dimensional subproblems. In the following we make a little abuse of notation using $S_{\Delta t}$ to denote (also) the approximate solution operators, in order to simplify the presentation, since it should not cause any confusion. As before, we choose the Lie-Trotter splitting as leading example. We have

$$u^{n+1} = S^y_{\Delta t} S^x_{\Delta t} u^n, \qquad \text{for } n = 0, \dots, N_T = T/\Delta t, \tag{1.31}$$

for the local splitting, which is usually used in numerical implementations, or

$$u(T, x, y) = S^y_T S^x_T u_0, \tag{1.32}$$

for the global splitting, where with S_T we mean a large-time step numerical method, such as for example, semi-lagrangian schemes (see [FF14] for a complete presentation of this class of schemes), or a "full approximation", in the sense of the time evolution, of the relative subproblem on the interval $(0, T)$.

It is worth noticing that using the presented methods, we can construct monotone (convergent) schemes for (1.24), using simple one-dimensional monotone schemes. Let us begin by expliciting the scheme (1.31),

$$\begin{cases} u_{i,j}^{n+\frac{1}{2}} &= S^x_{\Delta t}\left(u^n\right)_j = S^x_{\Delta t}\left(u^n_{i,j-r}, \dots, u^n_{i,j+r}\right) \\ u_{i,j}^{n+1} &= S^y_{\Delta t}\left(u^{n+\frac{1}{2}}\right)_i = S^y_{\Delta t}\left(u^{n+\frac{1}{2}}_{i-l,j}, \dots, u^{n+\frac{1}{2}}_{i+l,j}\right), \end{cases} \tag{1.33}$$

then we can prove the following lemma.

Lemma 1.22. *Let the one-dimensional schemes $S^x_{\Delta t}$, $S^y_{\Delta t}$ have differenced form with numerical hamiltonians h^M_x and h^M_y, and be consistent with H_1 and H_2, respectively. Then, the resulting scheme (1.33) can be written in differenced form and it is consistent with H. Moreover, if the one-dimensional schemes are monotone on $[-R, R]$, then the full scheme (1.33) is also monotone on $[-R, R]$.*

Proof. By hypothesis, we know that

$$u_{i,j}^{n+\frac{1}{2}} = u^n_{i,j} - \Delta t h^M_x\left(D^+_x u^n_{i,j-r}, \dots, D^+_x u^n_{i,j+r-1}\right)$$

and analogously,

$$u_{i,j}^{n+1} = u^n_{i,j} - \Delta t h^M_y\left(D^+_y u^{n+\frac{1}{2}}_{i-l,j}, \dots, D^+_y u^{n+\frac{1}{2}}_{i+l-1,j}\right).$$

Whence, composing the previous relations we have

$$\begin{aligned} u_{i,j}^{n+1} = u^n_{i,j} &- \Delta t h^M_x\left(D^+_x u^n_{i,j-r}, \dots, D^+_x u^n_{i,j+r-1}\right) - \\ &- \Delta t \widetilde{h}^M_y\left(D^+_x u^n_{i-l,j-r}, \dots, D^+_x u^n_{i+l-1,j+r}; D^+_y u^n_{i-l,j-r}, \dots, D^+_y u^n_{i+l,j+r-1}\right), \end{aligned} \tag{1.34}$$

with $\widetilde{h}_y^M$ such that $\widetilde{h}_y^M(u^n)_{i,j} = h_y^M(u^{n+\frac{1}{2}})_i$. To show that such $\widetilde{h}_y^M$ can be easily written, we explicit one of the arguments of h_y^M depending on u^n,

$$\begin{aligned} D_y^+ u_{i,j}^{n+\frac{1}{2}} &= \frac{u_{i+1,j}^{n+\frac{1}{2}} - u_{i,j}^{n+\frac{1}{2}}}{\Delta y} = \\ &= D_y^+ u_{i,j}^n - \lambda_y \left[h_x^M \left(D_x^+ u_{i+1,j-r}^n, \ldots, D_x^+ u_{i+1,j+r-1}^n \right) - \right. \\ &\quad \left. - h_x^M \left(D_x^+ u_{i,j-r}^n, \ldots, D_x^+ u_{i,j+r-1}^n \right) \right]. \end{aligned}$$

This shows that the scheme (1.33) has differenced form with numerical hamiltonian $h^M = h_x^M + \widetilde{h}_y^M$.

For the consistency, assume that $D_x^+ u_{i,j}^n = a$ and $D_y^+ u_{s,m}^n = b$ for $i - l < s < i + l$, $j - r < m < j + r$, then, exploiting the consistency of h_x^M with H_1, we have from (1.34)

$$D_y^+ u_{i-l,j-r}^{n+\frac{1}{2}} = \cdots = D_y^+ u_{i+l,j+r-1}^{n+\frac{1}{2}} = b.$$

Successively, using also the consistency of h_y^M con H_2, we get

$$h^M(a, \ldots, a; b, \ldots, b) = h_x^M(a, \ldots, a) + h_y^M(b, \ldots, b) = H_1(a) + H_2(b) = H(a, b),$$

that is h^M is consistent with H.

The monotonicity of (1.33) follows from the fact that the composition of nondecreasing functions is nondecreasing, and using, as proved in [CL84], $||\Delta_x^+ S_{\Delta t}^x(u^n)||_\infty \leq ||\Delta_x^+ u^n||_\infty$. □

The convergence of this class of schemes directly follows applying Theorem 1.17.

2. Smoothness indicators analysis

In this chapter we present a rather deep analysis of some well-known regularity indicators, based on the divided differences of a function f. In Section 2.1, we begin by recalling the theory of the regularity indicators defined in [JP00] for the construction of the Weighted Essentially Non Oscillatory (WENO) schemes. Briefly speaking, these indicators are defined as summations of weighted L^2 norms of appropriate approximations of the derivatives of the function. We will prove their properties depending on the scaling factor Δx^α, $\alpha \in \mathbb{Z}$, which depends on the order of the integrated derivative and gives the weight to the corresponding L^2 norm. Through this study we also wish to give general indications on how to correctly choose the regularity indicator, dependently on the order of the discontinuous derivative. This is mainly because, when looking for references for the present chapter, we have noticed that in various articles regarding *Hamilton-Jacobi equations*, which focus on the regularity of the first derivative of the solution, the scaling factor in the definition of the smoothness indicators is mistaken. More precisely, since in [JP00], where the indicators for HJ are first used (β^{HJ} in the sequel), a precise formula of the definition is missing, most authors usually use the definition in [JS96] for conservation laws (β^{CL} in the sequel). The idea of the proof is inspired by the proof of Corollary 2 in [ABM10] and it is quite general. In fact, it could be easily modified in order to get similar results on regularity indicators of higher degree derivatives. Then, in Sec. 2.1.1, we will show how to use those indicators as building blocks to define our function ϕ_j, $j \in \mathbb{Z}$, which measures the regularity of the solution in $I_j = (x_{j-1}, x_{j+1})$, needed for the construction of the Adaptive Filtered Scheme in Chapter 3. Finally, in Section 2.2 we will see how to generalize these constructions to the multidimensional case, defining some new smoothness indicators which are able to detect the discontinuities in the gradient of a function of two variables. At the best of our knowledge, this is a new contribution.

2.1. One-dimensional case. The regularity indicators of [JP00] and [JS96]

Briefly speaking, the idea behind the definitions below is to use the magnitude of the divided differences of a function f as indicator of regularity, similarly to the preceding Essentially Non Oscillatory (ENO) procedure. The main novelty, w.r.t. the ENO idea, is the scaling of the divided differences, which introduces a more direct dependence on the discretization parameter Δx.

More in detail, let us start the analysis by directly giving the definition of the smooth-

ness indicators for Hamilton-Jacobi equations (HJ) in the sense of [JP00] as

$$\beta_k^{HJ}[f;x_j] = \sum_{l=2}^{r} \int_{x_{j-1}}^{x_j} \Delta x^{2l-3} \left(P_k^{(l)}(x)\right)^2 dx, \tag{2.1}$$

and let us consider also the version for *conservation laws* (CL) of [JS96]

$$\beta_k^{CL}[f;x_j] = \sum_{l=1}^{r} \int_{x_{j-1}}^{x_j} \Delta x^{2l-1} \left(P_k^{(l)}(x)\right)^2 dx, \tag{2.2}$$

for $k = 0, \dots, r-1$, where $P_k(x)$ is the Lagrange polynomial of degree r interpolating a function f on the stencil $\mathcal{S}_{j+k} = \{x_{j+k-r}, \dots, x_{j+k}\}$. Notice that w.r.t. the definition in [JS96] we have shifted the domain of integration of $\Delta x/2$ in order to treat both definitions (2.1) - (2.2) together. In the sequel we will drop the dependence on f and on the point x_j for more clarity, since it should not cause any confusion. Moreover, we will often use the convention $h := \Delta x$.

Then as a first big step we prove our main result of this section.

Proposition 2.1. *Assume $f \in C^{r+1}(\Omega \setminus \{x_s\})$, with Ω a neighborhood of x_s, and $f'(x_s^-) \neq f'(x_s^+)$. Moreover, (just for simplicity) let $f^{(l)}(x) \neq 0$, $\forall x \in (\Omega \setminus \{x_s\})$ for $l = 1, 2$. Then, for $k = 0, \dots, r-1$ and $j \in \mathbb{Z}$, the following are true:*

(i) If $x_s \in \Omega \setminus \mathring{\mathcal{S}}_{j+k}$ $\Rightarrow \beta_k^{HJ} = O(\Delta x^2)$ and $\beta_k^{CL} = O(\Delta x^2)$;

(ii) If $x_s \in \mathring{\mathcal{S}}_{j+k}$ $\Rightarrow \beta_k^{CL} = O(\Delta x^2)$, while $\beta_k^{HJ} = O(1)$,

where $\mathcal{S}_{j+k} = \{x_{j-r+k}, \dots, x_{j+k}\}$ and $\mathring{\mathcal{S}}_{j+k} = (x_{j-r+k}, x_{j+k})$.

Proof. Let us take $r > 1$ and without loss of generality, let $x_s = 0$ (to simplify the notation).

Let us start by recalling that, using the Newton form of the interpolating polynomial, for $k = 0, \dots, r-1$ and $j \in \mathbb{Z}$, we get

$$P_k(x) = f(x_{j-r+k}) + \sum_{i=1}^{r} f[x_{j-r+k}, \dots, x_{j-r+k+i}]\omega_{i-1}(x), \tag{2.3}$$

where $\omega_i(x) = (x - x_{j-r+k}) \cdots (x - x_{j-r+k+i})$ and $f[\cdot]$ denotes the divided difference of f, which are defined recursively as

$$\begin{array}{ll} f[x_\nu] := f(x_\nu), & \nu = 0, \dots, r \\ f[x_\nu, \dots, x_{\nu+i}] := \frac{f[x_{\nu+1}, \dots, x_{\nu+i}] - f[x_\nu, \dots, x_{\nu+i-1}]}{x_{\nu+i} - x_\nu}, & \nu = 0, \dots, r-i,\ i = 1, \dots, r. \end{array} \tag{2.4}$$

Let us proceed with the proof of ***i)***. If we define the function $f_h(y) := f(x_j + hy)$, $y \in \mathbb{Z}$, we can write

$$f[x_{j-r+k}, \dots, x_{j-r+k+i}] = f[x_j + (k-r)h, \dots, x_j + (k-r+i)h] = \frac{f_h[k-r, \dots, k-r+i]}{i!\, h^i},$$

where $f_h[\cdot]$ denotes the undivided difference of f_h. At this point it is useful to notice that, for $l = 0, \dots, i$ (see Lemma A.1 in Appendix A for the proof),

$$f_h[k-r, \dots, k-r+i] = \sum_{n=0}^{i-l} \binom{i-l}{n} (-1)^{i-l-n} f_h[(k-r+n), \dots, (k-r+n+l)]. \quad (2.5)$$

Moreover, if $0 \notin \mathring{\mathcal{S}}_{j+k}$, expanding with Taylor, it can be shown that for $n = 0, \dots, i-l$ and $h \to 0$, $f_h[(k-r+n), \dots, (k-r+n+l)] \approx h^l f^{(l)}(x_j)$ (see Corollary A.3 in Appendix A). Then, we can infer

$$f_h[k-r, \dots, k-r+i] \to \begin{cases} o\left(h^l\right) & \text{for } l < i \leq r, \\ h^l f^{(l)}(x_j) & \text{for } i = l, \end{cases} \quad (2.6)$$

having exploited the relation $\sum_{j=0}^{i} \binom{i}{j} (-1)^{i-j} = 0$, for $i \geq 1$ (see Lemma A.1 in Appendix A).

Now, if we define the polynomial

$$Q_k(y) := P_k(x_j + hy) = f_h(k-r) + \sum_{i=1}^{r} f_h[(k-r), \dots, (k-r+i)] \frac{q_{i-1}(y)}{i!}, \quad (2.7)$$

where $q_i(y) = (y - (k-r)) \cdots (y - (k-r-i))$, $f_h(y) = f(x_j + hy)$, then we can rewrite

$$P_k^{(l)}(x) = \frac{d^l}{dx^l} \left(Q_k \left(\frac{x}{h} \right) \right) = \frac{1}{h^l} Q_k^{(l)}(y), \qquad l = 1, \dots, r. \quad (2.8)$$

Then, applying the change of variable $y = (x - x_j)/h$ in the integral in (2.1) or (2.2), we have

$$h^\alpha \int_{x_{j-1}}^{x_j} \left(P_k^{(l)}(x) \right)^2 dx = h^{\alpha - 2l + 1} \int_{-1}^{0} \left(Q_k^{(l)}(y) \right)^2 dy, \quad (2.9)$$

where $\alpha = 2l - 3$ or $2l - 1$, respectively, and

$$Q_k^{(l)}(y) = \sum_{i=l}^{r} f_h[(k-r), \dots, (k-r+i)] \frac{q_{i-1}^{(l)}(y)}{i!}. \quad (2.10)$$

Furthermore, keeping in mind (2.6), we can write

$$\begin{aligned} Q_k^{(l)}(y) &= f_h[(k-r), \dots, (k-r+l)] \frac{q_{l-1}^{(l)}(y)}{l!} + \sum_{i=l+1}^{r} f_h[(k-r), \dots, (k-r+i)] \frac{q_{i-1}^{(l)}(y)}{i!} \\ &= h^l f^{(l)}(x_j) + o(h^l), \end{aligned}$$

having noticed that, since $q_{l-1}(y)$ is a monic polynomial of degree l, we have $q_{l-1}^{(l)}(y) = l!$. From we have just seen, it follows

$$\begin{aligned} h^\alpha \int_{x_{j-1}}^{x_j} \left(P_k^{(l)}(x) \right)^2 dx &= h^{\alpha - 2l + 1} \int_{-1}^{0} \left(Q_k^{(l)}(y) \right)^2 dy \\ &= h^{\alpha+1} \left(f^{(l)}(x_j) \right)^2 + o(h^{\alpha+1}), \end{aligned}$$

as we wanted. In fact, having in mind that by hypothesis $f', f'' \neq 0$, if we consider, for example,

$$\begin{aligned}\beta_k^{HJ} &= \sum_{l=2}^{r} \int_{x_{j-1}}^{x_j} h^{2l-3} (P_k^{(l)})^2 dx \\ &= \sum_{l=2}^{r} h^{2l-2} \left(f^{(l)}(x_j) \right)^2 + o(h^{2l-2}) = O(h^2),\end{aligned}$$

and in the same way for the other indicator (notice that for β_k^{CL} the summation starts from $l = 1$).

***Remark* 2.1.** To prove this point, it could be sufficient to observe that using the regularity of f in $\Omega \setminus \{x_s\}$ and the properties of the interpolating polynomial (which we have just proved) we get

$$P_k^{(l)}(x) = f^{(l)}(x) + O(h^{r+1-l}), \qquad \text{for } x_{j-1} \leq x \leq x_j,\ k = 0, \ldots, r-1.$$

Moreover, expanding with Taylor, it holds

$$f^{(l)}(x) = f^{(m)}(x_j) O(h)^{m-l} + o(h^{m-l}), \tag{2.11}$$

where $m = \max\{s+1, l\}$ and $s = \max\{k : f^{(i)}(x_j) = 0,\ \forall i \leq k\}$ $(s \leq r)$. Then, integrating (remembering that by hypothesis $s = 0 \Rightarrow m = l$),

$$h^\alpha \int_{x_{j-1}}^{x_j} \left(P_k^{(l)}(x) \right)^2 dx = h^{\alpha+1} \left(f^{(l)}(x_j) \right)^2 + o(h^{\alpha+1}),$$

as before.

Let us continue with the proof of ***ii)***. In this case the proof is a little more complicated and it is better to treat separately the following two cases:

(a) 0 is a point of the grid $\{x_i\}$, $i \in \mathbb{Z}$;

(b) $0 \in I_i = (x_{i-1}, x_i)$ for some $i \in \mathbb{Z}$.

<u>*Case a.*</u> By assumption $0 \in \mathcal{S}_{j+k}$ for at least one $k = 0, \ldots, r-1$, then, for each fixed k, there exists an integer $j_s \in \{k-r+1, \ldots, k-1\}$ such that $x_j = -j_s h$ (for $j_s = k-r$ and $j_s = k$ we fall in the case treated previously). Substituting in (2.3),

$$P_k(x) = f((-j_s + k - r)h) + \sum_{i=1}^{r} f[(-j_s + k - r)h, \ldots, (-j_s + k - r + i)h] \omega_{i-1}(x),$$

with $\omega_i(x) = (x + (j_s - k + r)h) \cdots (x + (j_s - k + r - i)h)$. As we have done in the proof of ***i)***, if we define the function $f_h(y) := f(x_j + hy) = f(h(y - j_s))$, we can obtain the relations (2.8)-(2.9) with $Q_k^{(l)}(y)$ defined as in (2.10).

At this point, it is useful to notice that (2.5) for $l = 1$ reads, for $i = 1, \dots, r$,

$$f_h[k-r, \dots, (k-r+i)] = \sum_{j=0}^{i-1} \binom{i-1}{j} (-1)^{i-j-1} f_h[(k-r+j), (k-r+j+1)].$$

In order to simplify the notation let us define $i_s := j_s - k + r$, that is to say the index $i_s \in \{1, \dots, r-1\}$ such that $x_j + (k-r+i_s)h = 0$. Then, by assumption, we can write for all $i > t := \max\{i_s, l-1\}$,

$$\begin{aligned} f_h[k-r, \dots, (k-r+i)] = &\sum_{j=0}^{i_s-1} \binom{i-1}{j} (-1)^{i-j-1} f_h[(k-r+j), (k-r+j+1)] \\ &+ \sum_{j=i_s}^{i-1} \binom{i-1}{j} (-1)^{i-j-1} f_h[(k-r+j), (k-r+j+1)]. \end{aligned}$$

Since for h tending to 0 we have

$$f_h[z+j_s, z+j_s+1] = h\frac{f((z+1)h) - f(zh)}{h} \to \begin{cases} hf'(0^+) & \text{if } z \geq 0 \\ hf'(0^-) & \text{otherwise} \end{cases}$$

we can conclude that

$$\begin{aligned} f_h[k-r, \dots, (k-r+i)] &\approx h \left[\sum_{j=0}^{i_s-1} \binom{i-1}{j} (-1)^{i-j-1} f'(0^-) + \right. \\ &\qquad \left. + \sum_{j=i_s}^{i-1} \binom{i-1}{j} (-1)^{i-j-1} f'(0^+) \right] \\ &= h\left[f'(0^+) - f'(0^-)\right] \sum_{j=i_s}^{i-1} \binom{i-1}{j} (-1)^{i-j-1} \\ &= h\left[f'(0^+) - f'(0^-)\right] \binom{i-2}{i_s-1} (-1)^{i-i_s+1} \neq 0, \end{aligned}$$

where we have exploited the relations $\sum_{j=0}^{i} \binom{i}{j}(-1)^{i-j} = 0$ and $\sum_{j=0}^{n} \binom{i}{j}(-1)^{i-j} = \binom{i-1}{n}(-1)^{i-n}$, for $0 \leq n < i$ by Lemma A.1 (Appendix A).

Furthermore for $l \leq i \leq i_s$, as we have seen in the first point ***i)*** of the proof,

$$f_h[k-r, \dots, (k-r+i)] \approx h \sum_{j=0}^{i-1} \binom{i-1}{j} (-1)^{i-j-1} f'(0^-) = o(h),$$

which also follows directly from the first relation we have just reminded.

From what we have done so far we can deduce, recalling that $t := \max\{i_s, l-1\}$,

$$\begin{aligned} h^\alpha \int_{x_{j-1}}^{x_j} \left(P_k^{(l)}(x)\right)^2 dx &= h^{\alpha-2l+1} \int_{-1}^{0} \left(\sum_{i=t+1}^{r} f_h[k-r, \dots, (k-r+i)] \frac{q_{i-1}^{(l)}}{i!} \right)^2 \\ &\approx C_{rk} h^{\alpha-2l+3} \left[f'(0^+) - f'(0^-)\right]^2, \end{aligned}$$

where $C_{rk} = \int_{-1}^{0} \left(\sum_{i=t+1}^{r} \binom{i-2}{i_s-1} (-1)^{i-i_s+1} \frac{q_{i-1}^{(l)}}{i!} \right)^2$. Then, passing to the limit, we finally obtain

$$\lim_{h\to 0} h^{\alpha} \int_{x_{j-1}}^{x_j} \left(P_k^{(l)}(x) \right)^2 dx = \begin{cases} O(h^2) & \text{if } \alpha = 2l-1 \\ O(1) & \text{if } \alpha = 2l-3 \end{cases}$$

which is the thesis for case *a.*

Case b. By assumption there exists an integer $j_s \in \{k-r+1, \dots, k\}$ and a number $0 < a_s < 1$ such that $x_j = (-j_s + a_s)h$. It is clear now that we can repeat the same constructions of the previous case defining the function $f_{a_s,h}(y) := f(h(y - j_s + a_s))$ and using it in place of f_h. To obtain (2.9) it will suffice to apply the change of variables $y = \frac{x}{h} + j_s - a_s$. Then, defining $i_s = j_s - k + r$, for $i \geq t := \max\{i_s, l\}$, we get

$$\begin{aligned} f_{a_s,h}[k-r, \dots, (k-r+i)] = &\sum_{j=0}^{i_s-2} \binom{i-1}{j} (-1)^{i-j-1} f_{a_s,h}[(k-r+j), (k-r+j+1)] \\ &+ \binom{i-1}{i_s-1} (-1)^{i-i_s} f_{a_s,h}[j_s-1, j_s] \\ &+ \sum_{j=i_s}^{i-1} \binom{i-1}{j} (-1)^{i-j-1} f_{a_s,h}[(k-r+j), (k-r+j+1)]. \end{aligned} \tag{2.12}$$

Now, noticing that

$$\begin{aligned} f_{a_s,h}[j_s-1, j_s] &= f(a_s h) - f((a_s-1)h) \\ &= a_s h \left(\frac{f(a_s h) - f(0)}{a_s h} \right) + (1-a_s) h \left(\frac{f(0) - f((a_s-1)h)}{(1-a_s)h} \right) \\ &\approx a_s h f'(0^+) + (1-a_s) h f'(0^-) \\ &= a_s h \left[f'(0^+) - f'(0^-) \right] + h f'(0^-), \end{aligned}$$

and that

$$f_{a_s,h}[z + j_s - 1, z + j_s] \to \begin{cases} h f'(0^+) & \text{if } z \geq 1 \\ h f'(0^-) & \text{if } z \leq -1, \end{cases}$$

we can infer that if $i = i_s$ (in this case in (2.12) on the right side of the equation we have only the second term), then $f_{a_s,h}[k-r, \dots, (k-r+i)] \approx a_s h \left[f'(0^+) - f'(0^-) \right] \neq 0$, while if $i > i_s$,

$$\begin{aligned} f_{a_s,h}[k-r, \dots, (k-r+i)] \approx h \left[f'(0^+) - f'(0^-) \right] \Bigg[&\binom{i-2}{i_s-1} (-1)^{i-i_s+1} + \\ &+ a_s \binom{i-1}{i_s-1} (-1)^{i-i_s} \Bigg]. \end{aligned}$$

The last quantity, as it can be easily shown, it is null if and only if $a_s = \frac{i-i_s}{i-1}$; more precisely, for k fixed there exists an integer $i \geq t$ such that $f_{a_s,h}[k-r, \dots, (k-r+i)] \approx C h \left[f'(0^+) - f'(0^-) \right]$ with $C \neq 0$, whence the thesis even in the last case. □

***Remark* 2.2.** Notice that we could avoid the restrictions on f in the points of regularity by adding a small quantity $\sigma_h = \sigma h^2$, for some constant $\sigma > 0$, to the indicators β_k and consider instead

$$\widetilde{\beta}_k := \beta_k + \sigma_h, \tag{2.13}$$

as it has been done in [ABM10]. We will use this assumption in the sequel, choosing for the moment $\sigma = 1$.

2.1.1. Regularity indicator functions

Now we can proceed with the construction of our function ϕ, which will be used in the next chapter to measure the regularity of the approximated solution. In order to get a good smoothness indicator but trying to keep the stencil of the reconstructions as small as possible, we decided to stick with the simple case of $r = 2$. As we will show soon this choice it is not restrictive and will give us the results we are looking for. Clearly some adjustments will be needed and in doing so we will make use of the *mappings* introduced in [HAP05] (see (2.23)).

Our aim is to identify the points (or the intervals) in which there is a singularity in the first derivative of a function f. Let us focus the attention on a point x_j of the grid, the idea is to study the regularity of the function in the interval (x_{j-1}, x_{j+1}) using the smoothness indicators β_k^{HJ} with $r = 2$. In what follows we will drop the superscript *HJ* as it is now clear, by Proposition 2.1, which indicators we should use.

Let us consider separately the intervals $[x_{j-1}, x_j]$ and $[x_j, x_{j+1}]$ defining

$$\beta_k^- = \Delta x \int_{x_{j-1}}^{x_j} (P_k''(x))^2 dx, \tag{2.14}$$

for $k = 0, 1$, where P_0, P_1 are the polynomials interpolating the functions, respectively, on the stencils $\{x_{j-2}, x_{j-1}, x_j\}$ and $\{x_{j-1}, x_j, x_{j+1}\}$; and symmetrically

$$\beta_k^+ = \Delta x \int_{x_j}^{x_{j+1}} (P_k''(x))^2 dx, \tag{2.15}$$

for $k = 0, 1$, where now P_0, P_1 are the interpolating polynomials on $\{x_{j-1}, x_j, x_{j+1}\}$ and $\{x_j, x_{j+1}, x_{j+2}\}$. From the definition it is clear that $(\beta^+)_j = (\beta^-)_{j+1}$ so we have to compute the quantities just once. To be as clear as possible let us write the explicit form of the β_k and their Taylor expansion around the point x_j in the case of a uniform grid,

- $\beta_0^- = \left(\frac{f_j - 2f_{j-1} + f_{j-2}}{h}\right)^2 = h^2 (f_j'')^2 - 2h^3 f_j'' f_j''' + h^4 \left[(f_j''')^2 + \frac{7}{6} f_j'' f_j^{(4)}\right] - \frac{7}{6} h^5 f_j''' f_j^{(4)} + o(h^5)$,
- $\beta_1^- = \beta_0^+ = \left(\frac{f_{j+1} - 2f_j + f_{j-1}}{h}\right)^2 = h^2 (f_j'')^2 + \frac{h^4}{6} f_j'' f_j^{(4)} + o(h^5)$;
- $\beta_1^+ = \left(\frac{f_j - 2f_{j+1} + f_{j+2}}{h}\right)^2 = h^2 (f_j'')^2 + 2h^3 f_j'' f_j''' + h^4 \left[(f_j''')^2 + \frac{7}{6} f_j'' f_j^{(4)}\right] + \frac{7}{6} h^5 f_j''' f_j^{(4)} + o(h^5)$

Then, we define

$$\alpha_k^{\pm} := \frac{1}{(\beta_k^{\pm} + \sigma_h)^2}, \tag{2.16}$$

with $\sigma_h = \sigma h^2$ is the parameter we introduced in Remark 2.2. We focus on the informations given by the interpolating polynomial on $\{x_{j-1}, x_j, x_{j+1}\}$ defining

$$\omega_+ := \frac{\alpha_0^+}{\alpha_0^+ + \alpha_1^+},$$

to inspect the regularity on $[x_j, x_{j+1}]$ and in the same way for $[x_{j-1}, x_j]$,

$$\omega_- := \frac{\alpha_1^-}{\alpha_0^- + \alpha_1^-}.$$

By Proposition 2.1 and Remark 2.2 we know that $\widetilde{\beta}_k = O(h^2)$ if there is no singularity in the stencil, and $\widetilde{\beta}_k = O(1)$ otherwise, so in presence of a singularity we can only fall in one of the following cases:

- If $x_{j-2} < x_s \leq x_{j-1}$, then $\widetilde{\beta}_0^- = O(1)$, $\widetilde{\beta}_1^- = \widetilde{\beta}_0^+ = O(h^2)$, $\widetilde{\beta}_1^+ = O(h^2)$;
- If $x_{j-1} < x_s < x_j$, then $\widetilde{\beta}_0^- = O(1)$, $\widetilde{\beta}_1^- = \widetilde{\beta}_0^+ = O(1)$, $\widetilde{\beta}_1^+ = O(h^2)$;
- If $x_s = x_j$, then $\widetilde{\beta}_0^- = O(h^2)$, $\widetilde{\beta}_1^- = \widetilde{\beta}_0^+ = O(1)$, $\widetilde{\beta}_1^+ = O(h^2)$;
- If $x_j < x_s < x_{j+1}$, then $\widetilde{\beta}_0^- = O(h^2)$, $\widetilde{\beta}_1^- = \widetilde{\beta}_0^+ = O(1)$, $\widetilde{\beta}_1^+ = O(1)$;
- If $x_{j+1} \leq x_s < x_{j+2}$, then $\widetilde{\beta}_0^- = O(h^2)$, $\widetilde{\beta}_1^- = \widetilde{\beta}_0^+ = O(h^2)$, $\widetilde{\beta}_1^+ = O(1)$,

with x_s point of singularity. Now, we can compute

$$\frac{\alpha_1^{\pm} - \alpha_0^{\pm}}{\alpha_0^{\pm}} = \frac{(\beta_0^{\pm} + \sigma_h)^2 - (\beta_1^{\pm} + \sigma_h)^2}{(\beta_1^{\pm} + \sigma_h)^2} = \left(\frac{\beta_0^{\pm} - \beta_1^{\pm}}{\beta_1^{\pm} + \sigma_h}\right)\left(\frac{\beta_0^{\pm} + \beta_1^{\pm} + 2\sigma_h}{\beta_1^{\pm} + \sigma_h}\right), \tag{2.17}$$

which, noticing that, if the function is smooth in both stencils of $\beta_0^{\pm}$ and $\beta_1^{\pm}$, we have

$$\begin{aligned} \frac{\beta_0^{\pm} - \beta_1^{\pm}}{\beta_1^{\pm} + \sigma_h} &= -2h\frac{f_j'' f_j'''}{(f'')^2 + \sigma} + O(h^2) = O(h) \\ \frac{\beta_0^{\pm} + \beta_1^{\pm} + 2\sigma_h}{\beta_1^{\pm} + \sigma_h} &= 2 + O(h) = O(1), \end{aligned} \tag{2.18}$$

leads to

$$\alpha_1^{\pm} = \alpha_0^{\pm}(1 + O(h)). \tag{2.19}$$

Then, we can deduce that if the solution is regular enough in both stencils

$$\omega_{\pm} = \frac{\alpha_k^{\pm}}{\alpha_0^{\pm} + \alpha_1^{\pm}} = \frac{1}{2} + O(h), \tag{2.20}$$

with $k = 0$ for the superscript "+" and $k = 1$ for "−". On the contrary, if there is a singularity in at least one of the stencils

$$\alpha_k^{\pm} = \begin{cases} O(1) & \text{if } f \text{ is not smooth in } \mathcal{S}_{j+k} \\ O(h^{-4}) & \text{if } f \text{ is smooth in } \mathcal{S}_{j+k}, \end{cases} \tag{2.21}$$

then it is easy to verify that the behavior of our $\omega_{\pm}$ falls in one of the following cases:

- If $x_{j-2} < x_s \leq x_{j-1}$, then $\omega_- = 1 + O(h^4)$, $\omega_+ = 1/2 + O(h)$
- If $x_{j-1} < x_s < x_j$, then $\omega_- = O(1)$, $\omega_+ = O(h^4)$
- If $x_s = x_j$, then $\omega_- = O(h^4)$, $\omega_+ = O(h^4)$
- If $x_j < x_s < x_{j+1}$, then $\omega_- = O(h^4)$, $\omega_+ = O(1)$
- If $x_{j+1} \leq x_s < x_{j+2}$, then $\omega_- = 1/2 + O(h)$, $\omega_+ = 1 + O(h^4)$,

where with $\omega_{\pm} = O(1)$ we mean a number dependent on the jump of the derivative.

Now, defining $\omega_j := \min\{\omega_-, \omega_+\}$ we can rewrite

$$\omega_j = \begin{cases} O(h^4) & \text{if } x_{j-1} < x_s < x_{j+1} \\ \frac{1}{2} + O(h) & \text{otherwise.} \end{cases} \tag{2.22}$$

***Remark* 2.3.** It is interesting to point out that, without developing the algebra which follows almost the same lines, if we had used the indicators β_k^{CL} we would have concluded

- If $x_{j-2} < x_s \leq x_{j-1}$, then $\omega_- = O(1)$, $\omega_+ = 1/2 + O(h^3)$
- If $x_{j-1} < x_s < x_{j+1}$, then $\omega_- = O(1)$, $\omega_+ = O(1)$
- If $x_{j+1} \leq x_s < x_{j+2}$, then $\omega_- = 1/2 + O(h^3)$, $\omega_+ = O(1)$.

Then, it is clear that, using the indicators for conservation laws, we can not exploit the behavior of $\omega_{\pm}$ in order to precisely locate a point of singularity x_s. To be more precise, here the $O(1)$ comes from the fact that, by Proposition 2.1, we would not have the property (2.21) in presence of a singularity, so using again the expansion (2.17) we have to notice that

$$\frac{\beta_0^{\pm} - \beta_1^{\pm}}{\beta_1^{\pm} + \sigma_h} = O(1),$$

a number dependent on the jump of the derivative. Notice that in the case of β_k^{CL} this happens even if the solution is smooth in one of the two stencils.

Notice also that, when the solution is smooth, with the choice β_k^{CL} the information given by $\omega_{\pm}$ are more accurate, since the possible oscillations around the optimal value are of order $O(h^3)$ instead of $O(h)$, which follows expanding with Taylor as we have done for β_k^{HJ}.

Going back to our $\omega^{\pm}$, we noticed through numerical tests that the $O(h)$ term in regular regions may produce heavy oscillations around the optimal value $\overline{\omega} = 1/2$. To increase the accuracy, we can use higher order smoothness indicator ($r > 2$), as we will show briefly in the examples below, but we would need a bigger reconstruction stencil, or we can use the *mappings* defined in [HAP05],

$$g(\omega) = \frac{\omega(\overline{\omega} + \overline{\omega}^2 - 3\overline{\omega}\omega + \omega^2)}{\overline{\omega}^2 + \omega(1 - 2\overline{\omega})}, \qquad \overline{\omega} \in (0, 1) \tag{2.23}$$

which have the properties that $g(0) = 0$, $g(1) = 1$, $g(\overline{\omega}) = \overline{\omega}$, $g'(\overline{\omega}) = 0$ and $g''(\overline{\omega}) = 0$. Then, we define

$$\begin{aligned}
\omega_{\pm}^{*} &= g(\omega_{\pm}) \\
&= g(\overline{\omega}) + g'(\overline{\omega})(\omega_{\pm} - \overline{\omega}) + \frac{g''(\overline{\omega})}{2}(\omega_{\pm} - \overline{\omega})^2 + \frac{g'''(\overline{\omega})}{6}(\omega_{\pm} - \overline{\omega})^3 + O(h^4) \\
&= \overline{\omega} + \frac{(\omega_{\pm} - \overline{\omega})^3}{\overline{\omega} - \overline{\omega}^3} + O(h^4) \\
&= \overline{\omega} + O(h^3)
\end{aligned}$$

These mappings have been introduced to propose a modification of the original WENO procedure, termed *M-WENO*, which produces better approximations at critical points.

***Remark* 2.4.** Notice that with respect to the definition in [HAP05] we avoided the second weighting which seems unnecessary in our case. More explicitly, the mapping we use is

$$g(\omega) = 4\omega\left(\frac{3}{4} - \frac{3}{2}\omega + \omega^2\right). \tag{2.24}$$

Another useful technique to reduce the oscillations, which in particular does not require any mapping, has been proposed in the context of hyperbolic conservation laws in [BCCD08] and further generalized in [CCD11], leading to the definition of the so-called *WENO-Z* schemes. This procedure can be applied for our purpose without any relevant change. For example, for $r = 2$, it is implemented first defining

$$\tau^{\pm} := \left|\beta_0^{\pm} - \beta_1^{\pm}\right|, \tag{2.25}$$

which has the properties:

- If f is smooth in, respectively, $\mathcal{S}^{\pm} := \mathcal{S}_0^{\pm} \bigcup \mathcal{S}_1^{\pm}$, then $\tau^{\pm} = O(h^3)$;
- If f is smooth in some $\mathcal{S}_k^{\pm}$, but not in $\mathcal{S}^{\pm}$, then $\tau^{\pm} \gg \beta_k^{\pm}$;
- $\tau^{\pm} \leq \max_k \beta_k^{\pm}$.

Then, analogously to the usual WENO procedure, the final indicator is obtained computing

$$\alpha_k^{Z,\pm} = \frac{1}{2}\left(1 + \left(\frac{\tau^{\pm}}{\beta_k^{\pm} + \sigma_h}\right)^p\right), \qquad \omega_{\pm}^{Z} = \frac{\alpha_{\nu}^{Z,\pm}}{\alpha_0^{Z,\pm} + \alpha_1^{Z,\pm}}, \tag{2.26}$$

for $k = 0, 1$, with $p = 2$, $\nu = 1$ for the superscript '$-$' and $\nu = 0$ for '$+$'. Unfortunately, applying directly the 'WENO-Z' procedure to Hamilton-Jacobi equations, in particular in our context, we do not achieve improvements comparable to those obtained through the mapping (2.24). In fact, using the expansions and the same computations of the previous lines, it is straightforward to show that the resulting smoothness indicators are such that

$$\omega_{\pm}^{Z} = \frac{1}{2} + O(h^2),$$

and thus produce slightly wider oscillations around the optimal value. This problem would probably suggest to increase to power p in (2.44), but in doing so also the dependence on the magnitudes of successive derivatives of f is increased (see (2.18)), producing even wider oscillations for coarser grids, thus discouraging such an approach.

For these reasons we avoid the use of the indicator ω^Z in all the simulations of the present work, also in order to reduce the number of similar simulations.

***Remark* 2.5.** It seems that, at least in the case $r = 2$, making use of the full stencil $\{x_{j-2}, \ldots, x_{j+2}\}$ in defining $\tau^{\pm}$ in (2.25)-(2.44) gives very interesting results, even better than the mapped indicators. More precisely, we propose the following definition, valid for both cases,

$$\tau = \left|\beta_0^- - 2\beta_1^- + \beta_1^+\right|, \tag{2.27}$$

which gives

$$\tau = 2h^4 \left((f_j''')^2 f_j'' f_j^{(4)} \right) + O(h^5) = O(h^4), \tag{2.28}$$

if the function is smooth in the full stencil. Consequently, it can be shown that the indicators computed through (2.27)-(2.44), are such that

$$\omega_{new}^{Z} = \frac{1}{2} + O(h^4),$$

in regions of regularity and have similar behavior w.r.t. the previous constructions in the remaining cases. We only recently came up with this new formulation, thus it will not be further investigated in this thesis, but it will be most probably present in [FPTa].

The last step is to use such indicators to define our function $\phi(\omega)$, such that

$$\phi(\omega_j) = \begin{cases} 1 & \text{if the function } f \text{ is regular in } I_j, \\ 0 & \text{if } I_j \text{ contains a point of singularity,} \end{cases}$$

where $I_j = (x_{j-1}, x_{j+1})$, $j \in \mathbb{Z}$. From the previous computations, it is clear that we should define ϕ such that $\phi = 1$ if ω is close to $\frac{1}{2}$ and $\phi = 0$, otherwise. Notice that in the latter are included both cases in which the function has a singularity in the first derivative ($\omega = O(\Delta x^4)$) and when the second derivative is discontinuous ($\omega = O(1)$). The simplest choice is to take

$$\phi(\omega) = \chi_{\{\omega \geq M\}}, \tag{2.29}$$

with $M < \frac{1}{2}$, a number possibly dependent on Δx. Or we can choose a more regular function

$$\phi(\omega) = \frac{e^{-M\omega} - 1}{e^{-M} - 1},$$

where now M must be big enough to have a steep decent towards 0.

***Remark* 2.6.** Notice that to construct the function ϕ using the indicators (2.1) with $r = 2$ we need only five points to inspect the regularity in I_j.

Next, we show that if we make a particular choice for M we are able to prove the following result, which can be seen as an "inverse" of Proposition 2.1 for numerical solutions and (probably) gives a useful tool for the analysis of the next chapter. Unfortunately, at the moment, this result is valid only for indicators ω using the standard construction for $r = 2$, without the possibility to introduce any of the presented modifications, or higher order indicators. Moreover, as it will be briefly discussed in Remark 3.4, it introduces some limitations in the applicability even when using the standard indicators, testifying the necessity of some improvements in the argument used.

Before proceeding, let us remind that we are working with structured grids, then if we consider a one-parameter family of grid values $\{f_j(\Delta x)\}_{j\in J(\Delta x)}$, as Δx goes to 0, the indexed family of sets of indices $J(\Delta x)$ is *expanding*, in the sense that if $\Delta x_2 < \Delta x_1$, then $J(\Delta x_1) \subset J(\Delta x_2)$, where $J(\Delta x) \subseteq \mathbb{Z}$, for all $\Delta x > 0$. Moreover, we define $I_s(\Delta x)$ as the set of indices j such that $\phi_j = 0$ and assume, for simplicity, $|I_s(\Delta x)| < \infty$ and $I_s(\Delta x) \equiv I_s$, for $\Delta x > 0$.

Lemma 2.2. *Let ω be computed using (2.14)-(2.15) and ϕ be defined by (2.29) with $M(\Delta x) = \frac{1}{2} - C\Delta x$, for some constant C such that $0 < M(\Delta x) < \frac{1}{2}$. Consider a one-parameter family of sequences $\{f_j(\Delta x)\}_{j\in J(\Delta x)}$, and a partition $\{R_i\}_{i=0,\dots,|I_s|}$ of the regularity set $\mathcal{R} = \{j \in \mathbb{Z} : \phi_j = 1\} = \bigcup_i R_i$, and $\mathcal{R} = \mathbb{Z}$ if $I_s = \emptyset$. Then, if for all $i = 0, \dots, |I_s|$, there exists $j_i \in R_i$, such that $|D^2 f_{j_i}(\Delta x)| < \infty$, we have that*

$$|D^2 f_j(\Delta x)| = \frac{|f_{j+1}(\Delta x) - 2f_j(\Delta x) + f_{j-1}(\Delta x)|}{\Delta x^2} \leq B, \quad \forall j \in \mathcal{R}, \tag{2.30}$$

for a constant B independent of Δx.

Proof. It is clear that, since $|I_s| < \infty$ by hypothesis, it is enough to prove the assertion just for one $i \in I_s$, or more simply in the regular case $\mathcal{R} = \mathbb{Z}$. Let us assume then that the sequence is "regular" and, without loss of generality, that there exists $j_{\min} \in \mathcal{R}$ such that $D^2 f_{j_{\min}}(\Delta x) = 0$, which happens for example when the sequence has compact support. In the following, we drop the dependence on Δx for clarity of presentation, since it should not cause confusion.

Since we are considering the case of ω computed by non-mapped indicators, we have that, by definition of ϕ and ω, if $\phi_j = 1$ then at least one $\omega_\pm > M$. Without loss of generality, let us assume $\omega_+ > M$, with the other case being symmetrical. Then, by definition,

$$\omega_+ = \frac{(\beta_1^+ + \sigma_h)^2}{(\beta_1^+ + \sigma_h)^2 + (\beta_0^+ + \sigma_h)^2} > M,$$

which leads by simple computations to

$$|\beta_0^+| < \sqrt{\frac{1-M}{M}}|\beta_1^+| + \left(\sqrt{\frac{1-M}{M}} - 1\right)\sigma_h,$$

and dividing by Δx^2, we get

$$|D^2 f_j|^2 < \sqrt{\frac{1-M}{M}}|D^2 f_{j+1}|^2 + \left(\sqrt{\frac{1-M}{M}} - 1\right),$$

where we used the definition of $\sigma_h = \sigma \Delta x^2$, with $\sigma = 1$ for simplicity. Whence, iterating till $L_j \in \mathbb{N}$ such that $j + L_j = j_{\min}$, we have

$$\begin{aligned}
|D^2 f_j|^2 < \ldots &< \left(\frac{1-M}{M}\right)^{\frac{L_j}{2}} |D^2 f_{j_{\min}}|^2 + \left(\sqrt{\frac{1-M}{M}} - 1\right) \sum_{k=1}^{L_j} \left(\frac{1-M}{M}\right)^{\frac{k-1}{2}} \\
&= \left(\sqrt{\frac{1-M}{M}} - 1\right) \sum_{k=0}^{L_j - 1} \left(\frac{1-M}{M}\right)^{\frac{k}{2}} \\
&= \left(\sqrt{\frac{1-M}{M}} - 1\right) \frac{1 - \left(\frac{1-M}{M}\right)^{\frac{L_j}{2}}}{1 - \sqrt{\frac{1-M}{M}}} = \left(\frac{1-M}{M}\right)^{\frac{L_j}{2}} - 1.
\end{aligned}$$

Now, if notice that forall $j \in \mathcal{R}$ we can find a constant $L > 0$ independent on j such that $L_j \leq \frac{L}{\Delta x}$ and recall the hypothesis on $M = \frac{1}{2} - C\Delta x$, we can conclude

$$\begin{aligned}
|D^2 f_j|^2 &\leq \left(\frac{1-M}{M}\right)^{\frac{L_j}{2}} - 1 \leq \left(\frac{1}{M} - 1\right)^{\frac{L}{2\Delta x}} - 1 \\
&= \left(\frac{2}{1 - 2C\Delta x} - 1\right)^{\frac{L}{2\Delta x}} - 1 \\
&\approx (1 + 4C\Delta x)^{\frac{L}{2\Delta x}} - 1 \overset{\Delta x \to 0}{\longrightarrow} e^{2LC} - 1,
\end{aligned}$$

by the well known notable limit. Then, the statement follows taking $B := \sqrt{e^{2LC} - 1}$. □

***Remark* 2.7.** Notice that the previous lemma strongly relies on the fact that ω is computed using (2.14)-(2.15) without introducing the mappings (2.24). In fact, if we were to use (2.24), we could develop the algebra until the inequality

$$|D^2 f_j|^2 \leq \left(\frac{1}{g^{-1}(M)} - 1\right)^{\frac{L}{2\Delta x}} - 1,$$

but, by definition, g^{-1} can not be expanded in Taylor series around the point $\frac{1}{2}$, whence we could not use the notable limit to conclude.

On the other hand, through extensive numerical simulations on various critical situations, we could acknowledge that a weaker result seems to hold also for more general indicators. More precisely, we collected numerical evidence that, fixed $j \in \mathbb{Z}$, if the sequence of the second order increments

$$D^2 f_j(\Delta x) = \frac{f_{j+1}(\Delta x) - 2f_j(\Delta x) + f_{j-1}(\Delta x)}{\Delta x^2},$$

presents some kind of discontinuity, then we have

$$\omega_j = O(1).$$

Consequently, choosing $M(\Delta x) = \frac{1}{2} - C\Delta x$, for a constant C such that $0 < M(\Delta x) < \frac{1}{2}$, or even more simply, $M(\Delta x) \to 0$ as $\Delta x \to 0$, we could infer

$$j \notin \mathcal{R}, \quad \text{for } \Delta x \to 0,$$

Notice that this property can be proven almost directly in the case the discontinuity of the second order increment is "caused" by that of the first order finite difference, in the sense

$$D^+ f_j(\Delta x) := \frac{f_{j+1}(\Delta x) - f_j(\Delta x)}{\Delta x} \to f_x^+, \quad D^- f_j(\Delta x) := \frac{f_j(\Delta x) - f_{j-1}(\Delta x)}{\Delta x} \to f_x^-,$$

with $f_x^+ \neq f_x^-$. On the contrary, if the sequence of $D^{\pm} f_j(\Delta x)$ is "regular", the detection of a discontinuity in $D^2 f_j(\Delta x)$ is more involved. It is noteworthy to point out that we are interested mainly in detecting unbounded second order increments. Unfortunately, without any further assumption on the sequences $\{f_j(\Delta x)\}_{j \in J(\Delta x)}$, such a result would not suffice to infer that $|D^2 f_j| < B$ if $j \in \mathcal{R}$, for some $B > 0$, since we could not secure the boundedness of second order increments at points in some neighborhood of a cell (point) at which the sequence is "regular" but has unbounded second order increment.

2.1.2. Alternative constructions

In the following, we give some examples of other possible choices for the smoothness indicator ω, using again indicators in the sense of [JP00] but in slightly different versions, more appropriate for our purpose. The aim is mainly to show how higher order indicators ($r > 2$) can be used also in this context in order to have a better approximation of the singular region.

First we give a slightly different formulation for ω with $r = 2$, which allows to avoid the computation of both indicators to define ϕ.

Looking at the behavior of the $\beta_k^{\pm}$ around a singularity, if we consider the information given by the three stencils all together and remember that by definition $\alpha_0^+ = \alpha_1^-$, we can write

$$\omega = \frac{\theta \alpha_0^+}{\alpha_0^- + \theta \alpha_0^+ + \alpha_1^+} = \frac{\theta \alpha_1^-}{\alpha_0^- + \theta \alpha_1^- + \alpha_1^+}, \tag{2.31}$$

with θ a constant we can use to "decide" the optimal value $\overline{\omega}$ in regular regions, still ensuring easily that $\omega \in [0, 1]$. For example, $\theta = 1$ gives $\overline{\omega} = 1/3$, while $\theta = 2$ gives $\overline{\omega} = 1/2$. More precisely, reminding (2.19) and noticing that the exact same computation (with α_0^- in place of α_1^+) would lead to

$$\alpha_0^- = \alpha_0^+ (1 + O(h)), \tag{2.32}$$

we can see that if the function is smooth in $[x_{j-2}, x_{j+2}]$ then

$$\omega = \overline{\omega} + O(h) = \frac{\theta}{2 + \theta} + O(h).$$

On the other hand, if $x_s \in (x_{j-2}, x_{j+2})$,

- If $x_{j-2} < x_s \leq x_{j-1}$, then $\omega = \frac{\theta\alpha_0^+}{\alpha_0^+\left(\frac{\alpha_0^-}{\alpha_0^+}+\theta+(1+O(h))\right)} = \frac{\theta}{\theta+1} + O(h)$
- If $x_{j-1} < x_s < x_{j+1}$, then $\omega = O(h^4)$
- If $x_{j+1} \leq x_s < x_{j+2}$, then $\omega = \frac{\theta\alpha_0^+}{\alpha_0^+\left((1+O(h))+\theta+\frac{\alpha_1^+}{\alpha_0^+}\right)} = \frac{\theta}{\theta+1} + O(h)$.

Then we can use the mappings (2.23) if we want to reduce the effects of the $O(h)$ terms. Notice that in the case of $\omega = \theta/(\theta+1) + O(h)$ when the singularity enters or leaves the stencil of the regularity indicator, we have, at least in the limit for $h \to 0$,

$$g\left(\frac{\theta}{\theta+2} + O(h)\right) \leq g\left(\frac{\theta}{\theta+1} + O(h)\right) \Rightarrow \frac{\theta}{\theta+2} + O(h^3) \leq g\left(\frac{\theta}{\theta+1} + O(h)\right),$$

using the fact that g is a non decreasing function, and so the mapping will have little effects.

Another simple way to reduce the oscillations around the optimal value $\overline{\omega}$ is to increase the order of the reconstruction to compute the smoothness indicators. As an example we show how to construct the indicators $\omega_\pm$ using polynomials of degree $r = 3$. In order to remain as short as possible, we will focus only on the construction of ω_+, knowing also that the case of ω_- follows almost the same steps.

Let us start by considering the formula

$$\beta_k^+ = \sum_{l=2}^{3} \int_{x_j}^{x_{j+1}} \Delta x^{2l-3} \left(P_k^{(l)}\right)^2 dx,$$

which gives,

- $\beta_0^+ = \frac{13h^2}{12}\left(D^2 f_{j-1} - D^2 f_j\right)^2 + \frac{h^2}{4}\left(D^2 f_{j-1} - 3D^2 f_j\right)^2$,
- $\beta_1^+ = \beta_0^+ = \frac{13h^2}{12}\left(D^2 f_j - D^2 f_{j+1}\right)^2 + \frac{h^2}{4}\left(D^2 f_j + D^2 f_{j+1}\right)^2$,
- $\beta_2^+ = \beta_0^+ = \frac{13h^2}{12}\left(D^2 f_{j+1} - D^2 f_{j+2}\right)^2 + \frac{h^2}{4}\left(3D^2 f_{j+1} - D^2 f_{j+2}\right)^2$,

and, developing in Taylor series,

- $\beta_0^+ = h^2(f_j'')^2 + h^3 f_j'' f_j''' + h^4\left(\frac{4}{3}(f_j''')^2 - \frac{f_j'' f_j^{(4)}}{3}\right) + h^5\left(\frac{15}{12} f_j''' f_j^{(4)} + \frac{f_j'' f_j^{(5)}}{4}\right) + o(h^5)$,
- $\beta_1^+ = h^2(f_j'')^2 + h^3 f_j'' f_j''' + h^4\left(\frac{4}{3}(f_j''')^2 + \frac{2}{3} f_j'' f_j^{(4)}\right) + h^5\left(\frac{17}{12} f_j''' f_j^{(4)} + \frac{f_j'' f_j^{(5)}}{4}\right) + o(h^5)$,
- $\beta_2^+ = h^2(f_j'')^2 + h^3 f_j'' f_j''' + h^4\left(\frac{4}{3}(f_j''')^2 - \frac{f_j'' f_j^{(4)}}{3}\right) + h^5\left(\frac{37}{12} f_j''' f_j^{(4)} - \frac{3}{4} f_j'' f_j^{(5)}\right) + o(h^5)$.

Now, we define α_k^+ as in (2.16),

$$\alpha_k^+ = \frac{1}{(\beta_k^+ + \sigma_h)^2}, \qquad \text{for } k = 0, 1, 2,$$

with $\sigma_h = \sigma h^2$ and focus on the interval $[x_j, x_{j+1}]$ by defining

$$\omega_+ = \frac{\alpha_0^+ + \alpha_1^+}{\alpha_0^+ + \alpha_1^+ + \alpha_2^+}.$$

We recall that, by Proposition 2.1, $\widetilde{\beta}_k = O(h^2)$ if there is no singularity in the stencil and $\widetilde{\beta}_k = O(1)$ otherwise, so in presence of a singularity we can only fall in one of the following cases:

- If $x_{j-3} < x_s \leq x_{j-2}$, then $\widetilde{\beta}_0^+ = O(h^2)$, $\widetilde{\beta}_1^+ = O(h^2)$, $\widetilde{\beta}_2^+ = O(h^2)$;
- If $x_{j-2} < x_s \leq x_{j-1}$, then $\widetilde{\beta}_0^+ = O(1)$, $\widetilde{\beta}_1^+ = O(h^2)$, $\widetilde{\beta}_2^+ = O(h^2)$;
- If $x_{j-1} < x_s \leq x_j$, then $\widetilde{\beta}_0^+ = O(1)$, $\widetilde{\beta}_1^+ = O(1)$, $\widetilde{\beta}_2^+ = O(h^2)$;
- If $x_j \leq x_s < x_{j+1}$, then $\widetilde{\beta}_0^+ = O(1)$, $\widetilde{\beta}_1^+ = O(1)$, $\widetilde{\beta}_2^+ = O(1)$;
- If $x_{j+1} \leq x_s < x_{j+2}$, then $\widetilde{\beta}_0^+ = O(h^2)$, $\widetilde{\beta}_1^+ = O(1)$, $\widetilde{\beta}_2^+ = O(1)$,
- If $x_{j+2} \leq x_s < x_{j+3}$, then $\widetilde{\beta}_0^+ = O(h^2)$, $\widetilde{\beta}_1^+ = O(h^2)$, $\widetilde{\beta}_2^+ = O(1)$,

with x_s point of singularity. Now, we can repeat the computations we have done for the case $r = 2$ by writing for example

$$\frac{\alpha_k^+ - \alpha_2^+}{\alpha_2^+} = \frac{(\beta_2^+ + \sigma_h)^2 - (\beta_k^+ + \sigma_h)^2}{(\beta_k^+ + \sigma_h)^2} = \left(\frac{\beta_2^+ - \beta_k^+}{\beta_k^+ + \sigma_h}\right)\left(\frac{\beta_2^+ + \beta_k^+ + 2\sigma_h}{\beta_k^+ + \sigma_h}\right),$$

which, noticing that

$$\frac{\beta_2^+ - \beta_k^+}{\beta_k^+ + \sigma_h} = O(h^2)$$
$$\frac{\beta_2^+ + \beta_k^+ + 2\sigma_h}{\beta_k^+ + \sigma_h} = O(1)$$

leads to

$$\alpha_k^+ = \alpha_0^+(1 + O(h^2)), \tag{2.33}$$

for $k = 0$, 1, 2. Then we can deduce that if the solution is regular enough in all the three stencils

$$\omega_+ := \frac{\alpha_0^+ + \alpha_1^+}{\alpha_0^+ + \alpha_1^+ + \alpha_2^+} = \frac{2}{3} + O(h^2). \tag{2.34}$$

***Remark* 2.8.** At this point, to be as clear as possible, it is better to present also the definition of ω_-, for which similar computations easily hold. We use the following

$$\omega_- := \frac{\alpha_1^- + \alpha_2^-}{\alpha_0^- + \alpha_1^- + \alpha_2^-}. \tag{2.35}$$

On the other hand, if there is a singularity in at least one of the stencils

$$\alpha_k^+ = \begin{cases} O(1) & \text{if } f \text{ is not smooth in } \mathcal{S}_{j+k} \\ O(h^{-4}) & \text{if } f \text{ is smooth in } \mathcal{S}_{j+k}, \end{cases}$$

then it can be shown that the behavior of $\omega_\pm$ falls in the following cases:

- If $x_{j-3} < x_s \leq x_{j-2}$, then $\omega_- = 1 + O(h^4)$, $\omega_+ = 2/3 + O(h^2)$
- If $x_{j-2} < x_s \leq x_{j-1}$, then $\omega_- = 1 + O(h^4)$, $\omega_+ = 1/2 + O(h^2)$
- If $x_{j-1} < x_s < x_j$, then $\omega_- = O(1)$, $\omega_+ = O(h^4)$
- If $x_s = x_j$, then $\omega_- = O(h^4)$, $\omega_+ = O(h^4)$
- If $x_j < x_s < x_{j+1}$, then $\omega_- = O(h^4)$, $\omega_+ = O(1)$
- If $x_{j+1} \leq x_s < x_{j+2}$, then $\omega_- = 1/2 + O(h^2)$, $\omega_+ = 1 + O(h^4)$,
- If $x_{j+2} \leq x_s < x_{j+3}$, then $\omega_- = 2/3 + O(h^2)$, $\omega_+ = 1 + O(h^4)$,

whence, as in the case with $r = 2$, we have $\omega := \min\{\omega_-, \omega_+\} = O(h^4)$, if $x_s \in I_j = (x_{j-1}, x_{j+1})$, and ω away from 0, otherwise.

***Remark* 2.9.** Notice that in this case, we cannot repeat the constructions we have done for $r = 2$ because, as we can see developing β_0^- with Taylor,

$$\begin{aligned}\beta_0^- &= \frac{13h^2}{12}\left(D^2 f_{j-2} - D^2 f_{j-1}\right)^2 + \frac{h^2}{4}\left(D^2 f_{j-2} - 3D^2 f_{j-1}\right)^2 \\ &= h^2 (f_j'')^2 - h^3 f_j'' f_j''' + h^4 \left(\frac{4}{3}(f_j''')^2 - \frac{f_j'' f_j^{(4)}}{3}\right) - h^5 \left(\frac{37}{12} f_j''' f_j^{(4)} - \frac{3}{4} f_j'' f_j^{(5)}\right) + o(h^5),\end{aligned}$$

whence,

$$\frac{\beta_2^+ - \beta_0^-}{\beta_0^+ + \sigma_h} = O(h),$$

and we would lose the order of accuracy.

Finally, if needed we can use the mappings (2.23) to reduce the oscillations. Notice that repeating the same computations as before we get

$$\omega_\pm^* = g(\omega_\pm) = \overline{\omega} + O(h^6),$$

if the function is smooth in all the three stencils, while the mappings will have little effects in the remaining cases.

To conclude we just mention that, as it should be quite clear now, we can define higher order smoothness indicators using polynomials of order $r \geq 3$ and taking

$$\omega_- := \frac{\sum_{k=1}^{k=r-1} \alpha_k^-}{\sum_{k=0}^{k=r-1} \alpha_k^-}, \qquad \omega_+ := \frac{\sum_{k=0}^{k=r-2} \alpha_k^+}{\sum_{k=0}^{k=r-1} \alpha_k^+},$$

thus requiring a wider and wider stencil.

2.2. Regularity indicators in higher dimensions

In this section we will present a multidimensional extension of the smoothness indicators studied in the previous section. In order to keep the ideas as clear as possible, we will complete the construction only for the simplest case of two space dimensions and $r = 2$, with more general situations following directly.

Let us begin by defining, for simplicity, a uniform discretization of the plane (x, y), with mesh steps Δx and Δy, $-\infty < x_j < \infty$, $j \in \mathbb{Z}$ and $-\infty < y_i < \infty$, $i \in \mathbb{Z}$. We will also use the shorter notation $h_z := \Delta z$, for $z = x, y$.

Before proceeding with the construction, let us recall some important facts about multivariate interpolation. As it is well known, there are many possibilities to define polynomials in two dimensions, in fact for example we could fix the total degree r of the polynomial and consider a triangular array of points (then using polynomials in $\mathbb{P}_r(\mathbb{R}^2)$), or, as we will do in our approach, we can fix the degree r in each variable and define the 2D-polynomial as the tensor product of one-dimensional polynomials ($P \otimes Q \in \mathbb{Q}_r(\mathbb{R}^2)$, with $P, Q \in \mathbb{P}_r(\mathbb{R})$). Clearly, with the last approach, the number of points involved in the stencil of the reconstruction increases exponentially (considering a square grid, if n is the number of points of the one-dimensional stencils, then n^2 is the cardinality of the two-dimensional stencil).

Notice that in both cases the problem is well posed, in fact we can define a unique polynomial interpolating a given function $f(x, y)$ on the points of the stencil with the desired degree. Clearly some assumptions on the disposition of the points must be made, in fact, for example the points must not lie on the same line, but that is trivially the case for uniform cartesian grids. This is indeed our case of study, in which we will work on structured grids.

Let us consider the general case of a rectangular stencil $\mathcal{S} = \{x_0, \ldots, x_n\} \times \{y_0, \ldots, y_m\}$, then, using again the *Newton form*, we define the polynomial of degree $n + m$ interpolating a given function f as

$$P(x, y) = \sum_{s=0}^{n} \sum_{t=0}^{m} \omega_{t-1}(x) \omega_{s-1}(y) f[x_0, \ldots, x_t; y_0, \ldots, y_s], \tag{2.36}$$

where $\omega_k(z) = \omega_{k-1}(z - z_k)$, $\omega_{-1} = 1$ and the two-dimensional divided difference $f[\cdot; \cdot]$ are computed as in the one-dimensional case (2.4), keeping each time one of the two variables fixed and computing the divided difference with respect to the free variable,

that is, for example

$$
\begin{aligned}
&f[x_t, y_s] := f(x_t, y_s), \qquad\qquad t = 0, \dots, m,\ s = 0, \dots, n \\
&f[x_0, \dots, x_t; y_s] := \frac{f[x_1, \dots, x_t; y_s] - f[x_0, \dots, x_{t-1}; y_s]}{x_t - x_0}, \\
&f[x_0, \dots, x_t; y_0, \dots, y_s] := \frac{f[x_1, \dots, x_t; y_0, \dots, y_s] - f[x_0, \dots, x_{t-1}; y_0, \dots, y_s]}{x_t - x_0},
\end{aligned}
$$

and the same with respect to the second variable. Now, if we want to define a smoothness indicator of a function f, as in the one-dimensional case, we have to focus our attention on a single cell of the grid, say for example $(x_{j-1}, x_j] \times [y_{i-1}, y_i)$. Then, we propose the following

$$
\beta_{k,w} = \sum_{\alpha_1=0}^{n} \sum_{\alpha_2=0}^{m} \int_{x_{j-1}}^{x_j} \int_{y_{i-1}}^{y_i} \Delta x^{\gamma_1} \Delta y^{\gamma_2} \left(\frac{\partial^\alpha P_{k,w}(x,y)}{\partial x^{\alpha_1} \partial y^{\alpha_2}} \right)^2 dxdy, \tag{2.37}
$$

for $k = 0, \dots, n-1$ and $w = 0, \dots, m-1$, where $P_{k,w}$ is the interpolating polynomial on the stencil $\mathcal{S}_{k,w} = \{x_{j+k-n}, \dots, x_{j+k}\} \times \{y_{i+w-m}, \dots, y_{i+w}\}$, $\alpha = (\alpha_1, \alpha_2)$ is a multi-index and γ_1, γ_2 must be chosen (depending on α_1 and α_2, respectively) in order to have the property

- $\beta_{k,w} = O(\Delta x^2 + \Delta x \Delta y + \Delta y^2)$ if the function is smooth in $\mathcal{S}_{k,w}$;
- $\beta_{k,w} = O(1)$ if there is a singularity in $\overset{\circ}{\mathcal{S}}_{k,w}$,

which we proved in Proposition 2.1 for the one-dimensional indicators. Notice that, also in this situation, the computation of lower order derivatives can be avoided since they are not useful in detecting discontinuities in the gradient. Therefore, the summation in (2.37) should be restricted to multi-indices α such that $|\alpha| \geq 2$.

From here on, in order to obtain an easy and explicit formula, we will focus on the case $n = m = 2$, which is also enough for our needs. Notice that with this assumption we work with polynomials of second degree in each variable, constructed on stencils of nine points. Then, (2.37) simply reads

$$
\beta_{k,w} = \sum_{\substack{\alpha_1,\alpha_2=0 \\ |\alpha|\geq 2}}^{2} \int_{x_{j-1}}^{x_j} \int_{y_{i-1}}^{y_i} \Delta x^{2(\alpha_1-1)} \Delta y^{2(\alpha_2-1)} \left(\frac{\partial^\alpha P_{k,w}(x,y)}{\partial x^{\alpha_1} \partial y^{\alpha_2}} \right)^2 dxdy, \tag{2.38}
$$

where we have made the choice $\gamma_i = 2(\alpha_i - 1)$, $i = 1, 2$ in order to obtain the desired property (as will be shown in the numerical tests). We will soon change a little this notation in order to make it more comfortable for our approach, in which we will be changing the integration domain in each subcase. More precisely, referring to Figure 2.1, we will consider the splitting of the domain $[x_{j-1}, x_{j+1}] \times [y_{i-1}, y_{i+1}]$ into four subcells. For each subdomain we will then compare only the information given by the "outer" (light blue in the figure) and the "inner" stencil (yellow in the figure). Consequently, we will need only one index to denote the respective stencil, using '0' for the inner stencil,

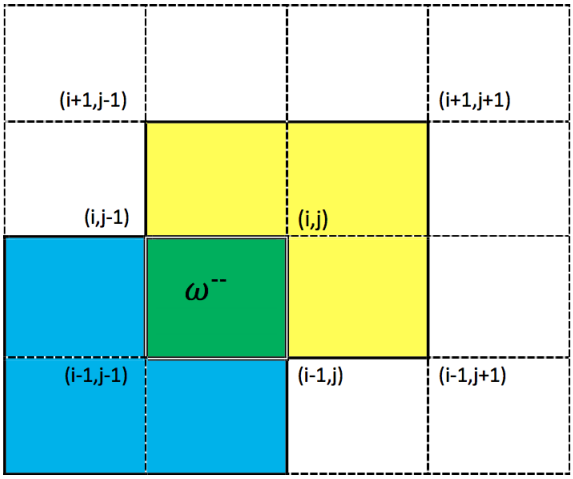

Figure 2.1.: Stencils of the polynomials needed to compute ω^{--}.

and '1' for the outer one. Moreover, we use the superscripts '$\pm\,\pm$, $\pm\,\mp$' in order to denote the shift with respect to the reference stencil, analogously to the one-dimensional case.

***Remark* 2.10.** To be precise, it is worth to point out that the formula (2.38) is an improvement of our first attempted extension, which we will also use in the numerical tests in Section 2.3. The previous definition is just slightly different, and it reads

$$\beta_{k,w} = \sum_{|\alpha|=2} \int_{x_{j-1}}^{x_j} \int_{y_{i-1}}^{y_i} \Delta x^{2(\alpha_1-1)} \Delta y^{2(\alpha_2-1)} \left(\frac{\partial^\alpha P_{k,w}(x,y)}{\partial x^{\alpha_1} \partial y^{\alpha_2}} \right)^2 dxdy. \tag{2.39}$$

Through extensive numerical tests we noticed that, extending the summation to include also the mixed derivative of order 4, the indicators give a more precise detection of singular regions. This behavior is rather surprising, since the two definitions have, in fact, the exact same computational cost and very similar explicit formulas, and will be object of future investigations.

Since we consider a function f smooth on a grid point (x_j, y_i) if it is so in the whole domain $I_{i,j} := [x_{j-1}, x_{j+1}] \times [y_{i-1}, y_{i+1}]$, following what we have done for the one-dimensional case, we have to inspect separately the four subcells around the point. We refer to the these subcells with the superscripts '$\pm\,\pm$, $\pm\,\mp$', according to the sign of the shift between the point (x_j, y_i) and its symmetrical edge with respect to the center of the considered cell, in both x and y directions, respectively. For example, if we focus on the cell $[x_{j-1}, x_j] \times [y_{i-1}, y_i]$ as in Figure 2.1, we use the superscript '$- -$'.

In this way we define four indicators $\omega^{\pm\pm}$ and $\omega^{\pm\mp}$ for each point (x_j, y_i) which quantify the regularity of the function in all the subcells around the point. Finally, we take $\omega = \min\{\omega^{\pm\pm}, \omega^{\pm\mp}\}$ as the smoothness indicator for the domain $I_{i,j}$. In order to compute these indicators, as Figure 2.1 shows, we always focus the attention on the polynomial constructed on the central stencil $\mathcal{S}_0 = \{x_{j-1}, x_j, x_{j+1}\} \times \{y_{i-1}, y_i, y_{i+1}\}$ and

compare it with the polynomial constructed on the symmetrical stencil with respect of the considered subcell (where they overlap), denoting the respective indicators by β_0 and β_1. Notice that with this approach even if we are using the 'central' polynomial for all the four indicators, we have to recompute β_0 for each case because of the change in the integration domain.

***Remark* 2.11.** Notice that the situation is slightly, but fundamentally, different from the one-dimensional case, in which through our procedure we are able to measure the regularity only in the *open* interval $I_j = (x_{j-1}, x_{j+1})$. That is indeed natural if we are focusing on the point x_j, since the integrals (2.14)-(2.15), with $k = 0$, are well-defined if the solution is regular in the open interval (x_{j-1}, x_{j+1}).

On the contrary, in the two-dimensional case, since the boundary of the domain $I_{i,j}$ is a curve, there are at least two possible directions, say v_1 and v_2, along which we can compute the directional derivative at any point $(x, y) \in I_{i,j}$. Then, if there exists a point $(x_s, y_s) \in I_{i,j}$ such that

$$\nabla_{v_1} f(x_s, y_s) \neq \nabla_{v_2} f(x_s, y_s),$$

it seems reasonable to consider the interpolating polynomial $P_0(x, y)$ not well defined in the whole domain, even if (x_s, y_s) coincides with one of the points of the boundary. Therefore, the integrals in (2.38), with $k, w = 0$, are not well defined and, consequently, the function f should be considered singular in $I_{i,j}$.

Now, we complete the construction for the case of ω^{--}, with the other three following the exact same lines. Before proceeding, in order to make the notation even lighter, assume without loss of generality that $(x_j, y_i) = (0, 0)$ and let us introduce again the function

$$f_h(\overline{x}, \overline{y}) := f(x_j + h_x\overline{x}, y_i + h_y\overline{y}),$$

using which we can rewrite the polynomials as

$$\begin{aligned} P_0^{--}(x,y) =& f_h(-1,-1) + (x+h_x) f_h[-1,0;-1] + (y+h_y) f_h[-1;-1,0] + \\ &+ x(x+h_x) f_h[-1,0,1;-1] + y(y+h_y) f_h[-1;-1,0,1] + \\ &+ (x+h_x)(y+h_y) f_h[-1,0;-1,0] + \\ &+ x(x+h_x)(y-h_y) f_h[-1,0,1;-1,0] + \\ &+ y(x+h_x)(y+h_y) f_h[-1,0;-1,0,1] + \\ &+ xy(x+h_x)(y+h_y) f_h[-1,0,1;-1,0,1], \end{aligned}$$

for the reference stencil, and

$$\begin{aligned} P_1^{--}(x,y) =& f_h(0,0) + x f_h[0,-1;0] + y f_h[0;0,-1] + x(x+h_x) f_h[0,-1,-2;0] + \\ &+ y(y+h_y) f_h[0;0,-1,-2] + xy f_h[0,-1;0,-1] + \\ &+ xy(x+h_x) f_h[0,-1,-2;0,-1] + xy(y+h_y) f_h[0,-1;0,-1,-2] + \\ &+ xy(x+h_x)(y+h_y) f_h[0,-1,-2;0,-1,-2], \end{aligned}$$

for the outer stencil, where now f_h denotes the divided difference of f_h (notice that the definition is different w.r.t. the one-dimensional case).

Whence, we plug this expressions in (2.38) and compute directly

$$\begin{aligned}\beta_k^{F,--} &= \sum_{\substack{\alpha_1,\alpha_2=0\\ |\alpha|\geq 2}}^{2} \int_{-\Delta x}^{0}\int_{-\Delta y}^{0} \Delta x^{2(\alpha_1-1)}\Delta y^{2(\alpha_2-1)} \left(\frac{\partial^\alpha P_k^{--}(x,y)}{\partial x^{\alpha_1}\partial y^{\alpha_2}}\right)^2 dxdy \\ &= \frac{1}{\Delta x\Delta y}\left[{f_{[3,1]}}^2 + {f_{[1,3]}}^2 + {f_{[2,2]}}^2 + \frac{17}{12}\left({f_{[3,2]}}^2 + {f_{[2,3]}}^2\right) + \frac{317}{720}{f_{[3,3]}}^2 + f_{[3,1]}f_{[3,2]} + \right. \\ &\qquad \left. + f_{[1,3]}f_{[2,3]} - \frac{1}{6}\left(f_{[3,1]}f_{[3,3]} + f_{[1,3]}f_{[3,3]}\right) - \frac{1}{12}\left(f_{[3,2]}f_{[3,3]} + f_{[2,3]}f_{[3,3]}\right)\right]\end{aligned} \tag{2.40}$$

where we have used the shorter notation $f_{[t,s]}$ to denote the multivariate undivided difference of f of order t in x and s in y. Notice that we avoided to specify the points on which the undivided difference are computed in order to obtain a unique formulation for both cases. If we want to write the explicit formula for a specific case, it is enough to substitute back the correct differences computed on the considered stencils. Notice also that we added due superscript 'F' to denote the *final* version of the indicators, mainly in order to help the discussion in the numerical tests of Section 2.3.2.

***Remark* 2.12.** Notice also that, in order to get a unique formula, we "used" the outer stencil in a smart way, writing the Newton form of the polynomial starting from the origin in both directions. More precisely, we have used the *ordered* stencils

$$\mathcal{S}_0^{--} = \{x_{j-1}, x_j, x_{j+1}\} \times \{y_{i-1}, y_i, y_{i+1}\}, \quad \mathcal{S}_1^{--} = \{x_j, x_{j-1}, x_{j-2}\} \times \{y_i, y_{i-1}, y_{i-2}\}.$$

Using the same approach, we can obtain the formula (2.40) for all the other indicators β_k^{++}, $\beta_k^{\pm\mp}$. In fact, it is enough to write the Newton form in a convenient way in each case. More precisely, we advise the use of the following *ordered* stencils:

- $\mathcal{S}_0^{+-} = \{x_{j+1}, x_j, x_{j-1}\}\times\{y_{i-1}, y_i, y_{i+1}\}$, $\mathcal{S}_1^{+-} = \{x_j, x_{j+1}, x_{j+2}\}\times\{y_i, y_{i-1}, y_{i-2}\}$;
- $\mathcal{S}_0^{++} = \{x_{j+1}, x_j, x_{j-1}\}\times\{y_{i+1}, y_i, y_{i-1}\}$, $\mathcal{S}_1^{++} = \{x_j, x_{j+1}, x_{j+2}\}\times\{y_i, y_{i+1}, y_{i+2}\}$;
- $\mathcal{S}_0^{-+} = \{x_{j-1}, x_j, x_{j+1}\}\times\{y_{i+1}, y_i, y_{i-1}\}$, $\mathcal{S}_1^{-+} = \{x_j, x_{j-1}, x_{j-2}\}\times\{y_i, y_{i+1}, y_{i+2}\}$.

Notice that we are changing also the ordering of the reference stencil in each case. Then, if we compute the integrals

$$\beta_k^{F,\vartheta_1\vartheta_2} = (-1)^{|\vartheta|}\sum_{\substack{\alpha_1,\alpha_2=0\\ |\alpha|\geq 2}}^{2} \int_{\vartheta_1\Delta x}^{0}\int_{\vartheta_2\Delta y}^{0} \Delta x^{2(\alpha_1-1)}\Delta y^{2(\alpha_2-1)} \left(\frac{\partial^\alpha P_k^{\vartheta_1\vartheta_2}(x,y)}{\partial x^{\alpha_1}\partial y^{\alpha_2}}\right)^2 dxdy, \tag{2.41}$$

where $|\vartheta|$ denotes the number of '$-$' in $(\vartheta_1, \vartheta_2)$, for $\vartheta_1, \vartheta_2 = +, -$, using the previous ordered stencils, we obtain the same formula (2.40).

From this point the construction follows the same steps we saw in the previous sections. At first we define

$$\alpha_k^{--} = \frac{1}{(\beta_k^{--} + \sigma_h)^2},$$

with $\sigma_h = \sigma(h_x^2 + h_y^2)$, $\sigma > 0$, then we take the informations given by the reference polynomial computing

$$\omega^{--} = \frac{\alpha_0^{--}}{\alpha_0^{--} + \alpha_1^{--}},$$

which will measure the regularity of the function in the cell $(x_{j-1}, x_j] \times (y_{i-1}, y_i]$. Once we have computed in the same way the other three indicators, we can finally define

$$\omega = \min\{\omega^{--}, \omega^{+-}, \omega^{-+}, \omega^{++}\}, \tag{2.42}$$

which has properties similar to its one-dimensional counterpart (which we have not proven yet). In fact, for example, we can use in the same way the mapping (2.24) to reduce the oscillations in regular regions.

Notice that also the 'WENO-Z' procedure can be directly generalized to the 2D-case. It is sufficient to substitute the superscript '$\pm$' with '$\vartheta_1\vartheta_2$' in (2.25)-(2.44), that is

$$\tau^{\vartheta_1\vartheta_2} := \left|\beta_0^{\vartheta_1\vartheta_2} - \beta_1^{\vartheta_1\vartheta_2}\right|, \qquad \text{for } \vartheta_1, \vartheta_2 = +, -, \tag{2.43}$$

$$\alpha_k^{Z,\vartheta_1\vartheta_2} = \frac{1}{2}\left(1 + \left(\frac{\tau^{\vartheta_1\vartheta_2}}{\beta_k^{\vartheta_1\vartheta_2} + \sigma_h}\right)^p\right), \qquad \omega^{Z,\vartheta_1\vartheta_2} = \frac{\alpha_0^{Z,\vartheta_1\vartheta_2}}{\alpha_0^{Z,\vartheta_1\vartheta_2} + \alpha_1^{Z,\vartheta_1\vartheta_2}}. \tag{2.44}$$

Finally, analogously to the one-dimensional case, if we want to define a function ϕ such that

$$\phi_{i,j} = \begin{cases} 1 & \text{if the function } f \text{ is regular in } I_{i,j}, \\ 0 & \text{if } I_{i,j} \text{ contains a point of singularity,} \end{cases} \tag{2.45}$$

with $I_{i,j} = [x_{j-1}, x_{j+1}] \times [y_{i-1}, y_{i+1}]$, it is enough to apply one of the functions ϕ defined in Section 2.1.1.

***Remark* 2.13.** For completeness of presentation, we give also the explicit formula for the indicators as defined in Remark 2.10, that is

$$\begin{aligned}\beta_k^{P,--} &= \sum_{|\alpha|=2} \int_{-\Delta x}^{0} \int_{-\Delta y}^{0} \Delta x^{2(\alpha_1-1)} \Delta y^{2(\alpha_2-1)} \left(\frac{\partial^\alpha P_k^{--}(x,y)}{\partial x^{\alpha_1} \partial y^{\alpha_2}}\right)^2 dxdy \\ &= \frac{1}{\Delta x \Delta y}\left[f_{[3,1]}{}^2 + f_{[1,3]}{}^2 + f_{[2,2]}{}^2 + \frac{5}{12}\left(f_{[3,2]}{}^2 + f_{[2,3]}{}^2\right) + \frac{17}{720} f_{[3,3]}{}^2 + f_{[3,1]} f_{[3,2]} + \right. \\ &\qquad \left. + f_{[1,3]} f_{[2,3]} - \frac{1}{6}\left(f_{[3,1]} f_{[3,3]} + f_{[1,3]} f_{[3,3]}\right) - \frac{1}{12}\left(f_{[3,2]} f_{[3,3]} + f_{[2,3]} f_{[3,3]}\right)\right]\end{aligned} \tag{2.46}$$

which is evidently very similar to (2.40). As can be observed from the numerical tests in Sec. 2.3.2, the indicators constructed using (2.46) seem to be able to correctly locate a singularity in the domain

$$\text{int}\,(I_{i,j}) \bigcup \{(x_{j-1}, y_{i-1}), (x_{j+1}, y_{i-1}), (x_{j+1}, y_{i+1}), (x_{j-1}, y_{i+1})\}.$$

This reminds of the behavior of the indicators in the one-dimensional case, which we recall do not detect singularities at the edge points. Although it is still not clear to us which one is the most natural definition of regularity around a point (x_j, y_i) of the grid, following the discussion in Remark 2.11, we choose the definition given by (2.37) as the correct 2D-extension of the smoothness indicators. Moreover, when we applied the 2D-Adaptive Filtered Scheme, defined in Sec. 3.4, to the problem of front propagation (see Example 6 in Section 3.5), the formula (2.37) gave more stable results, helping reducing the oscillations in the approximated solution. This and similar considerations, eventually, led us to use (2.37) in the definition of the 2D extension of our scheme.

It is important to point out that we could have chosen many different ways to extend the indicators in two dimensions, some of which would have been even less expensive. For example, we could have constructed the polynomials on rectangular stencils overlapping on a whole half of $I_{i,j}$, requiring less points and computations (there are only three polynomials and four integrals), thus losing the symmetry in both directions. Another possibility would have been to consider triangular stencils and compute again four indicators, thus requiring just six points to construct each polynomial (which total degree would have been at most 2). Although this last case seems plausible and it is clearly less expensive than our choice, we do not believe that it would be effective, or at least as effective as the one we defined, for the detection of discontinuities in the gradient (it would be interesting to find out if it is true).

2.3. Numerical tests

In this section we will present some numerical tests on the smoothness indicators we have defined in Chapter 2. In the first part we will focus on the one-dimensional case, trying on one hand to confirm the thesis of Proposition 2.1 and on the other to show the improvements we can get by using the mappings of [HAP05] or higher order indicators. Then, we will pass to the multidimensional case, aiming to show the same good properties, even if they are not proven true yet.

2.3.1. One-dimensional examples

In the following examples we will test the properties of the one-dimensional smoothness indicators and, at least in the case of functions with discontinuities on the first derivative, we will consider both the cases in which the singularities fall on points of the grid or they do not. For all the tests we use the discontinuous function ϕ (2.29) with constant $M = 0.15$.

Test 1. Linear function with more than one singularity

As a first example, let us consider the function

$$f(x) = \begin{cases} \frac{1}{2} - |x| & \text{if } -1 \leq x \leq 1 \\ 0 & \text{otherwise} , \end{cases}$$

which is a linear function with three points of singularity, located at $x = -0.5$, 0 and 0.5. We will consider both the cases in which the singularities fall on a grid point or inside a cell. More precisely, in the latter we consider a situation in which the singularities do not fall at the center of a cell, rather inside its "left half". In this first test we just compare the results obtained with the indicators β_k^{CL} and β_k^{HJ} with $r = 2$ in their standard definition, without adding any mapping or modification.

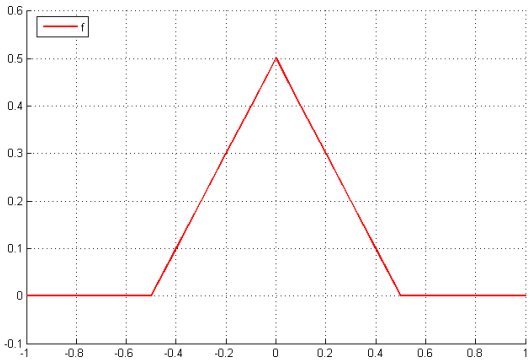

Figure 2.2.: Test 1. Linear function with singularities in $x = -0.5$, 0 and 0.5.

We start testing the indicators in the case in which the singularities fall on grid points.

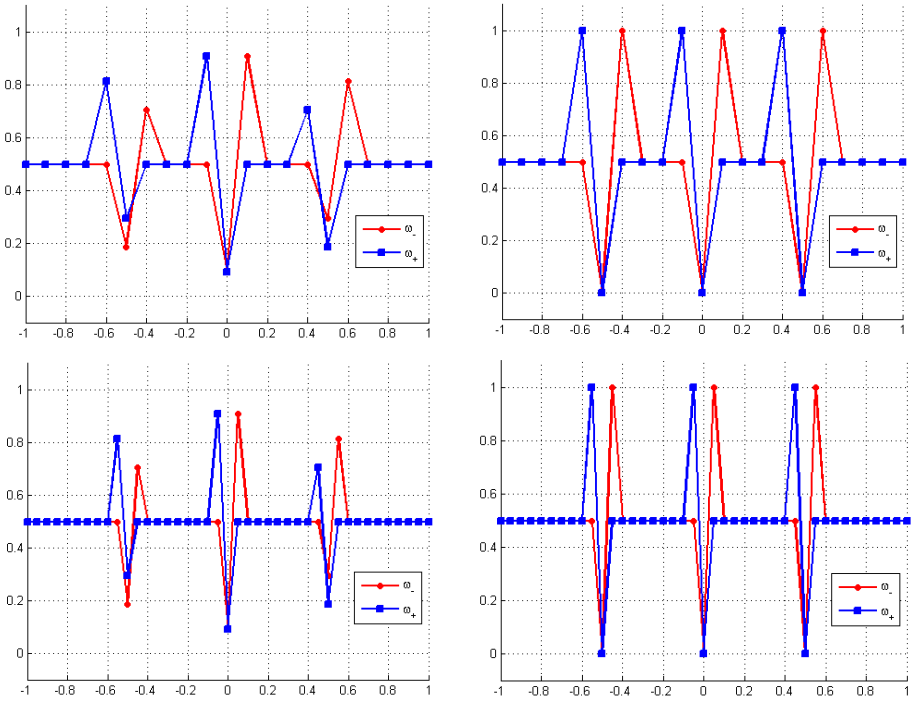

Figure 2.3.: Test 1. Singularities on grid points. Results obtained using β^{CL} (on the left) and β^{HJ} (on the right) with $r = 2$ for $\Delta x = 0.1$ and $\Delta x = 0.05$.

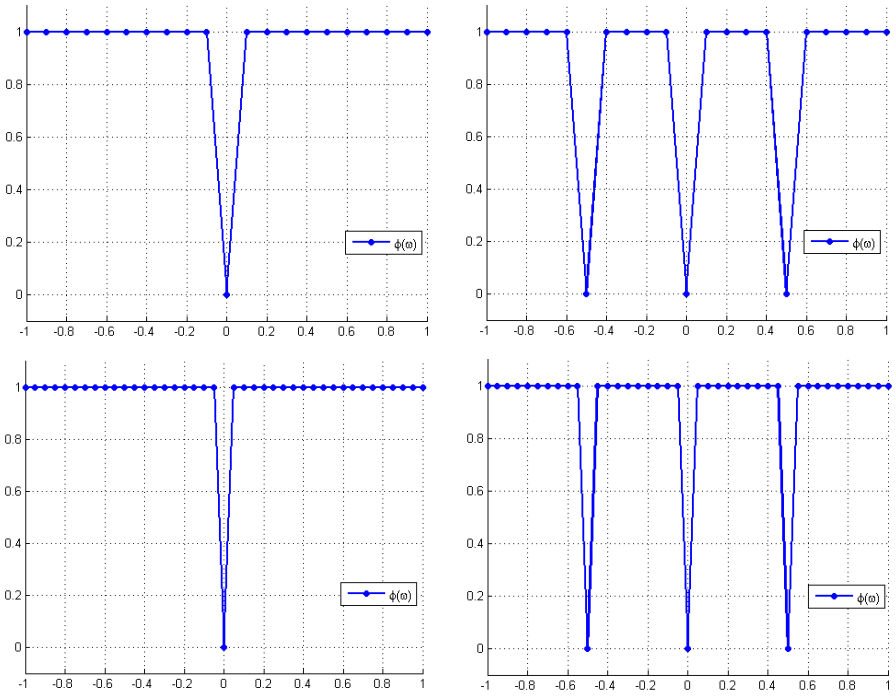

Figure 2.4.: Test 1. Singularities on grid points. Results obtained using β^{CL} (on the left) and β^{HJ} (on the right) with $r = 2$ for $\Delta x = 0.1$ and $\Delta x = 0.05$.

From Figure 2.3, we can clearly see that if there is a singularity in the stencil, then $\omega_{\pm}^{CL} = O(1)$, with the constant depending on the jump of the derivative, while $\omega_{\pm}^{HJ} \to 0$ as $\Delta x \to 0$, as we expected. Notice that if the singularity falls on a grid points both $\omega_{\pm}^{HJ}$ vanish on such points. Consequently, as Figure 2.4 shows, the smoothness indicator function ϕ is able to detect all the singularities only using β^{HJ}, while on the other case it detects only the singularity in the center (by coincidence).

Then, we test the case in which the singularities fall inside an interval $C_j = (x_{j-1}, x_j)$. In particular, we test the case in which they are always located closer to the left node. This second test case is even more emblematic, in fact in Figure 2.5 we can see the good behavior of the smoothness indicators obtained using β^{HJ}. We have that if the singularity is inside $\overset{\circ}{\mathcal{S}}$, then one of the two $\omega_{\pm}^{HJ} = O(h^4)$ and the other is equal to a constant. On the other hand we have that $\omega_{\pm}^{CL} = O(1)$ close to a singularity. As a consequence, the ϕ^{CL} function is not able to detect any node close to a singularity, while ϕ^{HJ} selects precisely all the cells which contain a point of singularity (Figure 2.6).

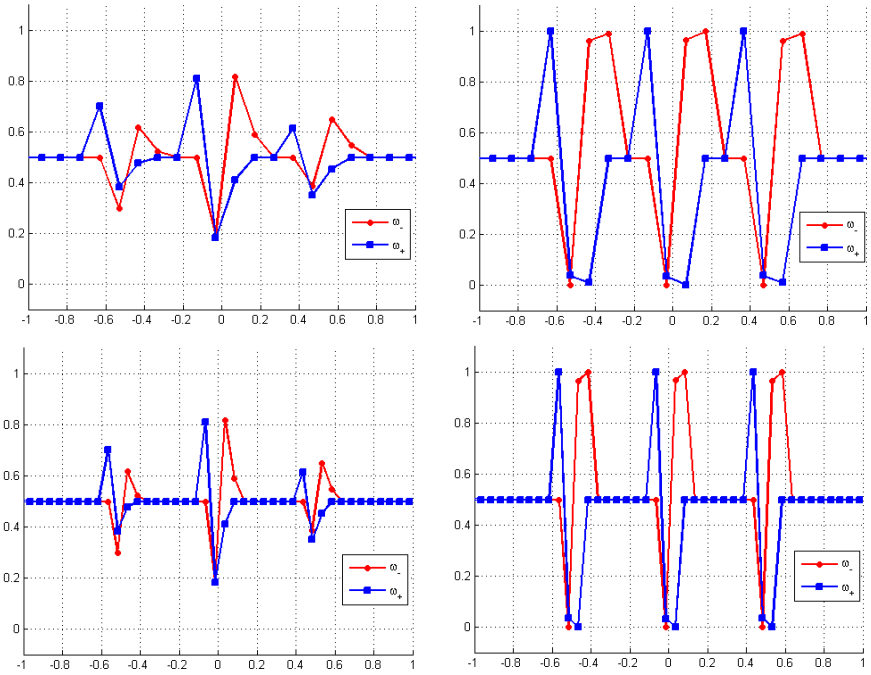

Figure 2.5.: Test 1. Singularities not on grid points. Results obtained using β^{CL} (on the left) and β^{HJ} (on the right) with $r = 2$ for $\Delta x = 0.1$ and $\Delta x = 0.05$.

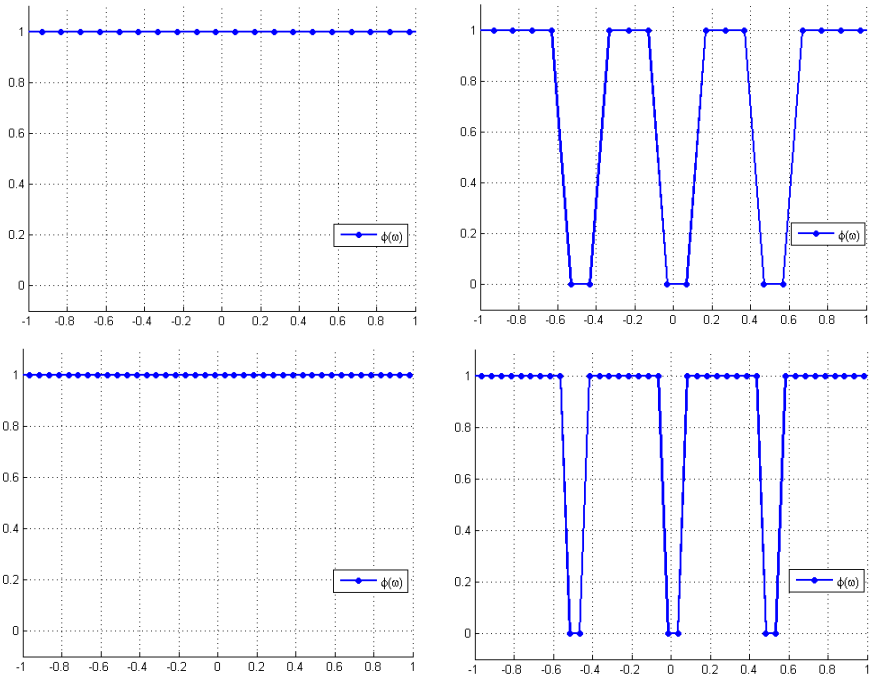

Figure 2.6.: Test 1. Singularities not on grid points. Results obtained using β^{CL} (on the left) and β^{HJ} (on the right) with $r = 2$ for $\Delta x = 0.1$ and $\Delta x = 0.05$.

Test 2. Nonlinear function with singularity

In order to check the behavior of the indicators in both singular and regular regions at the same time, in the second test we consider the function

$$f(x) = \begin{cases} \frac{(1-|x|)^2}{1.2} & \text{if } -1 \leq x \leq 1 \\ 0 & \text{otherwise} , \end{cases}$$

which has one point of singularity located at $x = 0$. Notice that this function is $C^1(\mathbb{R}) \setminus \{0\}$, but the second derivative is discontinuous in $x = -1$ and $x = 1$, at the edge of its support. This will expose a limit of the smoothness indicators β^{HJ}, which we try to suppress introducing some modifications. We will consider both the cases in which the singularity falls on a grid point or inside a cell. For this test at first we compare again the results obtained with the indicators β_k^{CL} and β_k^{HJ} with $r = 2$ in their standard definition, then we present some test obtained adding two modifications, the mapping (2.24) for β^{HJ} with $r = 2$ and the indicators (2.34)-(2.35) using β^{HJ} with $r = 3$.

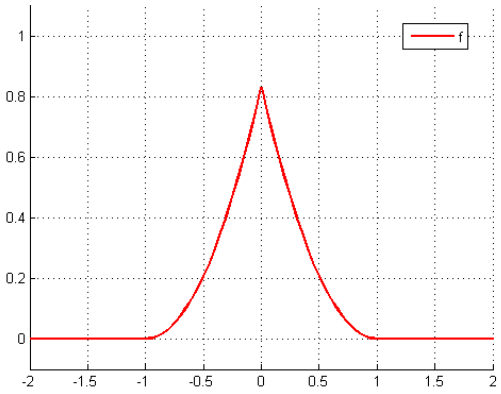

Figure 2.7.: Test 2. Nonlinear function with a point of singularity in $x = 0$.

We begin by testing the indicators when the singularity falls on a grid point. As Figure 2.8 shows, both indicators are able to detect the singularity in the origin, but using β^{HJ} we have that $\omega_\pm = O(1)$ in $x = -1$ and $x = 1$. This is the problem we mentioned previously, the presence of a discontinuity in the second derivative (which causes a "jump" between $\alpha_0^\pm$ and $\alpha_1^\pm$, respectively) affects the results by causing the formation of fluctuations around the optimal value $\frac{1}{2}$, which do not vanish as $\Delta x \to 0$.

For this test, although the two indicators present some problems, the ϕ function is able to detect the singularity in the origin and recognize the regularity on the rest of domain as we can see from Figure 2.9. Notice that these results are rather fortunate, in fact the "random" $O(1)$ terms (in the origin for β^{CL} and at the support boundary for β^{HJ}) do not cause any real problem. In the worst case scenario, as the next test will expose, we could have not detected the singularity in the first case and obtained some false negatives in $x = -1$ and $x = 1$ for the latter. Notice that the first indicator has a better behavior in regularity regions, in particular it is not affected by the discontinuity in the second derivative.

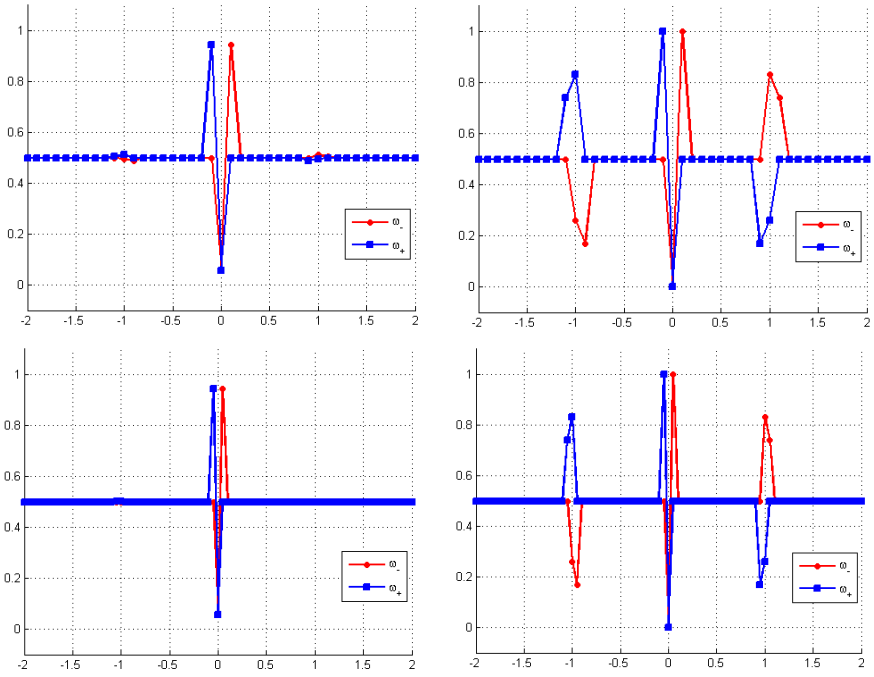

Figure 2.8.: Test 2. Singularity on grid point. Results obtained using β^{CL} (on the left) and β^{HJ} (on the right) with $r = 2$ for $\Delta x = 0.1$ and $\Delta x = 0.05$.

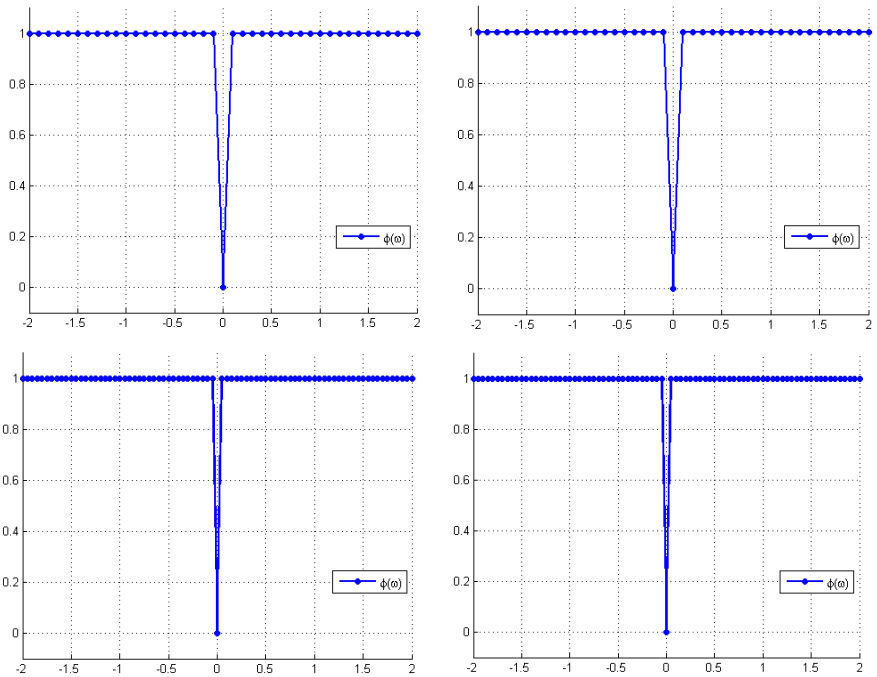

Figure 2.9.: Test 2. Singularity on grid point. Results obtained using β^{CL} (on the left) and β^{HJ} (on the right) with $r = 2$ for $\Delta x = 0.1$ and $\Delta x = 0.05$.

Then, let us repeat the test in the case in which the singularity falls inside an interval $C_j = (x_{j-1}, x_j)$, again in the situation in which it is located closer to the left node. This simulation comprehends all the situations we analyzed previously, in fact in Figure 2.10 we can notice on one hand the good behavior of ω^{CL} in regularity regions and the random jump close to the origin (which is way more evident in the left node of the singular cell), while ω^{HJ} is able to isolate perfectly the nodes of the singular cell, although presents oscillations around $x = -1$ and $x = 1$, even wider than the previous test case.

Consequently, the results given by our ϕ function are affected by the problems we just described. The indicators on the left of Figure 2.11 are able to detect just the left node of the singular cell (again by coincidence), while for those on the right we notice some false negatives located at the edge of the support, although the singular cell is precisely detected.

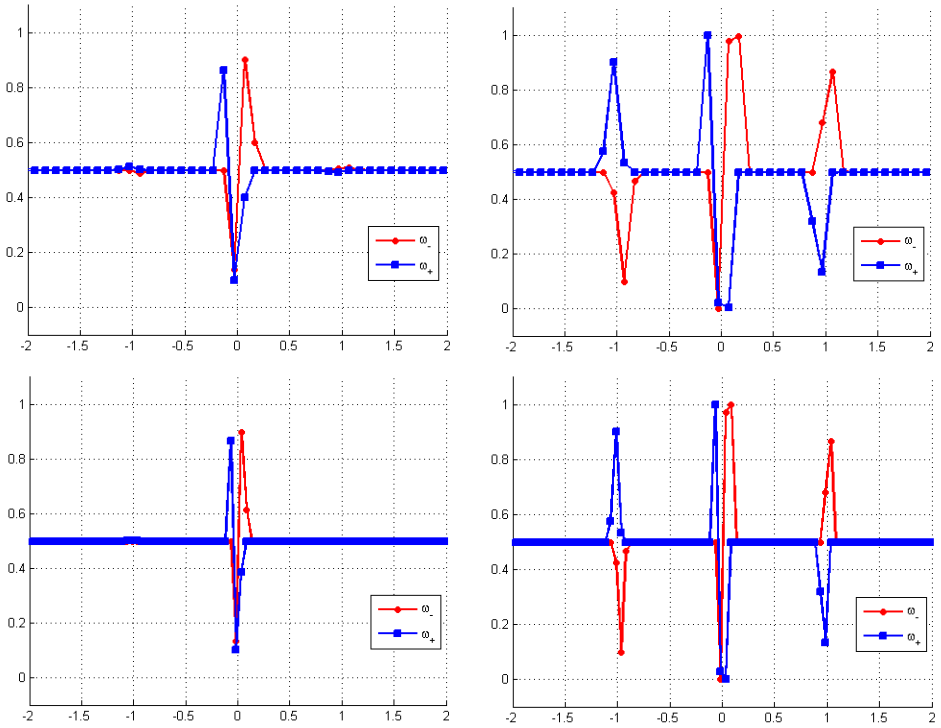

Figure 2.10.: Test 2. Singularity not on grid point. Results obtained using β^{CL} (on the left) and β^{HJ} (on the right) with $r = 2$ for $\Delta x = 0.1$ and $\Delta x = 0.05$.

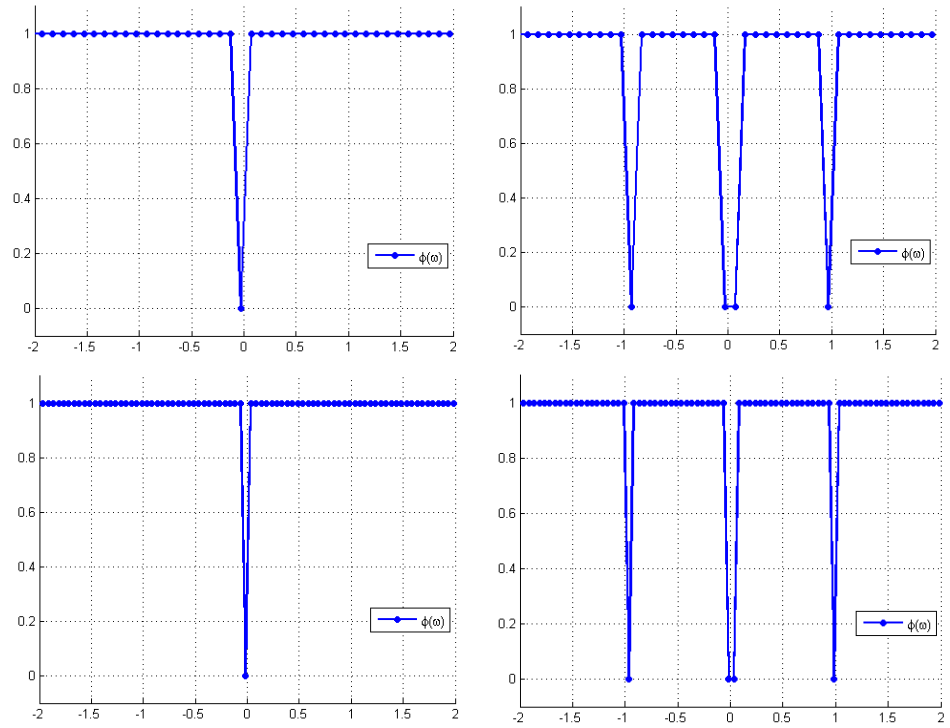

Figure 2.11.: Test 2. Singularity not on grid point. Results obtained using β^{CL} (on the left) and β^{HJ} (on the right) with $r = 2$ for $\Delta x = 0.1$ and $\Delta x = 0.05$.

We conclude this second test case by trying to solve, at least partially, the problems just mentioned. We are interested in modifying only the indicators ω^{HJ}, because, as it is already pretty clear, ω^{CL} are not well suited for this kind of functions. Then we repeat the test adding the mapping (2.24) to ω^{HJ} with $r = 2$ and consider also the case of ω^{HJ} with $r = 3$, for both the previous test cases but only for the first refinement of the grid, which is enough to show the improvements. As Figures 2.12 and 2.13 clearly show, the two modifications are able to reduce the random oscillations around $x = -1$ and $x = 1$, then improving the results given by the ϕ function. It is good to point out that, as we can see from the right column, even increasing the order of the reconstructions, the discontinuities in the second derivative still affects the results of the indicator ω^{HJ} (the oscillations again do not vanish as $\Delta x \to 0$, but for brevity we avoided to show the second refinement of the grid).

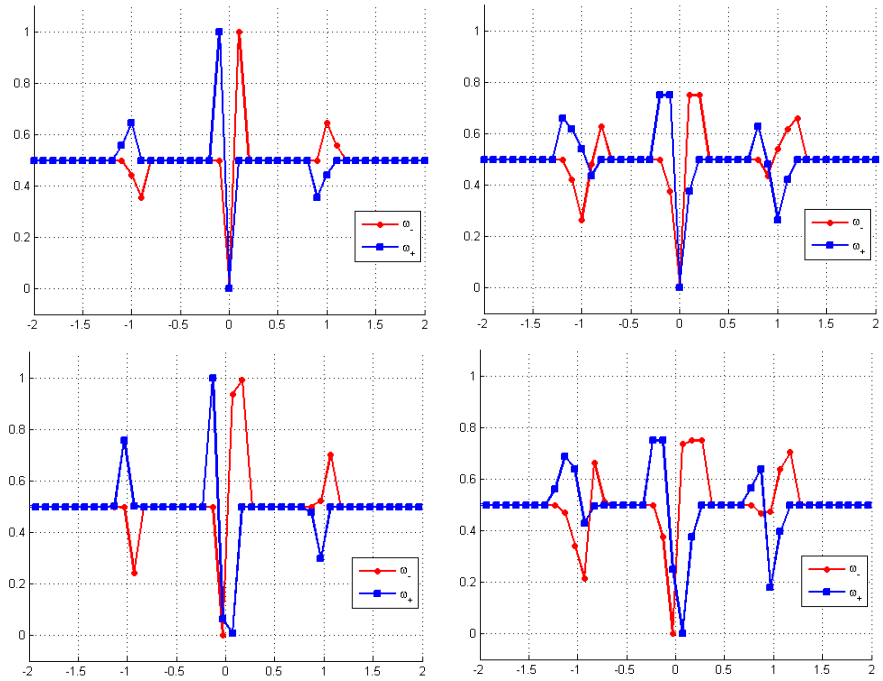

Figure 2.12.: Test 2. Some modifications. Results obtained using β^{HJ} with $r = 2$ and the the mapping (2.24) (left) and β^{HJ} with $r = 3$ (right), in the case the singularity falls on a grid point (top) or does not (bottom) with $\Delta x = 0.1$.

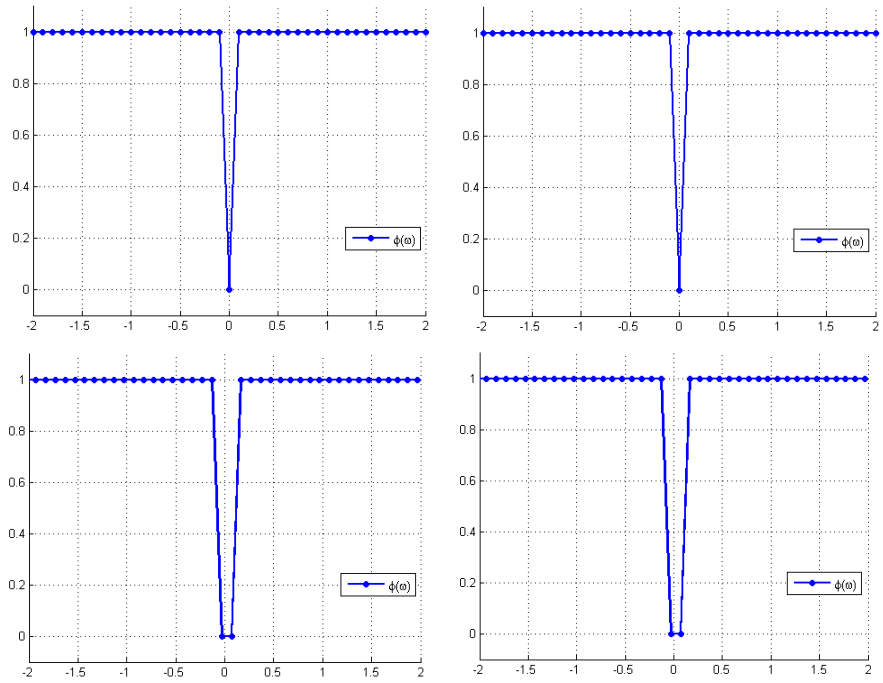

Figure 2.13.: Test 2. Some modifications. Results obtained using β^{HJ} with $r = 2$ and the the mapping (2.24) (left) and β^{HJ} with $r = 3$ (right), in the case the singularity falls on a grid point (top) or does not (bottom) with $\Delta x = 0.1$.

Test 3. Smooth function

In the last test we focus on the behavior of the indicators in regions of regularity, so we consider the function

$$f(x) = \begin{cases} (1-x^2)^4 & \text{if } -1 \leq x \leq 1 \\ 0 & \text{otherwise ,} \end{cases}$$

which is at least $C^2(\mathbb{R})$, but it has several points in which the concavity changes. Also in this test we begin comparing the results obtained with the indicators β_k^{CL} and β_k^{HJ} with $r = 2$ in their standard definition, then, to reduce the oscillations suffered by the latter, we add the same modification we introduced in the previous case.

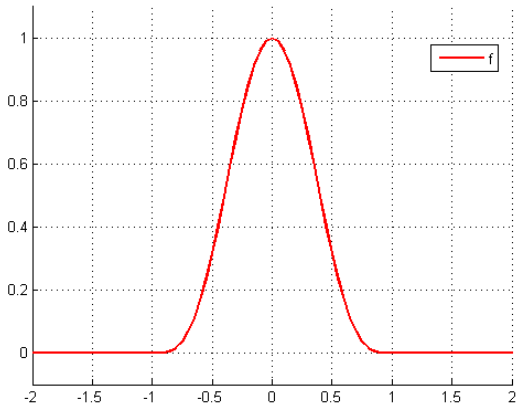

Figure 2.14.: Test 3. Smooth function.

The difference between the two indicators is evident in Figure 2.15, in fact $\omega_{\pm}^{HJ}$ present wide oscillations around the optimal value, due to the $O(h)$ term in (2.20) and the changes in the magnitude of the second derivative, which vanish very slowly. On the other hand, $\omega_{\pm}^{CL}$ presents way smaller oscillations which completely vanish in one mesh refinement. Then, as shown in Figure 2.16, the function ϕ recognizes the whole regularity of the f just using ω^{CL}, while on the other case it presents several false negatives, even with the second refinement. This behavior is not surprising since, as highlighted in Remark 2.3, the indicators for conservation laws have better accuracy in regions of regularity when using the same degree r. This is because the CL indicator uses more information on the function f, especially those coming from the integral of the first derivative of the interpolating polynomials, which can not be used when dealing with singular functions (in the derivative).

We can try to suppress the oscillations and at same time increase the convergence of $\omega_{\pm}^{HJ}$ towards the optimal value $\frac{1}{2}$ by adding the mapping (2.24) or by increasing the degree of the polynomials, as shown in Figures 2.17 and 2.18. Using the modifications we are able to obtain good results for $\phi(\omega^{HJ})$, in fact although some false negatives are still visible in the first choice of the grid for both indicators, we get the desired behavior in just one mesh refinement.

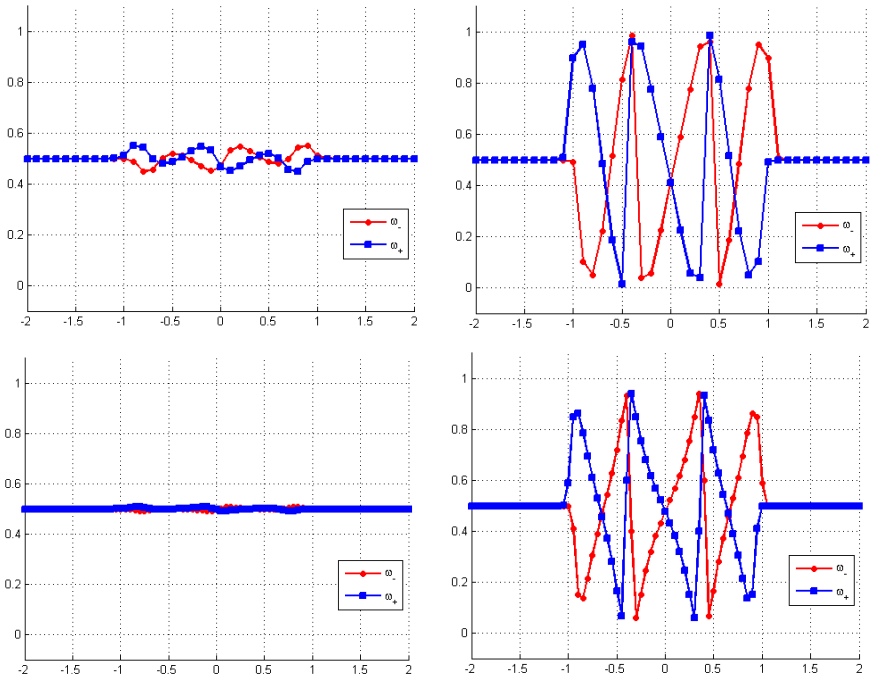

Figure 2.15.: Test 3. Results obtained using β^{CL} (on the left) and β^{HJ} (on the right) with $r = 2$ for $\Delta x = 0.1$ and $\Delta x = 0.05$.

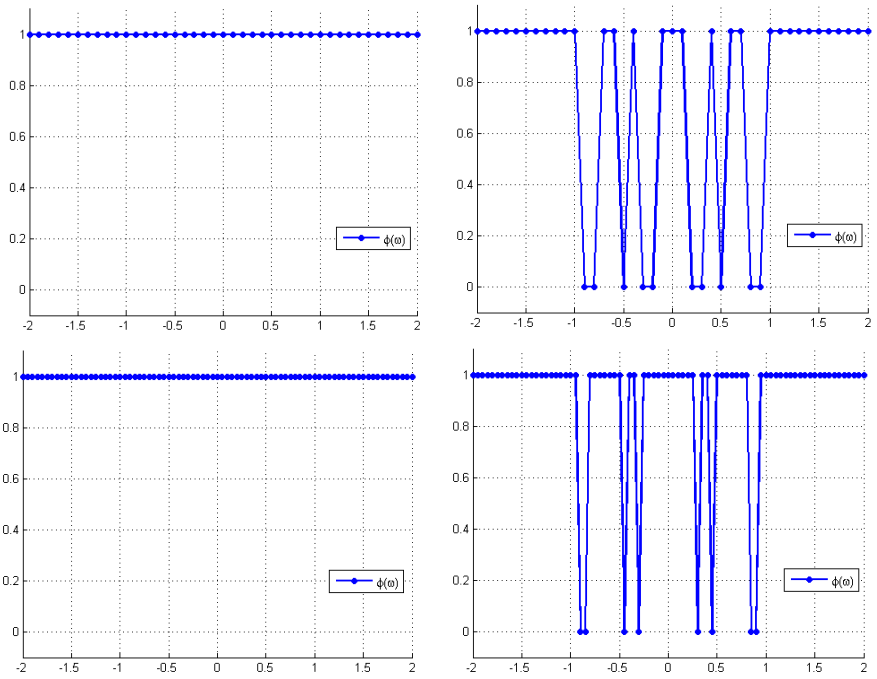

Figure 2.16.: Test 3. Results obtained using β^{CL} (on the left) and β^{HJ} (on the right) with $r = 2$ for $\Delta x = 0.1$ and $\Delta x = 0.05$.

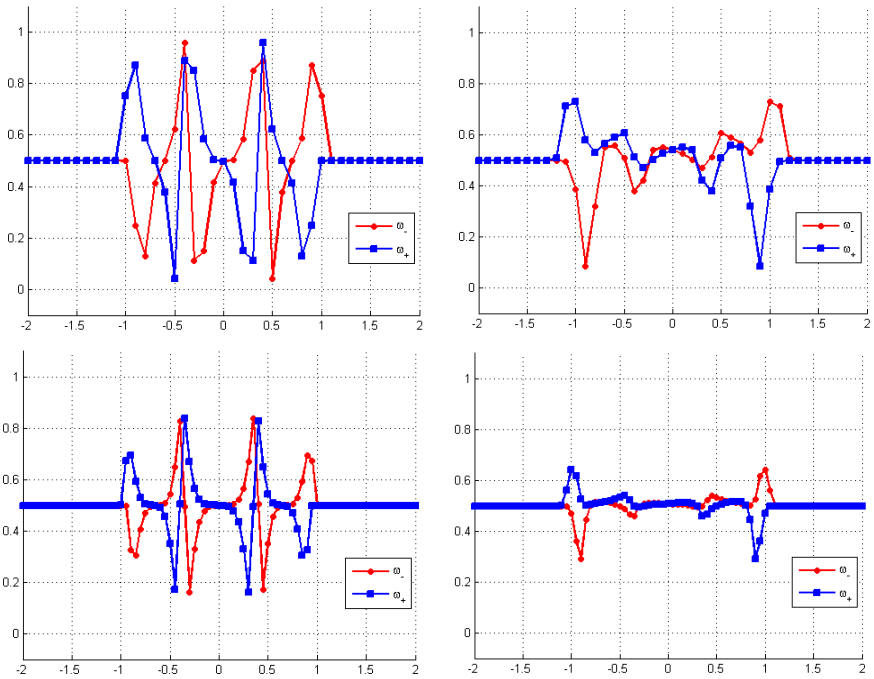

Figure 2.17.: Test 3. Results obtained using β^{HJ} with $r = 2$ and the the mapping (2.24) (on the left) and β^{HJ} with $r = 3$ (on the right), for $\Delta x = 0.1$ and $\Delta x = 0.05$.

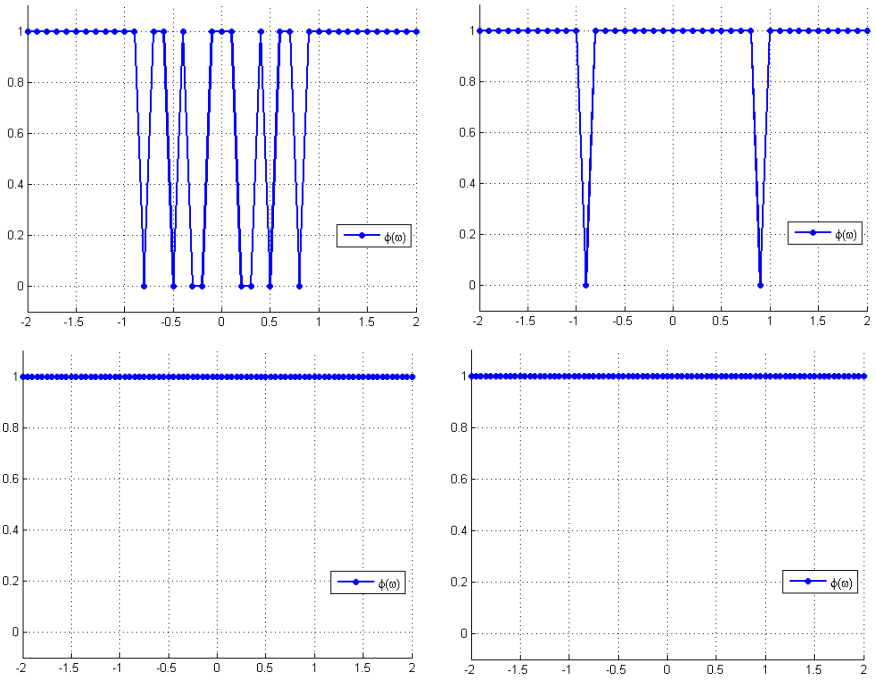

Figure 2.18.: Test 3. Results obtained using β^{HJ} with $r = 2$ and the the mapping (2.24) (on the left) and β^{HJ} with $r = 3$ (on the right), for $\Delta x = 0.1$ and $\Delta x = 0.05$.

2.3.2. Two-dimensional examples

The aim of the section is to investigate the behavior of the 2D smoothness indicators defined in Section 2.2 in various situations, as for the one-dimensional case. In fact, all the following examples are direct generalizations of those of the previous section. One of the main topic here is to compare the results obtained using the two very similar formulas (2.40) and (2.46), trying to show the reasons of our preference in the construction of the 2D-Adaptive Filtered Scheme in Section 3.4. For better clarity, we denote the two indicators using the symbols, respectively, β_{2D}^F (full) and β_{2D}^P (partial). Moreover, in order to reduce the number of redundant simulations, for both constructions we consider only the case $r = 2$ and add directly the mapping (2.24) to reduce the oscillations.

It is noteworthy that an easy and direct $2D$-extension of the smoothness indicator can be obtained by dimensional splitting, that is,

$$\omega_{split} = \min\{\omega_x, \omega_y\}, \tag{2.47}$$

where ω_x and ω_y are the 1D smoothness indicators in each direction, computed fixing each time the other variable. Notice that for the "genuine" 2D indicator (2.40) we use polynomials in $\mathbb{Q}_2(\mathbb{R}^2)$, while for the *splitting indicator* only couples of polynomials in $\mathbb{P}_2(\mathbb{R})$). By extensive numerical simulations, which we do not report in the sequel, but will be most probably present in [FPTb], we observed that the latter approach, although very simple and fast, has clearly some drawbacks in terms of reliability with respect to β_{2D}, especially because of the oscillations around the optimal value in regions of regularity and the problems in localizing singularities which do not fall on grid points.

For all the tests we use the discontinuous function ϕ (2.29) with constant $M = 0.2$. Notice that we slightly increased the value of M with respect to the one-dimensional case, that is because the indicators seem to give smoother results in regular regions, having smaller oscillations around the optimal value 0.5.

Test 1. Function with singularities in one point and on a curve

Let us begin these two-dimensional examples by considering the analogous of the first test of the previous section,

$$f(x,y) = \begin{cases} 1 - \sqrt{x^2+y^2} & \text{if } x^2+y^2 \le 1 \\ 0 & \text{otherwise ,} \end{cases}$$

which clearly has a point of singularity located in the origin and a singularity circle at the base of the cone. We consider both the cases in which the point of singularity in the origin falls on a grid point or inside a cell (again not in the center but closer to the south/west node).

First we analyze the case of the central singularity falling on a grid point, using the contour plots to show precisely the behavior of the indicators with respect to the location of the singularities. In order to highlight this fact, we plot also the points of singularity, in green in Figure 2.20 and in red in Figure 2.21.

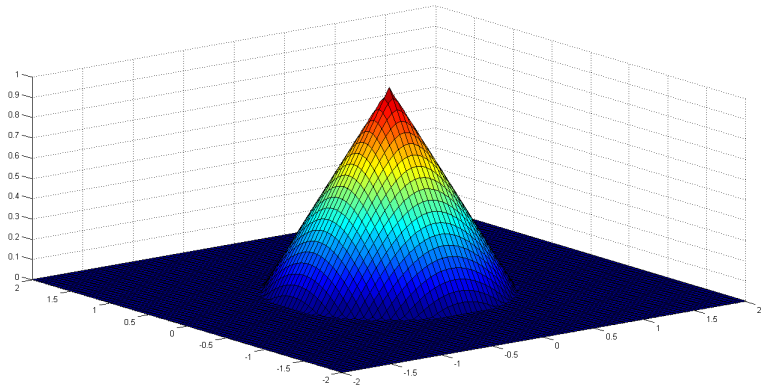

Figure 2.19.: Test 1. Function with singularities in a point and on a circle.

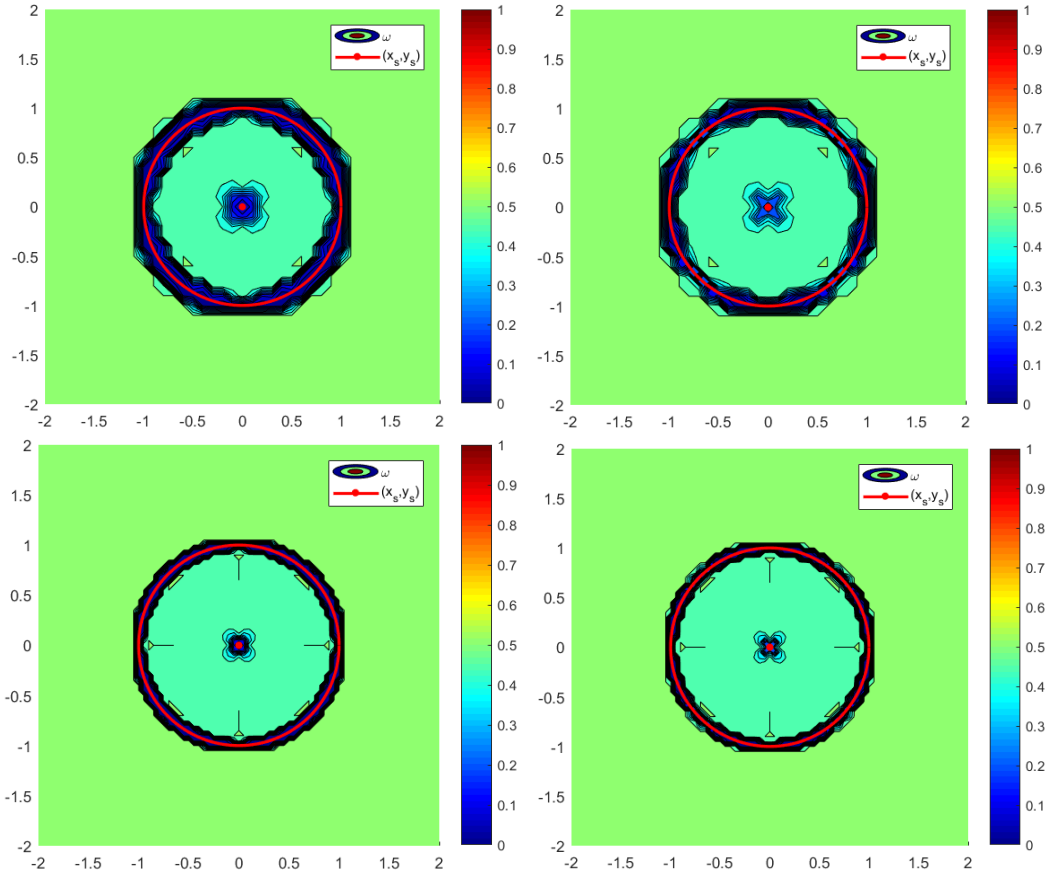

Figure 2.20.: Test 1. Singularity on a grid point. Results obtained using β_{2D}^{F} (on the left) and β_{2D}^{P} (on the right), with $r = 2$ and the mapping (2.24), for $\Delta x = \Delta y = 0.1$ and $\Delta x = \Delta y = 0.05$.

From Figure 2.20 we can see that both indicators seem to detect the right regions of singularity and have a very good behavior in regular regions (in light green), although

ω_{2D}^F is evidently more precise. Looking at the neighborhood of the origin we notice that ω_{2D}^P has wide fluctuations only in the diagonal direction, whereas ω_{2D}^F has a more uniform behavior, also around the singularities on the circle.

Consequently, looking at Figure 2.21 in which we have highlighted only the 0.1-level and the 1-level, we have that ϕ_{2D}^F recognizes with extreme precision all the cells and the grid nodes containing a singularity, while ϕ_{2D}^P seems to miss some points of the circle and localizes the singularity in the center, consisting in this case of just one point. This behavior is rather typical and will be investigated throughout all the simulations. The full indicator ω_{2D}^F is able to detect singular cells even when the singularity just barely intersects the considered region, whereas ω_{2D}^P recognizes only "strong" singularities, that are close to a grid point or situated around the center of the considered cell. Moreover, looking at the behavior of ϕ_{2D}^P around the singularity in the center, which enlarges the detected singular region in the second refinement, we are led to believe that the results of ϕ_{2D}^F are more precise and, evidently, more stable w.r.t. mesh refinements.

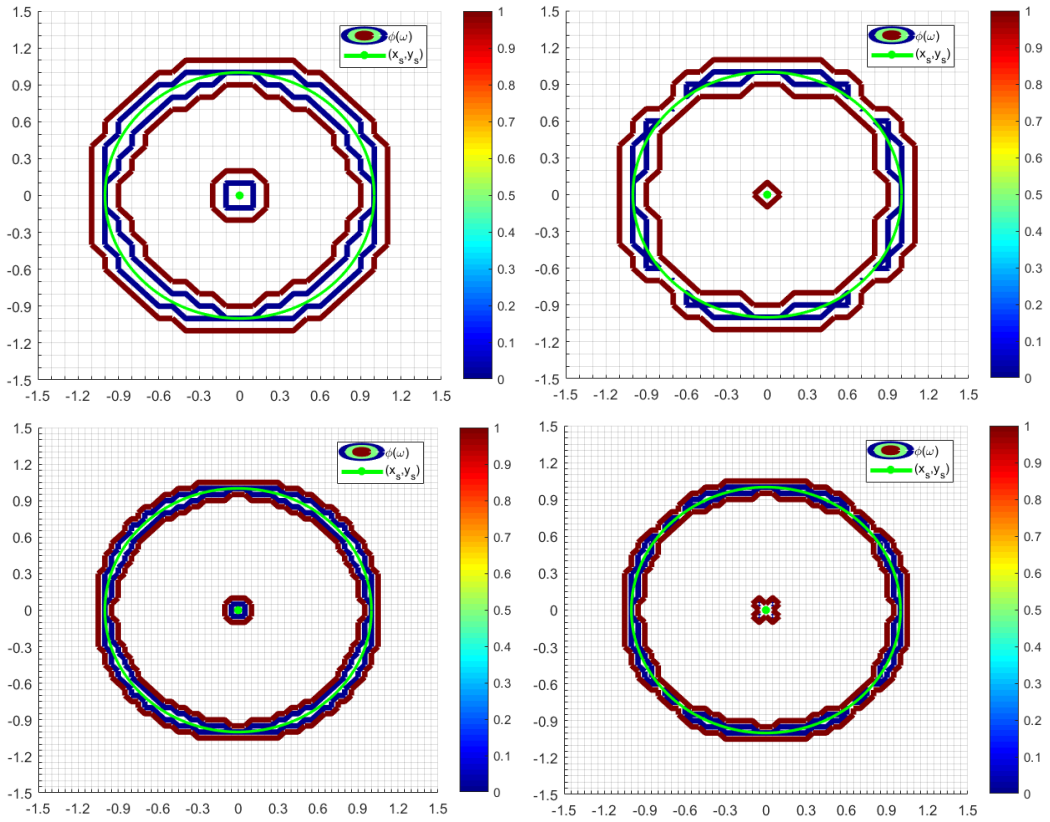

Figure 2.21.: Test 1. Singularity on a grid point. Results obtained using β_{2D}^F (on the left) and β_{2D}^P (on the right), with $r = 2$ and the mapping (2.24), for $\Delta x = \Delta y = 0.1$ and $\Delta x = \Delta y = 0.05$.

Then, we repeat the test using a grid staggered w.r.t the singularity in the origin.

Figures 2.22 and 2.23 confirm the impressions given by the previous simulation, in fact the indicator ω_{2D}^F and the function ϕ_{2D}^F are able to recognize all the cells containing

a singularity, in particular those around the circle, which is always inside the 0-level set of ϕ^F_{2D} (Fig. 2.23 on the left), while ω^P_{2D} has a rather asymmetrical behavior on the circle. The portion of the circle in the "South-West" direction is well detected, while on the other three directions the detected regions degenerate into points (Fig 2.23 on the right). Notice that in this case is the full indicator that has a particular behavior around the origin in the second refinement, spreading the detected singular region in the direction of the singularity.

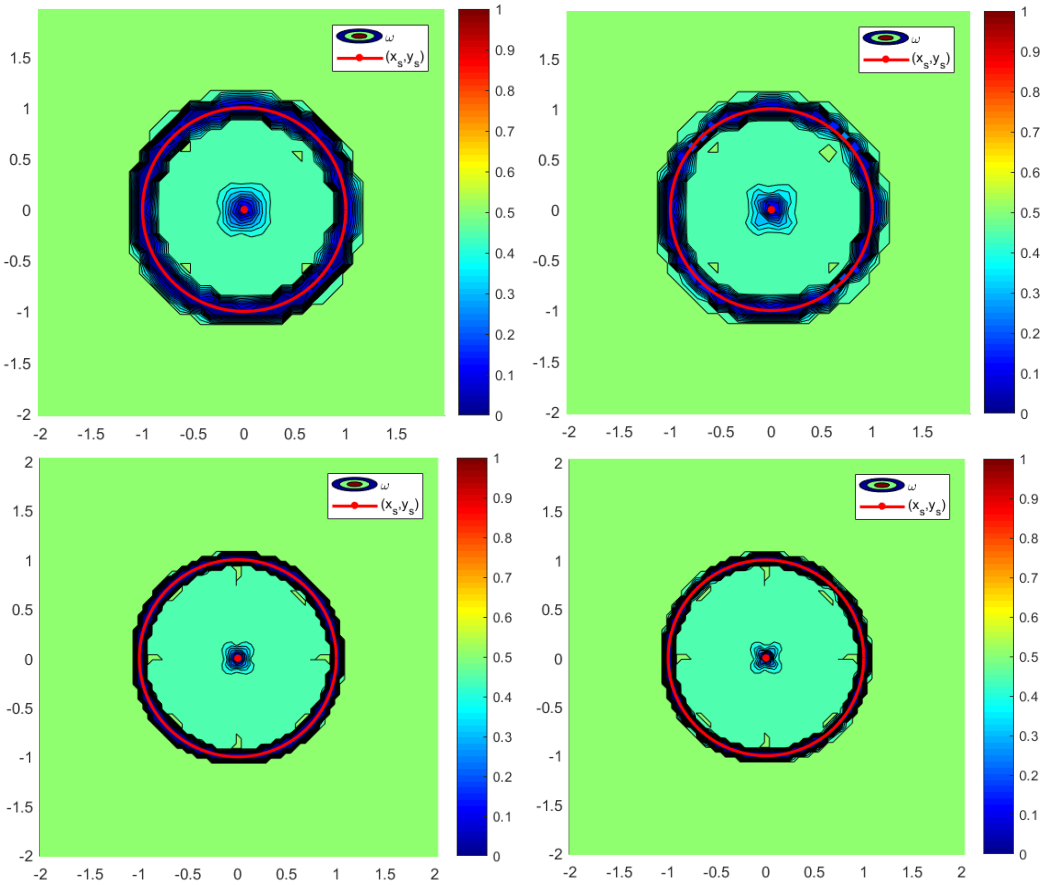

Figure 2.22.: Test 1. Singularity inside a cell. Results obtained using β^F_{2D} (on the left) and β^P_{2D} (on the right), with $r = 2$ and the mapping (2.24), for $\Delta x = \Delta y = 0.1$ and $\Delta x = \Delta y = 0.05$.

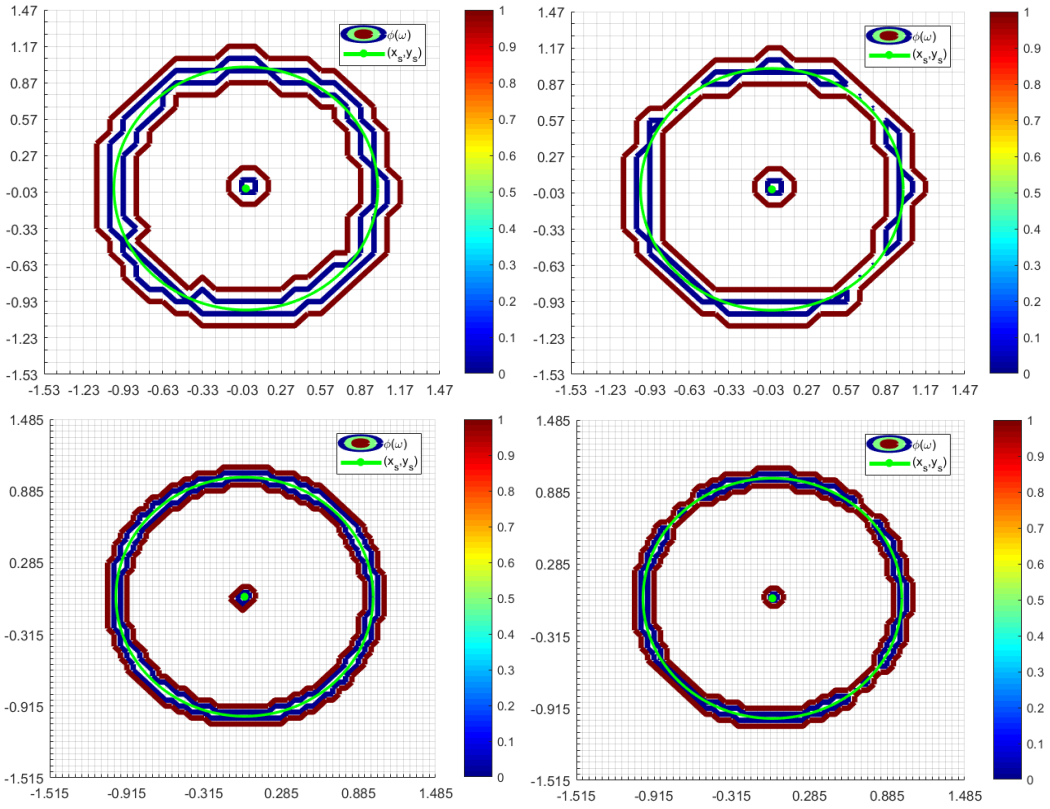

Figure 2.23.: Test 1. Singularity on a grid point. Results obtained using β^F_{2D} (on the left) and β^P_{2D} (on the right), with $r = 2$ and the mapping (2.24), for $\Delta x = \Delta y = 0.1$ and $\Delta x = \Delta y = 0.05$.

Test 2. Nonlinear function with singularity in one point

For the second test we consider a nonlinear function with a singularity in the origin (again analogous to the 1D-case),

$$f(x,y) = \begin{cases} \left(1 - \sqrt{x^2+y^2}\right)^2 & \text{if } \sqrt{x^2+y^2} \leq 1 \\ 0 & \text{otherwise .} \end{cases}$$

We repeat the tests as in the previous example, first with the point of singularity on a grid node, then the staggered grid.

Figures 2.25 and 2.26 (in which we have zoomed the neighborhood of the singularity) follow the same line of the first test, at least regarding the point of singularity in the origin. In fact, we can observe that ω^F_{2D} has again a more uniform behavior with respect to ω^P_{2D} in all directions, whereas the results on the regular regions are practically the same.

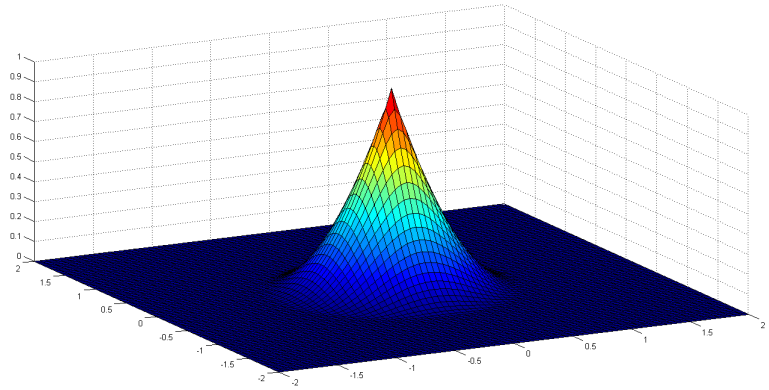

Figure 2.24.: Test 2. Function with a singularity in the origin.

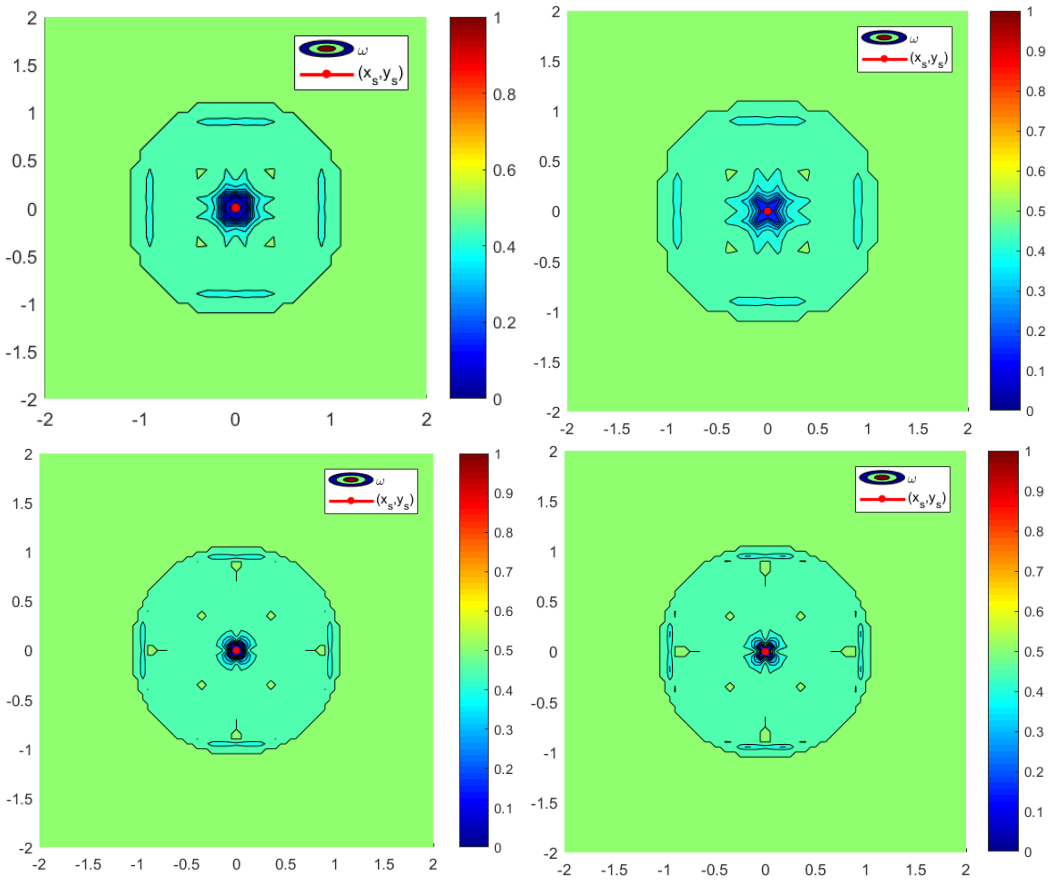

Figure 2.25.: Test 2. Singularity on a grid point. Results obtained using β_{2D}^{F} (on the left) and β_{2D}^{P} (on the right), with $r = 2$ and the mapping (2.24), for $\Delta x = \Delta y = 0.1$ and $\Delta x = \Delta y = 0.05$.

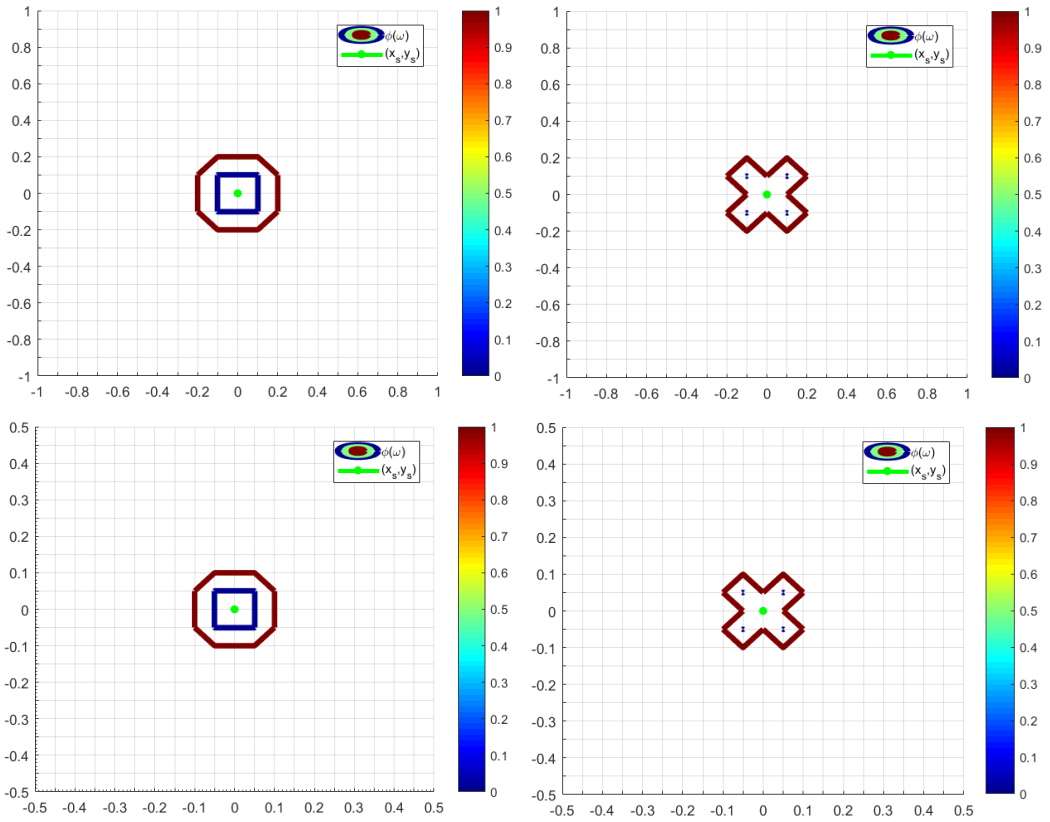

Figure 2.26.: Test 2. Singularity on a grid point. Results obtained using β^F_{2D} (on the left) and β^P_{2D} (on the right), with $r = 2$ and the mapping (2.24), for $\Delta x = \Delta y = 0.1$ and $\Delta x = \Delta y = 0.05$.

In the case of the staggered grid, in Figures 2.27 and 2.28, we can observe that, as in the previous example, ω^P_{2D} localizes better the singular regions, while ω^F_{2D} again spreads asymmetrically the region of singularity, although in the correct direction given by the position of the singularity. This is in fact natural, since from a numerical point of view, if a grid point is very close to a singularity, then at that point the function behaves similarly. Consequently, we can consider the function singular even at such points.

Summarizing the previous observations, we can deduce that the formula (2.46) is more suited to localize singularities with high precision, since the corresponding indicator is able to select the correct grid points or cells characterized by a strong discontinuity in the gradient. On the other hand, using the formula (2.40), we are able to detect the regularity of the function in the whole domain $I_{i,j} = [x_{j-1}, x_{j+1}] \times [y_{i-1}, y_{i+1}]$, boundary included. It is clear then, since when developing genuinely 2D-numerical schemes we usually need at least a nine-point stencil $\mathcal{S}_{i,j} = \{x_{j-1}, x_j, x_{j+1}\} \times \{y_{i-1}, y_i, y_{i+1}\}$, that the correct indicators for the construction of our Adaptive Filtered Scheme should be based on (2.40), instead of (2.46). Moreover, when working with second order schemes in 2D, as the ones we will define in Sec. 3.4, in order to verify the high-order consistency property using Taylor expansion, we have to require the regularity of the function in the whole domain $I_{i,j}$.

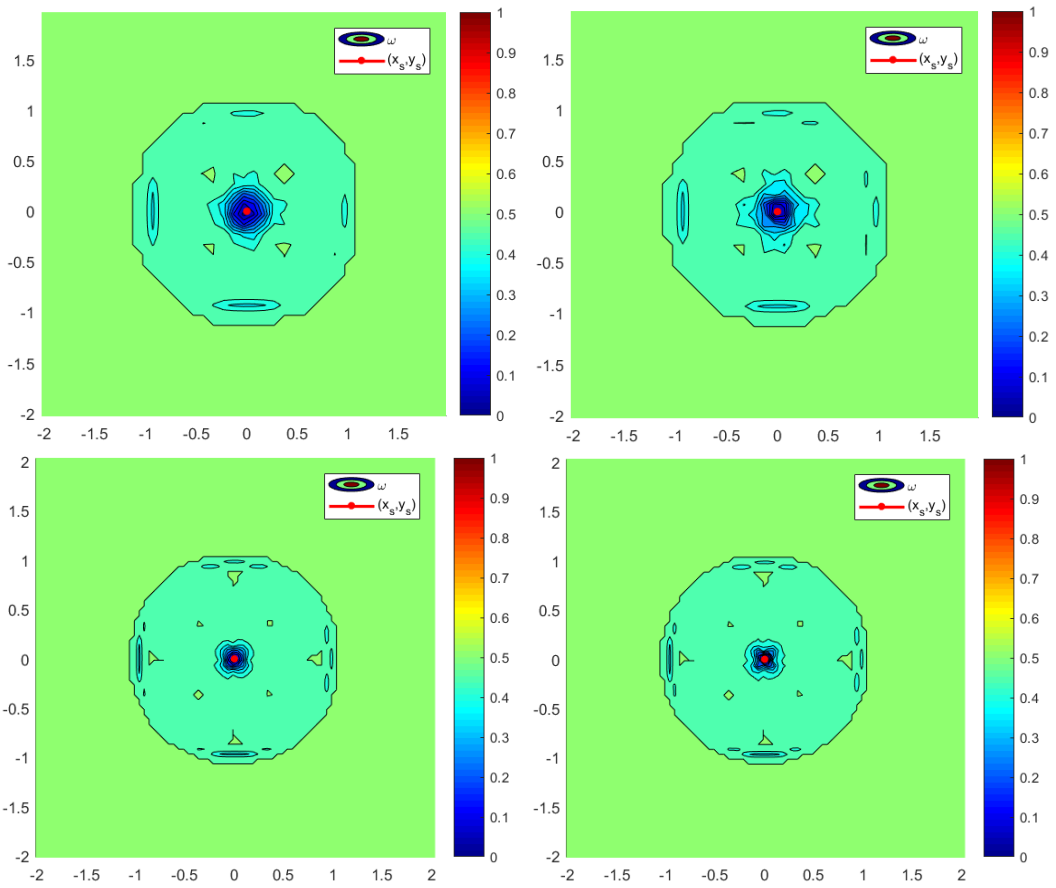

Figure 2.27.: Test 2. Singularity inside a cell. Results obtained using β^F_{2D} (on the left) and β^P_{2D} (on the right), with $r = 2$ and the mapping (2.24), for $\Delta x = \Delta y = 0.1$ and $\Delta x = \Delta y = 0.05$.

Test 3. Smooth function

For completeness of presentation, we conclude this section with the analysis of the behavior in regions of regularity, whence we consider the smooth function

$$f(x,y) = \begin{cases} \left(1-(x^2+y^2)\right)^4 & \text{if } x^2+y^2 \leq 1 \\ 0 & \text{otherwise .} \end{cases}$$

As we can see from Figure 2.30, both indicators give the desired responses, recognizing the regularity of the function in the whole domain. Moreover, we can observe that the results are very similar, presenting the exact same (small) oscillations around the optimal value, with minor differences only at the boundary of the support. We avoid to present also the results for ϕ since it is clear that it is a constant function ($\phi \equiv 1$) in all four cases.

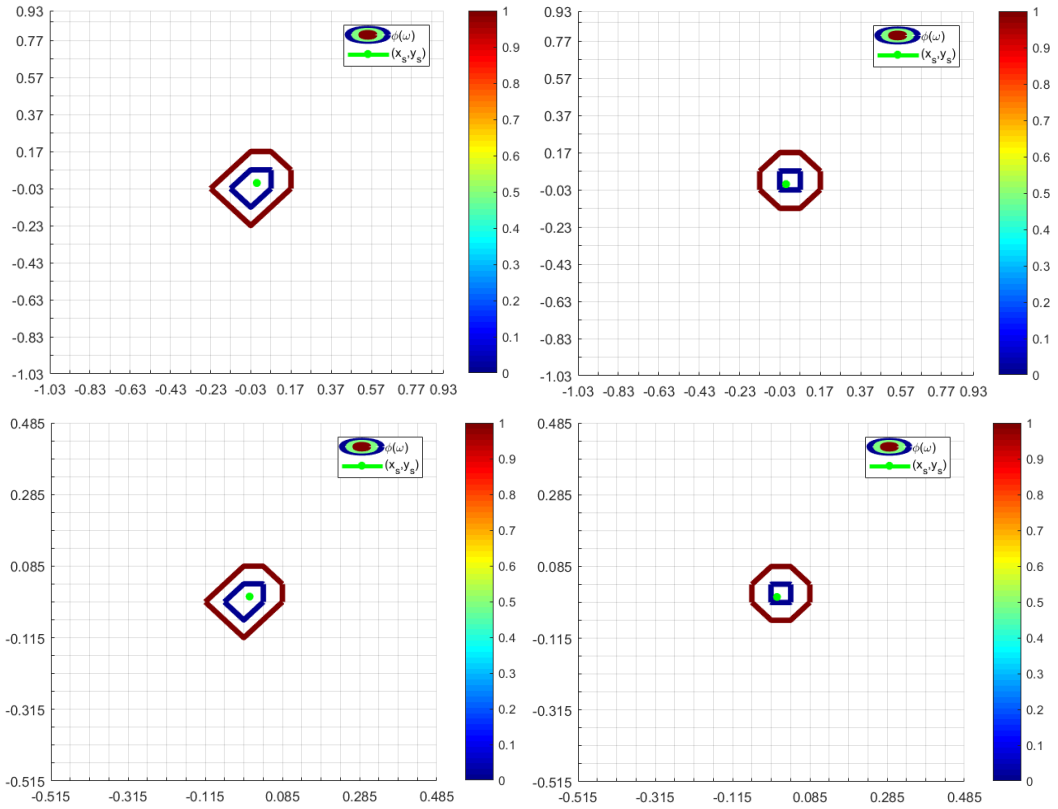

Figure 2.28.: Test 2. Singularity inside a cell. Results obtained using β_{2D}^{F} (on the left) and β_{2D}^{P} (on the right), with $r = 2$ and the mapping (2.24), for $\Delta x = \Delta y = 0.1$ and $\Delta x = \Delta y = 0.05$.

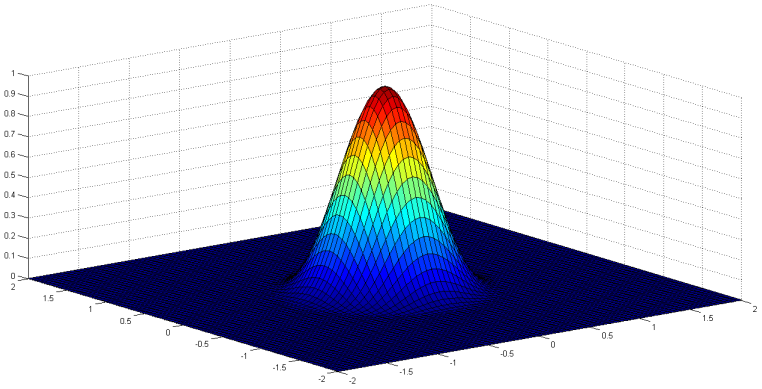

Figure 2.29.: Test 3. Smooth function.

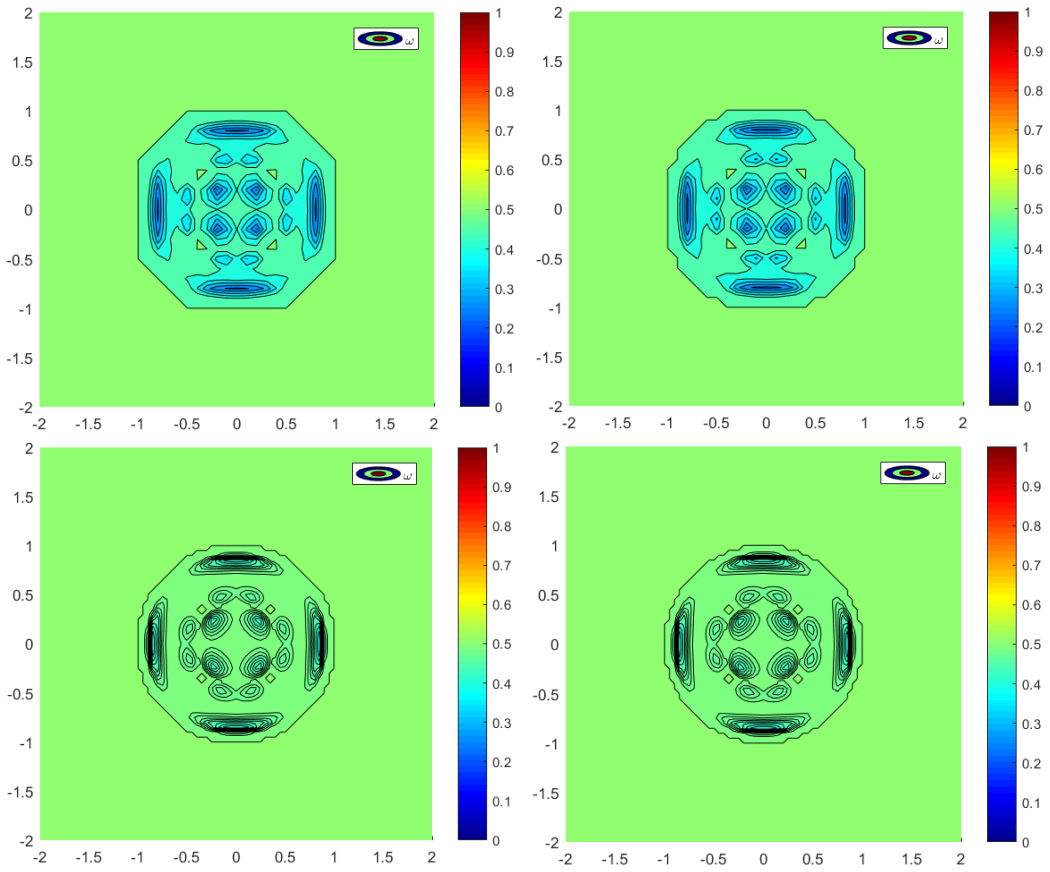

Figure 2.30.: Test 3. Accuracy in regions of regularity. Results obtained using β^F_{2D} (on the left) and β^P_{2D} (on the right), with $r = 2$ and the mapping (2.24), for $\Delta x = \Delta y = 0.1$ and $\Delta x = \Delta y = 0.05$.

2.4. Conclusions

We have presented a detailed analysis of the smoothness indicators of [JP00], showing that they are the correct definition to use when working with continuous function with discontinuous derivative. Consequently, we showed a slightly (but fundamentally) different application in order to analyze the regularity of a function on a symmetric interval, presenting also various possible constructions. In the second part, we proposed a new genuinely 2D-extension of the studied approach, giving also a very compact explicit formula for the computation. Moreover, we have tested several different one-dimensional smoothness indicators, trying to show the properties proved in Proposition 2.1. As we expected, we observed that the indicators β^{HJ} are the right indicators to localize discontinuities in the first derivative, although they present some problems in regularity regions due to direct dependence on the value (and regularity) of the second derivative. We were able to overcome these limitation, at least partially, adding the mapping 2.24 and increasing the order of the polynomials involved in the computation of the indi-

cators. We also acknowledged the better behavior of the indicators β^{CL} in regions of regularity, but also the fact that they are not able to localize the singularities in general situations. Finally, we have tested the proposed possible approaches to extend the one-dimensional ideas to the multidimensional case. We were able to infer the good behavior of both indicators which detect the singularity, localized in one point or on a curve, and recognize the regions of regularity, although the "full" indicator (2.37) has given more promising responses in terms of applicability to the adaptive filtering idea, developed in Chapter 3.

To summarize the above facts, when working with Hamilton-Jacobi equations, we advise the use of the β^{HJ}, with $r = 2$, in one space dimension and β^{F} in the two dimensional case, both augmented by the mappings (2.24). In this way it is possible to devise an efficient and reliable indicator using only five grid points (5×5 in 2D). Moreover, in order to avoid any unnecessary complication, we suggest the use of the discontinuous function ϕ with $M > 0$, a small positive constant, say $M = 0.15$. This is, in fact, the implementation that we will be using in all the remaining numerical tests of the thesis.

Future investigation will focus on the improvement of Lemma 2.2, in order to include also more useful definition of the indicators and to help the analysis of the adaptive filtering approach. Moreover, we will inspect the possible advantages brought by the use of the 'WENO-Z' procedure, which, as already noted in Remark 2.5, seems to give more possibilities with respect to the mapping of [HAP05]. Finally, the precise analysis of the multidimensional indicators, in the aim of Proposition 2.1, is another possible interesting development.

3. Adaptive Filtered Schemes

In this chapter we will show how to define a class of convergent schemes for first order Hamilton-Jacobi equations, the *Adaptive Filtered Schemes*. We will start by giving some hints on the construction of the schemes and then we will specify the hypotheses needed to ensure their properties. Consequently, we will prove our main result, concerning the consistency and convergence properties of our class of schemes. Finally, we will generalize the construction to more space dimensions, defining the schemes used in Chapter 4 for the problem of the segmentation of an image.

The chapter is organized as follows: in Section 3.2 we define our new filtered scheme and present in detail all its basic components, then in Section 3.3 we state our convergence result. Section 3.4 is focused on the multidimensional extension of the proposed constructions and, finally, in 3.5 we conclude the chapter by presenting some one- and two-dimensional numerical tests.

3.1. Introduction and first definitions

The accurate numerical solution of Hamilton-Jacobi (HJ) equations is a challenging topic of growing importance in many fields of application, e.g. control theory, KAM theory, image processing and material science. Due to the lack of regularity of viscosity solutions, this issue is delicate and the construction of high-order methods is usually rather difficult (e.g. ENO, WENO). In recent years a general approach to the construction of high-order methods using filters has been proposed by Lions and Souganidis in [LS95] for a class of implicit schemes and reinterpreted by Augoula and Abgrall ([AA00]) in the context of explicit schemes, leading to the first definition of ε-monotonicity. More recently, a general framework for ε-monotone schemes has been introduced by Froese and Obermann in [FO13], for stationary Hamilton-Jacobi equations, by Obermann and Salvador in [ObSa15], for particular second order equations and finally applied to time-dependent first order Hamilton-Jacobi equations by Bokanowski et al. in [BFS16].

Following the approach in [BFS16], we will present a procedure to define a class of "filtered" schemes for time-dependent Hamilton-Jacobi equation of the form

$$\begin{cases} v_t + H(\nabla v) = 0, & (t,x) \in [0,T] \times \mathbb{R}^N, \\ v(0,x) = v_0(x), & x \in \mathbb{R}^N, \end{cases} \tag{3.1}$$

where the hamiltonian H and the initial data v_0 are Lipschitz continuous functions. It is well known that with these assumptions we have an existence and uniqueness result for the viscosity solution (see [FF14] or [B98]). Notice that, at least for the moment, we are considering the most simple case with the hamiltonian depending only on the gradient of

the solution. This is in fact a rather usual approach, since the main numerical difficulties come from the nonlinear dependence of H on ∇v. Moreover, more general cases usually follow with minor modifications. Our aim here is to present a rather simple way to construct convergent schemes to the viscosity solution v of (3.3) with the property to be of high-order in the region of regularity.

A typical feature of a filtered scheme S^F is that at the node x_j it is a mixture of a high-order scheme S^A and a monotone scheme S^M according to a filter function F. The scheme is written as

$$u_j^{n+1} \equiv S^F(u^n)_j := S^M(u^n)_j + \varepsilon \Delta t F\left(\frac{S^A(u^n)_j - S^M(u^n)_j}{\varepsilon \Delta t}\right), \quad j \in \mathbb{Z}, \tag{3.2}$$

where $\varepsilon = \varepsilon_{\Delta t, \Delta x} > 0$ is a parameter going to 0 as $(\Delta t, \Delta x)$ is going to 0 and does not depend on n. Filtered schemes are high-order accurate where the solution is smooth, monotone otherwise, and this feature is crucial to prove a convergence result for viscosity solutions as in [BFS16].
In this chapter we improve the filtered scheme (3.2) introducing an adaptive and automatic choice of the parameter $\varepsilon = \varepsilon^n$ at every iteration. To this end, we use a smoothness indicator in order to select the regions where we can compute the regularity threshold ε^n. Our smoothness indicators are based on the ideas of Jiang and Peng [JP00], but other indicators with similar properties can be used.

3.2. A new Adaptive Filtered scheme

In this section we will complete the construction of the one-dimensional scheme, presenting in detail all its basic components. Let us consider the one-dimensional time-dependent Hamilton-Jacobi equation

$$\begin{cases} v_t + H(v_x) = 0, & (t,x) \in [0,T] \times \mathbb{R} \\ v(0,x) = v_0(x), & x \in \mathbb{R}, \end{cases} \tag{3.3}$$

where the hamiltonian H and the initial data v_0 are Lipschitz continuous functions.

Starting from the ideas of [BFS16] on *filtered schemes*, we proceed in this study introducing a procedure to compute the regularity threshold ε in an automatic way, in order to exploit the local regularity of the solution.

Let us begin defining a uniform grid in space $x_j = j\Delta x$, $j \in \mathbb{Z}$, and in time $t_n = t_0 + n\Delta t$, $n \in [0, N_T]$, with $(N_T - 1)\Delta t < T \leq N_T \Delta t$. Then, we compute the numerical approximation $u_j^n = u(t_n, x_j)$ with the simple formula

$$u_j^{n+1} = S^{AF}(u^n)_j := S^M(u^n)_j + \phi_j^n \varepsilon^n \Delta t F\left(\frac{S^A(u^n)_j - S^M(u^n)_j}{\varepsilon^n \Delta t}\right), \tag{3.4}$$

where $u_j^{n+1} := u(t_{n+1}, x_j)$, S^M and S^A are, respectively, the monotone and the high-order scheme, F is the *filter function* needed to switch between the two schemes, ε^n is the switching parameter at time t_n and ϕ_j^n is the *smoothness indicator function* at the

node x_j and time t_n, defined as shown in Section 2.1.1. More details on the components of the schemes will be given in the following sections.

Notice that if $\varepsilon^n \equiv \varepsilon\Delta x$, with $\varepsilon > 0$ and $\phi_j^n \equiv 1$, we get the Filtered Schemes of [BFS16].

3.2.1. Assumptions on the schemes

In this section we present in detail the basic components of our scheme, which are a monotone finite difference scheme S^M and a high-order scheme S^A, possibly unstable. Let us begin by giving the assumptions on the monotone scheme.

Assumptions on S^M.

(M1) The scheme can be written in *differenced form*

$$u_j^{n+1} \equiv S^M(u^n)_j := u_j^n - \Delta t\ h^M(D^- u_j^n, D^+ u_j^n)$$

for a function $h^M(p^-, p^+)$, with $D^\pm u_j^n := \pm\frac{u_{j\pm1}^n - u_j^n}{\Delta x}$;

(M2) h^M is a Lipschitz continuous function;

(M3) (Consistency) $\forall v$, $h^M(v, v) = H(v)$;

(M4) (Monotonicity) for any functions u, v,

$$u \le v \quad \Rightarrow \quad S^M(u) \le S^M(v).$$

Under assumption (M2), the consistency property (M3) is equivalent to say that for all functions $v \in C^2([0,T] \times \mathbb{R})$, there exists a constant $C_M \ge 0$ independent on $\Delta = (\Delta t, \Delta x)$ such that

$$\mathcal{E}_M(v)(t,x) := \left| \frac{v(t+\Delta t, x) - S^M(v(t,\cdot))(x)}{\Delta t} \right| \le C_M \left(\Delta t ||v_{tt}||_\infty + \Delta x ||v_{xx}||_\infty\right), \quad (3.5)$$

where $\mathcal{E}_M$ is the consistency error. The last relation clearly shows the bound on the accuracy of the monotone schemes, which are at most first order accurate even for regular solutions.

***Remark* 3.1.** As pointed out in [BFS16], under the Lipschitz assumption (M2) the monotonicity property (M4) can be restated in terms of some quantities that can be easily computed. In fact, it is enough to require, for a.e. $(p^-, p^+) \in \mathbb{R}^2$,

$$\frac{\partial h^M}{\partial p^-}(p^-, p^+) \ge 0, \qquad \frac{\partial h^M}{\partial p^+}(p^-, p^+) \le 0, \quad (3.6)$$

and the *CFL condition*

$$\frac{\Delta t}{\Delta x}\left(\frac{\partial h^M}{\partial p^-}(p^-, p^+) - \frac{\partial h^M}{\partial p^+}(p^-, p^+)\right) \le 1. \quad (3.7)$$

We call the *CFL number*, dependent on the hamiltonian of the considered problem, the constant ratio $\lambda := \frac{\Delta t}{\Delta x}$ such that (3.7) is satisfied. Notice that working with explicit finite difference schemes this number can always be computed.

Example 3.1. We give some examples of monotone schemes in differenced form which satisfy (M1)-(M4). Other examples may be found in the pioneering work [CL84] or in [Sh98].

- For the *eikonal equation*,
$$v_t + |v_x| = 0,$$
we can use the simple numerical hamiltonian
$$h^M(p^-, p^+) := \max\{p^-, -p^+\}. \tag{3.8}$$
- For general equations, instead, we recall the *Central Upwind scheme* of [KNP01]
$$h^M(p^-, p^+) := \frac{1}{a^+ - a^-}\left[a^- H(p^+) - a^+ H(p^-) - a^+ a^- (p^+ - p^-)\right], \tag{3.9}$$
with $a^+ = \max\{H_p(p^-), H_p(p^+), 0\}$ and $a^- = \min\{H_p(p^-), H_p(p^+), 0\}$.
- Another numerical hamiltonian we could use is the *Local Lax-Friedrichs hamiltonian*
$$h^M(p^-, p^+) := H\left(\frac{p^- + p^+}{2}\right) - \frac{\alpha(p^-, p^+)}{2}(p^+ - p^-), \tag{3.10}$$
with
$$\alpha(p^-, p^+) = \max_{p \in I(p^-, p^+)} |H_p(p)|,$$
where $I(a, b)$ represents the interval with endpoints a and b.The scheme is monotone under the restriction $\lambda < \max_p |H_p(p)|$.

Next, we define the requirements on the high-order scheme.

Assumptions on S^A.

(A1) The scheme can be written in *differenced form*
$$u_j^{n+1} = S^A(u^n)_j := u_j^n - \Delta t h^A(D^{k,-}u_j, \dots, D^- u_j^n, D^+ u_j^n, \dots, D^{k,+}u_j^n),$$
for some function $h^A(p^-, p^+)$ (in short), with $D^{k,\pm}u_j^n := \pm\frac{u_{j\pm k}^n - u_j^n}{k\Delta x}$;

(A2) h^A is a Lipschitz continuous function;

(A3) (High-order consistency) Fix $k \geq 2$ order of the scheme, then for all $l = 1, \dots, k$ and for all functions $v \in C^{l+1}$, there exists a constant $C_{A,l} \geq 0$ such that
$$\begin{aligned}\mathcal{E}_A(v)(t,x) &:= \left|\frac{v(t+\Delta t, x) - S^A(v(t,\cdot))(x)}{\Delta t}\right| \\ &\leq C_{A,l}\left(\Delta t^l ||\partial_t^{l+1} v||_\infty + \Delta x^l ||\partial_x^{l+1} v||_\infty\right).\end{aligned}$$

It is interesting to notice that we are not making any assumption on the stability of the high-order scheme, that is because filtered schemes are able to stabilize a possibly unstable scheme.

Before giving some examples of high-order schemes satisfying (A1)-(A3), let us state an interesting property of the solution v of (3.3) in case of sufficient regularity. Notice that we are considering the simplest case of H dependent only on the gradient of v.

Lemma 3.2. *Let v be the solution of (3.3). Then, if $v \in C^r\left(\Omega_{(t,x)}\right)$, $r \geq 2$, where $\Omega_{(t,x)}$ is a neighborhood of a point $(t,x) \in \Omega := [0,T] \times \mathbb{R}$, it holds*

$$\begin{aligned} \frac{\partial^k v(t,x)}{\partial t^k} &= (-1)^k \frac{\partial^{k-2}}{\partial x^{k-2}} \left(H_p^k(v_x(t,x)) v_{xx}(t,x) \right) \\ &= (-1)^k \frac{\partial^{k-2}}{\partial x^{k-2}} \left(H_p^{k-1}(v_x(t,x)) \frac{\partial}{\partial x} H(v_x(t,x)) \right), \end{aligned} \tag{3.11}$$

for $k = 2, \ldots, r$.

Proof. Let us proceed by induction on $2 \leq k \leq r$, omitting the dependence on (t,x) to simplify the notation. For $k = 2$, we have

$$v_{tt} = \frac{\partial}{\partial t}(-H(v_x)) = -H_p(v_x) v_{xt} = -H_p(v_x) \frac{\partial}{\partial x}(-H(v_x)) = H_p^2(v_x) v_{xx},$$

and the statement holds in this case. Suppose now that (3.11) holds for $2 < k < r-1$, then we can compute

$$\begin{aligned} \frac{\partial^{k+1} v}{\partial t^{k+1}} &= \frac{\partial}{\partial t} \left(\frac{\partial^k v}{\partial t^k} \right) \\ &= \frac{\partial}{\partial t} \left((-1)^k \frac{\partial^{k-2}}{\partial x^{k-2}} \left(H_p^k(v_x)) v_{xx} \right) \right) && \text{by inductive hypothesis} \\ &= (-1)^k \frac{\partial^{k-2}}{\partial x^{k-2}} \left(\frac{\partial}{\partial t} \left(H_p^k(v_x)) v_{xx} \right) \right) \\ &= (-1)^k \frac{\partial^{k-2}}{\partial x^{k-2}} \left(\frac{\partial}{\partial p} \left(H_p^k(v_x) \right) v_{xt} v_{xx} + H_p^k(v_x) v_{xxt} \right) \\ &= (-1)^k \frac{\partial^{k-2}}{\partial x^{k-2}} \left(\frac{\partial}{\partial x} (H_p^k(v_x)) v_{xt} + H_p^k(v_x) \frac{\partial}{\partial x}(v_{xt}) \right) \\ &= (-1)^k \frac{\partial^{k-1}}{\partial x^{k-1}} \left(H_p^k(v_x) v_{tx} \right) \\ &= (-1)^{k+1} \frac{\partial^{k-1}}{\partial x^{k-1}} \left(H_p^{k+1}(v_x) v_{xx} \right), \end{aligned}$$

as we wanted. □

Let us now consider the value of the solution at $v(t + \Delta t, x)$, with $\Delta t > 0$ and its Taylor expansion of order $r \geq 2$ around the point (t,x). Using Lemma 3.2, we can

rewrite

$$\begin{aligned} v(t+\Delta t,x) &= v(t,x) + \Delta t v_t(t,x) + \sum_{k=2}^{r} \frac{\Delta t^k}{k!} \frac{\partial^k v(t,x)}{\partial t^k} + O(\Delta t^{r+1}) \\ &= v(t,x) - \Delta t H(v_x(t,x)) + \\ &\quad + \sum_{k=2}^{r} \frac{(-\Delta t)^k}{k!} \frac{\partial^{k-2}}{\partial x^{k-2}} \left(H_p^k(v_x(t,x)) v_{xx}(t,x) \right) + O(\Delta t^{r+1}), \end{aligned} \tag{3.12}$$

which for $r = 2$ simply reads

$$v(t+\Delta t,x) = v(t,x) - \Delta t H(v_x(t,x)) + \frac{\Delta t^2}{2} H_p^2(v_x(t,x)) v_{xx}(t,x) + O(\Delta t^3). \tag{3.13}$$

***Remark* 3.2.** Using this last relation we could show that, assuming (A1)-(A2), the consistency property is equivalent to require that for $l = 2, \dots, k$, and for all $v \in C^{l+1}$,

$$\begin{aligned} \mathcal{E}_A(v)(t,x) &:= \left| h^A(D^-v, D^+v) - H(v_x) + \frac{\Delta t}{2} H_p^2(v_x) v_{xx} \right| \\ &\leq C_{A,l} \left(\Delta t^l ||\partial_t^{l+1} v||_\infty + \Delta x^l ||\partial_x^{l+1} v||_\infty \right). \end{aligned} \tag{3.14}$$

Now, let us give some examples of high-order schemes satisfying (A1)-(A3) with $l = 2$.

Example 3.3. As a first example let us consider the class of schemes obtained combining a *high-order in space numerical hamiltonian* h_*^A and the second order *Runge-Kutta SSP* (or *Heun scheme*). To explain the simple procedure, let us consider the semidiscrete problem

$$u_t = h_*^A(D^-u(t,x), D^+u(t,x))),$$

where h_*^A, is a high-order in space numerical hamiltonian of second order,

$$h_*^A(D^-v_j^n, D^+v_j^n) = H(v_x(t^n, x_j)) + O(\Delta x^2), \tag{3.15}$$

such as the simple second order *central approximation*

$$h_*^A(D^-u_j^n, D^+u_j^n) = H\left(\frac{D^-u_j^n + D^+u_j^n}{2} \right), \tag{3.16}$$

then to obtain the same accuracy in time we discretize using the second order SSP Runge-Kutta scheme,

$$\begin{cases} u^* = u^n - \Delta t h_*^A(D^-u^n, D^+u^n) \\ u^{n+1} = \frac{1}{2}u^n + \frac{1}{2}u^* - \frac{\Delta t}{2} h_*^A(D^-u^*, D^+u^*). \end{cases} \tag{3.17}$$

The scheme can be written in differenced form in the sense of (A1)-(3.14) defining

$$h^A(D^-u^n, D^+u_j^n) = \frac{1}{2} \left[h_*^A(D^-u^n, D^+u^n) + h_*^A(D^-u^*, D^+u^*) \right]. \tag{3.18}$$

To verify that the scheme is second order we can use the Taylor expansion to see that

$$\begin{aligned} h_*^A(D^-v_j^*, D^+v_j^*) &= H\left(v_x^n(x_j) - \Delta t \frac{\partial}{\partial x} H(v_x^n(x_j))\right) + O(\Delta x^2) \\ &= H(v_x^n(x_j)) - \Delta t \left[H_p(v_x^n(x_j)) v_{xx}^n(x_j)\right] H_p(v_x^n(x_j)) + O(\Delta x^2), \end{aligned}$$

having exploited the relation $v^* = v^n - \Delta t[H(v_x^n(x_j) + O(\Delta x^2)]$, the Lipschitz continuity of H and having assumed a CFL condition $\lambda = \frac{\Delta t}{\Delta x} = const$; whence, again using the consistency property (3.15)

$$h^A(D^-v^n, D^+v^n) = H(v_x^n(x_j)) - \frac{\Delta t}{2} H_p^2(v_x^n(x_j)) v_{xx}^n(x_j) + O(\Delta x^2),$$

as we wanted. Notice that through this procedure the stencil of the scheme (3.15) becomes doubled for h^A.

Notice also that this procedure can be easily extended to the case of hamiltonian dependent on the space variable x.

Example 3.4. Then we propose a couple of numerical hamiltonians h^A obtained discretizing directly the formula (3.13) or, equivalently, obtained from the same *Lax-Wendroff schemes* for conservation laws by the substitution $u_j^n = \frac{v_{j+1}^n - v_j^n}{\Delta x}$. The first is the original *Lax-Wendroff* scheme

$$\begin{aligned} h^A(D^-u_j^n, D^+u_j^n) = \; & \tfrac{1}{2}\left\{H\left(D^+u_j^n\right) + H\left(D^-u_j^n\right) + \right. \\ & \left. - \tfrac{\Delta t}{\Delta x} H_p\left(\tfrac{D^-u_j^n + D^+u_j^n}{2}\right)\left[H\left(D^+u_j^n\right) - H\left(D^-u_j^n\right)\right]\right\}, \end{aligned} \tag{3.19}$$

and the second is its variation proposed by *Richtmyer*,

$$h^A(D^-u_j^n, D^+u_j^n) = H\left(\frac{D^-u_j^n + D^+u_j^n}{2} - \frac{\Delta t}{2\Delta x}\left[H\left(D^+u_j^n\right) - H\left(D^-u_j^n\right)\right]\right). \tag{3.20}$$

Example 3.5. Following the approach of the Lax-Wendroff schemes and making use of the expansion (3.12), we can easily write higher order schemes, in both space and time, using very compact stencils. The idea is simply to discretize directly the above expansion using finite difference approximations of the right order. For example, if we want to write a *fourth order Lax-Wendroff scheme* using only five points, one of the possibilities is to define

$$\begin{aligned} H_1 &= H\left(\tfrac{u_{j-2} - 8u_{j-1} + 8u_{j+1} - u_{j+2}}{12\Delta x}\right), \\ H_2 &= H_p^2\left(\tfrac{u_{j-2} - 8u_{j-1} + 8u_{j+1} - u_{j+2}}{12\Delta x}\right)\left(\tfrac{-u_{j-2} + 16u_{j-1} - 30u_j + 16u_{j+1} - u_{j+2}}{12\Delta x^2}\right), \\ H_3 &= \tfrac{1}{2\Delta x}\left[H_p^3\left(\tfrac{u_{j+2} - u_j}{2\Delta x}\right)\left(\tfrac{u_{j+2} - 2u_{j+1} + u_j}{\Delta x^2}\right) - H_p^3\left(\tfrac{u_j - u_{j-2}}{2\Delta x}\right)\left(\tfrac{u_j - 2u_{j-1} + u_{j-2}}{\Delta x^2}\right)\right], \\ H_4 &= \tfrac{1}{\Delta x^2}\left[H_p^4\left(\tfrac{u_{j+2} - u_j}{2\Delta x}\right)\left(\tfrac{u_{j+2} - 2u_{j+1} + u_j}{\Delta x^2}\right) - 2H_p^4\left(\tfrac{u_{j+1} - u_{j-1}}{2\Delta x}\right)\left(\tfrac{u_{j+1} - 2u_j + u_{j-1}}{\Delta x^2}\right)\right. \\ &\qquad \left. + H_p^4\left(\tfrac{u_j - u_{j-2}}{2\Delta x}\right)\left(\tfrac{u_j - 2u_{j-1} + u_{j-2}}{\Delta x^2}\right)\right], \end{aligned}$$

and then compute

$$h^A(D^-u_j^n, D^+u_j^n) = H_1 - \frac{\Delta t}{2}\left[H_2 - \frac{\Delta t}{3}\left(H_3 - \frac{\Delta t}{4}H_4\right)\right]. \tag{3.21}$$

It is straightforward to verify that, if the solution v is regular enough, using Taylor expansion we have

- $H_1 = H(v_x) + O(\Delta x^4)$,
- $H_2 = H_p^2(v_x)v_{xx} + O(\Delta x^4)$,
- $H_3 = \frac{\partial}{\partial x}\left(H_p^3(v_x)v_{xx}\right) + O(\Delta x^2)$,
- $H_4 = \frac{\partial^2}{\partial x^2}\left(H_p^4(v_x)v_{xx}\right) + O(\Delta x^2)$,

and that the resulting scheme satisfies (A1)-(A3) with $l = 4$. Notice that to obtain fourth order it would have been enough to have approximations of one order lower for H_2 and H_4, but thanks to the symmetry of the discretizations we can get higher orders without increasing the number of points in the stencil.

3.2.2. Filter function

In order to couple the schemes and their properties, we need to define a function F, called *filter function* F, such that

(F1) $F(x) \approx x$ for $|x| \leq 1$,

(F2) $F(x) = 0$ for $|x| > 1$,

which implies that

- If $|S^A - S^M| \leq \Delta t\varepsilon^n$ and $\phi_j^n = 1 \Rightarrow S^{AF} \approx S^A$
- If $|S^A - S^M| > \Delta t\varepsilon^n$ or $\phi_j^n = 0 \Rightarrow S^{AF} = S^M$.

It is clear that, with just these two requirements, we are left with several possible choices for F. In the following, we present some examples of filter functions satisfying the previous relations, which differ especially for regularity properties. We number the functions in order to be clearer in Figure 3.1.

Example 3.6. As a first example let us present the filter function we use in our numerical tests, defined in [BFS16] as

$$F_1(x) = \begin{cases} x & \text{if } |x| \leq 1 \\ 0 & \text{otherwise,} \end{cases} \tag{3.22}$$

which is clearly discontinuous at $x = -1, 1$ and satisfies trivially the properties (F1)-(F2).

Example 3.7. As a second possibility we propose the family of regular filter functions given by the formula

$$F(x) = x \exp\left(-c(|x|-a)^b\right),$$

for appropriate choices of the parameters a, b and c.

***Remark* 3.3.** We give some hints on how to chose the parameters. We notice that

- a controls the amplitude of the transition phase around 1 and -1;
- b controls the slope of the transition phase;
- c can be used to make the exponent approach 0 faster when $x \approx 1, -1$.

In particular, in Figure 3.1 we represent two choices for the parameters,

$$F_2(x) = x \exp\left(-4(|x|-0.25)^{20}\right) \qquad (a = 0.25,\ b = 20,\ c = 4) \tag{3.23}$$

and a variant graphically more similar to F_1,

$$F_3(x) = x \exp\left(-\frac{(|x|-0.01)^{50}}{100}\right) \qquad (a = 0.01,\ b = 50,\ c = 0.01). \tag{3.24}$$

These functions are very regular ($F \in C^\infty$) and developing with Taylor we can see that they satisfy (F1)-(F2).

Example 3.8. As last examples let us consider some functions which satisfy (F1)-(F2) and are continuous, but are not necessarily differentiable. First, let us consider the family of functions

$$F(x) = \begin{cases} x \exp\left(-\frac{a}{b-|x|}\right) & \text{if } |x| \leq b \\ 0 & \text{otherwise,} \end{cases} \tag{3.25}$$

varying the parameters a and b. We propose the choice

$$F_4(x) = \begin{cases} x \exp\left(-\frac{0.001}{1.05-|x|}\right) & \text{if } |x| \leq 1.05 \\ 0 & \text{otherwise,} \end{cases} \tag{3.26}$$

where we chose the value $b = 1.05$ in order to make the function approach better the value 1 for $x = -1, 1$. Finally, we recall also the filter defined in [ObSa15] as

$$F_5(x) = \begin{cases} x & |x| \leq 1 \\ 0 & |x| \geq 2 \\ -x + 2 & 1 \leq x \leq 2 \\ -x - 2 & -2 \leq x \leq -1. \end{cases} \tag{3.27}$$

After extensive computations, we noticed that the results obtained with our AF scheme are not sensitive with respect to changes in regularity of the filter function, even with very large transition phases. That is probably because, as we will see in the next section, the parameter ε^n is designed to obtain the property (F1) whenever possible, then in regions of regularity of the solution the argument of F lies most probably in $[-1, 1]$, where all the filter functions are practically the same. Some major differences, instead, can be seen in the results obtained with the SF scheme, for which the threshold ε is fixed at the beginning.

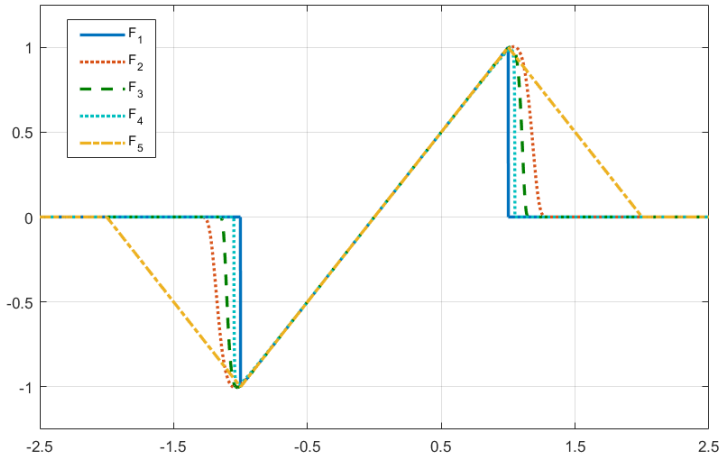

Figure 3.1.: Possible choices for the filter function F.

3.2.3. Tuning of the parameter ε^n

The last step is to show how to compute the switching parameter ε^n, which is the real core of the adaptivity of our scheme. Then, if we want the scheme (3.4) to switch to the high-order scheme when some regularity is detected, we have to choose ε^n such that

$$\left|\frac{S^A(v^n)_j - S^M(v^n)_j}{\varepsilon^n \Delta t}\right| = \left|\frac{h^A(\cdot) - h^M(\cdot)}{\varepsilon^n}\right| \leq 1, \qquad \text{for } (\Delta t, \Delta x) \to 0, \tag{3.28}$$

in the *region of regularity at time* t_n, that is

$$\mathcal{R}^n = \left\{x_j : \phi_j^n = 1\right\}. \tag{3.29}$$

For the definition of a function ϕ such that

$$\phi_j^n = \begin{cases} 1 & \text{if the solution } u^n \text{ is regular in } I_j, \\ 0 & \text{if } I_j \text{ contains a point of singularity,} \end{cases} \tag{3.30}$$

we refer to Section 2.1.1, where we show how it can be computed. To be precise, following the discussion in Remark 2.7, in the actual formulation we are forced to add a "technical" assumption in order to justify the proof of Proposition 3.9. In the end, we define the *region of regularity* $\mathcal{R}^n$ detected by the function $\widetilde{\phi}$ as the set

$$\mathcal{R}^n = \left\{j \in \mathbb{Z} : \widetilde{\phi}(\omega_j^n) = 1\right\}, \quad \text{with} \quad \widetilde{\phi}_j^n = \begin{cases} 1 & \text{if } \phi(\omega_j^n) = 1 \text{ and } |D^2 u_j^n| < B, \\ 0 & \text{otherwise,} \end{cases} \tag{3.31}$$

for some constant $B \gg 0$. Notice that with this definition, which, we recall, is needed only for theoretical reasons, it is not necessary to require $M(\Delta x) \to 0$, then we can

simply choose a constant $M > 0$ small enough (e.g. $M = 0.15$), as we will do in the numerical tests of Section 3.5.

Therefore, computing directly by Taylor expansions, we have for the monotone scheme

$$h^M(D^- v_j^n, D^+ v_j^n) = H(v_x^n(x_j)) + \frac{\Delta x}{2} v_{xx}^n(x_j) \left(\partial_{p_+} h_j^M - \partial_{p_-} h_j^M\right) + O(\Delta x^2),$$

where we used the relation

$$D^{\pm} v_j^n = v_x^n(x_j) \pm \frac{\Delta x}{2} v_{xx}^n(x_j) + O(\Delta x^2),$$

while for the high-order scheme, by the consistency property,

$$h^A(D^- v_j^n, D^+ v_j^n) = H(v_x^n(x_j)) - \frac{\Delta t}{2} H_p^2(v_x^n) v_{xx} + O(\Delta t^2) + O(\Delta x^2).$$

Whence, from (3.28) we obtain

$$\varepsilon^n \geq \left| \frac{\Delta x}{2} v_{xx}^n \left(\partial_{p_+} h_j^M - \partial_{p_-} h_j^M + \lambda H_p^2(v_x^n)\right) + O(\Delta t^2) + O(\Delta x^2) \right|. \tag{3.32}$$

Finally, we use a numerical approximation of the lower bound on the right hand side of the previous inequality to obtain the following formula for ε^n,

$$\begin{aligned} \varepsilon^n = \max_{x_j \in \mathcal{R}^n} K \Big| & H\left(D\, u_j^n\right) - H\left(D\, u_j^n - \lambda \left[H(D^+ u_j^n) - H(D^- u_j^n)\right]\right) \\ & + \left[h^M(D\, u_j^n, D^+ u_j^n) - h^M(D\, u_j^n, D^- u_j^n)\right] \\ & - \left[h^M(D^+ u_j^n, D\, u_j^n) - h^M(D^- u_j^n, D\, u_j^n)\right] \Big|, \end{aligned} \tag{3.33}$$

with $K > \frac{1}{2}$, $\lambda := \frac{\Delta t}{\Delta x}$ and $D\, u_j^n := \frac{u_{j+1}^n - u_{j-1}^n}{2\Delta x}$. Notice that if we assume enough regularity on the solution v, then (3.33) gives a second order approximation of the right hand side of (3.32) multiplied by $2K$.

It is worth to mention that, using very similar computations, in [BFS16] the authors propose a "simplified" approach to compute the parameter $\varepsilon = c_1 \Delta x$, where the constant c_1 has to chosen roughly such that

$$c_1 \geq \frac{1}{2} ||v_{xx}||_\infty \left|\left| \frac{\partial h^M}{\partial p_+}(v_x, v_x) - \frac{\partial h^M}{\partial p_-}(v_x, v_x) \right|\right|_\infty, \tag{3.34}$$

with the range of values of v_x and v_{xx} to be estimated in terms of the values of $(v_0)_x$ and $(v_0)_{xx}$ and the hamiltonian function H.

3.3. Convergence result

We are now able to present our main result, but before doing so let us state a useful proposition about the numerical solution and the parameter ε^n.

Proposition 3.9. *Let u^n be the solution obtained by the scheme (3.4)-(3.33) and assume that v_0 and H are Lipschitz continuous functions. Assume also that $\mathcal{R}^n$ is defined by (3.31), with ϕ given by (2.29), and that $\lambda = \Delta t/\Delta x = $ constant. Then, ε^n is well defined and u^n satisfies, for any i and j, the discrete Lipschitz estimate*

$$\frac{|u_i^n - u_j^n|}{\Delta x} \leq L$$

for some constant $L > 0$, for $0 \leq n \leq T/\Delta t$. Moreover, there exists a constant $C > 0$ such that

$$\varepsilon^n \leq C\Delta x.$$

Proof. Before proceeding with the proof let us notice that, if u^n satisfies (3.9) for a constant $L_n > 0$, calling for brevity

$$D^* u_j := D\, u_j^n - \lambda \left[H(D^+ u_j^n) - H(D^- u_j^n)\right],$$

we have that

$$\begin{aligned}
\varepsilon^n = & \max_{x_j \in \mathcal{R}^n} K \left| H\left(D\, u_j^n\right) - H\left(D^* u_j\right) + \left[h^M(D\, u_j^n, D^+ u_j^n) - h^M(D\, u_j^n, D^- u_j^n)\right] \right. \\
& \left. - \left[h^M(D^+ u_j^n, D\, u_j^n) - h^M(D^- u_j^n, D\, u_j^n)\right] \right| \\
= & \max_{x_j \in \mathcal{R}^n} K \left| \left[\Delta t \left(\frac{H\left(D\, u_j^n\right) - H\left(D^* u_j\right)}{D\, u_j^n - D^* u_j} \right) \left(\frac{H(D^+ u_j^n) - H(D^- u_j^n)}{D^+ u_j^n - D^- u_j^n} \right) \right. \right. \\
& + \Delta x \left(\frac{h^M(D\, u_j^n, D^+ u_j^n) - h^M(D\, u_j^n, D^- u_j^n)}{D^+ u_j^n - D^- u_j^n} \right) \\
& \left. \left. - \Delta x \left(\frac{h^M(D^+ u_j^n, D\, u_j^n) - h^M(D^- u_j^n, D\, u_j^n)}{D^+ u_j^n - D^- u_j^n} \right) \right] \left(\frac{D^+ u_j^n - D^- u_j^n}{\Delta x} \right) \right| \\
= & \max_{x_j \in \mathcal{R}^n} K \left| \left[\Delta t \left(\frac{H\left(D\, u_j^n\right) - H\left(D^* u_j\right)}{D\, u_j^n - D^* u_j} \right) \left(\frac{H(D^+ u_j^n) - H(D^- u_j^n)}{D^+ u_j^n - D^- u_j^n} \right) \right. \right. \\
& + \Delta x \left(\frac{h^M(D\, u_j^n, D^+ u_j^n) - h^M(D\, u_j^n, D^- u_j^n)}{D^+ u_j^n - D^- u_j^n} \right) \\
& \left. \left. - \Delta x \left(\frac{h^M(D^+ u_j^n, D\, u_j^n) - h^M(D^- u_j^n, D\, u_j^n)}{D^+ u_j^n - D^- u_j^n} \right) \right] \frac{\sqrt{\beta_0^+ [u_j^n]}}{\Delta x} \right| \\
\leq & K \left| (\Delta t L_{H2} L_H + 2\Delta x L_{h^M}) B \right| \\
= & K B \left(\lambda L_{H2} L_H + 2 L_{h^M}\right) \Delta x, \qquad (3.35)
\end{aligned}$$

where L_H and L_{H2} are the local Lipschitz constant of H on $[-L_n, L_n]$ and $[-2L_n - \Delta t L_H B, 2L_n + \Delta t L_H B]$, respectively, and L_n is the Lipschitz constant of u^n. Notice that if the function H is globally Lipschitz continuous we have the same estimate with

$L_{H2} = L_H$, where now L_H is the global Lipschitz constant of H. Notice also that we have used the fact that, by definition,

$$\frac{\sqrt{\beta_0^+[u_j^n]}}{\Delta x} = \frac{D^+u_j^n - D^-u_j^n}{\Delta x} = D^2u_j^n,$$

and that $x_j \in \mathcal{R}^n \Rightarrow D^2u_j^n < B$, for some constant $B > 0$ independent on n, by Lemma 2.2 or by the definition of $\mathcal{R}^n$ (3.31). Then, the last statement would follow with $C = KB(\lambda L_H^2 + 2L_{h^M})$.

Let us now prove the main statement proceeding, as usual, by induction on $n \geq 0$ and noticing that it is sufficient to prove (3.9) for i and j such that $i = j \pm 1$.

For $n = 0$, as we take $u_j^0 = v_0(x_j)$ for $j \in \mathbb{Z}$, we have that (3.9) is satisfied by the Lipschitz continuity assumption on v_0, with constant L_0.

Now, assuming that (3.9) is satisfied for $n - 1 > 0$ so that ε^k for $k = 0, \ldots, n-1$ are bounded by (3.35), we can compute

$$\begin{aligned}
\frac{|u_i^n - u_j^n|}{\Delta x} &= \frac{1}{\Delta x}\left|S^M(u^{n-1})_i + \phi_i\varepsilon^{n-1}\Delta t F(\cdot)_i - S^M(u^{n-1})_j - \phi_j\varepsilon^{n-1}\Delta t F(\cdot)_j\right| \\
&\leq \frac{1}{\Delta x}\left(|S^M(u^{n-1})_i - S^M(u^{n-1})_j| + \varepsilon^{n-1}\Delta t|\phi_i F(\cdot)_i - \phi_j F(\cdot)_j|\right) \\
&\leq \frac{|u_i^{n-1} - u_j^{n-1}|}{\Delta x} + \frac{2\Delta t}{\Delta x}\varepsilon^{n-1}
\end{aligned}$$

then, iterating back and using the same arguments,

$$\begin{aligned}
\frac{|u_i^n - u_j^n|}{\Delta x} &\leq \frac{|u_i^{n-1} - u_j^{n-1}|}{\Delta x} + 2\Delta t C \leq \ldots \\
&\leq \frac{|u_i^1 - u_j^1|}{\Delta x} + 2(n-1)\Delta t C \\
&\leq \frac{|u_i^0 - u_j^0|}{\Delta x} + 2n\Delta t C \\
&\leq L_0 + 2\frac{T}{\Delta t}\Delta t C = L,
\end{aligned}$$

where C is well defined by (3.35). Notice that we have used the monotonicity of S^M and the fact that $|F| \leq 1$. □

Therefore, it is clear that by construction our scheme is *ε-monotone*, in the sense of the following

Definition 3.10 (ε-monotonicity)**.** A numerical scheme S is *ε-monotone* if for any functions u, v,

$$u \leq v \Rightarrow S(u) \leq S(v) + C\varepsilon\Delta t,$$

where C is constant and $\varepsilon \to 0$ as $\Delta = (\Delta t, \Delta x) \to 0$.

Finally, we conclude this section giving our convergence result for the Adaptive Filtered Schemes.

Theorem 3.11. *Let the assumptions on S^M and S^A be satisfied. Assume that v_0 and H are Lipschitz continuous functions, u_j^{n+1} is computed by (3.4)-(3.33), with $K > 1/2$ and $\lambda = \frac{\Delta t}{\Delta x}$, a constant such that (3.7) is satisfied. Let us denote by $v_j^n := v(t^n, x_j)$ the values of the viscosity solution on the nodes of the grid. Then,*

i) the AF scheme (3.4) satisfies Crandall-Lions estimate [CL84]

$$||u^n - v^n||_\infty \leq C_1\sqrt{\Delta x}, \quad \forall\ n = 0, \dots, N_T,$$

for some constant $C_1 > 0$ independent of Δx.

ii) (First order convergence for regular solutions) Moreover, if $v \in C^2([0,T] \times \mathbb{R})$, then

$$||u^n - v^n||_\infty \leq C_2\Delta x, \quad \forall\ n = 0, \dots, N_T,$$

for some constant $C_2 > 0$ independent of Δx.

iii) (High-order local consistency) Let $k \geq 2$ be the order of the scheme S^A. If $v \in C^{l+1}$, with $1 \leq l \leq k$, in some neighborhood of a point $(t,x) \in [0,T] \times \mathbb{R}$, then

$$\mathcal{E}_{AF}(v^n)_j = \mathcal{E}_A(v^n)_j = O(\Delta x^l) + O(\Delta t^l)$$

for $t^n - t$, $x_j - x$, Δt, Δx sufficiently small.

Proof. ***i)*** Let us proceed as has been done in [BFS16] defining $w_j^{n+1} = S^M(w^n)_j$, the solution computed with the monotone scheme alone with $w_j^0 = v_0(x_j)$. Then by definition,

$$u_j^{n+1} - w_j^{n+1} = S^M(u^n)_j - S^M(w^n)_j + \phi_j\varepsilon^n\Delta t F\left(\frac{S^A(u^n)_j - S^M(u^n)_j}{\varepsilon^n\Delta t}\right), \tag{3.36}$$

whence, exploiting the nonexpansivity in L^∞ of S^M, the definition of ε^n and that $|F| \leq 1$,

$$\max_j |u_j^{n+1} - w_j^{n+1}| \leq \max_j |u_j^n - w_j^n| + \varepsilon^n\Delta t. \tag{3.37}$$

Then, proceeding recursively on $n \leq N_T$ and recalling that by Proposition 3.9 there exists a constant $C > 0$ such that $\varepsilon^n \leq C\Delta x := \varepsilon$ for each n,

$$\max_j |u_j^n - w_j^n| \leq \sum_{k=0}^{n-1} \varepsilon^k\Delta t \leq n\varepsilon\Delta t \leq T\varepsilon. \tag{3.38}$$

At this point, by the triangular inequality

$$\max_j |u_j^{n+1} - v_j^{n+1}| \leq \max_j |u_j^{n+1} - w_j^{n+1}| + \max_j |w_j^{n+1} - v_j^{n+1}|, \tag{3.39}$$

whence we have that

$$\max_j |u_j^{n+1} - v_j^{n+1}| \leq \max_j |w_j^n - v_j^n| + \varepsilon T \leq (C_{CL} + CT)\sqrt{\Delta x}, \tag{3.40}$$

with $C_{CL} > 0$ given by the Crandall-Lions estimate for S^M.
ii) Let us recall that by (3.5), in the case of $v \in C^2$ the consistency error for the monotone scheme is such that $\mathcal{E}_M(v^n)_j \leq C_M(\Delta t + \Delta x)$. Then we can compute

$$\begin{aligned} |u_j^{n+1} - v_j^{n+1}| &= |S^M(u^n)_j + \phi_j \varepsilon^n \Delta t F(\cdot) - v_j^{n+1}| \\ &\leq |S^M(u^n)_j - S^M(v^n)_j| + |S^M(v^n)_j - v_j^{n+1}| + \varepsilon^n \Delta t \\ &\leq ||u^n - v^n||_\infty + \Delta t \left(\mathcal{E}_M(v^n) + \varepsilon^n\right), \end{aligned}$$

whence, by recursion on $n \leq N_T$ and recalling what we have done in the previous point,

$$||u^n - v^n||_\infty \leq ||u^0 - v^0||_\infty + T\left(\max_{k=0,\dots,n-1} ||\mathcal{E}_M(v^k)||_\infty + \varepsilon\right). \tag{3.41}$$

To conclude this proof what is left is to use the estimate on $\mathcal{E}_M$ and Proposition 3.9.

iii) In order to show that $S^{AF}(v^n)_j = S^A(v^n)_j$ for Δt e Δx small enough it is sufficient to prove that

$$\frac{|S^A(v^n)_j - S^M(v^n)_j|}{\varepsilon^n \Delta t} \leq 1, \qquad \text{for } (\Delta t, \Delta x) \to 0, \tag{3.42}$$

which follows directly from the computation we have done in section 3.2.3 for the tuning of the parameter ε^n and the definition of the regularity set $\mathcal{R}^n$ (3.31), with ϕ given by (2.29). In fact, if we plug (3.33) inside the previous inequality, we can deduce that

$$\frac{|S^A(v^n)_j - S^M(v^n)_j|}{\varepsilon^n \Delta t} \leq \frac{1}{2K} + O(\Delta x) + O(\Delta t),$$

which, using that $K > 1/2$ by assumption, leads to the thesis as $(\Delta t, \Delta x) \to 0$. Notice that we have used the property $\varepsilon^n = O(\Delta x)$ and exploited the CFL condition. □

***Remark* 3.4.** Notice that the assumption $M(\Delta x) = \frac{1}{2} - C\Delta x$, for some constant $C > 0$ such that $M(\Delta x) > 0$, needed to apply Lemma 2.2, may give some problems in the proof of third assertion of the previous theorem. In fact, applying the standard definition (2.20) to the viscosity solution v at a point x_j and recalling the computations that led to (2.18), we get that

$$\omega_j^{\pm} = \frac{1}{2} \mp 4\Delta x \frac{v_j'' v_j'''}{(v_j'')^2 + \sigma} + O(\Delta x^2).$$

Consequently, in order to be sure that if $v \in C^3$, then $j \in \mathcal{R}$, we have to choose the constant C such that

$$C \geq \left| \frac{v_j'' v_j'''}{(v_j'')^2 + \sigma} \right|,$$

or require additional smoothness assumptions on v, for example $v_j''' \ll v_j''$. This in fact poses a strong limitation on the applicability of Lemma 2.2, at least in the present formulation. It would be interesting to see if a similar argument can be applied using indicators with $r = 3$ or the 'WENO-Z' procedure, which would definitely solve the previous problem. This issue will be object of future investigations.

3.4. Adaptive Filtered Scheme in 2D

The aim of this section is to show how to properly generalize the ideas presented in the previous section in order to solve problems in more space dimensions. To keep the ideas as clear as possible we will focus on the two dimensional case and, again, we consider the simplest situation in which the hamiltonian depends only on the gradient of the solution, with more general situations following with minor modifications. In particular, we give some details on how to modify the formulas in the case of hamiltonians depending also on the space variables, which will be used in Chapter 4.

Let us begin by considering the problem

$$\begin{cases} v_t + H(v_x, v_y) = 0, & (t,x,y) \in [0,T] \times \mathbb{R}^2, \\ v(0,x,y) = v_0(x,y), & x \in \mathbb{R}^2, \end{cases} \tag{3.43}$$

where the hamiltonian H and the initial data v_0 are Lipschitz continuous functions, as usual, in order to ensure the existence and uniqueness of the viscosity solution. Notice that, if we consider problems involving hamiltonians depending also on the position (x,y) or the solution v ($H(x,y,v,\nabla v)$), clearly we have to make more assumptions on the behavior of H w.r.t. the other variables. For a detailed presentation of uniqueness and existence results in more general situations we refer the reader to [CL83] and [B98].

Now, let us define a uniform grid in space $(x_j, y_i) = (j\Delta x, i\Delta y)$, j,$i \in \mathbb{Z}$, and in time $t_n = t_0 + n\Delta t$, $n \in [0, N_T]$, with $(N_T - 1)\Delta t < T \leq N_T \Delta t$. Then, we compute the numerical approximation $u_{i,j}^n = u(t_n, x_j, y_i)$ with the simple formula

$$u_{i,j}^{n+1} = S^{AF}(u^n)_{i,j} := S^M(u^n)_{i,j} + \phi_{i,j}^n \varepsilon^n \Delta t F\left(\frac{S^A(u^n)_{i,j} - S^M(u^n)_{i,j}}{\varepsilon^n \Delta t}\right), \tag{3.44}$$

where S^M and S^A are respectively the monotone and the high-order scheme, now dependent on both space variables. F is the same *filter function* of the one-dimensional version, needed to switch between the two schemes, ε^n is the switching parameter at time t_n and $\phi_{i,j}^n$ is the *smoothness indicator function* at the node (x_j, y_i) and time t_n, based on the 2D-smoothness indicators defined in Section 2.2.

3.4.1. Assumptions on the schemes

For convenience of presentation in this section we recall the assumptions we have to require on the composing schemes S^M and S^A, which are clearly direct generalizations in two space dimensions of those presented in Section 3.2.1. Let us begin by restating

the assumptions on the monotone scheme.

Assumptions on S^M.

(M1) The scheme can be written in *differenced form*

$$u_{i,j}^{n+1} \equiv S^M(u^n)_{i,j} := u_{i,j}^n - \Delta t\ h^M\left(D_x^- u_{i,j}^n, D_x^+ u_{i,j}^n; D_y^- u_{i,j}^n, D_y^+ u_{i,j}^n\right)$$

for a function $h^M(p^-, p^+; q^-, q^+)$, with $D_x^{\pm} u_{i,j}^n := \pm\frac{u_{i,j\pm1}^n - u_{i,j}^n}{\Delta x}$ and $D_y^{\pm} u_{i,j}^n := \pm\frac{u_{i\pm1,j}^n - u_{i,j}^n}{\Delta y}$;

(M2) h^M is a Lipschitz continuous function;

(M3) (Consistency) $\forall u, v$, $h^M(u, u, v, v) = H(u, v)$;

(M4) (Monotonicity) for any functions u, v,

$$u \le v \quad \Rightarrow \quad S^M(u) \le S^M(v).$$

Under assumption (M2), the consistency property (M3) is equivalent to say that for all functions $v \in C^2([0,T] \times \mathbb{R})$, there exists a constant $C_M \ge 0$ independent on $\Delta = (\Delta t, \Delta x)$ such that

$$\begin{aligned}\mathcal{E}_M(v)(t,x,y) := &\left|\frac{v(t+\Delta t, x, y) - S^M(v(t,\cdot,\cdot))(x,y)}{\Delta t}\right| \\ &\le C_M\left(\Delta t||v_{tt}||_\infty + \Delta x||v_{xx}||_\infty + \Delta y||v_{yy}||_\infty\right), \end{aligned} \tag{3.45}$$

where $\mathcal{E}_M$ is the consistency error. The last relation highlights the well-known first order bound on the accuracy of the monotone schemes for regular solutions.

***Remark* 3.5.** Using the same arguments of the one-dimensional case it can be easily shown that, under the Lipschitz assumption (M2), the monotonicity property (M4) is equivalent to require, for a.e. $(p^-, p^+) \in \mathbb{R}^2$,

$$\frac{\partial h^M}{\partial p^-} \ge 0, \qquad \frac{\partial h^M}{\partial p^+} \le 0, \qquad \frac{\partial h^M}{\partial q^-} \ge 0, \qquad \frac{\partial h^M}{\partial q^+} \le 0, \tag{3.46}$$

and the *CFL condition*

$$\frac{\Delta t}{\Delta x}\left(\frac{\partial h^M}{\partial p^-} - \frac{\partial h^M}{\partial p^+}\right) + \frac{\Delta t}{\Delta y}\left(\frac{\partial h^M}{\partial q^-} - \frac{\partial h^M}{\partial q^+}\right) \le 1. \tag{3.47}$$

Analogously, we define the constant ratios $\lambda_z := \frac{\Delta t}{\Delta z}$, for $z = x, y$ such that (3.47) is satisfied and call the *CFL number* the maximum $\lambda = \max\{\lambda_x, \lambda_y\}$.

Example 3.12. In this example we recall a couple of monotone schemes in differenced form satisfying (M1)-(M4), which will be used in the numerical tests.

- For the *eikonal equation*,
$$v_t + \sqrt{v_x^2 + v_y^2} = 0,$$
we can use the simple numerical hamiltonian
$$h^M(p^-, p^+; q^-, q^+) := \sqrt{\max\{p^-, -p^+, 0\}^2 + \max\{q^-, -q^+, 0\}^2}. \tag{3.48}$$

- For general equations, instead, we can use the 2D-version of the *Local Lax-Friedrichs hamiltonian*
$$\begin{aligned} h^M(p^-, p^+; q^-, q^+) :=& H\left(\frac{p^+ + p^-}{2}, \frac{q^+ + q^-}{2}\right) \\ &- \frac{\alpha_x(p^-, p^+)}{2}(p^+ - p^-) - \frac{\alpha_y(q^-, q^+)}{2}(q^+ - q^-), \end{aligned} \tag{3.49}$$
with
$$\alpha_x(p^-, p^+) = \max_{p \in I(p^-, p^+)} |H_p(p, q)|, \qquad \alpha_y(q^-, q^+) = \max_{q \in I(q^-, q^+)} |H_q(p, q)|,$$
where the maximum are computed uniformly in q and p, respectively, and $I(a, b)$ represents the interval with endpoints a and b. The scheme is monotone under the restrictions $\lambda_x \leq \max|H_p|^{-1}$ and $\lambda_y \leq \max|H_q|^{-1}$.

- Another interesting (and easy to compute) numerical hamiltonian we would like to use is the 2D-extension of the *Central Upwind scheme* of [KNP01]
$$\begin{aligned} h^M(p^-, p^+; q^-, q^+) :=& \frac{a^- b^- H(p^+, q^+) - a^- b^+ H(p^+, q^-)}{(a^+ - a^-)(b^+ - b^-)} \\ &+ \frac{a^+ b^+ H(p^-, q^-) - a^+ b^- - H(p^-, q^+)}{(a^+ - a^-)(b^+ - b^-)} \\ &+ \frac{a^+ a^-}{a^+ - a^-}(p^+ - p^-) + \frac{b^+ b^-}{b^+ - b^-}(q^+ - q^-), \end{aligned} \tag{3.50}$$
with $a^\pm$ and $b^\pm$ defined as the positive part of the maximum (superscript "+") and the negative part of the minimum (superscript "-"), respectively, of H_p and H_q in the cell $C_{i,j} = [x_{j-\frac{1}{2}}, x_{j+\frac{1}{2}}] \times [y_{i-\frac{1}{2}}, y_{i+\frac{1}{2}}]$. These quantities can be estimated as
$$\begin{aligned} a^+ &= \max_\pm \{H_p(p^\pm, q^\pm), 0\}, & a^- &= \min_\pm \{H_p(p^\pm, q^\pm), 0\}, \\ b^+ &= \max_\pm \{H_q(p^\pm, q^\pm), 0\}, & b^- &= \min_\pm \{H_q(p^\pm, q^\pm), 0\}. \end{aligned}$$
Unfortunately, at least to our knowledge, this numerical hamiltonian has still to be proven monotone, whence we avoid to use it in the numerical tests.

***Remark* 3.6.** It is straightforward to adapt the monotone schemes defined in the previous example in order to handle hamiltonian dependent also on the space variables. In fact, it is enough to consider such dependence of H, H_p and H_q in the formulas for $h^M(x_j, y_i, p^-, p^+; q^-, q^+)$.

Next, we recall the assumptions on the high-order scheme.

Assumptions on S^A.

(A1) The scheme can be written in *differenced form*

$$u_j^{n+1} = S^A(u^n)_j := u_j^n - \Delta t h^A\left(D_x^{k,-}u_{i,j}, \dots, D_x^- u_{i,j}^n, D_x^+ u_{i,j}^n, \dots, D_x^{k,+}u_{i,j}^n; \right.$$
$$\left. D_y^{k,-}u_{i,j}, \dots, D_y^- u_{i,j}^n, D_y^+ u_{i,j}^n, \dots, D_y^{k,+}u_{i,j}^n\right), \tag{3.51}$$

for some function $h^A(p^-, p^+; q^-, q^+)$ (in short), with $D_x^{k,\pm}u_{i,j}^n := \pm\frac{u_{i,j\pm k}^n - u_{i,j}^n}{k\Delta x}$ and $D_y^{k,\pm}u_{i,j}^n := \pm\frac{u_{i\pm k,j}^n - u_{i,j}^n}{k\Delta y}$;

(A2) h^A is a Lipschitz continuous function.

(A3) (High-order consistency) Fix $k \geq 2$ order of the scheme (for all the variables), then for all $l = 1, \dots, k$ and for all functions $v \in C^{l+1}$, there exists a constant $C_{A,l} \geq 0$ such that

$$\mathcal{E}_A(v)(t,x,y) := \left|\frac{v(t+\Delta t, x, y) - S^A(v(t,\cdot))(x,y)}{\Delta t}\right|$$
$$\leq C_{A,l}\left(\Delta t^l ||\partial_t^{l+1} v||_\infty + \Delta x^l ||\partial_x^{l+1} v||_\infty + \Delta y^l ||\partial_y^{l+1} v||_\infty\right).$$

Again, it is worth to notice that no further hypothesis are needed, neither for stability reasons.

In order to restate the consistency property in a more useful form, we first compute the second order term of the expansion

$$v(t+\Delta t, x, y) = v(t,x,y) + \Delta t v_t(t,x,y) + \frac{\Delta t^2}{2} v_{tt}(t,x,y) + O\left(\Delta t^3\right), \tag{3.52}$$

which gives, dropping the dependance on (x, y) for brevity,

$$\begin{aligned} v_{tt} &= \frac{\partial}{\partial t}(-H(v_x, v_y)) \\ &= -H_p(v_x, v_y)v_{xt} - H_q(v_x, v_y)v_{yt} \\ &= H_p(v_x, v_y)\frac{\partial}{\partial x}\left(H(v_x, v_y)\right) + H_q(v_x, v_y)\frac{\partial}{\partial y}\left(H(v_x, v_y)\right) \\ &= H_p^2(v_x, v_y)v_{xx} + H_q^2(v_x, v_y)v_{yy} + 2H_p(v_x, v_y)H_q(v_x, v_y)v_{xy}, \end{aligned} \tag{3.53}$$

then it is straightforward, assuming (A1)-(A2), to write the consistency property in terms of the numerical hamiltonian h^A, that is

(A3′) (High-order consistency) Fix $k \geq 2$ order of the scheme (for all the variables), then for all $l = 1, \dots, k$ and for all functions $v \in C^{l+1}$, there exists a constant $C_{A,l} \geq 0$ such that

$$\begin{aligned}\mathcal{E}_A(v)(t,x,y) := \Big| & h^A(D_x^- v, D_x^+ v; D_y^- v, D_y^+ v) - H(v_x, v_y) \\ & + \frac{\Delta t}{2}\left[H_p(v_x, v_y)\frac{\partial}{\partial x}\left(H(v_x, v_y)\right) + H_q(v_x, v_y)\frac{\partial}{\partial y}\left(H(v_x, v_y)\right)\right]\Big| \\ & \leq C_{A,l}\left(\Delta t^l ||\partial_t^{l+1} v||_\infty + \Delta x^l ||\partial_x^{l+1} v||_\infty + \Delta y^l ||\partial_y^{l+1} v||_\infty\right). \quad (3.54)\end{aligned}$$

In the following examples we present some simple high-order schemes satisfying (A1)-(A3) with $l = 2$, dropping the dependence on (i, j) (and also on n) in order to lighten the notations.

Example 3.13. Analogously to the one-dimensional case, the easiest way to construct an high-order scheme satisfying (A1)-(A3) with $l = 2$, is to consider a *second order in space numerical hamiltonian* h_*^A

$$h_*^A(D_x^- v, D_x^+ v; D_y^- v, D_y^+ v) = H(v_x, v_y) + O\left(\Delta x^2\right) + O\left(\Delta y^2\right), \quad (3.55)$$

such as the simple second order *Centered approximation*

$$h_*^A(D_x^- u, D_x^+ u; D_y^- u, D_y^+ u) = H\left(\frac{D_x^- u + D_x^+ u}{2}, \frac{D_y^- u + D_y^+ u}{2}\right), \quad (3.56)$$

and combine it with the second order *Runge-Kutta SSP* (or *Heun scheme*) (3.17).

Example 3.14. In this example we propose a series of numerical hamiltonians h^A obtained discretizing directly the formula (3.53).

- The first is the most direct and simple discretization, to which we will refer as the *Lax-Wendroff* scheme

$$\begin{aligned}h^A(D_x^\pm u; D_y^\pm u) = & H(D_x u, D_y u) - \frac{\Delta t}{2}\left[H_p^2(D_x u, D_y u) D_x^2 u + \right. \\ & \left. + H_q^2(D_x u, D_y u) D_y^2 u + 2 H_p(D_x u, D_y u) H_q(D_x u, D_y u) D_{xy}^2 u\right], \quad (3.57)\end{aligned}$$

  where $D_z^\pm u$, $D_z u$, $D_z^2 u$ are, respectively, the usual one-sided and centered one-dimensional finite difference approximations of the first and second derivative in the z-direction, while for the mixed derivative we use

$$D_{xy}^2 u_{i,j} = \frac{u_{i+1,j+1} - u_{i-1,j+1} - u_{i+1,j-1} + u_{i-1,j-1}}{4\Delta x \Delta y}.$$

- Another possibility, which is more closely related to the one-dimensional Lax-Wendroff scheme (3.19), is the following

$$h^A(D_x^{\pm}u; D_y^{\pm}u) = H(D_xu, D_yu) - \frac{\Delta t}{2\Delta x} H_p(D_xu, D_yu) H_x^* - \frac{\Delta t}{2\Delta y} H_q(D_xu, D_yu) H_y^*, \tag{3.58}$$

where we have defined

$$\begin{aligned} H_x^* =& H\left(\frac{u_{i,j+1} - u_{i,j}}{\Delta x}, \frac{u_{i+1,j+1} - u_{i-1,j+1} + u_{i+1,j} - u_{i-1,j}}{4\Delta y}\right) \\ & - H\left(\frac{u_{i,j} - u_{i,j-1}}{\Delta x}, \frac{u_{i-1,j} - u_{i-1,j} + u_{i+1,j-1} - u_{i-1,j-1}}{4\Delta y}\right) \end{aligned} \tag{3.59}$$

and

$$\begin{aligned} H_y^* =& H\left(\frac{u_{i+1,j+1} - u_{i+1,j-1} + u_{i,j+1} - u_{i,j-1}}{4\Delta x}, \frac{u_{i+1,j} - u_{i,j}}{\Delta y}\right) \\ & - H\left(\frac{u_{i,j+1} - u_{i,j-1} + u_{i-1,j+1} - u_{i-1,j-1}}{4\Delta x}, \frac{u_{i-1,j} - u_{i,j}}{\Delta y}\right). \end{aligned} \tag{3.60}$$

This can be seen as a discretization of the third relation in (3.53).

- The last example we propose is the *Richtmyer* form,

$$h^A(D_x^{\pm}u; D_y^{\pm}u) = H\left(D_xu - \frac{\Delta t}{2\Delta x} H_x^*, D_yu - \frac{\Delta t}{2\Delta y} H_y^*\right), \tag{3.61}$$

which, in particular, does not require any computation of H_p or H_q.

Example 3.15. Finally, we would like to show a simple way to define a scheme satisfying (A1)-(A3) with $l = 4$, reminding that, in our approach, the high-order scheme has no need to be stable, in any sense. Then, generalizing the construction of Example 3.13, we can define a fourth-order scheme by combining the simple fourth-order central approximation

$$h_*^A(D_x^-u, D_x^+u; D_y^-u, D_y^+u) = H\left(D_x^*u, D_y^*u\right), \tag{3.62}$$

where the approximated partial derivative are computed as

$$D_x^*u = \frac{u_{i,j-2} - 8u_{i,j-1} + 8u_{i,j+1} - u_{i,j+2}}{12\Delta x}, \qquad D_y^*u = \frac{u_{i-2,j} - 8u_{i-1,j} + 8u_{i+1,j} - u_{i+2,j}}{12\Delta y},$$

with the classic *fourth-order Runke-Kutta scheme*

$$\begin{cases} u^{(1)} = u^n + \frac{\Delta t}{2} h_*^A\left(D^{\pm}u^n\right) \\ u^{(2)} = u^n + \frac{\Delta t}{2} h_*^A\left(D^{\pm}u^{(1)}\right) \\ u^{(3)} = u^n + \Delta t h_*^A\left(D^{\pm}u^{(2)}\right) \\ u^{n+1} = u^n + \frac{\Delta t}{6}\left[h_*^A\left(D^{\pm}u^n\right) + 2h_*^A\left(D^{\pm}u^{(1)}\right) + 2h_*^A\left(D^{\pm}u^{(2)}\right) + h_*^A\left(D^{\pm}u^{(3)}\right)\right] \\ \qquad = \frac{1}{3}\left[\frac{\Delta t}{2}\left(h_*^A\left(D^{\pm}u^n\right) + h_*^A\left(D^{\pm}u^{(3)}\right)\right) + 2u^{(2)} + u^{(3)}\right]. \end{cases} \tag{3.63}$$

Notice that, differently from the usual approach, used when working with Hamilton-Jacobi equations (e.g. WENO schemes of higher order), we do not require the time discretization to be *strong stability preserving*, thus we can use the more efficient formula (3.63), which is also easier to implement w.r.t. the SSP version (see [Sh98] for more details). It is worth to point out that, differently from the Lax-Wendroff construction (3.21), the stencil of the scheme is not very compact, requiring 17×17 points in total. Further investigation will focus on the extension of Lemma 3.2 to the multidimensional case, which would most probably lead to the definition of fourth-order schemes with compact stencils of 5×5 points.

3.4.2. Tuning of ε^n

Let us explain how to generalize the formulas to compute the switching parameter ε^n.Then, if we want the scheme (3.44) to switch to the high-order scheme when some regularity is detected, we have to choose ε^n such that

$$\left| \frac{S^A(v^n)_{i,j} - S^M(v^n)_{i,j}}{\varepsilon^n \Delta t} \right| = \left| \frac{h^A(\cdot;\cdot) - h^M(\cdot;\cdot)}{\varepsilon^n} \right| \leq 1, \qquad \text{for } (\Delta t, \Delta x, \Delta y) \to 0, \quad (3.64)$$

in the region of regularity at time t_n, that is $\mathcal{R}^n = \left\{ (x_j, y_i) : \phi^n_{i,j} = 1 \right\}$. We recall that, in Section 2.2, we have defined a function ϕ such that

$$\phi^n_{i,j} = \begin{cases} 1 & \text{if the solution } u^n \text{ is regular in } I_{i,j}, \\ 0 & \text{if } I_{i,j} \text{ contains a point of singularity,} \end{cases} \quad (3.65)$$

with $I_{i,j} = (x_{j-1}, x_{j+1}) \times (y_{i-1}, y_{i+1})$.

Following the same computations of the one-dimensional case, we proceed by Taylor expansions, then for the monotone scheme we have

$$\begin{aligned} h^M(D_x^- v^n, D_x^+ v^n; D_y^- v^n, D_y^+ v^n) =& H(v_x^n, v_y^n) + \frac{\Delta x}{2} v_{xx}^n \left(\partial_{p_+} h^M_{i,j} - \partial_{p_-} h^M_{i,j} \right) \\ &+ \frac{\Delta y}{2} v_{yy}^n \left(\partial_{q_+} h^M_{i,j} - \partial_{q_-} h^M_{i,j} \right) (\Delta x^2) + O\left(\Delta y^2\right), \end{aligned}$$

while for the high-order scheme, by the consistency property,

$$\begin{aligned} h^A(D_x^{\pm} v^n_{i,j}; D_y^{\pm} v^n_{i,j}) =& H(v_x^n, v_y^n) - \frac{\Delta t}{2} \left[H_p^2(v_x^n, v_y^n) v_{xx}^n + H_q^2(v_x^n, v_y^n) v_{yy}^n \right. \\ &\left. + 2 H_p(v_x^n, v_y^n) H_q(v_x^n, v_y^n) v_{xy} \right] + O\left(\Delta t^2\right) + O\left(\Delta x^2\right) + O\left(\Delta y^2\right). \end{aligned}$$

Whence, from (3.64) we obtain

$$\begin{aligned} \varepsilon^n \geq \Big| & \frac{\Delta x}{2} v_{xx}^n \left(\partial_{p_+} h^M - \partial_{p_-} h^M + \lambda_x H_p^2(v_x^n, v_y^n) \right) + \frac{\Delta y}{2} v_{yy}^n \left(\partial_{q_+} h^M - \partial_{q_-} h^M + \lambda_y H_q^2(v_x^n, v_y^n) \right) \\ & + \Delta t v_{xy}^n H_p(v_x^n, v_y^n) H_q(v_x^n, v_y^n) + O(\Delta t^2) + O(\Delta x^2) + O(\Delta y^2) \Big|, \end{aligned}$$

that has to be satisfied in the *region of regularity at time* t^n, detected by the smoothness indicator function ϕ, that is

$$\mathcal{R}^n = \left\{(x_j, y_i) : \phi_{i,j}^n = 1\right\}. \tag{3.66}$$

Then, we use a numerical approximation of the lower bound on the right hand side of the previous inequality to obtain the formula for ε^n. In order to devise a simple formula, we introduce the notation

$$\widetilde{h}_{p^+}^M = h^M\left(D_x u^n, D_x^+ u^n, D_y u^n, D_y u^n\right) - h^M\left(D_x u^n, D_x^- u^n, D_y u^n, D_y u^n\right),$$

with the other cases following analogously. Finally, the simplest discretization, which we use in the numerical examples, is

$$\varepsilon^n = \max_{(x_j,y_i)\in\mathcal{R}^n} K\left|\frac{\Delta t}{2}\left(H_p^2 D_x^2 u^n + H_q^2 D_y^2 u^n + 2H_p H_q D_{xy}^2 u^n\right) + \right.$$
$$\left. + \left(\widetilde{h}_{p^+}^M - \widetilde{h}_{p^-}^M\right) + \left(\widetilde{h}_{q^+}^M - \widetilde{h}_{q^-}^M\right)\right|, \tag{3.67}$$

where all the derivatives of H are computed at $(D_x u^n, D_y u^n)$ and the finite difference approximations around the point (i, j), while $K > \frac{1}{2}$. Another possibility, which does not require the computation of the derivatives of H is the following

$$\varepsilon_2^n = \max_{(x_j,y_i)\in\mathcal{R}^n} K\left|H\left(D_x u^n, D_y u^n\right) - H\left(D_x u^n - \frac{\lambda_x}{2}H_x^*, D_y u^n - \frac{\lambda_y}{2}H_y^*\right) + \right.$$
$$\left. + \left(\widetilde{h}_{p^+}^M - \widetilde{h}_{p^-}^M\right) + \left(\widetilde{h}_{q^+}^M - \widetilde{h}_{q^-}^M\right)\right|, \tag{3.68}$$

where H_x^* and H_y^* are defined by (3.59)-(3.60). The only reason we avoid the use of the formula for ε_2^n, which is a more direct generalization of (3.33), is because it is not easy to keep the same formulation when adding also the dependence on the space variables (x, y), as we will be doing in Chapter 4.

***Remark* 3.7.** It is worth to note that all the formulas derived from (3.53), such as the 2D-Lax-Wendroff schemes and the definition of ε^n, easily extend to the case in which the hamiltonian depends also on the space variables $H(x, y, v_x, v_y)$. In fact, we can repeat the exact same computations of (3.53) till the third equation, then we simply have

$$\begin{aligned} v_{tt} &= H_p(x, y, v_x, v_y)\frac{\partial}{\partial x}\left(H(x, y, v_x, v_y)\right) + H_q(x, y, v_x, v_y)\frac{\partial}{\partial y}\left(H(x, y, v_x, v_y)\right) \\ &= H_p\left(H_x + H_p v_{xx}\right) + H_q\left(H_y + H_q v_{yy}\right) + 2H_p H_q v_{xy}, \end{aligned} \tag{3.69}$$

where in the last equality we have dropped all the dependence for brevity. Therefore, to adapt the previous formulas it is enough to add a centered approximation for H_x and H_y or, again, to use a direct approximation of the space derivatives in the first equation of (3.69). Notice also that the procedure described in Example 3.13, extends to this situation without relevant modifications. In fact, as for the monotone schemes, it is enough to consider the dependence of H on (x, y).

***Remark* 3.8.** Notice that, assuming to define the region of regularity $\mathcal{R}^n$ analogously to (3.31), we could state and prove a convergence result equivalent to Proposition 3.9 and Theorem 3.11 for the 2D-Adaptive Filtered Scheme defined in this section. We avoid to present the results since we would follow the exact same lines of the one-dimensional counterparts, at least for the proof of the convergence theorem. The analogous of Proposition 3.9 would need only minor modifications when proving the estimate for ε^n, which now involves the second order increments in both direction $D_x^2 u$, $D_y^2 u$ and the mixed finite difference $D_{xy}^2 u$.

3.5. Numerical tests

In this section we present some one-dimensional and two-dimensional examples designed to show the properties of our scheme, stated by Theorem 3.11. Our goal is also to compare the performances of our Adaptive Filtered Schemes S^{AF} with those of the Filtered Schemes S^F scheme of [BFS16] and of the WENO scheme of second/third order of [JP00]. Here we use the same efficient implementation suggested in [JP00] (Remark pp. 6-7) with $\sigma = 10^{-8}$, independent on Δx, avoiding any of the improvements proposed in [ABM10], [HAP05] or [CCD11]. To be precise, for the basic Filtered Scheme we use the implementation suggested in [BFS16] using also different high-order schemes, but avoiding the use of the limiter correction in all simulations. This is mainly because we want to show that the problems that were fixed through the introduction of the limiter in [BFS16], can be solved, at least partially, by the adaptive procedure and the function ϕ. Moreover, since the main aim is to use even higher order schemes (fourth order schemes) and to show the reliability of the adaptive tuning, the use of the limiter would be counterproductive, since it would inevitably limit also the full accuracy of the resulting scheme. This fact has been proved in various forms in literature (the scheme is TVD) and can be easily confirmed through some easy numerical tests, such as the transport of a regular function.

For each test we specify the monotone and high-order schemes composing the filtered schemes, as well as the the CFL number λ, and compute the errors and orders in L^∞ and L^1 norm. In particular, the CFL number will be chosen to satisfy the standard CFL condition

$$\lambda \max|H_p(p)| \leq 1, \tag{3.70}$$

which is in fact equivalent to (3.7) and more easily computed. Moreover, in two dimensions it simply becomes

$$\max\{\lambda_x \max|H_p(p,q)|, \lambda_y \max|H_q(p,q)|\} \leq \frac{1}{2}. \tag{3.71}$$

3.5.1. One-dimensional examples

Example 1. Transport equation.

In order to test the capability of our scheme to handle both regular and singular regions, let us begin with a simple linear example and consider the problem

$$\begin{cases} v_t(t,x) + v_x(t,x) = 0 & \text{in } (0,T)\times\Omega \\ v(0,x) = v_0(x), \end{cases}$$

with periodic boundary conditions, in two different situations. At first, aiming to test the full accuracy of the schemes, we consider the regular initial data (*Case a*),

$$v_0(x) = \sin(\pi x), \qquad x \in \Omega \tag{3.72}$$

with $\Omega = [-2,2]$ ans $T = 0.9$. Then, as a second test, we take the mixed initial data (*Case b*),

$$v_0(x) = \begin{cases} \min\{(1-x)^2, (1+x)^2\} & \text{if } -1 \le x \le 1, \\ \sin^2(\pi(x-2)) & \text{if } 2 \le x \le 3, \\ 0 & \text{otherwise,} \end{cases} \tag{3.73}$$

with $\Omega = [-1.5, 3.5]$ and $T = 2$. The latter problem models the transport of a function composed by two peaks, the first with one point of singularity while the second is in C^2. For these tests we use the *Central Upwind scheme* (3.9) as monotone scheme and the simple *Heun-Centered* (HC) scheme (3.16)-(3.17) as high-order scheme, with $\lambda = 0.9$ for *Case a* and $\lambda = 0.4$ for *Case b*. We also compare the results obtained using S^{AF} with the *4th order Lax-Wendroff* scheme (3.21) as high-order scheme. We recall that the latter high-order scheme has a very compact 5-points stencil, while the WENO scheme of second/third order (coupled with the *third order Runge Kutta scheme*) has a stencil of nine points.

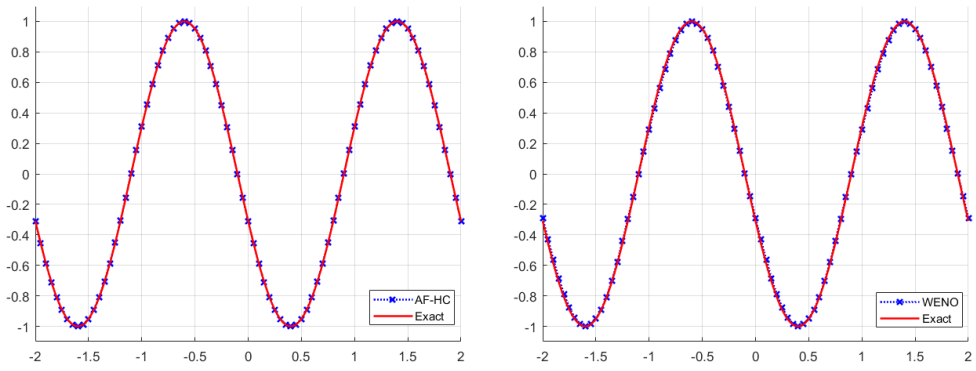

Figure 3.2.: (Example 1a.) Plots at time $T = 0.9$ with the AF-HC scheme on the left and WENO on the right for $\Delta x = 0.05$.

In the first test all the schemes are very accurate and achieve optimal order in both norms, as shown in Figure 3.2 and Table 3.1. In this case both filtered schemes have

Table 3.1.: (Example 1a.) Errors and orders in L^∞ and L^1 norms.

		F-HC ($5\Delta x$)		**AF-HC**		**AF-LW4ord**		**WENO 2/3**	
N_x	N_t	L^∞ Err	Ord	L^∞ Err	Ord	L^∞ Err	Ord	L^∞ Err	Ord
40	10	1.36e-02		1.36e-02		1.37e-04		8.02e-02	
80	20	2.56e-03	2.41	2.56e-03	2.41	8.66e-06	3.98	2.62e-02	1.62
160	40	5.76e-04	2.15	5.76e-04	2.15	5.43e-07	4.00	4.50e-03	2.54
320	80	1.40e-04	2.04	1.40e-04	2.04	3.40e-08	4.00	1.95e-04	4.52
N_x	N_t	L^1 Err	Ord	L^1 Err	Ord	L^1 Err	Ord	L^1 Err	Ord
40	10	3.58e-02		3.58e-02		3.62e-04		2.07e-01	
80	20	6.66e-03	2.43	6.66e-03	2.43	2.25e-05	4.01	4.14e-02	2.32
160	40	1.48e-03	2.17	1.48e-03	2.17	1.40e-06	4.01	5.09e-03	3.02
320	80	3.57e-04	2.05	3.57e-04	2.05	8.69e-08	4.01	3.08e-04	4.05

the same numerical results and coincide with the HC high-order scheme (we avoided to report the same results), as wanted. Moreover, we can see that our fourth order scheme is much more accurate even than the WENO scheme, despite the smaller stencil required.

Table 3.2.: (Example 1b.) Errors and orders in L^∞ and L^1 norms.

		F-HC ($10\Delta x$)		**AF-HC**		**AF-LW4ord**		**WENO 2/3**	
N_x	N_t	L^∞ Err	Ord	L^∞ Err	Ord	L^∞ Err	Ord	L^∞ Err	Ord
50	50	3.46e-01		2.98e-01		2.65e-01		3.47e-01	
100	100	1.41e-01	1.29	1.78e-01	0.75	1.56e-01	0.77	2.07e-01	0.75
200	200	9.69e-02	0.54	1.12e-01	0.66	9.08e-02	0.78	1.28e-01	0.70
400	400	7.29e-02	0.41	7.05e-02	0.67	5.06e-02	0.84	7.66e-02	0.74
N_x	N_t	L^1 Err	Ord	L^1 Err	Ord	L^1 Err	Ord	L^1 Err	Ord
50	50	4.34e-01		2.94e-01		2.21e-01		3.62e-01	
100	100	1.41e-01	1.63	9.77e-02	1.59	4.26e-02	2.38	1.39e-01	1.39
200	200	4.24e-02	1.73	3.06e-02	1.67	9.22e-03	2.21	3.83e-02	1.86
400	400	1.38e-02	1.62	1.01e-02	1.60	2.61e-03	1.82	8.39e-03	2.19

For the second case, looking at Figure 3.3 we can observe that the adaptive tuning of ε^n is able to contain the oscillations behind the peaks produced by the unstable HC scheme, which are clearly visible instead in the case of S^F with $\varepsilon = 5\Delta x$. We can also see that our scheme coupled with the fourth order scheme produces again the best results in terms of errors and orders in both norms (see Table 3.2) and gives the best resolution of the peaks, preserving better the kink of the singularity and the feet of the regular part, without introducing any oscillation.

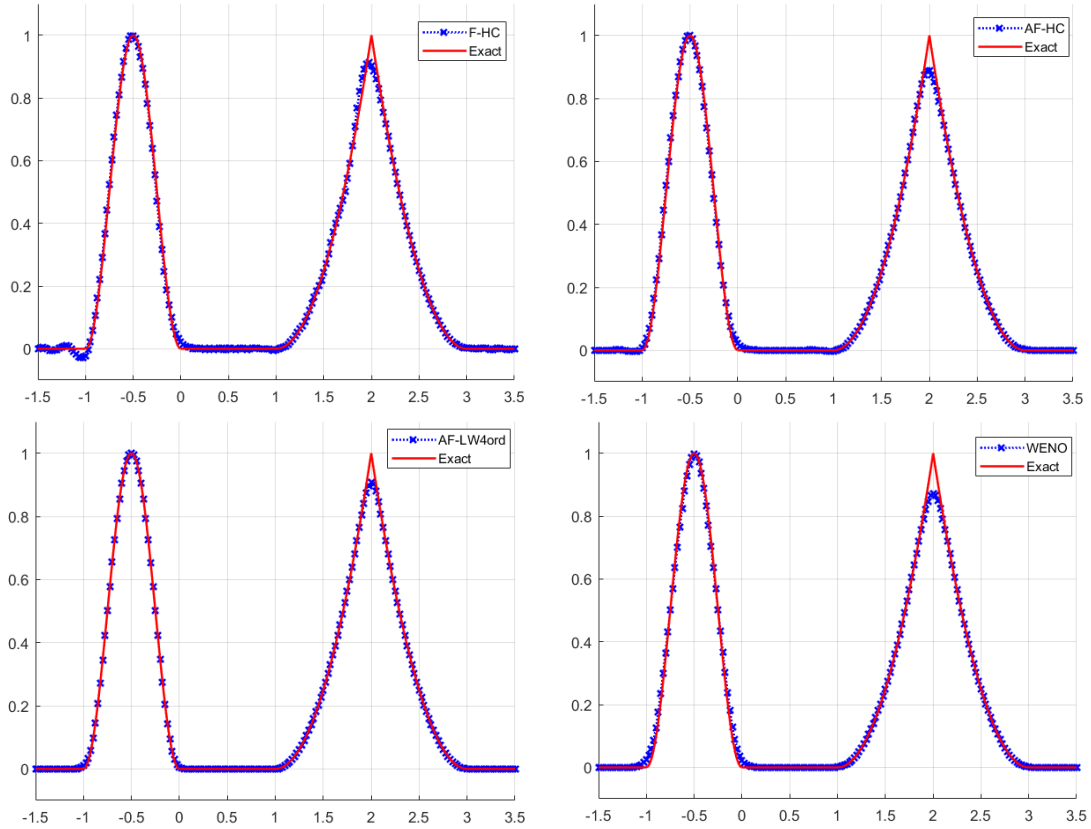

Figure 3.3.: (Example 1b.) Plots of the solution at time $T = 2$ with $\Delta x = 0.025$. Top: simple filtered scheme with HC on the left, adaptive on the right. Bottom: fourth order scheme AF scheme on the left and WENO on the right.

Example 2: Eikonal equation.

As a first nonlinear problem let us consider the eikonal equation

$$\begin{cases} v_t(t,x) + |v_x(t,x)| = 0 & \text{in } (0,0.3) \times (-2,2), \\ v_0(x) = \max\{1-x^2, 0\}^4, \end{cases} \tag{3.74}$$

where v_0 is a Lipschitz continuous initial data with high regularity. Then, we repeat the simulation with the "reversed" initial data

$$v_0(x) = -\max\{1-x^2, 0\}^4, \tag{3.75}$$

which presents also a major problem in the origin because of the saddle point in the hamiltonian, where two directions of propagation occur. Here the aim is mainly to compare the results obtained by the unfiltered high-order schemes with their filtered versions, in order to show the stabilization property of the filtering process. For the monotone scheme we use the numerical hamiltonian (3.8), while to achieve high-order we use the *Lax-Wendroff-Richtmyer (LWR)* scheme (3.20). Moreover, as in the previous

example, we present also the results obtained with the AF scheme coupled with the fourth order LW scheme. The CFL number is set to 0.375 for both simulations.

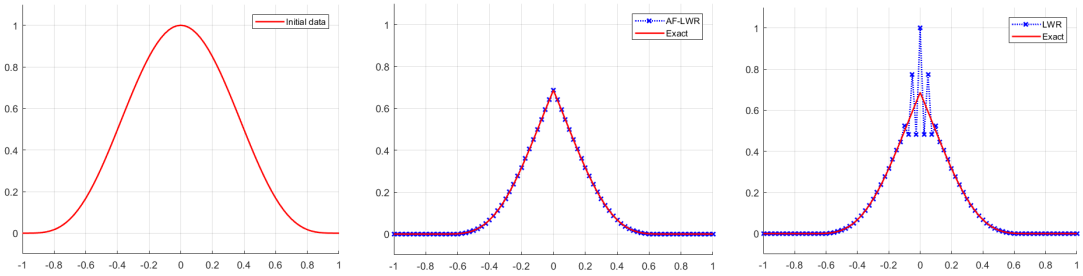

Figure 3.4.: (Example 2a.) Initial data (left) and plots of the solution at time $T = 0.3$ with the AF scheme (center) and the LWR scheme (right) for $\Delta x = 0.025$.

Table 3.3.: (Example 2a.) Errors and orders in L^∞ and L^1 norms.

		F-LWR ($5\Delta x$)		**AF-LWR**		**AF-LW4ord**		**WENO 2/3**	
N_x	N_t	L^∞ Err	Ord	L^∞ Err	Ord	L^∞ Err	Ord	L^∞ Err	Ord
40	8	1.96e-02		1.64e-02		2.18e-02		6.81e-02	
80	16	4.48e-03	2.13	4.00e-03	2.04	9.98e-03	1.13	3.42e-02	1.00
160	32	1.06e-03	2.08	1.11e-03	1.85	1.35e-03	2.89	1.62e-02	1.08
320	64	2.56e-04	2.05	2.56e-04	2.12	2.31e-04	2.55	7.52e-03	1.11
N_x	N_t	L^1 Err	Ord	L^1 Err	Ord	L^1 Err	Ord	L^1 Err	Ord
40	8	1.52e-02		1.16e-02		1.11e-02		2.05e-02	
80	16	3.78e-03	2.01	3.71e-03	1.65	1.05e-03	3.40	4.68e-03	2.13
160	32	8.94e-04	2.08	8.96e-04	2.05	7.28e-05	3.85	9.55e-04	2.29
320	64	2.09e-04	2.09	2.09e-04	2.10	7.14e-06	3.35	1.40e-04	2.78

Let us first point out that, as Figures 3.4-3.5 clearly show, the LWR scheme is unstable in the origin in both situations, while the AF scheme (and the simple filtered scheme) is not. Then, for the first case, looking at Table 3.3 we can see that the filtered-LWR schemes give almost the same results, are of high-order in both norms and have better errors even than the WENO scheme in almost all simulations. Moreover, we can recognize the typical, as will be shown also by the following examples, improvements and drawbacks of the fourth order LW scheme, which has a slightly wider stencil. In fact, the scheme has bigger errors in the L^∞ norm in the first three refinements of the grid, while has way better errors and orders in the L^1 norm, achieving almost optimal order, which testifies the overall improvement.

For Case b, as testified by Table 3.4, we can repeat almost the same considerations, but this time the improvements given by the adaptive filtering are evident. The AF-LWR scheme is again of high-order especially in the L^1 norm, without the need to introduce any limiter as has been done in [BFS16], and the numerical results are always comparable to those obtained by the WENO scheme of second/third order, while the

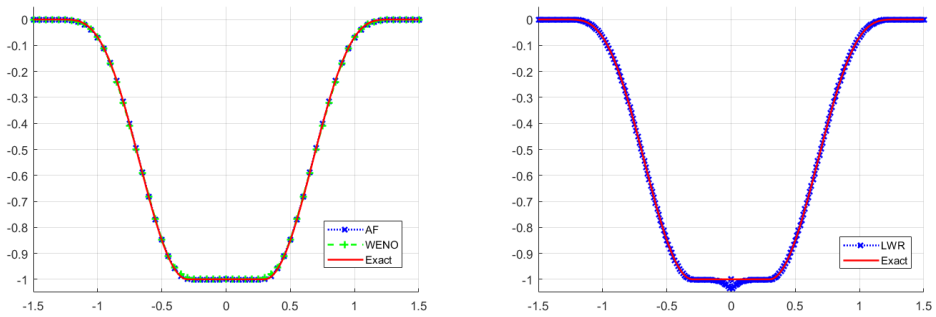

Figure 3.5.: (Example 2b.) Plots at time $T = 0.3$ with the AF and WENO schemes for $\Delta x = 0.05$ (left) and LWR scheme for $\Delta x = 0.0125$.

Table 3.4.: (Example 2b.) Errors and orders in L^∞ and L^1 norms.

		F-LWR $(5\Delta x)$		**AF-LWR**		**AF-LW4ord**		**WENO 2/3**	
N_x	N_t	L^∞ Err	Ord	L^∞ Err	Ord	L^∞ Err	Ord	L^∞ Err	Ord
40	8	1.91e-02		1.40e-02		1.63e-02		2.33e-02	
80	16	9.24e-03	1.04	3.37e-03	2.06	7.51e-03	1.11	1.02e-02	1.19
160	32	5.77e-03	0.68	1.58e-03	1.09	2.14e-03	1.81	4.10e-03	1.32
320	64	3.46e-03	0.74	7.09e-04	1.16	6.92e-04	1.63	1.22e-03	1.75
N_x	N_t	L^1 Err	Ord	L^1 Err	Ord	L^1 Err	Ord	L^1 Err	Ord
40	8	2.38e-02		2.01e-02		1.29e-02		2.96e-02	
80	16	8.48e-03	1.49	5.70e-03	1.82	2.05e-03	2.65	7.04e-03	2.07
160	32	3.41e-03	1.32	1.82e-03	1.65	3.20e-04	2.68	1.43e-03	2.30
320	64	1.52e-03	1.17	5.84e-04	1.64	6.38e-05	2.33	2.82e-04	2.34

AF-LW4ord scheme has again worse L^∞ errors for the first discretizations and better errors and orders in the L^1 norm.

Example 3: Burgers' equation.

Let us consider now the Burgers' equation for HJ with a regular initial data

$$\begin{cases} v_t(t,x) + \frac{1}{2}(v_x(t,x)+1)^2 = 0 & \text{in } (0,T)\times(0,2), \\ v_0(x) = -\cos(\pi x), \end{cases} \tag{3.76}$$

which is a test case widely used in literature. In order to test the full accuracy of the schemes even in the nonlinear case, we first run the simulation for $T = \frac{4}{5\pi^2}$, when the solution is still regular, with $\lambda = \frac{2}{\pi^2} \approx 0.2 < \max|H_p|^{-1} = 0.5$. Then, we consider the final time $T = \frac{3}{2\pi^2}$ when a moving (to the right) singularity appears, taking $\lambda = \frac{15}{8\pi^2} \approx 0.19$. For both simulations we use the the Central Upwind monotone scheme and the LWR scheme for both the filtered schemes and compare the results as before with the WENO scheme and the fourth order AF scheme.

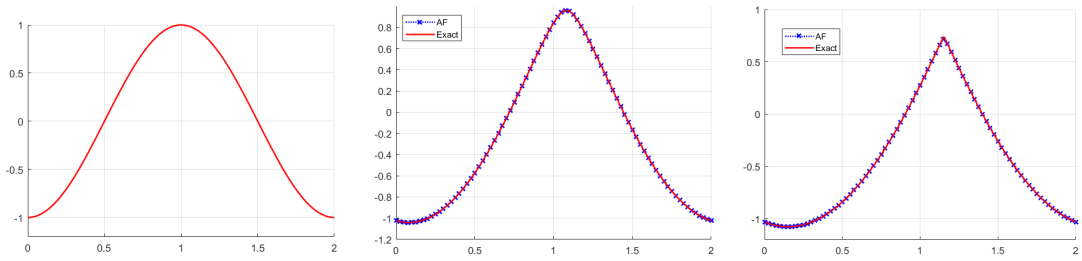

Figure 3.6.: (Example 3.) From left to right: initial data of problem (3.76) and plots of the solution at time $T = 4/(5\pi^2)$ and $T = 3/(2\pi^2)$ for $\Delta x = 0.025$.

Table 3.5.: (Example 3.) $T = 4/(5\pi^2)$. Errors and orders in L^∞ and L^1 norms.

		F-LWR $(10\Delta x)$		**AF-LWR**		**AF-LW4ord**		**WENO 2/3**	
N_x	N_t	L^∞ Err	Ord	L^∞ Err	Ord	L^∞ Err	Ord	L^∞ Err	Ord
40	8	1.30e-02		9.61e-03		1.89e-03		1.04e-02	
80	16	8.67e-03	0.59	2.77e-03	1.79	2.84e-04	2.73	2.12e-03	2.30
160	32	5.07e-03	0.77	7.24e-04	1.94	2.68e-05	3.41	1.82e-04	3.54
320	64	2.66e-03	0.93	1.83e-04	1.99	1.89e-06	3.83	2.05e-05	3.15
N_x	N_t	L^1 Err	Ord	L^1 Err	Ord	L^1 Err	Ord	L^1 Err	Ord
40	8	3.76e-03		3.30e-03		2.76e-04		3.67e-03	
80	16	1.29e-03	1.54	8.20e-04	2.01	1.97e-05	3.81	6.57e-04	2.48
160	32	4.49e-04	1.52	2.04e-04	2.01	1.50e-06	3.71	5.43e-05	3.60
320	64	1.82e-04	1.30	5.09e-05	2.00	1.04e-07	3.86	2.98e-06	4.19

This example summarizes all the behaviors already seen in the previous cases. In fact, as displayed by Tables 3.5-3.6, if the solution is still regular the fourth order AF scheme gives the best results and achieves the optimal order in both norms, while when the singularity appears has the usual problems in the L^∞ norm and better orders (than the second order scheme) in the L^1 norm. Here we have to notice that the WENO scheme has better errors and orders in the second simulation w.r.t. all the filtered schemes. Moreover, we can clearly see that the simple filtered scheme depends heavily on the choice of ε, in fact after extensive computations we noticed that choosing for example $\varepsilon = 5\Delta x$ we get worse results in both cases, while if we increase the constant we get better results in the regular case and worse in the latter. In the tables we presented the results for the choice that gives the best results in the singular case, while it has clearly problems in the first situation. This is the main advantage of the adaptive ε^n which is able to tune itself in the right way depending on the local (in time) regularity of the solution.

Table 3.6.: (Example 3.) $T = 3/(2\pi^2)$. Errors and orders in L^∞ and L^1 norms.

		F-LWR $(10\Delta x)$		**AF-LWR**		**AF-LW4ord**		**WENO 23**	
N_x	N_t	L^∞ Err	Ord	L^∞ Err	Ord	L^∞ Err	Ord	L^∞ Err	Ord
40	16	4.88e-02		5.30e-02		6.31e-02		3.89e-02	
80	32	2.47e-02	0.98	2.47e-02	1.10	2.87e-02	1.13	1.61e-02	1.27
160	64	9.81e-03	1.33	9.95e-03	1.31	1.03e-02	1.48	5.12e-03	1.65
320	128	2.57e-03	1.93	2.59e-03	1.94	2.69e-03	1.94	8.40e-04	2.61
N_x	N_t	L^1 Err	Ord	L^1 Err	Ord	L^1 Err	Ord	L^1 Err	Ord
40	16	5.17e-03		5.28e-03		3.83e-03		3.69e-03	
80	32	1.26e-03	2.03	1.27e-03	2.06	8.89e-04	2.11	6.94e-04	2.41
160	64	2.86e-04	2.14	2.87e-04	2.14	1.43e-04	2.64	8.67e-05	3.00
320	128	5.68e-05	2.33	5.68e-05	2.34	1.82e-05	2.97	6.40e-06	3.76

3.5.2. Two-dimensional examples

Example 4: Rotation in two dimensions.

In this example we show how to use our one-dimensional schemes to solve simple problems in two dimensions. We consider the problem

$$\begin{cases} v_t - yv_x + xv_y = 0 & \text{in } (0, 2\pi) \times \Omega, \\ v(0, x, y) = \max\left\{0, \frac{r_0 - (x+1)^2 - y^2}{1 - r_0^2}\right\}^2, \end{cases} \tag{3.77}$$

with $r_0 = 0.5$ and $\Omega = [-2.5, 2.5]^2$, which models the rotation of a C^1 function around the origin. In this situation, since the hamiltonian $H(x, y, v_x, v_y) = -yv_x + xv_y$ can be expressed as a sum of one-dimensional hamiltonian, depending on the evolution along the x and y direction, respectively, we can use a dimensional splitting to solve the problem. More precisely, if we write $H(x, y, v_x, v_y) = H_1(y, v_x) + H_2(x, v_y)$, we can approximate the solution by solving sequentially the problems in one space dimension

$$v_t + H_1(y, v_x) = 0 \qquad \text{and} \qquad v_t + H_2(x, v_y) = 0,$$

keeping each time the other space variable constant. Although the hamiltonians do not commute, has can be seen computing the Lie bracket $[H_1, H_2] = -xv_x - yv_y \neq 0$, we can still use the dimensional splitting has shown previously, but we should not expect the simple *Lie-Trotter* splitting

$$u^{n+1} = S_y^{\Delta t}\left(S_x^{\Delta t}(u^n)\right), \tag{3.78}$$

where $S_x^{\Delta t}$ and $S_y^{\Delta t}$ are numerical schemes of time step Δt for the problems in the x and y direction, respectively, to be more than first order accurate in time. Nevertheless, we perform the simulations using the one-dimensional second order filtered schemes of the previous tests.

For this test we use the Central Upwind monotone scheme and the Heun-Centered scheme as high-order scheme and compare the results obtained by the one-dimensional

WENO scheme in both directions. We use a square uniform grid with the CFL number set to $\lambda = \frac{\pi}{8} \approx 0.39 < 0.4 = \max\{H_p\}^{-1} = \max\{H_q\}^{-1}$.

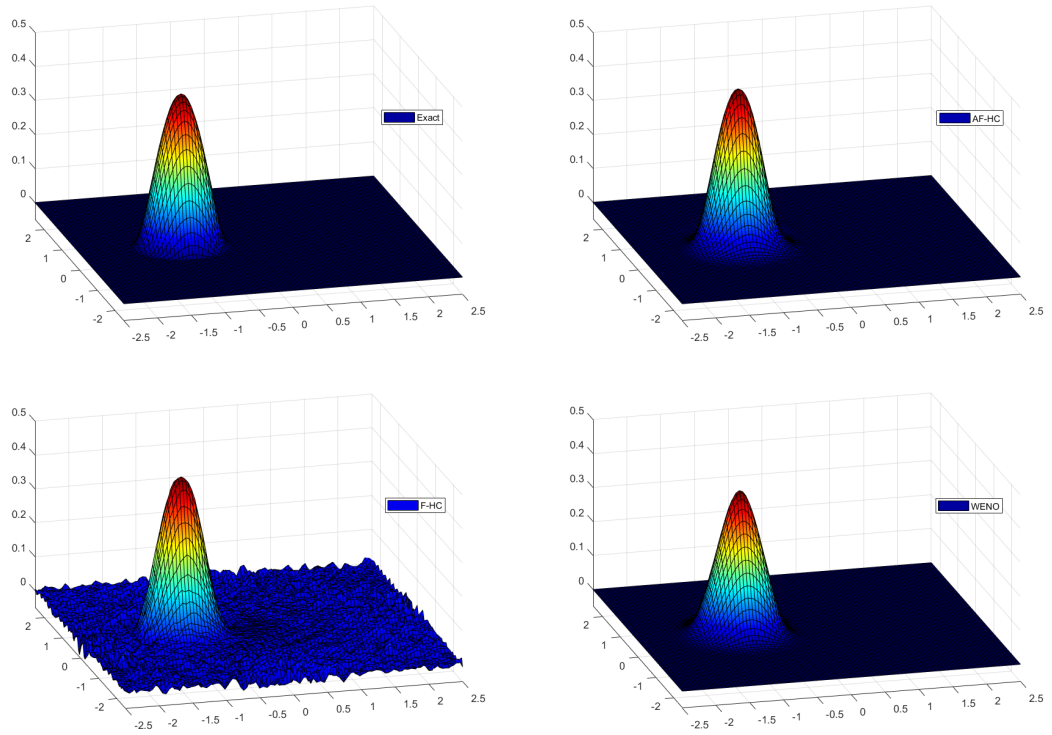

Figure 3.7.: (Example 4.)Plots of exact and computed solution (AF top-right, F bottom-left, WENO bottom-right) at time $T = 2\pi$ with $\Delta x = 0.05$.

If we look at Figure 3.7 we can clearly see the advantages brought by the automatic tuning of the parameter ε^n and the stabilizing properties of the ϕ function. In fact, our AF scheme is able to almost completely nullify the oscillations caused by the unstable HC scheme (see Table 3.7), while the simple F scheme with $\varepsilon = 5\Delta x$ cannot. Moreover, in the latter case the oscillations keep on being amplified as time goes on, reducing the effective accuracy of the scheme. Despite this graphically evident improvement, the errors and orders of the filtered schemes are rather close, but the adaptive version has always better results, which are also close to those of the WENO scheme, especially in the L^1 norm.

Moreover, reminding the discussion at the beginning of the example and looking at Table 3.7, we can observe rather surprising high orders for all the schemes. This is probably due to radial symmetry of the evolution and of the function and does not happen when using the AF-LW4ord scheme, which instead gives very poor results.

Table 3.7.: (Example 4.) Errors and orders in L^∞ and L^1 norms.

		HC		**F-HC** ($5\Delta x$)		**AF-HC**		**WENO 2/3**	
N_x	N_t	L^∞ Err	Ord	L^∞ Err	Ord	L^∞ Err	Ord	L^∞ Err	Ord
50	160	2.44e-01		1.78e-01		1.53e-01		1.82e-01	
100	320	9.10e+03	−15.18	6.26e-02	1.51	5.71e-02	1.43	4.49e-02	2.02
200	640	7.30e+16	−42.87	2.02e-02	1.64	1.95e-02	1.55	1.32e-02	1.76
400	1280	1.87e+60	−144.20	8.64e-03	1.22	6.34e-03	1.62	3.32e-03	1.99
N_x	N_t	L^1 Err	Ord	L^1 Err	Ord	L^1 Err	Ord	L^1 Err	Ord
50	160	4.98e-01		1.32e-01		1.21e-00		1.42e-01	
100	320	2.36e+03	−12.21	4.42e-02	1.58	3.96e-02	1.61	4.77e-02	1.57
200	640	3.99e+15	−40.62	1.35e-02	1.71	1.16e-02	1.77	1.02e-02	2.22
400	1280	3.79e+58	−142.77	3.99e-03	1.75	3.32e-03	1.81	1.22e-03	3.07

Example 5: Generating singularities in two dimensions.

In this example we consider a problem similar to the Burgers' equation in two dimensions, which is strictly connected to (3.76),

$$\begin{cases} v_t + (v_x + 1)^2 + (v_y + 1)^2 = 0 & \text{in } (0,T) \times \Omega, \\ v(0,x,y) = -0.5\left(\cos(\pi x) + \cos(\pi y)\right), \end{cases} \tag{3.79}$$

with $\Omega = [0,2]^2$ and periodic boundary conditions. As for problem (3.76), we consider the final time $T = \frac{4}{5\pi^2}$, when the solution is still smooth, and then $T = \frac{3}{2\pi^2}$, time at which an interesting set of singularities develops. It is clear that we can use the dimensional splitting also in this situation and solve the problem using one-dimensional schemes and the Lie-Trotter splitting, since the hamiltonians H_1 and H_2 commute. We use the same schemes as in the previous example and a slightly more restrictive CFL number w.r.t. problem (3.76) in order to use coarser grids, which is set to $\lambda = \frac{4}{5\pi^2} \approx 0.08$ for the first test, and $\lambda = \frac{3}{4\pi^2} \approx 0.076$ for the latter.

The exact solution is computed by the *Hopf-Lax* formula,

$$v(t,x,y) = \left(\min_{a \in A} -\frac{1}{2}\cos(\pi(x - at)) + \frac{1}{4}a^2 - a + \min_{b \in A} -\frac{1}{2}\cos(\pi(y - bt)) + \frac{1}{4}b^2 - b \right),$$

with $A = [-5,5]$, where we used the fact that the evolution can be seen as the sum of separate one-dimensional evolutions.

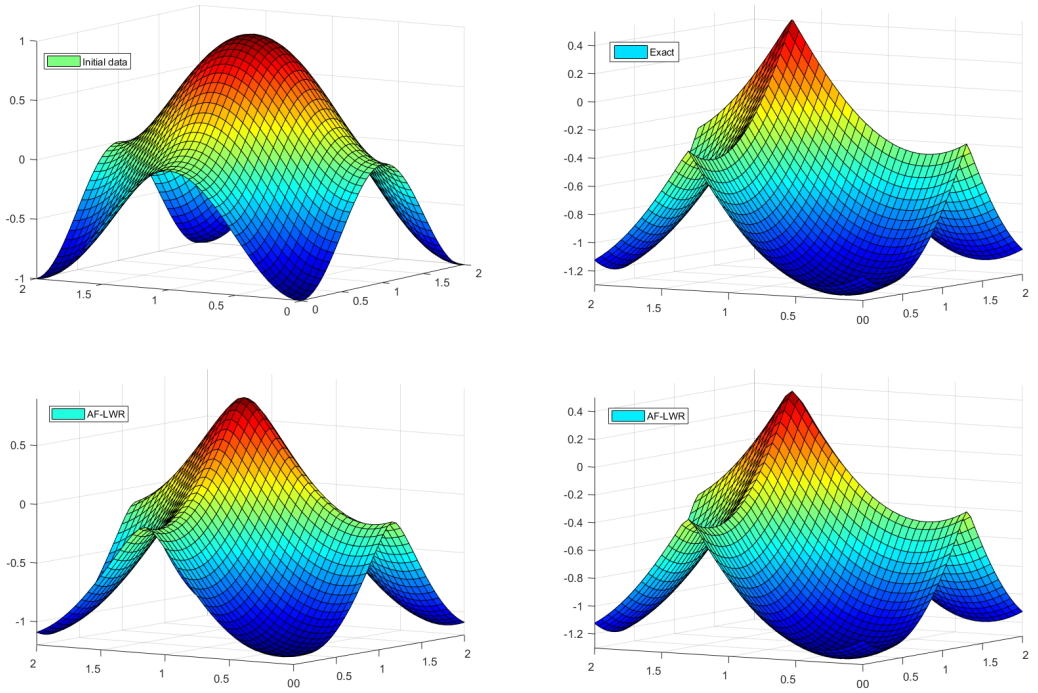

Figure 3.8.: (Example 5.) Top: Initial data (left) and exact solution at $T = 3/(2\pi^2)$ (right). Bottom: solution at $T = 4/(5\pi^2)$ (left) and $T = 3/(2\pi^2)$ (right) computed by the AF scheme with $\Delta x = 0.1$

Table 3.9.: (Example 5.) $T = 3/(2\pi^2)$. Errors and orders in L^∞ and L^1 norms.

		LWR		**F-LWR** $(10\Delta x)$		**AF-LWR**		**WENO 2/3**	
N_x	N_t	L^∞ Err	Ord	L^∞ Err	Ord	L^∞ Err	Ord	L^∞ Err	Ord
20	20	1.69e-01		9.49e-02		1.04e-01		8.68e-02	
40	40	6.39e-02	1.40	3.67e-02	1.37	3.80e-02	1.45	2.27e-02	1.93
80	80	3.23e-02	0.98	1.41e-02	1.38	1.47e-02	1.37	9.08e-03	1.32
160	160	2.64e-02	0.29	3.73e-03	1.91	3.90e-03	1.92	2.22e-03	2.03
N_x	N_t	L^1 Err	Ord	L^1 Err	Ord	L^1 Err	Ord	L^1 Err	Ord
20	20	6.18e-02		4.62e-02		4.62e-02		3.60e-02	
40	40	1.74e-02	1.83	8.19e-03	2.50	8.24e-03	2.49	4.68e-03	2.94
80	80	4.54e-03	1.94	1.88e-03	2.13	1.80e-03	2.20	6.92e-04	2.76
160	160	1.13e-03	2.00	3.73e-04	2.33	3.78e-04	2.25	7.62e-05	3.18

As we could expect, in this example we have analogous result w.r.t. Example 3, with the AF scheme performing well in both situations and better than the F scheme in the regular case (see Tables 3.8-3.9). Here again the simple filtered scheme has slightly

Table 3.8.: (Example 5.) $T = 4/(5\pi^2)$. Errors and orders in L^∞ and L^1 norms.

		LWR		**F-LWR** $(10\Delta x)$		**AF-LWR**		**WENO 2/3**	
N_x	N_t	L^∞ Err	Ord	L^∞ Err	Ord	L^∞ Err	Ord	L^∞ Err	Ord
20	10	7.45e-02		7.75e-02		8.32e-02		8.66e-02	
40	20	3.38e-02	1.14	5.12e-02	0.60	2.77e-02	1.59	3.59e-02	1.27
80	40	1.49e-02	1.18	3.25e-02	0.66	6.58e-03	2.08	1.30e-02	1.47
160	80	6.42e-03	1.22	1.94e-02	0.75	1.78e-03	1.89	4.87e-03	1.41
N_x	N_t	L^1 Err	Ord	L^1 Err	Ord	L^1 Err	Ord	L^1 Err	Ord
20	10	3.67e-02		4.42e-02		3.72e-02		3.71e-02	
40	20	9.53e-03	1.94	1.21e-02	1.87	8.50e-03	2.13	1.00e-02	1.89
80	40	2.28e-03	2.06	4.29e-03	1.49	2.05e-03	2.05	1.95e-03	2.36
160	80	6.51e-04	1.81	1.70e-03	1.33	5.42e-04	1.92	4.50e-04	2.11

better results after the singularities develop, due to the action of the ϕ function in the regions of singularity, but the loss of accuracy is in fact minimal. Moreover, our scheme performs as good as the WENO scheme when the solution is still regular, while the latter performs much better in the second case.

To conclude the section, in the following examples we perform the analysis of the schemes defined in Section 3.4, with the same aim of the previous tests. We compare the 2D-AFS scheme with the basic 2D-F scheme with $\varepsilon = 20\Delta x$ using the same composing schemes, again avoiding the use of the limiting correction of [BFS16], as well as with the 2D-WENO scheme of second/third order, which uses the same monotone numerical hamiltonian. All the following simulations will be performed using the *local Lax-Friedrichs scheme* (3.49) as monotone scheme, while for the high-order schemes we focus on the centered schemes of second (3.56) and fourth order (3.62)-(3.63), named *fourth-order centered Runke-Kutta* (RKC4 in short). The comparison is conducted also in terms of elapsed CPU-time, since when working with multidimensional problems it becomes a major issue.

Example 6. Transport of a regular function in 2D

As in the one-dimensional case, we begin the analysis by testing the full accuracy of the schemes on the transport of a very regular function at constant velocity. That is, we solve the following problem

$$\begin{cases} v_t + v_x + v_y = 0, & \text{in } (0,T)\times\Omega, \\ v(0,x,y) = v_0(x,y), & \text{in } \Omega \end{cases} \tag{3.80}$$

where $\Omega = [-2,2]^2$ and $T = 0.9$, with the regular initial condition

$$v_0(x,y) = \max\left\{0, 1 - x^2 - y^2\right\}^4. \tag{3.81}$$

The CFL number is $\lambda = 0.2 < \frac{1}{2}$. The results showed in Table 3.10 clearly testify the success of the filtering process, with the filtered schemes achieving the optimal order of

the high-order scheme in both norms. Only the fourth-order scheme seems to have some problems in the L^∞ norm, although the results obtained are evidently superior to all the other schemes. Notice also that, as it is expected, the AF-HC basically coincides with HC when the solution is regular. Moreover, in Table 3.11 we can see that the adaptive filtering almost doubles the computational cost of the scheme, while the further increase produced by the higher-order scheme seems almost negligible. On the other hand, the implementation of the WENO scheme is evidently more demanding.

Table 3.10.: (Example 6.) Errors and orders in L^∞ and L^1 norms.

		HC		**AF-HC**		**AF-RKC4**		**WENO 2/3**	
N_x	N_t	L^∞ Err	Ord	L^∞ Err	Ord	L^∞ Err	Ord	L^∞ Err	Ord
40	30	7.14e-02		7.14e-02		4.69e-03		2.88e-02	
80	60	1.63e-02	2.13	1.63e-02	2.13	5.28e-04	3.15	2.82e-03	3.35
160	120	4.02e-03	2.02	4.02e-03	2.02	5.38e-05	3.29	3.46e-04	3.03
320	240	9.99e-04	2.01	9.99e-04	2.01	6.44e-06	3.06	4.33e-05	3.00
N_x	N_t	L^1 Err	Ord	L^1 Err	Ord	L^1 Err	Ord	L^1 Err	Ord
40	30	6.84e-02		6.88e-02		7.94e-03		2.79e-02	
80	60	1.66e-02	2.05	1.66e-02	2.05	6.34e-04	3.65	3.50e-03	2.99
160	120	4.03e-03	2.04	4.03e-03	2.04	4.64e-05	3.77	4.37e-04	3.00
320	240	9.96e-04	2.02	9.96e-04	2.02	3.28e-06	3.82	5.40e-05	3.02

Table 3.11.: (Example 6.) CPU times in seconds.

N_x	N_t	**HC**	**AF-HC**	**AF-RKC4**	**WENO 2/3**
40	30	0.053 s	0.087 s	0.098 s	0.41 s
80	60	0.385 s	0.648 s	0.723 s	2.96 s
160	120	2.797 s	5.108 s	5.675 s	43.38 s
320	240	22.18 s	40.85 s	46.11 s	674.5 s

Example 7: Eikonal equation in two dimensions.

In the last example we consider the equivalent of equation (3.74) in two dimensions,

$$\begin{cases} v_t + \sqrt{v_x^2 + v_y^2} = 0 & \text{in } (0,T)\times\Omega, \\ v(0,x,y) = v_0(x,y), \end{cases} \tag{3.82}$$

where $\Omega = [-3,3]^2$ and $T = 0.6$, in four different situations. Let us first recall that this problem comes out in front propagation problems through the level set method as briefly shown in Section 1.3. Here we are considering a simple expansion with constant velocity. Our aim is to compare the schemes in terms of the computed error and of the resolution of the 0-level set, in both cases of a single front and of a merging, also varying the regularity of the representation function v_0. The CFL number is set to $\lambda = 0.25 < \frac{1}{2}$

for all the simulations. In the first two tests we perform the evolution of a single front, for a regular initial datum (Case a)

$$v_0(x,y) = 0.5 - 0.5\left(\max\left\{0, \frac{1-x^2-y^2}{1-r_0^2}\right\}\right)^4,$$

and for a sharper initial condition (Case b)

$$v_0(x,y) = \min\left\{|x|, |y|, 0.2\right\}.$$

Our objective is to inspect the behavior of the schemes when varying the "number of singularities" in the evolution. One is brought directly by the hamiltonian, since it is only Lipschitz continuous and presents a saddle point in the origin, then, others may be already present in the initial datum or caused by some merging. In this context, the first simulation is similar to an accuracy test, in fact, looking at Table 3.12 we can acknowledge the better results given by the fourth order scheme in the L^1 norm, whereas the more compact AF-HC scheme performs better in L^∞, analogously to what happens in the one-dimensional case (see Example 2b). Moreover, looking at the results of the basic filtered scheme, we may have a first confirmation of the stabilizing properties of the ϕ function, even in the two space dimensions, without the need to introduce any limiter correction. This fact is recognizable also in Case b, in the sense that it prevents the oscillating behavior (of the numerical order of convergence) of the basic filtered scheme in Table 3.14, especially in the L^∞ norm, although the latter seems to produce better results for the first refinements. Nevertheless, the overall behavior is very similar to the Case a, with the standard WENO schemes losing in terms of both norms w.r.t. the adaptive filtered schemes.

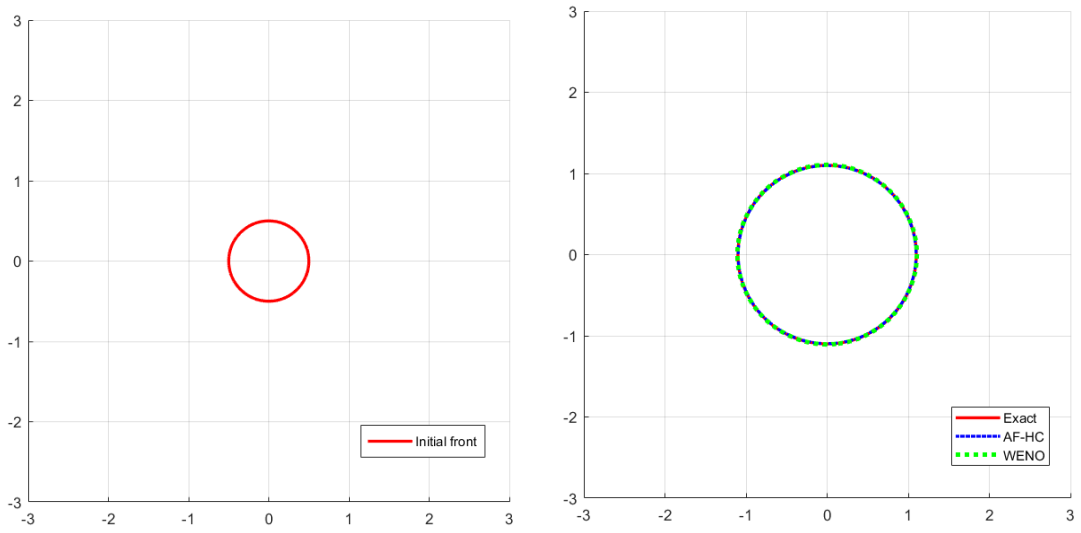

Figure 3.9.: (Example 7a.) Left: the initial front. Right: comparison of the 0-level sets at $T = 0.6$ using the AF-HC and WENO scheme with $\Delta x = \Delta y = 0.05$.

Table 3.12.: (Example 7a.) Errors and orders in L^∞ and L^1 norms.

		F-HC ($20\Delta x$)		**AF-HC**		**AF-RKC4**		**WENO 2/3**	
N_x	N_t	L^∞ Err	Ord	L^∞ Err	Ord	L^∞ Err	Ord	L^∞ Err	Ord
30	12	1.90e-01		2.33e-01		2.22e-01		2.22e-01	
60	24	1.05e-01	0.86	4.80e-02	2.28	4.62e-02	2.27	7.05e-02	1.65
120	48	6.83e-02	0.62	1.78e-02	1.43	1.88e-02	1.30	2.54e-02	1.47
240	96	4.66e-02	0.55	8.56e-03	1.06	8.77e-03	1.10	9.81e-03	1.37
N_x	N_t	L^1 Err	Ord	L^1 Err	Ord	L^1 Err	Ord	L^1 Err	Ord
30	12	6.44e-01		7.90e-01		6.14e-01		7.39e-01	
60	24	2.25e-01	1.52	1.92e-01	2.04	6.95e-02	3.14	2.48e-01	1.58
120	48	7.51e-02	1.58	5.32e-02	1.85	1.73e-02	2.01	5.67e-02	2.13
240	96	2.83e-02	1.41	1.55e-02	1.78	6.20e-03	1.48	1.01e-02	2.49

Table 3.13.: (Example 7a.) CPU times in seconds.

N_x	N_t	**F-HC**	**AF-HC**	**AF-RKC4**	**WENO 2/3**
30	12	0.015 s	0.034 s	0.027 s	0.042 s
60	24	0.077 s	0.157 s	0.172 s	0.509 s
120	48	0.569 s	1.194 s	1.317 s	7.577 s
240	96	4.338 s	9.219 s	10.04 s	109.1 s

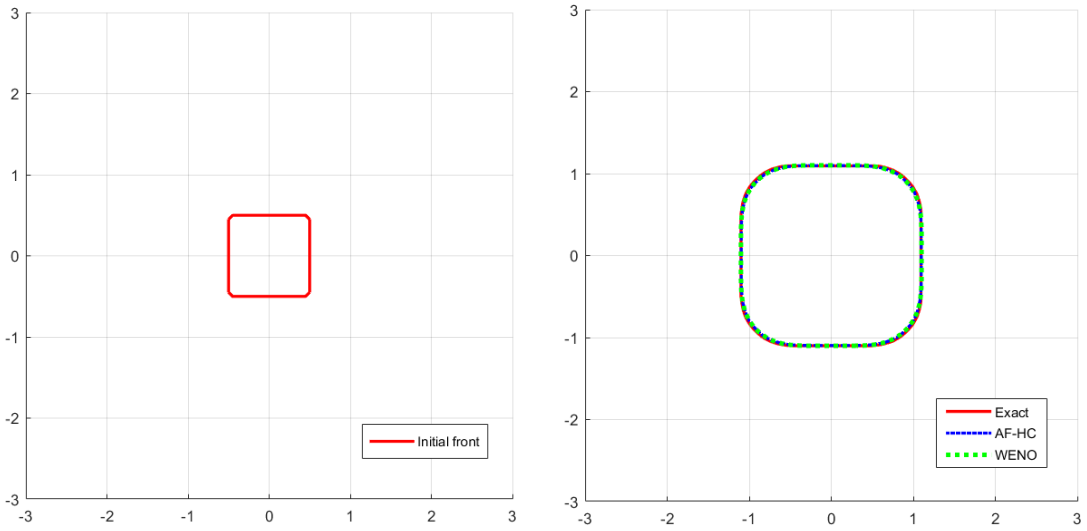

Figure 3.10.: (Example 7b.) Left: the initial front. Right: comparison of the 0-level sets at $T = 0.6$ using the AF-HC and WENO scheme with $\Delta x = \Delta y = 0.05$.

Finally, we complete our analysis inspecting the case of merging fronts, first with two collapsing regular representations (Case c), which 0-level set is composed by two circles,

Table 3.14.: (Example 7b.) Errors and orders in L^∞ and L^1 norms.

		F-HC ($20\Delta x$)		**AF-HC**		**AF-RKC4**		**WENO 2/3**	
N_x	N_t	L^∞ Err	Ord	L^∞ Err	Ord	L^∞ Err	Ord	L^∞ Err	Ord
30	12	8.11e-02		8.88e-02		1.00e-01		1.25e-01	
60	24	6.22e-02	0.38	6.72e-02	0.40	6.29e-02	0.67	7.34e-02	0.76
120	48	7.75e-02	−0.32	3.92e-02	0.78	4.04e-02	0.64	4.42e-02	0.73
240	96	5.38e-02	0.53	2.30e-02	0.77	2.39e-02	0.76	2.63e-02	0.75
N_x	N_t	L^1 Err	Ord	L^1 Err	Ord	L^1 Err	Ord	L^1 Err	Ord
30	12	2.58e-01		3.63e-01		2.50e-01		4.52e-01	
60	24	1.82e-01	0.51	2.04e-01	0.83	1.80e-01	0.47	2.15e-01	1.07
120	48	9.38e-02	0.95	9.76e-02	1.06	9.01e-02	1.00	9.60e-02	1.16
240	96	5.23e-02	0.84	3.86e-02	1.34	3.79e-02	1.25	3.98e-02	1.27

Table 3.15.: (Example 7b.) CPU times in seconds.

N_x	N_t	**F-HC**	**AF-HC**	**AF-RKC4**	**WENO 2/3**
30	12	0.013 s	0.023 s	0.025 s	0.041 s
60	24	0.071 s	0.167 s	0.172 s	0.501 s
120	48	0.479 s	1.143 s	1.274 s	7.887 s
240	96	3.825 s	9.048 s	9.89 s	110.3 s

that is

$$v_0(x,y) = 0.5 - 0.5\max\left(\max(0,f_-)^4, \max(0,f_+)^4\right),$$
$$f_\pm = \frac{1 - \left(x \pm \frac{\sqrt{2}}{2}\right)^2 - \left(y \pm \frac{\sqrt{2}}{2}\right)^2}{1 - r_0^2}, \qquad r_0 = 0.5,$$

then, we consider also a couple of sharp representations (Case d), which 0-level sets, composed by two squares, presents some marked corners,

$$v_0(x,y) = \min\left\{f_1 - r_0, f_2 - r_0, \frac{1}{2}r_0^2\right\}, \quad \text{with} \quad f_1 = \max\left\{\left|x - \frac{\sqrt{2}}{2}\right|, \left|y - \frac{\sqrt{2}}{2}\right|\right\},$$
$$f_2 = \max\left\{\left|\left(\sqrt{r_0}x + \frac{\sqrt{2}}{2}\right) + \left(\sqrt{r_0}y + \frac{\sqrt{2}}{2}\right)\right|, \left|\left(\sqrt{r_0}x + \frac{\sqrt{2}}{2}\right) - \left(\sqrt{r_0}y + \frac{\sqrt{2}}{2}\right)\right|\right\},$$

where $r_0 = 0.5$ is a parameter mainly needed to control magnitude of the square 0-level front. For Case b the solution is computed for $T = 0.7$, in order to have the two fronts merge. The results obtained are collected in Tables 3.16-3.18 and Figures 3.11, 3.12 and 3.13, in which we visualize the merging through both the representations and 0-level sets. It is clear that we can repeat the same observations of the previous cases, following almost the same lines, also in presence of a merging. The most interesting results are given by Table 3.18 and Figure 3.12, in which the AF-RKC4 proves its reliability in

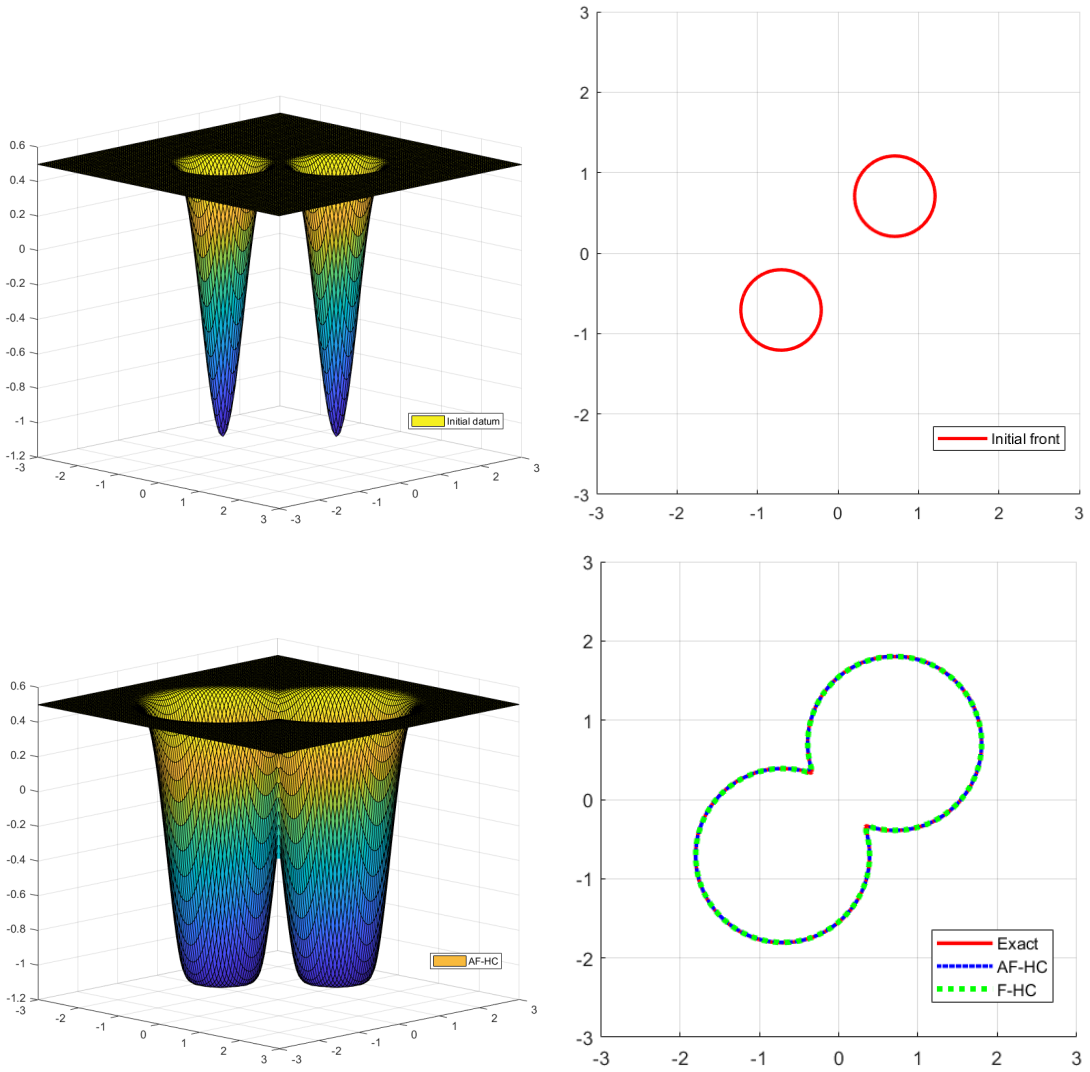

Figure 3.11.: (Example 7c.) From left to right. Top: initial datum and initial front. Bottom: plot of the solution using the AF-HC scheme and fronts using the filtered schemes at $T = 0.6$ with $\Delta x = \Delta y = 0.05$.

terms of error and resolution of the front, being able to keep the sharpness of the edges also w.r.t the WENO scheme. Finally, as Figure 3.13 clearly shows, all the schemes produce very sharp results, but the AF-RKC4 seems considerably sharper, at least in the region where the representation connects with the flat part.

Table 3.16.: (Example 7c.) Errors and orders in L^∞ and L^1 norms.

		F-HC ($20\Delta x$)		**AF-HC**		**AF-RKC4**		**WENO 2/3**	
N_x	N_t	L^∞ Err	Ord	L^∞ Err	Ord	L^∞ Err	Ord	L^∞ Err	Ord
30	12	2.31e-01		2.08e-01		1.96e-01		2.16e-01	
60	24	8.84e-02	1.38	6.45e-02	1.69	1.06e-01	0.88	9.08e-02	1.25
120	48	5.42e-02	0.71	4.93e-02	0.39	6.02e-02	0.82	4.49e-02	1.02
240	96	4.71e-02	0.20	1.78e-02	1.47	3.34e-02	0.85	2.31e-02	0.96
N_x	N_t	L^1 Err	Ord	L^1 Err	Ord	L^1 Err	Ord	L^1 Err	Ord
30	12	1.29e+00		1.20e+00		9.25e-01		1.37e+00	
60	24	4.07e-01	1.67	3.41e-01	1.82	1.57e-01	2.55	4.46e-01	1.62
120	48	1.27e-01	1.67	8.78e-02	1.96	5.68e-02	1.47	1.08e-01	2.04
240	96	5.24e-02	1.28	3.35e-02	1.39	3.39e-02	0.75	3.40e-02	1.67

Table 3.17.: (Example 7c.) CPU times in seconds.

N_x	N_t	**F-HC**	**AF-HC**	**AF-RKC4**	**WENO 2/3**
30	12	0.012 s	0.023 s	0.033 s	0.042 s
60	24	0.075 s	0.172 s	0.21 s	0.499 s
120	48	0.506 s	1.227 s	1.518 s	7.525 s
240	96	3.979 s	9.513 s	11.87 s	124.7 s

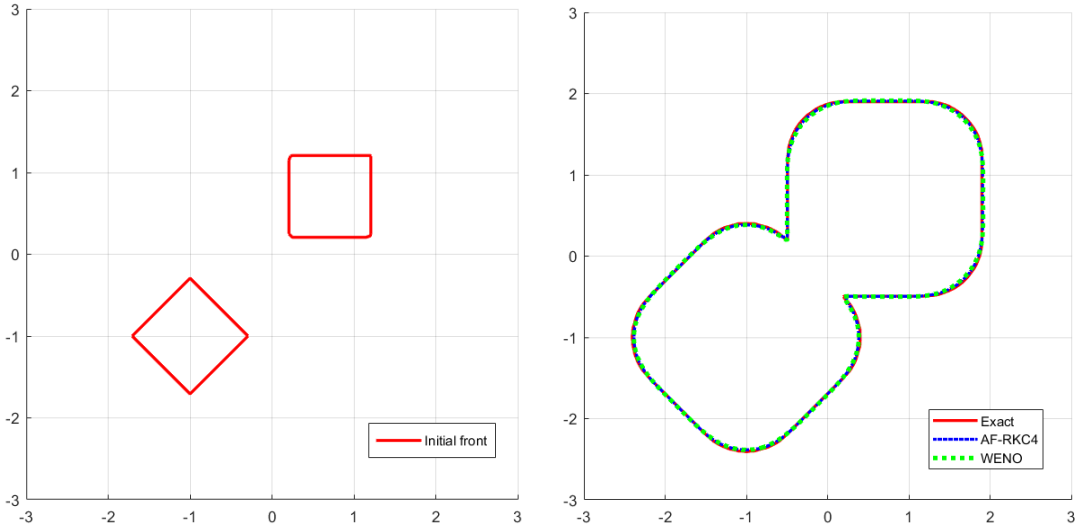

Figure 3.12.: (Example 7d.) Initial front (left) and fronts at $T = 0.7$ using WENO and AF-RKC4 scheme (right).

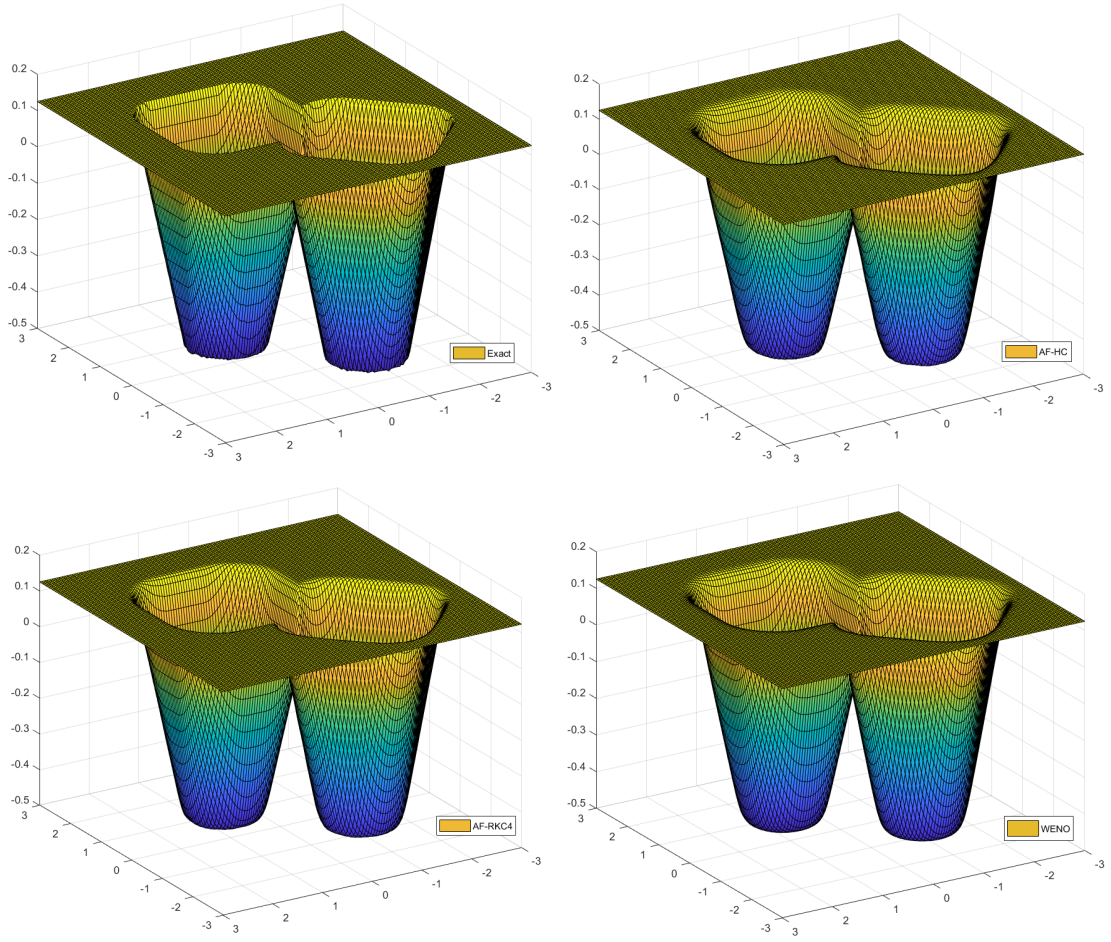

Figure 3.13.: (Example 7d.) Top: exact solution (left) and plot of the solution at $T = 0.7$ computed by the AF-HC scheme(right). Bottom: plots of the solution at $T = 0.7$ computed by the AF-RKC4 (left) and WENO (right) schemes.

Table 3.19.: (Example 7d.) CPU times in seconds.

N_x	N_t	**F-HC**	**AF-HC**	**AF-RKC4**	**WENO 2/3**
30	14	0.013 s	0.031 s	0.03 s	0.048 s
60	28	0.085 s	0.195 s	0.22 s	0.596 s
120	56	0.623 s	1.469 s	1.626 s	8.488 s
240	112	4.544 s	11.29 s	12.35 s	129.8 s

The conclusions of these last two examples are rather promising, testifying the good properties of the Adaptive Filtered Schemes also in more space dimensions. It is interesting to notice that, although the wideness of the stencil seems to limit excessively the accuracy in the L^∞ norm in presence of some singularity, the simple AF-RKC4 gave very

Table 3.18.: (Example 7d.) Errors and orders in L^∞ and L^1 norms.

		F-HC		**AF-HC**		**AF-RKC4**		**WENO 2/3**	
N_x	N_t	L^∞ Err	Ord	L^∞ Err	Ord	L^∞ Err	Ord	L^∞ Err	Ord
30	14	1.64e-01		1.06e-01		1.06e-01		1.26e-01	
60	28	6.93e-02	1.24	8.35e-02	0.35	7.31e-02	0.54	7.28e-02	0.79
120	56	6.73e-02	0.04	4.76e-02	0.81	4.24e-02	0.79	4.18e-02	0.80
240	112	8.57e-02	−0.35	2.79e-02	0.77	2.41e-02	0.82	2.71e-02	0.63
N_x	N_t	L^1 Err	Ord	L^1 Err	Ord	L^1 Err	Ord	L^1 Err	Ord
30	14	6.11e-01		8.33e-01		5.20e-01		8.53e-01	
60	28	3.18e-01	0.94	3.98e-01	1.07	3.60e-01	0.53	4.31e-01	0.98
120	56	1.65e-01	0.95	1.89e-01	1.08	1.73e-01	1.05	1.78e-01	1.27
240	112	1.22e-01	0.43	8.48e-02	1.15	7.57e-02	1.19	7.22e-02	1.30

good responses, especially in terms of sharpness of the representation. It would be really interesting to see the results obtained through a more compact Lax-Wendroff scheme, in the aim of (3.21), but such a definition seems to be rather difficult to reach, at least in the form of Lemma 3.2. Nevertheless, it will surely be object of future investigations.

3.6. Conclusions

We have presented a rather simple way to construct convergent schemes, which are of high-order in the regions of regularity for the solution. The filter method is able to stabilize an otherwise unstable (high-order) scheme, still preserving its accuracy. The novelty here is the adaptive and automatic choice of the parameter ε^n which improves the scheme in [BFS16]. The computation of ε^n, although more expensive, is still affordable in low dimension. The adaptive scheme is able to reduce the oscillations which may appear choosing a constant ε and, as shown by the numerical tests, gives always better results. Finally, we note that the adaptive filtered scheme, with a wise choice for the high-order scheme, has results close to the WENO scheme in terms of errors but seems to have a better accuracy on the kinks.

All the construction directly generalize to treat multidimensional problems, still keeping the good properties of the one-dimensional scheme. The main advantage of the filtered schemes relies in their simplicity of implementation and in the extreme generality allowed for the choice of the high-order scheme. In the latter context, the adaptive definition of the parameter ε^n and the stabilization properties of the function ϕ, represent a relevant improvement w.r.t. the basic scheme in [BFS16]. That is because no further limiting correction is needed, thus preventing the risk of losing accuracy when using schemes of order of accuracy higher than 2. The general applicability of our procedure has been testified by the successful implementation of simple and efficient fourth-order schemes in one and two space dimensions.

4. Segmentation of images via the Level Set Method

Our aim is to present a rather simple approach to the problem of the segmentation of a real image, i.e. we are interested in the detection of the boundaries of objects represented in a picture. This problem appears in many different fields such as, for example, biomedicine, astronomy and security.

The chapter is organized as follows: in Section 4.1 we begin by giving a brief introduction to the problem and to the state-of-the-art literature, while in Section 4.2, we recall the level set method for front propagation, focusing on its application to the segmentation problem and presenting the modification to be made w.r.t. the classical approach. Then, in Section 4.3, we present the numerical procedure used to solve the segmentation problem, based on our new multidimensional Adaptive Filtered Scheme and finally, in Section 4.4, we show the applicability of the proposed ideas through some numerical tests on synthetic and real images.

4.1. Introduction

In the last decades many authors faced the problem proposing different approaches, based mainly on variational formulations or on appropriates PDE models. Regarding the variational formulations, there is already a huge variety of models based on energy minimization, that have been applied to segmentation problems, for example, by Kass et al. in [KWT88]. They use energy-minimizing *snakes*, also termed *active contours*, that are attracted to the image relevant features, whereas internal spline forces impose a smoothness constraint. Since this method relies on the construction of functionals, which local minima may have several possible solutions, it usually leads to good results in some situations but presents various problems in terms of general applicability, being heavily dependent on the knowledge of the specific situation (e.g. the results generally change varying the initial guess). Moreover, despite a good initialization, the active contour model cannot be forced to extrude trough any significant protrusion, thus it cannot be used to recover very complex shapes (see [MSV93] for a more detailed discussion and some numerical examples of this behavior).

On the other hand, the PDE approach, which was first presented in [MSV93] and inspired by the pioneering works of Sethian ([Se85], [Se90]) and Osher ([OSe88]) on the level set method for front evolution, seems to overcome most of the limitations imposed by the variational techniques and can be applied to general cases. However, depending on the complexity of the proposed model, further assumptions on the "smoothness" of the picture to be segmented are needed, but that is rather natural (e.g. second order

equations coming from the Mean Curvature Motion problem, where the curvature is not well defined if the boundary of the object is not smooth). The solution of the PDE model is an evolving surface that is able to follow the topological changes caused by the relevant edges to be detected, still keeping its precise mathematical reasoning (as viscosity solution to some Hamilton-Jacobi type equation) and that of the segmentation front (as the 0-level set of the solution). The only real drawback of this idea, as we will see in the sequel, is that the problem (reconstruction of a 2D closed curve or of a finite collection of such curves) has to be immersed in a higher dimensional space, leading to more heavy computations.

In this work we will follow the PDE approach in its most simple form, which leads to a first order time-dependent Hamilton-Jacobi equation of eikonal type. In this situation, the front evolves in the normal direction with velocity dependent only on the information given by the considered image. Here we renew the ideas proposed in [MSV93], mainly w.r.t. the technique used to track back the evolving 0-level set, achieving pretty good results without the need to look for the curvature information, which would instead lead to a second order Hamilton-Jacobi equation.

To approximate the solution of the proposed model we use the new Adaptive Filtered Scheme in two space dimensions, that has been presented in Section 3.4 and first successfully applied to 1D problems in Chapter 3 and by the authors in [FPTa]. We will also give some hints on the construction of the scheme for the image segmentation.

4.2. Level Set Method for Segmentation

Let us begin by recalling that, as shown in Section 1.3, the problem of following a front Γ_t evolving in its normal direction can be restated in terms of the 0-level set of the solution of the problem

$$\begin{cases} v_t + c(x,y)|\nabla v| = 0, & (t,x,y) \in (0,T) \times \mathbb{R}^2 \\ v(0,x,y) = v_0(x,y) & (x,y) \in \mathbb{R}^2, \end{cases} \tag{4.1}$$

where $c(x,y)$ is the velocity of the front in the normal direction $\eta(t,x,y) = \frac{\nabla v(t,x,y)}{|\nabla v(t,x,y)|}$ and v_0 is a *proper representation* of the front Γ_0 (see Definition 1.19). Since in this case the velocity c depends only on (x,y), we obtain a first order time-dependent Hamilton-Jacobi equation of eikonal type, which solution remains a function at any time $t > 0$ as long as the velocity c is smooth enough. On the other hand, the front Γ_t can have complex shapes (due to breaking or merging) and also develop sharp corners. This feature is due to the definition of viscosity solution, which allows the selection of the correct solution even when singularities appear.

We recall also that the choice of the level set is arbitrary, and in fact all the level sets of v are moving according to the same law. This fact represents a major problem in the application to shape recovery due to the discontinuity of the velocity field c, whence a modification to the classical model is needed.

After defining the model equation, we are now able to describe how the level set formulation for front propagation can be used for the segmentation of an image. The

basic idea behind this technique is that the boundaries of a specific object inside a given image, described by the brightness function $I(x, y)$, are characterized by an abrupt change of the values of I, so that the magnitude of $|\nabla I|$ can be used as an indication of the edges. In order to make use of this intuition, we have to define the velocity $c(x, y)$ accordingly.

First, we notice that the 0-level set represents the boundary of an evolving shape that should stop in the proximity of the edges of the desired objects. As a first consequence, depending on the orientation of the evolution, we have to choose the initial front in an appropriate manner. More precisely, if we want the front to expand towards the desired shape, then the initial configuration must lie entirely inside the object, with more separate fronts in the case of more than one object of interest. Whenever we want the front to shrink and envelope all the interesting objects, we have to choose a configuration that contains all of them (for example, we can use the frame of the picture). Next, in the case of an expanding (shrinking) front we should require the velocity $c(x, y)$ to be basically a positive (negative) constant in regions where the value of I varies smoothly, and to be close to 0 when the front is close to an edge, where the evolution should stop. In the sequel, for brevity, we will consider only the case of an expansion, since the opposite can be obtained by simply taking $-c(x, y)$, or inverting the sign of the representation funtcion v_0.

Following the classical models, we use and compare two different definitions of c with similar properties, both presented in [MSV93], and then propose a slight modification in order to make the computations more stable, with a very little increase in the computational cost. We recall that, in order to reduce the possible noise of the image and the relative importance of natural changes of I inside the objects of interest (that could lead to the detection of wrong edges), the input data are usually smoothed through the convolution with a Gaussian filter. This can be obtained by evolving the function I according to the heat equation for an appropriate time interval (usually very small). Notice that, in general, the function $I(x, y)$ is usually disontinuous at the boundary of the objects in the image, then, at such points, the "smoothing effect" of the Gaussian filter has a smaller impact w.r.t. that on the noise.

The first velocity we consider is

$$c_1(x, y) = \frac{1}{(1 + |\nabla(G * I(x, y))|^p)}, \qquad p \geq 1, \tag{4.2}$$

where p is used to give more weight to the changes in the gradient, if necessary. According to this definition, the velocity has values in $[0, 1]$ and is ,at least locally, a strictly positive constant when I is smooth, otherwise it is close to 0 if there is a rapid change in the values of I. Another possibility, which we will not use since it is very similar to c_1, could be

$$c(x, y) = e^{-|\nabla(G*I(x,y))|}. \tag{4.3}$$

The second choice we consider, defined in [MSV93], has the form

$$c_2(x, y) = 1 - \frac{|\nabla(G * I(x, y))| - M_2}{M_1 - M_2}, \tag{4.4}$$

where M_1 and M_2 are the maximum and minimum values of $|\nabla(G * I(x,y))|$. This latter velocity has similar properties w.r.t. the previous one, having values in $[0,1]$ and being close to 0 if the magnitude of the image gradient is close to its maximal value, and basically equal to 1 otherwise. If we want the value of the velocity closer to 0 near the edges also in this case, we can simply take the p-power of c_2, having care of the possible change in the sign for negative velocities (shrinking case). More precisely, when we want the front to shrink, in order to keep the negative sign of the velocity, we should use

$$c_2(x,y) = (-1)^{p-1}\left(1 - \frac{|\nabla(G * I(x,y))| - M_2}{M_1 - M_2}\right)^p. \tag{4.5}$$

It is clear that both definitions have the desired properties, but with slightly different features. More precisely, in the first case the velocity depends more heavily on the changes in the magnitude of the gradient, thus giving an easier detection of the edges but also possibly producing false edges inside the object (think of light reflection, as an example). On the other hand, the latter velocity is smoother inside the objects, being less dependent on the relative changes in the gradient, but might present some problems in the detection of all the edges if at least one of those is "more marked".

4.2.1. Extension of the velocity function

In this section we recall the problem, first addressed in [MSV93], regarding the extension of the image-based velocity function $c(x,y)$ to all the level sets of the representation v. As pointed out in the cited paper, the speed function c defined in the previous section has meaning only on the front Γ_t, since it was designed precisely to force the 0-level set to stop in the proximity of an edge. Consequently, it derives its meaning not on the geometry of v but only on the configuration of the front Γ_t. Therefore, it is necessary to give a physical meaning also to the speed used to make all the other (infinite) level sets evolve. That is because, if we apply this discontinuous (around the edges) speed to all the level sets of v, then as time evolves discontinuous solutions will appear as soon as some level set reaches the region of the plane characterized by the object boundaries. As a consequence, the problem would not fit anymore in the classical definition of viscosity solution and most classical numerical methods for Hamilton-Jacobi equations would need at least a limiting correction in order to select the correct solution, or otherwise would produce unstable results (only monotone methods would probably be able to treat such situations without the limiting process).

Here we follow the idea discussed in [MSV93], proposing a simple way to extend the velocity function, which depends only on the initial condition v_0 and let us avoid all the heavy computations required by the first solution proposed by the authors, still obtaining stable results. Thus, recalling their approach, the first property that has to be required on the velocity function is the following

Property 1. Any external (image-based) speed function that is used in the equation of motion written for the function v should not cause the level sets to collide and cross each other during the evolutionary process.

As we previously stated, the appropriate extension depends on the choice of the initial representation, then to present the basic idea, as a first (and fundamental) example let us take the distance (to the initial 0-level set) function, that is

$$v_0(x,y) = dist\left\{(x,y), \Gamma_0\right\}. \tag{4.6}$$

Therefore, with this choice we can define the velocity extension as follows

Property 2. The value of the speed function c at a point P lying on a level set $\{v = C\}$ is exactly the value of c at a point Q, such that the point Q is a distance C away from P and lies on the level set $\{v = 0\}$.

In order to apply this construction, in [MSV93] the authors propose a simple but heavy procedure to track the point Q on the 0-level set associated to each point P of any level set. These computations clearly lead to the necessity of some modifications, such as the reinitialization for stability purposes and the *narrow band* approach to reduce the computational cost.

In this work we try to avoid such problems in tracking the 0-level set, making use of the knowledge on the evolution and on the initial condition. The idea is straightforward and it is based on the fact that the evolution is oriented in the normal direction to the front, whence if the reciprocal disposition of the level sets is also known (that is why we must choose wisely the initial condition) and we make all the points in the normal direction to the 0-level set evolve according to the same law, then it is reasonable to expect that all such points will keep their relative distance unchanged as time flows.

To present our modification, let us still consider the distance to Γ_0 as initial condition, then by construction all the C-level set are at a distance C from the 0-level set, as stated by Property 2. Whence, if we consider a generic point (x_c, y_c) on a C-level set, then it is reasonable to assume that the closest point on Γ_0 should be

$$(x_0, y_0) = (x_c, y_c) - v(t, x_c, y_c)\frac{\nabla v(t, x_c, y_c)}{|\nabla v(t, x_c, y_c)|}.$$

Therefore, it seems natural to define the extended velocity $\tilde{c}(x,y)$ as

$$\tilde{c}(x, y, v, v_x, v_y) = c\left(x - v\frac{v_x}{|\nabla v|}, y - v\frac{v_y}{|\nabla v|}\right), \tag{4.7}$$

which coincides with $c(x,y)$ on the 0-level set, as it is needed. The exact same approach can be applied as long as the initial distance between the level sets is known, then, if we want higher regularity to the evolving surface, which would be preferable in the case of high-order schemes such as the Adaptive Filtered scheme that we use in the numerical tests, we can define an appropriate initial condition, for example, by simply rotating a regular function in one space dimension. More precisely, let us consider a regular function $\overline{v}_0 : \mathbb{R}^+ \to \mathbb{R}$ such that $\overline{v}_0(r_0) = 0$, where r_0 is the radius of the initial circle Γ_0 (e.g. the right branch of a parabola centered in the origin), and define $v_0(x,y)$ rotating its profile, that is

$$v_0(x,y) = \overline{v}_0\left(\sqrt{x^2 + y^2}\right). \tag{4.8}$$

Then, it is clear that the C-level set of v_0 are located at a distance

$$d(C) := \overline{v}_0^{-1}(C) - r_0, \qquad \text{with } \overline{v}_0^{-1}(C) \geq 0, \tag{4.9}$$

from the 0-level set and, according to the previous reasoning, they should keep this property as time evolves. Consequently, also in this case we can define

$$\tilde{c}(x, y, v, v_x, v_y) = c\left(x - d(v)\frac{v_x}{|\nabla v|}, y - d(v)\frac{v_y}{|\nabla v|}\right). \tag{4.10}$$

More details on the function $d(v)$ will be given in the next section. Notice that in the last construction we assumed, for simplicity of presentation, the representation function to be centered in the origin, but it is straightforward to extend the same procedure to more general situations. Notice also that if we have only one object to be segmented (or we are considering the shrinking from the frame of the picture) we can always use a representation function centered in the origin since we can freely choose the domain of integration, which represents the pixels of the image.

***Remark* 4.1.** We noticed through extensive numerical tests that this simple structure of the model, governed essentially by the speed function $\tilde{c}$, needs a slight modification in order to stabilize the asymptotic solution. More precisely, since the velocity does not exactly vanish on the edges of the objects (but it is only fairly close to 0), in some cases the front may keep expanding even if the edge has been reached. This happens especially if some portion of the boundary is not heavily marked, then to stabilize the results we can simply introduce a cutting level for the values of $\tilde{c}$, considering equal to 0 all the velocities below this prescribed level. In the case of an expansion, for example, the velocity becomes

$$\overline{c}(x, y, v, v_x, v_y) := \begin{cases} \tilde{c}(x, y, v, v_x, v_y) & \text{if } \tilde{c} > c_{min} \\ 0 & \text{otherwise,} \end{cases} \tag{4.11}$$

with an analogous definition in the shrinking case. Recalling the discussion at the end of the previous section, it is not surprising that this problem is more evident if the velocity is defined initially by c_2, while c_1 usually gives stable results without the need of this last tuning.

4.2.2. Motivations of the new velocity function

Since the idea behind the modification of the velocity $c(x, y)$ into

$$\tilde{c}(x, y, v, v_x, v_y) = c\left(x - d(v)\frac{v_x}{|\nabla v|}, y - d(v)\frac{v_y}{|\nabla v|}\right),$$

with $d(v) = 0$ if $v = 0$, is to follow the evolution of the 0-level set and then to define accordingly the evolution on the other level sets, we can see the new definition, in some sense, as a *characteristic based velocity*. Consequently, in order to justify our approach, as a first step we inspect the characteristics of the equation, assuming the regularity necessary for the computations.

Therefore, we have to assume $v \in C^2(\Omega)$ (or at least C^2 in space and C^1 in time) and $c(x,y) \in C^1(\Omega)$, although the original problem does not satisfy (in general) these requirements, and clearly that the characteristics do not cross each other during the evolution. We will try to loosen these assumptions in a second moment.

***Remark* 4.2.** To avoid to introduce new variables we will use the familiar notation

$$H(\overline{x}, v, p) = \tilde{c}(\overline{x}, v, \overline{p})|\overline{p}|, \tag{4.12}$$

for the arguments of the hamiltonian, since it should not cause confusion in the computations. We will also drop the line over x and p, which highlights the vector nature of the arguments, if not strictly necessary.

Let us begin, by the method of characteristics, writing the usual system (see (1.6))

$$\begin{cases} \dot{x}(s) = \nabla_p H \\ \dot{v}(s) = \nabla_p H \cdot p - H \\ \dot{p}(s) = -\nabla_x H - H_v p, \end{cases} \tag{4.13}$$

which, since in our case

$$\begin{aligned} \frac{\partial H}{\partial v_x} =& - d(v)\tilde{c}_x \left(\frac{|\nabla v| - \frac{v_x^2}{|\nabla v|}}{|\nabla v|^2} \right) |\nabla v| - d(v)\tilde{c}_y v_y \left(- \frac{\frac{v_x}{|\nabla v|}}{|\nabla v|} \right) |\nabla v| + \tilde{c}\frac{v_x}{|\nabla v|} \\ =& - d(v)\tilde{c}_x \frac{v_y^2}{|\nabla v|^2} + d(v)\tilde{c}_y \frac{v_x v_y}{|\nabla v|^2} + \tilde{c}\frac{v_x}{|\nabla v|}, \end{aligned}$$

and analogously for $\frac{\partial H}{\partial v_y}$, we have that

$$\nabla_p H = \begin{pmatrix} \frac{d(v)v_y}{|\nabla v|^2}\left(v_x\tilde{c}_y - v_y\tilde{c}_x\right) + \tilde{c}\frac{v_x}{|\nabla v|} \\ \frac{d(v)v_x}{|\nabla v|^2}\left(v_y\tilde{c}_x - v_x\tilde{c}_y\right) + \tilde{c}\frac{v_y}{|\nabla v|} \end{pmatrix} \quad \Rightarrow \quad \nabla_p H \cdot p = \tilde{c}(x,v,p)|p|, \tag{4.14}$$

it simply reads

$$\begin{cases} \dot{x}(s) = \nabla_p H \\ \dot{v}(s) = \tilde{c}(x,v,p)|p| - \tilde{c}(x,v,p)|p| = 0 \\ \dot{p}(s) = -\nabla_x \tilde{c}(x,v,p)|p| + d'(v)\nabla_x \tilde{c}(x,v,p)|p|^2 = \nabla_x \tilde{c}(x,v,p)|p|(d'(v)|p| - 1) \end{cases} \tag{4.15}$$

then, from the third equation, if we now define $d(v)$ such that

$$d'(v) = |p|^{-1}, \tag{4.16}$$

we have the final system

$$\begin{cases} \dot{x}(s) = \nabla_p H \\ \dot{v}(s) = 0 \\ \dot{p}(s) = 0 \end{cases} \tag{4.17}$$

which states that, as long as the function $\tilde{c}$ remains smooth enough ($\tilde{c}_x \approx 0$ and $\tilde{c}_y \approx 0$), the characteristics are basically directed in the normal direction and that along them

both the height and the gradient are preserved. These last two properties remain valid even when $\tilde{c}$ is no longer smooth. Moreover, a deeper inspection into the third relation of (4.17) and (4.16) reveals that, since $p(s) \equiv p_0 = \nabla v_0$ along the characteristics, we can choose more simply

$$d'(v) = |\nabla v_0|^{-1}, \tag{4.18}$$

which is indeed the case, trivially, for the function $d(v) = v$ and also for $d(v)$ given by (4.9) defined in the previous section. In fact, using the inverse function theorem we have

$$d'(v) = \frac{d}{dv}\left(\overline{v}_0^{-1}(v)\right) = \frac{1}{\overline{v}_0'(z)},$$

with z such that $\overline{v}_0(z) = v$. Moreover, recalling the definition (4.8) we can compute,

$$\begin{aligned}
|\nabla v_0(x,y)| &= \left|\nabla \overline{v}_0\left(\sqrt{x^2+y^2}\right)\right| \\
&= \left|\left(\frac{\overline{v}_0'\left(\sqrt{x^2+y^2}\right)x}{\sqrt{x^2+y^2}}, \frac{\overline{v}_0'\left(\sqrt{x^2+y^2}\right)y}{\sqrt{x^2+y^2}}\right)\right| \\
&= \frac{\overline{v}_0'\left(\sqrt{x^2+y^2}\right)}{\sqrt{x^2+y^2}}\sqrt{x^2+y^2} = \overline{v}_0'\left(\sqrt{x^2+y^2}\right),
\end{aligned}$$

then it is enough to consider (x, y) such that $z = \sqrt{x^2+y^2}$.

***Remark* 4.3.** From a numerical point of view the equations (4.16)-(4.18) give two different means to compute the velocity at each time step. If we prefer to compute the function $d(v)$ analytically, through the knowledge of the initial datum v_0, we have to use (4.18), while if we prefer to compute $d(v)$ independently on v_0 we can use, for example, a numerical integration for

$$d(v) = \int \frac{1}{|\nabla v|} dx, \tag{4.19}$$

where the integral is taken on the projected characteristic $x(s)$. The latter choice would probably produce an even more stable scheme, assuming to be able to compute a reasonably good approximation of (4.19).

Thanks to the previous computations, we reached a good understanding of the nature of the evolution given by (4.1)-(4.10), but we still have not justified the main motivation that led us to define (4.10), that is to make all the level sets of v evolve according to the same law. More precisely, we have to show that, if we consider the evolution of two points on the same characteristic but on two different levels sets, say the 0-level set $x_0(s)$ and a generic level set $x_v(s)$, then their relative distance (along the characteristic) does not change during the evolution. This fact would imply that, if we choose the level sets of v_0 to be such that

$$x_0(0) = x_v(0) - d(v_0)\frac{\nabla v_0}{|\nabla v_0|}, \tag{4.20}$$

than the points $\underline{x}(s) = x_v(s) - d(v)\frac{\nabla v}{|\nabla v|}$ are always on the 0-level set of v. In order to prove this last statement, let us proceed by a simple differentiation

$$\begin{aligned}\dot{\underline{x}}(s) =& \dot{x}_v(s) - \frac{d}{ds}\left(d(v)\frac{p}{|p|}\right)\\ =& \dot{x}_v(s) - \left[d'(v)\dot{v}(s)\frac{p}{|p|} + \frac{d(v)}{|p|^2}\left(\dot{p}(s)|p| - \frac{d}{ds}(|p(s)|)p\right)\right]\\ =& \dot{x}_v(s) + \frac{d(v)}{|p|^2}\left(\frac{p\cdot\dot{p}}{|p|}\right)p\\ =& \dot{x}_v(s),\end{aligned} \tag{4.21}$$

where we have used the second and third relation of (4.17) in the last two equalities of the chain. This last computation states that all the level sets evolve according to the same law along characteristics. As a consequence, if we assume (4.20), then $\underline{x}(s) \equiv x_0(s)$ till the characteristics do not cross, as we wanted.

One of the main consequences of this property, which will be very useful in the numerical implementation, is that the points $\left(x - d(v)\frac{\nabla v}{|\nabla v|}\right)$ are on the 0-level set of v as long as the gradient is preserved. Then, assuming to have a coherent way to recover the point on the associated characteristic, we can approximate the problem by previously computing the points $\underline{x} = \left(x - d(v)\frac{\nabla v}{|\nabla v|}\right)$ and then updating the solution considering the *simplified problem* with (locally in time) *isotropic velocity*

$$v_t + c(\underline{x})|\nabla v| = 0, \qquad (t, x, y) \in (t^n, t^{n+1}) \times \mathbb{R}^2. \tag{4.22}$$

This can be done in a very simple and direct manner, with only a slight increase in the computational cost, as we will see in the next section. Otherwise, we should consider the full problem (4.1)-(4.10) and treat numerically all the dependence of $H(x, v, \nabla v)$.

4.3. Numerical implementation

In this section we discuss the problems that can arise when appoximating the solution of (4.1)-(4.10) and how we can partially avoide them by assuming the simplificed problem (4.22). The main problems come from the fact that we do not have an analytical expression for the velocity $c(x, y)$, so we can not compute a priori the derivatives of $\tilde{c}$ with respect to the variables x, y and v.

Before illustrating the numerical tests, let us first give some comments on the numerical schemes composing the Adaptive Filtered Scheme adopted for the tests in Sect. 4.4. The main issue concerning the *local Lax-Friedrichs* and the *Lax-Wendroff schemes* defined by (3.49) and (3.57), respectively, is the need to compute the one-directional velocities H_p and H_q which depend also on $\tilde{c}_x$ and $\tilde{c}_y$, as visible in (4.14). Moreover, in order to implement the local Lax-Friedrichs scheme we should be able to compute the maximum of $|H_p|$ (resp. $|H_q|$) uniformly w.r.t. v_y (resp. v_x), which is a very intricate matter due to the (possible) low regularity of $\tilde{c}$. In fact, if we focus on the usual behavior of $\tilde{c}(x, y)$ in the proximity of a relevant edge, we can expect the derivatives $\tilde{c}_x$

and $\tilde{c}_y$ to be really big. This is not surprising since the front decelerates rapidly in the neighborhood of an edge. Therefore, we could not use the simple relation

$$\max_{v_x} \max_{v_y} |H_p(\cdot, v_x, v_y)| = \max_{v_x} |H_p(\cdot, v_x, 0)|, \tag{4.23}$$

where the maximums are taken over all the possible values of v_x and v_y, which is instead valid in the classical model with isotropic velocity. Analogous comment holds for H_q.

More importantly, from the numerical point of view this simplification brings another fundamental consequence. In fact, when we apply the numerical schemes to solve (4.22), we are considering, formally, a problem with bounded velocities $\max\{|H_p|, |H_q|\} \leq 1$. This implies that we can choose the following CFL condition:

$$\lambda := \max\left\{\frac{\Delta t}{\Delta x}, \frac{\Delta t}{\Delta y}\right\} \leq \frac{1}{2} \max\left\{|H_p|^{-1}, |H_q|^{-1}\right\}, \tag{4.24}$$

using the relation (4.23). This condition is less restrictive with respect to the original one, for which it is necessary to compute $\max\{|H_p|, |H_q|\}$ with the full formula (4.14). Consequently, as already noted, λ could be excessively small due to the low regularity of $\tilde{c}(x, y)$. In the latter case we would clearly need an adaptive mesh refinement technique to reduce the computational cost.

Lastly, when using the full model (4.1)-(4.10), we should take into account also the remaining dependence of $H(x, y, v, \cdot, \cdot)$ when deriving the second-order Lax-Wendroff scheme and, clearly, the formula to compute the threshold ε^n.

In order to avoid most of these complications in the numerical implementation, we choose to approximate the solution of the simplified problem (4.22), adjusting the velocity $\tilde{c}$ according to (4.10) at each time step. In the following we will use the same notations introduced in Sect. 3.4, except for the number of time steps N_T, which will be replaced by the total number of iterations N_i used by the scheme, since now we are looking for an asymptotic solution (in some stationary sense). The maximum number of iterations, which is fixed at the beginning of the procedure, will be denoted by $N_{\max}$.

Let use give some details on the precise numerical implementation, commenting the main procedures involved in the (sketched) *Pseudo-code* 1.

Fixed the parameters of the simulation, which are the power p in (4.2) or in (4.5), the number of iterations N of the heat equation for the Gaussian filter, the tolerance $tol > 0$ of the stopping rule, the amplitude of the pixels $(\Delta x, \Delta y)$ and, subsequently, the time step Δt according to the CFL condition (4.24).

Then, at each iteration $n = 0, \ldots, N_i$, which has to be interpreted in the sense "until convergence" (notice that N_i is not known a priori, but depends on the stopping rule described in Step 3), we repeat the following steps.

Step 1. We precompute the matrix $\tilde{c}(x_j, y_i, u_{i,j}, \nabla u_{i,j})$ at the beginning of every iteration using central finite difference approximations for the first order derivatives u_x and u_y (notice that the quantities depend only on (i, j) also through u). Clearly, this method is valid only as long as the representation u remains smooth at all the level sets, and

Pseudo-code 1 Segmentation via the LSM

Input: p, N, tol, $N_{\max}$, I, u_0;
$E^0 = 1$, $n = 0$;
regularize the matrix I (apply the Gaussian filter);
compute the velocity matrix c using (4.2) or (4.5);
store the position of the front in the matrix F^0;
while $(E^n > tol)$ and $(n < N_{\max})$ **do**
Step 1: compute the modified velocity matrix $\tilde{c}^n$ using (4.10);
Step 2: update the solution $u^n \to u^{n+1}$;
$n = n + 1$;
Step 3: store the front F^n;
compute the error E^n;
$N_i = n$;
Output: N_i, u^{N_i}.

should be justified in the case of singular edges (although we will not pursue this precise matter). Moreover, in general the point

$$(x_{j_u}, y_{i_u}) := \left(x_j - d(u_{i,j}) \frac{D_x u}{\sqrt{(D_x u)^2 + (D_y u)^2}}, y_i - d(u_{i,j}) \frac{D_x u}{\sqrt{(D_x u)^2 + (D_y u)^2}} \right)$$

is not a point of the grid (x_j, y_i). To reconstruct the correct value (or at least a reasonable approximation) we propose two different implementations. The first one is a simple *bilinear reconstruction* from the neighboring values

$$\mathcal{N}_u := \left\{ \left(x_{\lfloor j_u \rfloor}, y_{\lfloor i_u \rfloor}\right), \left(x_{\lceil j_u \rceil}, y_{\lceil i_u \rceil}\right), \left(x_{\lceil j_u \rceil}, y_{\lfloor i_u \rfloor}\right), \left(x_{\lfloor j_u \rfloor}, y_{\lceil i_u \rceil}\right) \right\},$$

where we used the notation,

$$\lceil j_u \rceil := j - \left\lceil \frac{x_{j_u} - x_j}{\Delta x} \right\rceil \quad \text{and} \quad \lfloor i_u \rfloor := i - \left\lfloor \frac{y_{i_u} - y_i}{\Delta y} \right\rfloor,$$

with the other cases following an analogous definition. The second possibility, which we use in the numerical examples since it seems to give nicer results in terms of the shape of the approximate representation u, consists only in taking as (x_{j_u}, y_{i_u}) the point such that

$$|u_{i_u, j_u}| := \min_{(x_j, y_i) \in \mathcal{N}_u} |u_{i,j}|.$$

Notice that this construction is well defined only if $|\nabla u_{i,j}| \neq 0$. Therefore, we define the updated velocity matrix as

$$\tilde{c}^n_{i,j} = \begin{cases} c_{i_u, j_u} & \text{if } |\nabla u_{i,j}| \neq 0, \\ c_{i,j} & \text{otheriwse,} \end{cases} \tag{4.25}$$

and we use $\tilde{c}^n$ as an isotropic velocity in the next step.

Step 2. We approximate the problem

$$v_t + \tilde{c}^n(x,y)|\nabla v| = 0, \qquad (t,x,y) \in (t^n, t^{n+1}) \times \mathbb{R}^2, \tag{4.26}$$

using the Adaptive Filtered Scheme (3.44), with the local Lax-Friedrichs scheme (3.49) as S^M and the Lax-Wendroff scheme (3.57) as S^A.

Step 3. The last step of the iteration is to "test" the approximated solution according to the prescribed stopping criterion. In our implementation we use one of the following *stopping rules*.

The idea is simply to stop the iterations as soon as the 0-level, or more precisely a neighborhood of the front Γ_t of radius $\delta = \max\{\Delta x, \Delta y\}$, ceases to move. In order to apply this procedure, at each time step we store in a matrix F^n the values of the points (x_j, y_i) such that $u^n_{i,j}$ changes sign (we use the closest points on the grid, that are $(i \pm 1, j \pm 1)$ and $(i \pm 1, j \mp 1)$), and set $F^n_{i,j} = 0$ otherwise. In this way we automatically store also the position of the front, or more precisely its disposition with an error of order $\delta = \max\{\Delta x, \Delta y\}$. Consequently, the iterations stop whenever

$$E_\infty := ||u^{n+1} - u^n||_{L^\infty(\theta_\delta)} = \max_{i,j} |F^n_{i,j} - F^{n-1}_{i,j}| < \tau(\Delta x, \Delta y),$$

where $\tau > 0$ is the prescribed tolerance, which depends also on the discretization parametrs, and $\delta > 0$ the radius of the neighborhood θ_δ around the front Γ_t. Another possible choice for the stopping rule, which seems also preferable numerically, is to compute the error between two consecutive iterations in the L^1-norm and then stop the iterations as soon as

$$E_1 := ||u^{n+1} - u^n||_{L^1(\theta_\delta)} = \Delta x \Delta y \sum_{i,j} |F^n_{i,j} - F^{n-1}_{i,j}| < \tau(\Delta x, \Delta y).$$

4.4. Numerical tests

In this section we present a series of numerical experiments on both synthetic and real images, comparing the results obtained with the Adaptive Filtered scheme and with the basic monotone scheme. The first aim is to show the possible improvements of the modified model w.r.t. the classical formulation. After extensive numerical simulations in fact, we noticed that the classical model is not well defined when using high-order schemes, since they can produce heavy oscillations as soon as a discontinuity in the representation appears. This effect causes the stopping rule to be practically ineffective (in both norms) in most cases tested when using the Adaptive Filtered Scheme, while the simple monotone scheme seems to give always stable results. Notice that, when the singularity develops, the "neighborhood" of the front becomes more and more vertical as time flows, creating additional difficulties in tracking the 0-level set.

In the case of the classical model, consequently, we present only the results obtained with the choices of the parameters for which we were able to achieve convergence and, for comparison reasons, most of the tests with the modified model will be presented

using the same parameters. We would like to remark that, using the new model, we were able to obtain convergence for the Adaptive Filtered Schemes in almost all cases tested. Therefore, the most complex (and interesting) cases with real images will be approximated using exclusively the modified model.

The following tests will be focused also on the comparison of the results varying the initial datum, when using the modified velocity, or varying the norm of the stopping rule, when using the classical model, and finally, in the case of synthetic images, also varying the size $\Delta x = \Delta y$ (and number) of the pixels.

***Remark* 4.4.** Since for real images the number of pixels is fixed and we would be forced to change also the "artificial" range of (x, y), there is not much sense in doing computations varying the size Δx of the pixels. It is clear that, in such a case, we would need also to tune again the parameters p, N and tol (also in the L^∞ norm) according to the new setting. Notice that the values of Δx and Δy control the sensitivity of $|\nabla I(x, y)|$ as smoothness indicator of the data.

Next, we specify the initial condition used in each case. When the velocity is defined by the classical model, in the expansion case (*Case a*) we use the paraboloid

$$u_0(x, y) = \min\left\{x^2 + y^2 - r^2, \frac{1}{2}r^2\right\}, \tag{4.27}$$

where $r > 0$ is the radius of the initial circle and $\frac{1}{2}r^2$ a value chosen in order to cut the surface from above (therefore we have a flat surface at the numerical boundary), while in the shrinking case (*Case b*) we use the truncated pyramid (tent) with a square (rectangular) base, that is

$$u_0(x, y) = \min\{2(x - b_x), 2(a_x - x), 2(y - b_y), 2(a_y - y), -0.2\}, \tag{4.28}$$

where $[a_x, b_x] \times [a_y, b_y]$ is the the frame of the image, -0.2 is the value at which we truncate the pyramid (tent) and 2 is the steepness of the faces of the surface (higher the value the more steep the faces). With this choice we use a slightly more regular front w.r.t. the discontinuous representation that simply changes value crossing the frame of the image, still being able to keep the whole surface outside the region occupied by the objects to segment.

On the other hand, when using the modified velocity, in the expansion case (*Case a*) we use again the paraboloid (4.27) (*Datum 1*) or the following distance function (*Datum 2*)

$$u_0(x, y) = dist\{(x, y), \Gamma_0\}, \tag{4.29}$$

with Γ_0 the usual circle centered in $(0, 0)$ with radius $r = 0.5$, while in the shrinking case (*Case b*) we use only the distance function (4.29), with Γ_0 representing the frame of the image.

In the reported tables we will compare the results in terms of *number of iterations* N_i and *relative error*, defined as

$$P\text{-}Err_{rel} = \frac{|P_{ex} - P_a|}{P_{ex}}, \tag{4.30}$$

where P_{ex} and P_a are the number of pixels inside, respectively, the exact and approximated boundaries. Notice that we can compute the "exact" boundary only if the background is really smooth (in the numerical tests it is always uniform), because we usually use a comparison with a "threshold" for the values of $I(x, y)$ in order to select the regions occupied by the object (exact boundary). Whereas, for the approximated boundary we will clearly consider the region such that $u_{i,j} \leq 0$. Moreover, we measure the error also with a closely related quantity, that is

$$P\text{-}Err_1 = |P_{ex} - P_a|\Delta x \Delta y, \tag{4.31}$$

in order to show some dependence on the discretization parameters. This latter error will be of some interest only in the synthetic case, for which we have the freedom to choose the number of pixels (therefore their size, fixing the ranges for the x and y values).

If the schemes do not arrive to convergence in the fixed maximum number of iterations, we will put a "$-$" inside the tables, in place of N_i. For all our tests, we will fix $N_{\max} = 2000$. Moreover, in the case the front does not stop correctly on the boundary of the object, thus giving an unstable and unusable result, we will put an "X" in correspondence of the errors column. For each test we specify the main parameters involved (p, N, tol and $\Delta x = \Delta y$), the norm used in the stopping rule and the chosen velocity function. More precisely, we use the notation c for the velocity function in the classical model and $\tilde{c}$ for the modified velocity. For all numerical tests presented in this thesis, we used the velocity function c_1 (4.2), with $c_{\min}$ or $c_{\max} = 0$ in (4.11), and CFL number $\lambda = \max\{\lambda_x, \lambda_y\} = \frac{1}{2}$.

4.4.1. Synthetic images

Let us begin by presenting some simple synthetic examples, in which we compare the performance of the filtered and monotone scheme for both the velocities. The main aim here is to show the convergence of the scheme varying the size Δx of the pixels. This is indeed the case for all the test cases, as can be seen in the following tables, with the relative error decreasing with Δx. For the synthetic tests below we decided to perform the simulations only in the case of an expansion, since these situations are even more easily handled by a shrinking front and the results do not vary evidently changing the schemes or the velocity.

Comparing the results obtained by the two schemes, we can see that, if we use the classical model, both schemes give the exact same results, while using the modified velocity more differences can be noted, with the filtered scheme performing usually better in terms of number of iterations N_i or, more often, in terms of computed error. More evidently, if we compare the results varying the model, we can see that the modified velocity $\widetilde{c}$ gives usually more stable results. In fact, the model that uses $\tilde{c}$ converges to the solution sooner and in a more accurate way with the same tolerance chosen in all simulations. The only exception is the synthetic vase (Test 3), where we probably should use a more restrictive tolerance when using $\widetilde{c}$.

Test 1. Ellipse

The first two tests that we propose are constructed by simple discontinuous functions I (that define the information given by the image) that are equal to 1 inside the geometrical figure considered, and 0 otherwise. In the first test the figure is an *ellipse* centered in the origin, that is

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1, \qquad (x, y) \in [-2, 2]^2, \tag{4.32}$$

with $a = \frac{3}{2}$ and $b = \frac{3}{4}$. For this very regular example, in Tables 4.1, 4.2 and 4.3 we can recognize the behavior described in the introduction of the section. When using the classical model we have the same results using the filtered and the monotone schemes, whereas some differences are visible when using the modified velocity. In the latter case, the AF-LW scheme, shown in Figure 4.1, performs usually better than the monotone scheme and we obtain better results w.r.t. the classical model in terms of number of iterations and errors. The only exception is the case $\Delta x = 0.04$ with Datum 1 (see Table 4.2), where the monotone schemes performs particularly well and much better than all the other cases tested.

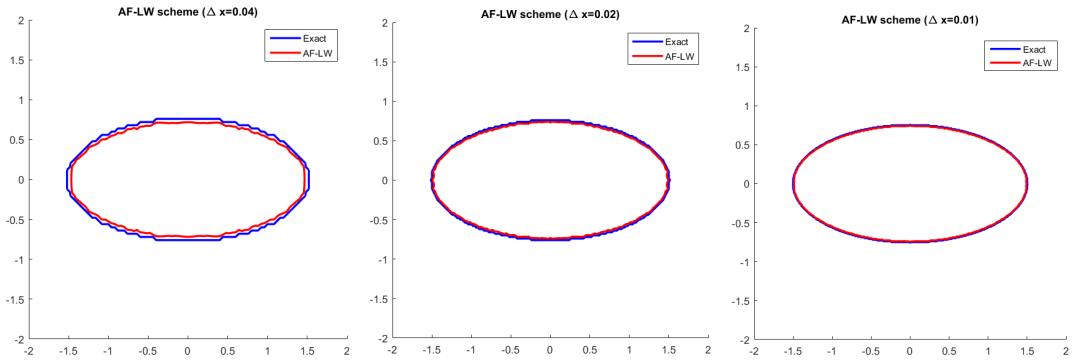

Figure 4.1.: (Test 1a) Datum 1. Plots of the final front using the AF-LW scheme with velocity $\tilde{c}$ varying Δx.

Table 4.1.: (Test 1a) Errors and number of iterations varying Δx.

c	**Norm** $\|\|\ \|\|_\infty$			**Monotone**			**AF-LW**		
Δx	tol	p	N	N_i	$P\text{-}Err_{rel}$	$P\text{-}Err_1$	N_i	$P\text{-}Err_{rel}$	$P\text{-}Err_1$
0.04	0.0001	2	0	83	0.0744	0.2624	74	0.0744	0.2624
0.02	0.0005	2	0	156	0.0376	0.1328	146	0.0376	0.1328
0.01	0.001	2	0	316	0.0189	0.0668	300	0.0189	0.0668

Table 4.2.: (Test 1a) Errors and number of iterations varying Δx (L^∞ norm).

$\widetilde{c}$	**Datum 1**			**Monotone**			**AF-LW**		
Δx	tol	p	N	N_i	$P\text{-}Err_{rel}$	$P\text{-}Err_1$	N_i	$P\text{-}Err_{rel}$	$P\text{-}Err_1$
0.04	0.00005	2	0	51	0.0545	0.1920	52	0.0654	0.2304
0.02	0.0001	2	0	103	0.0321	0.1136	102	0.0292	0.1032
0.01	0.0005	2	0	202	0.0160	0.0564	200	0.0155	0.0548

Table 4.3.: (Test 1a) Errors and number of iterations varying Δx (L^∞ norm).

$\widetilde{c}$	**Datum 2**			**Monotone**			**AF-LW**		
Δx	tol	p	N	N_i	$P\text{-}Err_{rel}$	$P\text{-}Err_1$	N_i	$P\text{-}Err_{rel}$	$P\text{-}Err_1$
0.04	0.00005	2	0	51	0.0599	0.2112	50	0.0581	0.2048
0.02	0.0001	2	0	102	0.0297	0.1052	106	0.0285	0.1008
0.01	0.0005	2	0	202	0.0145	0.0512	200	0.0155	0.0548

Test 2. Rhombus

The second test presents similar behaviors to the previous one. The picture visible in Figure 4.2 is given by the equation

$$\frac{|x|}{2} + |y| = \frac{3}{4}, \qquad (x, y) \in [-2, 2]^2, \tag{4.33}$$

that produces a final front which presents some heavily marked corners, where the normal direction is not well defined. This causes some serious problems when using the classical model (see Tables 4.4 and 4.5), at least if the filtered scheme is involved. In fact, the AF-LW scheme is clearly unstable, as it can be seen in Table 4.4, not being able to stop correctly at the boundary of the object for any refinement.This is because, as the representation approaches the region of the domain characterized by the boundary of the object, discontinuities start to develop and the surface becomes more and more vertical. Consequently, small oscillations begin to appear and cause the front to keep expanding, since they prevent the fulfillment of the condition on consecutive iterations, even though the solution seems to remain overall stable. In fact, when using the L^1 the scheme seems to converge for some appropriate choice of the tolerance and gives also pretty nice results.

On the other hand, if we use the modified velocity $\tilde{c}$, both the schemes always converge and give again better results in terms of both errors and number of iterations w.r.t. the correspondent results using the classical velocity. We notice also that the filtered scheme performs usually better, at least when using the paraboloid (Datum 1) as initial condition (see Table 4.6).

Table 4.4.: (Test 2a) Errors and number of iterations varying Δx.

c	**Norm** $\|\|\ \|\|_\infty$			**Monotone**			**AF-LW**		
Δx	tol	p	N	N_i	$P\text{-}Err_{rel}$	$P\text{-}Err_1$	N_i	$P\text{-}Err_{rel}$	$P\text{-}Err_1$
0.04	0.005	2	1	67	0.1208	0.2720	479	X	X
0.02	0.001	2	1	203	0.0600	0.1368	1390	X	X
0.01	0.0005	2	1	436	0.0396	0.0896	1378	X	X

Table 4.5.: (Test 2a) Errors and number of iterations varying Δx.

c	**Norm** $\|\|\ \|\|_1$			**Monotone**			**AF-LW**		
Δx	tol	p	N	N_i	$P\text{-}Err_{rel}$	$P\text{-}Err_1$	N_i	$P\text{-}Err_{rel}$	$P\text{-}Err_1$
0.04	0.0016	2	1	40	0.1692	0.3808	40	0.1663	0.3744
0.02	0.0004	2	1	93	0.0859	0.1960	130	0.0782	0.1784
0.01	0.0001	2	1	182	0.0481	0.1090	180	0.0480	0.1088

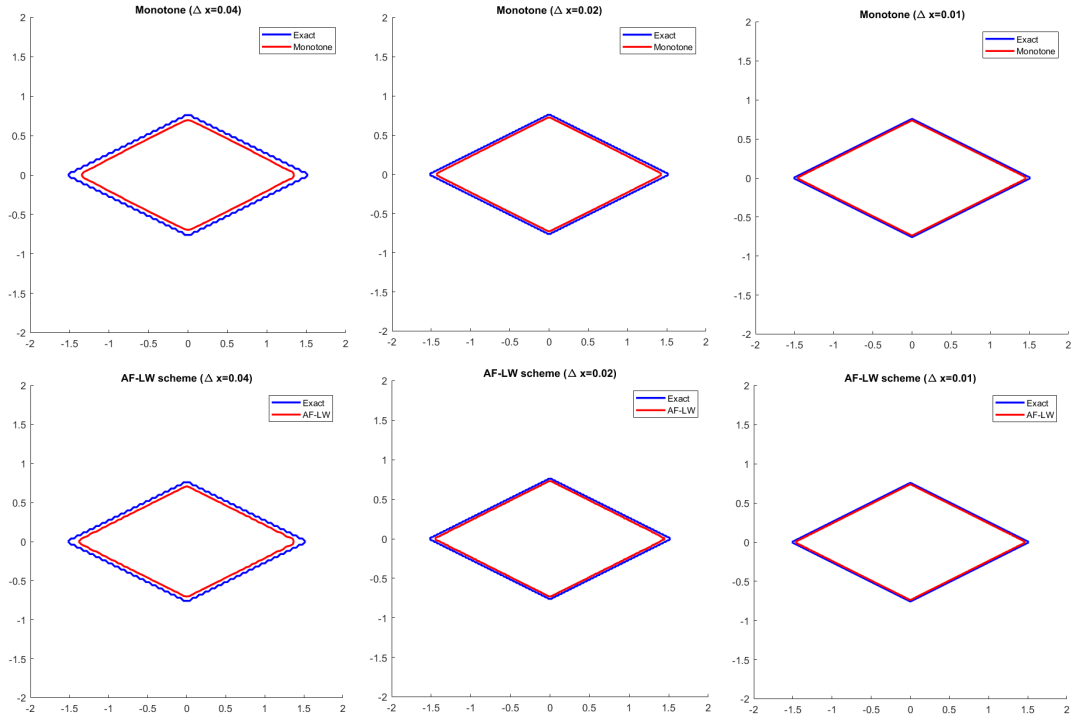

Figure 4.2.: (Test 2a) Datum 2. Plots of the final front using the Monotone scheme (top) and the AF-LW scheme (bottom) with velocity $\tilde{c}$ varying Δx.

Table 4.6.: (Test 2a) Errors and number of iterations varying Δx (L^∞ norm).

$\widetilde{c}$	**Datum 1**			**Monotone**			**AF-LW**		
Δx	tol	p	N	N_i	$P\text{-}Err_{rel}$	$P\text{-}Err_1$	N_i	$P\text{-}Err_{rel}$	$P\text{-}Err_1$
0.04	0.0005	2	1	60	0.1137	0.2560	80	0.1059	0.2384
0.02	0.0001	2	1	143	0.0607	0.1384	130	0.0635	0.1448
0.01	0.00005	2	1	207	0.0389	0.0881	242	0.0380	0.0861

Table 4.7.: (Test 2a) Errors and number of iterations varying Δx (L^∞ norm).

$\widetilde{c}$	**Datum 2**			**Monotone**			**AF-LW**		
Δx	tol	p	N	N_i	$P\text{-}Err_{rel}$	$P\text{-}Err_1$	N_i	$P\text{-}Err_{rel}$	$P\text{-}Err_1$
0.04	0.0005	2	1	56	0.1137	0.2560	84	0.1052	0.2368
0.02	0.0001	2	1	140	0.0614	0.1400	200	0.0558	0.1272
0.01	0.00005	2	1	226	0.0392	0.0889	283	0.0306	0.0830

Test 3. Synthetic vase

Finally, we conclude our synthetic simulations by considering a more intricate example, that is the synthetic vase. To define the brightness function I associated to the figure,

we first compute the function

$$f(x,y) = \begin{cases} \sqrt{P(\bar{y})^2 - x^2} & \text{if } P(\bar{y})^2 > x^2 \text{ and } |y| < 2, \\ 0 & \text{otherwise,} \end{cases} \tag{4.34}$$

where $\bar{y} = y/4$, $(x,y) \in [-1.5, 1-5] \times [-3,3]$ and

$$P(\bar{y}) = 4(-10.8\bar{y}^6 + 7.2\bar{y}^5 + 6.6\bar{y}^4 - 3.8\bar{y}^3 - 1.375\bar{y}^2 + 0.5\bar{y} + 0.25).$$

Then, we simply take $I(x,y) = \left(\sqrt{1 + |\nabla f(x,y)|^2}\right)^{-1}$, which comes from the shape from shading problem when considering a Lambertian model with a vertical light source. See [FT16], [DFS08] for more details on Shape-from-Shading.

For this particular test the comparison between the models is rather more interesting. In fact, as we can observe in Tables 4.8, 4.9 and 4.10, the classical model gives always better result than the modified velocity, although it usually requires more iterations to achieve convergence. It is worth to notice also that using the modified velocity the difference between the monotone and the filtered scheme is more marked, with the latter performing better for all refinements. The overall results are satisfying using both models.

In Figure 4.4 we show the good properties of the new model, by some contour plots of the initial conditions and the final representations obtained using the monotone scheme. In fact, it is clear that the representations preserve the steepness of the gradient and that the level sets do not collide during the evolutionary process, as desired.

Table 4.8.: (Test 3a) Errors and number of iterations varying Δx.

c	**Norm** $\|\|\ \|\|_\infty$			**Monotone**			**AF-LW**		
Δx	tol	p	N	N_i	$P\text{-}Err_{rel}$	$P\text{-}Err_1$	N_i	$P\text{-}Err_{rel}$	$P\text{-}Err_1$
0.04	0.005	3	2	149	0.1305	0.8160	159	0.1310	0.8192
0.02	0.001	3	2	505	0.0657	0.4088	653	0.0502	0.3128
0.01	0.0005	3	2	1240	0.0433	0.2692	1507	0.0399	0.2482

Table 4.9.: (Test 3a) Errors and number of iterations varying Δx (L^∞ norm).

$\widetilde{c}$	**Datum** 1			**Monotone**			**AF-LW**		
Δx	tol	p	N	N_i	$P\text{-}Err_{rel}$	$P\text{-}Err_1$	N_i	$P\text{-}Err_{rel}$	$P\text{-}Err_1$
0.04	0.001	3	2	132	0.1561	0.9760	163	0.1244	0.7776
0.02	0.0002	3	2	375	0.0923	0.5748	507	0.0669	0.4164
0.01	0.00004	3	2	1162	0.0487	0.3024	1276	0.0439	0.2726

4.4.2. Real images

Here, we present some real cases, repeating the comparison between the two models. A quantitative comparison is done for Test 4 (real vase), Test 5 (grains) and Test 6 (chess horse), for which we are also able to compute the errors, and for Test 7 (brain) and

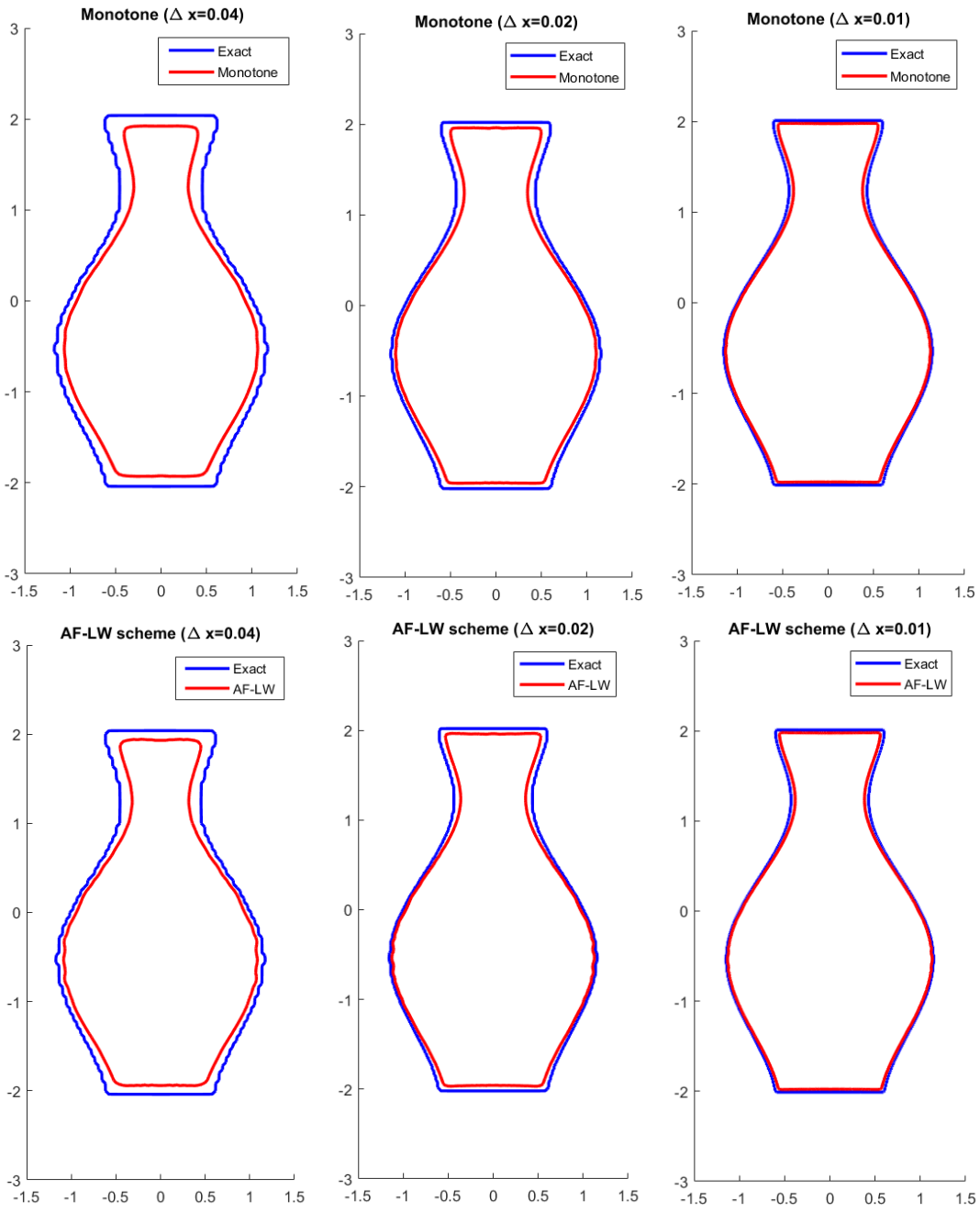

Figure 4.3.: (Test 3a) Datum 1. Plots of the final front varying Δx with error computed in L^∞ norm, using the monotone scheme (top) and the AF-LW scheme (bottom) and velocity $\tilde{c}$.

Test 11 (hip bone fracture), for which we use an artificial mask to compute analogous quantities. In such cases, looking at the relative tables, we can observe again that the

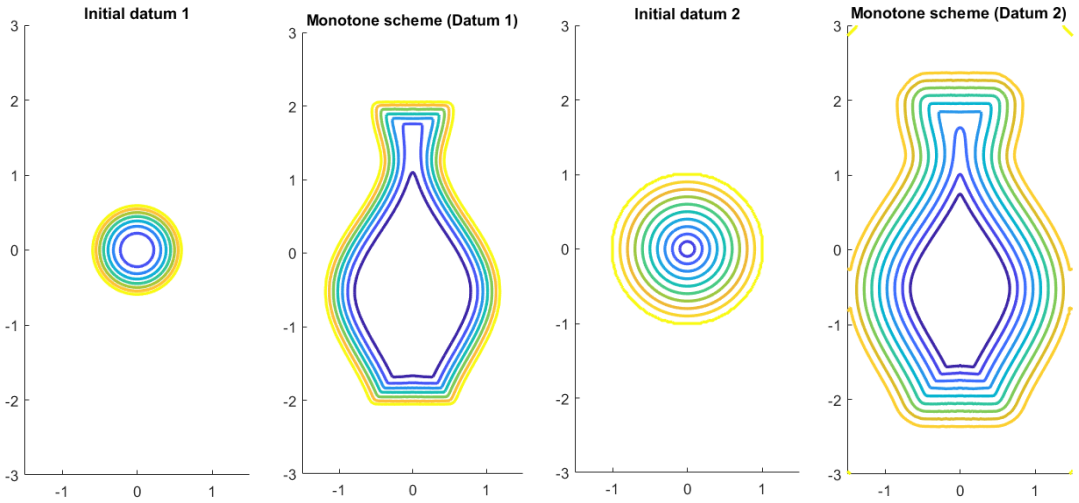

Figure 4.4.: (Test 3a) Contour plots of the initial datum and of the final representations, using the monotone scheme and velocity $\tilde{c}$, for both choices of the initial datum.

Table 4.10.: (Test 3a) Errors and number of iterations varying Δx (L^∞ norm).

$\widetilde{c}$	**Datum 2**			**Monotone**			**AF-LW**		
Δx	tol	p	N	N_i	$P\text{-}Err_{rel}$	$P\text{-}Err_1$	N_i	$P\text{-}Err_{rel}$	$P\text{-}Err_1$
0.04	0.001	3	2	143	0.1433	0.8960	161	0.1254	0.7840
0.02	0.0002	3	2	505	0.0893	0.5560	496	0.0677	0.4216
0.01	0.00004	3	2	1176	0.0477	0.2962	1237	0.0454	0.2824

modified velocity gives always better results (especially in terms of stability, since both schemes converge sooner) and that the adaptive filtered scheme has lower errors than the monotone scheme in most situations.

Finally, we conclude this section with some examples using biomedical images. These last tests will be performed only using the modified model, since we were not able to achieve convergence with the classical model, also considering a wide range of possible choices for the parameters N, p and tol.

Test 4. Real vase

The first real image tested is rather common in the literature on the topic, that is the real vase. Here we perform the simulation for both cases of an expanding (Case a) and a shrinking front (Case b), referring to the introduction at the beginning of the section for the choices of the initial conditions w.r.t. the model.

It it evident from the tables below that the schemes achieve good (but different) results in all simulations, for both velocities. In the case of the expanding front, for example, the classical model gives better results in terms of error, but it usually needs many more iterations to achieve convergence w.r.t. the modified velocity, which produces more

stable solutions, especially with the adaptive filtered scheme (see Tables 4.11-4.12 and Figure 4.6). Moreover, in Table 4.12 we can see that the AF-LW scheme has lower errors w.r.t. the monotone scheme, with the exception of the first choice for the tolerance using Datum 2, when it probably converges too soon. On the other hand, in Case b (Tables 4.13-4.14 and Figure 4.7), the new velocity improves the results in both errors and N_i, in particular those of the monotone scheme, which performs even better than the AF-LW scheme.

Moreover, as displayed in Figure 4.8, the properties of the modification are preserved even with non-analytical data, although some little oscillations (noise) are visible.

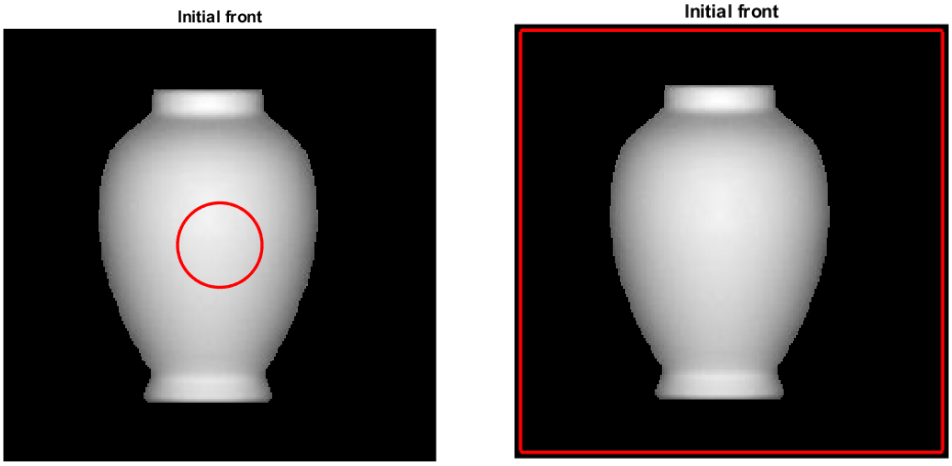

Figure 4.5.: (Test 4) Initial fronts for the two cases tested.

Table 4.11.: (Test 4a) Errors and number of iterations varying the stopping rule.

c	$\Delta x = 0.02$			**Monotone**			**AF-LW**		
Norm	tol	p	N	N_i	$P\text{-}Err_{rel}$	$P\text{-}Err_1$	N_i	$P\text{-}Err_{rel}$	$P\text{-}Err_1$
$\|\|\ \|\|_\infty$	0.0005	3	3	672	0.0332	0.2536	1359	0.0122	0.0956
$\|\|\ \|\|_1$	0.0001	3	3	539	0.0361	0.2756	835	0.0285	0.2180

Table 4.12.: (Test 4a) Errors and number of iterations using L^∞ norm.

$\widetilde{c}$	$\Delta x = 0.02$			**Monotone**			**AF-LW**		
Datum	tol	p	N	N_i	$P\text{-}Err_{rel}$	$P\text{-}Err_1$	N_i	$P\text{-}Err_{rel}$	$P\text{-}Err_1$
1	0.0005	3	3	334	0.0451	0.3444	324	0.0454	0.3472
1	0.0001	3	3	454	0.0384	0.2932	482	0.0372	0.2844
2	0.0005	3	3	338	0.0445	0.3400	328	0.0449	0.3432
2	0.0001	3	3	463	0.0381	0.2916	479	0.0371	0.2836

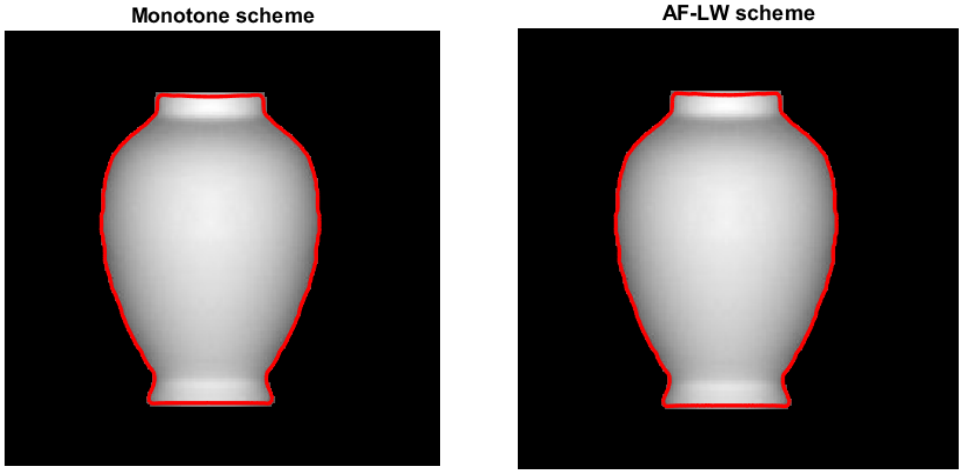

Figure 4.6.: (Test 4a) Plots of the final front using the monotone scheme (left) and the AF-LW scheme (right) and velocity $\widetilde{c}$, with L^∞ norm and $tol = 0.0001$, Datum 1, $\Delta x = 0.02$, $p = 3$ and $N = 3$.

Table 4.13.: (Test 4b) Errors and number of iterations varying the stopping rule.

c	$\Delta x = 0.02$			**Monotone**			**AF-LW**		
Norm	tol	p	N	N_i	$P\text{-}Err_{rel}$	$P\text{-}Err_1$	N_i	$P\text{-}Err_{rel}$	$P\text{-}Err_1$
$\|\|\ \|\|_\infty$	0.005	3	0	176	0.0311	0.2376	174	0.0267	0.2040
$\|\|\ \|\|_1$	0.00001	3	0	186	0.0308	0.2352	176	0.0267	0.2040

Table 4.14.: (Test 4b) Errors and number of iterations using L^∞ norm.

$\widetilde{c}$	$\Delta x = 0.02$			**Monotone**			**AF-LW**		
	tol	p	N	N_i	$P\text{-}Err_{rel}$	$P\text{-}Err_1$	N_i	$P\text{-}Err_{rel}$	$P\text{-}Err_1$
	0.0001	3	0	165	0.0218	0.1664	161	0.0235	0.1796

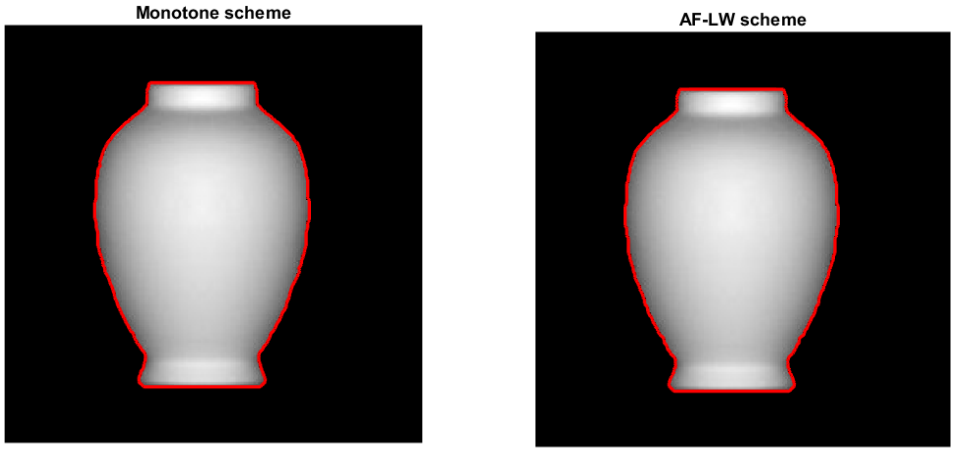

Figure 4.7.: (Test 4b) Plots of the final front using the monotone scheme (left) and the AF-LW scheme (right) with velocity $\widetilde{c}$, L^∞ norm, $tol = 0.0001$, $\Delta x = 0.02$, $p = 3$, $N = 0$.

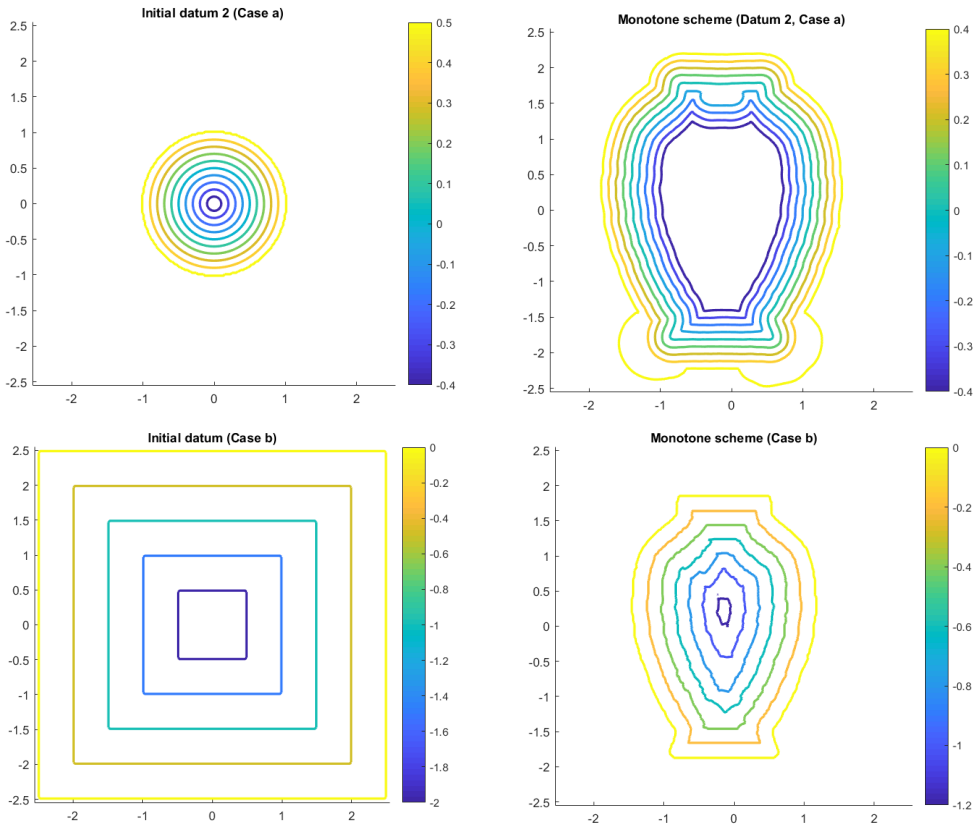

Figure 4.8.: (Test 4) Contour plots of the initial datum (left) and the final representations (right), using the monotone scheme and velocity $\tilde{c}$. Top: Case a, bottom: Case b.

Test 5. Grains

The following test is focused on the problem of approximating the boundaries of various separate objects in a picture. Clearly, using our (basic) approach, we can face the problem only in the case of a shrinking from the frame of the picture, since we need an initial condition such that the front envelopes all the objects. Looking at Tables 4.15-4.16 and Figure 4.9 we can see that both schemes perform well in all the simulations. Notice that in the case of the classical model we avoided to regularize the data, since otherwise is very difficult to have the scheme that achieves convergence, whereas using the velocity $\tilde{c}$ we are forced to filter the image, at least a little, in order to nullify some spurious oscillations otherwise visible. We believe that this phenomenon is connected to the definition of ∇u in the precomputation of the velocity $\tilde{c}^n$, which should be adapted to treat more coherently also singular (or separating) fronts. It is good to point out that, when using the velocity c, the AF-LW scheme stops correctly on the boundary of the objects only choosing at least $p \geq 5$, even when not applying the gaussian filter. In fact, in all the other cases tested, it normally crosses the boundary of some less marked grain and keeps shrinking.

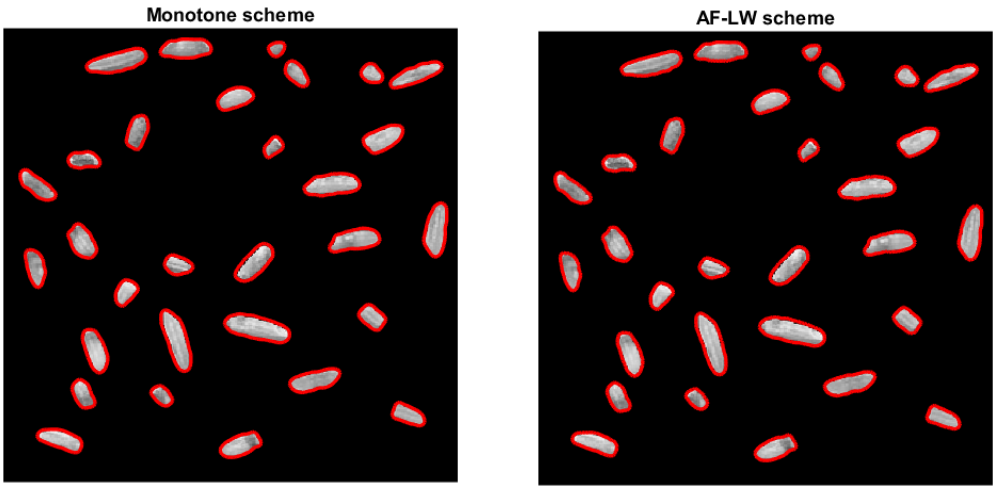

Figure 4.9.: (Test 5b) Plots of the final front using the monotone scheme (left) and the AF-LW scheme (right) and velocity $\widetilde{c}$, with L^1 norm and $tol = 0.0001$, $\Delta x = 0.02$, $p = 2$ and $N = 2$.

Table 4.15.: (Test 5b) Errors and number of iterations varying the stopping rule.

c	$\Delta x = 0.02$			**Monotone**			**AF-LW**		
Norm	tol	p	N	N_i	$P\text{-}Err_{rel}$	$P\text{-}Err_1$	N_i	$P\text{-}Err_{rel}$	$P\text{-}Err_1$
$\|\|\ \|\|_\infty$	0.002	5	0	329	0.0807	0.2628	319	0.0507	0.1652
$\|\|\ \|\|_1$	0.0001	5	0	315	0.0861	0.2804	308	0.0511	0.1664

Table 4.16.: (Test 5b) Errors and number of iterations varying the stopping rule.

$\widetilde{c}$	$\Delta x = 0.02$			**Monotone**			**AF-LW**		
Norm	tol	p	N	N_i	$P\text{-}Err_{rel}$	$P\text{-}Err_1$	N_i	$P\text{-}Err_{rel}$	$P\text{-}Err_1$
$\|\|\ \|\|_\infty$	0.001	2	2	319	0.0141	0.0460	354	0.0372	0.1212
$\|\|\ \|\|_1$	0.0001	2	2	312	0.0097	0.0316	299	0.0074	0.0240

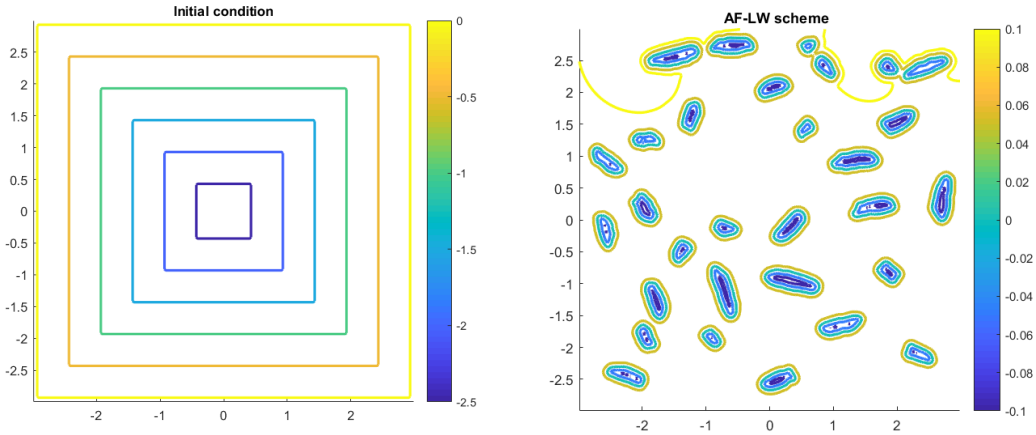

Figure 4.10.: (Test 5) Contour plots of the initial datum (left) and the final representations (right), using the AF-LW scheme and velocity $\tilde{c}$.

Test 6. Chess horse

The following test is devised in order to show the capability of the schemes (and of the modified model) of handling more complex shapes. We choose a picture of the horse piece in the game of chess and we approximate its boundary from inside (Case a), varying the initial datum, and outside (Case b). The simulations are performed using only the modified model, because of the usual stability problems of the solutions obtained by the classical model (the main issue seems to be the approximation of the mouth of the horse).

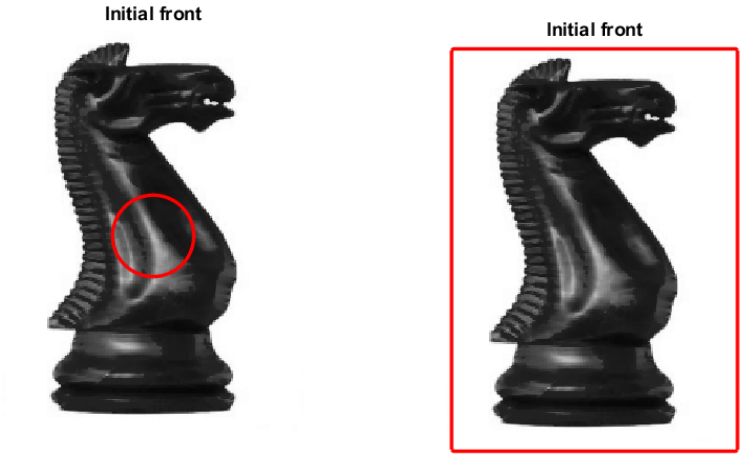

Figure 4.11.: (Test 6) Initial fronts for the two cases tested.

From Tables 4.17 and 4.18, it is clear the advantage of the adaptive filtered scheme over the monotone scheme, since it produces a more accurate segmentation whenever it achieves convergence, in particular using the same tolerance. Notice that in the last test of Table 4.17 the filtered scheme does not give usable results, since the front overcomes the boundary (in proximity of the neck) and keeps expanding, but it has better errors than the monotone scheme even with similar number of iterations. We would like to remark also that, if we reduce the tolerance even more ($tol = 0.00001$) then neither the monotone scheme gives usable results, for the same reason.

Table 4.17.: (Test 6a) Errors and number of iterations using L^1 norm.

$\widetilde{c}$	$\Delta x = 0.02$			**Monotone**			**AF-LW**		
Datum	tol	p	N	N_i	$P\text{-}Err_{rel}$	$P\text{-}Err_1$	N_i	$P\text{-}Err_{rel}$	$P\text{-}Err_1$
1	0.00005	2	5	433	0.0794	0.6424	396	0.0672	0.5436
1	0.000025	2	5	508	0.0599	0.4848	544	0.0438	0.3544
2	0.00005	2	5	400	0.0925	0.7488	532	0.0511	0.4136
2	0.000025	2	5	501	0.0613	0.4960	1248	X	X

Table 4.18.: (Test 6b) Errors and number of iterations using L^1 norm.

$\widetilde{c}$	$\Delta x = 0.02$			**Monotone**			**AF-LW**		
	tol	p	N	N_i	$P\text{-}Err_{rel}$	$P\text{-}Err_1$	N_i	$P\text{-}Err_{rel}$	$P\text{-}Err_1$
	0.00005	2	3	187	0.0303	0.2452	197	0.0273	0.2212

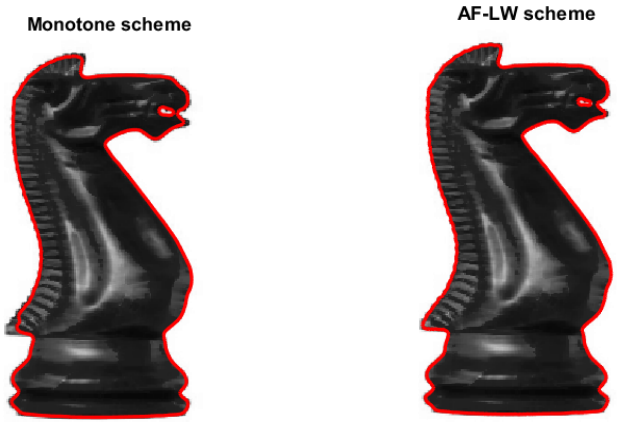

Figure 4.12.: (Test 6a) Datum 1. Plots of the final front using the monotone scheme, $N_i = 508$ (left), and the AF-LW scheme, $N_i = 548$ (right), with L^1 norm and $tol = 0.000025$, $p = 2$ and $N = 5$, $\Delta x = 0.02$ and velocity $\widetilde{c}$.

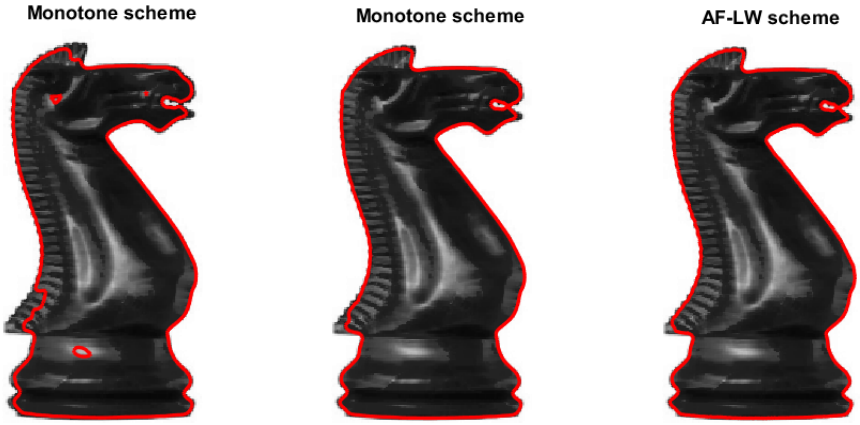

Figure 4.13.: (Test 6a) Datum 2. Plots of the final front using the monotone scheme with $tol = 0.00005$, $N_i = 400$ (left), and with $tol = 0.000025$, $N_i = 501$ (middle), and using the AF-LW scheme with $tol = 0.00005$, $N_i = 497$ (right), in L^1 norm, with $p = 2$ and $N = 5$, $\Delta x = 0.02$ and velocity $\widetilde{c}$.

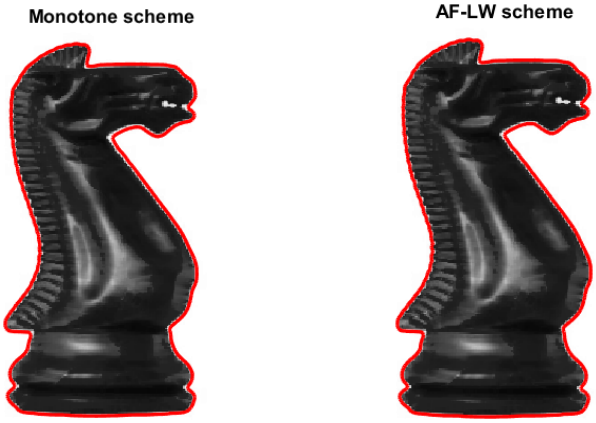

Figure 4.14.: (Test 6b) Plots of the final front using the monotone scheme (left) and the AF-LW scheme (right) with velocity $\widetilde{c}$, with L^1 norm and $tol = 0.0005$, $\Delta x = 0.02$, $p = 2$ and $N = 3$.

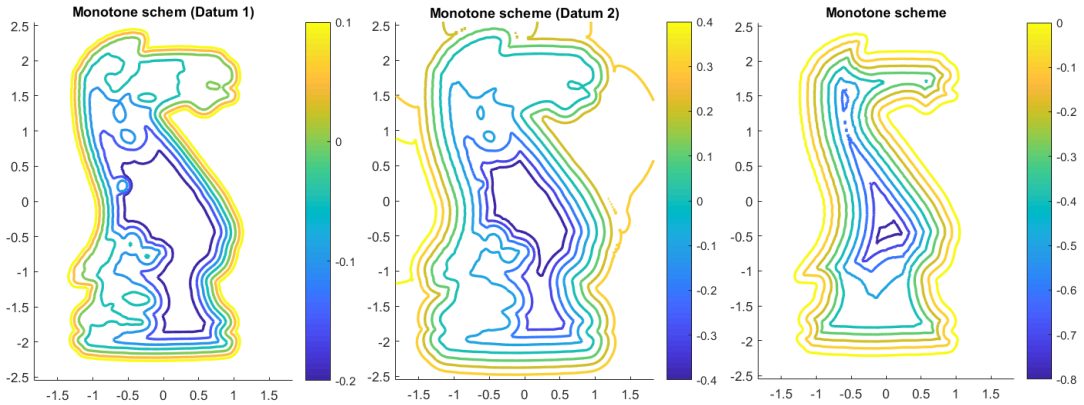

Figure 4.15.: (Test 6) Contour plots of the the final representations using the monotone scheme and velocity $\tilde{c}$, for Case a with Datum 1 (left) and Datum 2 (middle), and for Case b (right).

We conclude the section related to numerical tests by giving examples of possible application of the proposed method in the biomedical field. We test the modified model on a series of biomedical images of different nature, which we collected on the internet, comparing the results obtained by the schemes mainly in terms of number of iterations N_i and, clearly, qualitatively looking at the different figures reported. Whenever possible we will also present the comparison in terms of error, where the precise sense will be specified case by case. In most cases, since we are not able to compute the error, we avoid to collect the results in the tables, writing the details of the simulation in the caption of the relative figure.

Test 7. Brain

The first biomedical test proposed is a picture of a human brain. We approximate the relevant boundary from both the inside (Case a) and the outside (Case b), varying also the initial datum in the case of the expanding front. Moreover, in Case b, we are also able to compute a "reasonable" error by constructing artificially (by hand) the mask depicted in Figure 4.21.

The effectiveness of the modified model is testified by the figures proposed below, in which are clearly visible also the better results obtained with the filtered scheme over the monotone scheme. In particular, Figure 4.19 seems rather emblematic. Although the expanding front may have a complex evolution due to non-smooth real data, the gradient of the initial datum seems to keep being preserved, even if new relevant fronts arise inside the surface. Even more clearly, this behavior can be seen also in Figure 4.21, where the distance function to the 0-level set is still recognizable.

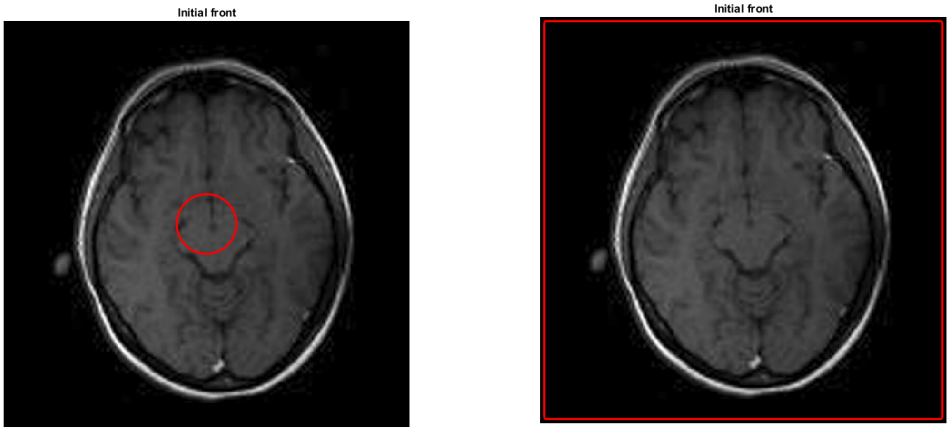

Figure 4.16.: (Test 7) Initial fronts for the two cases tested.

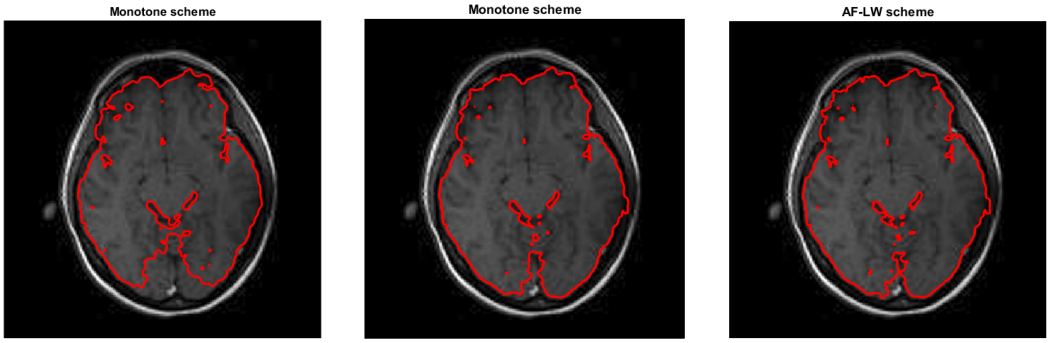

Figure 4.17.: (Test 7a) Datum 1. Plots of the final front using the monotone scheme with $tol = 0.00001$, $N_i = 376$ (left), and with $tol = 0.000005$, $N_i = 468$ (middle), and using the AF-LW scheme with $tol = 0.00001$, $N_i = 403$ (right), all using L^1 norm in the stopping criterion, with $p = 4$ and $N = 5$, $\Delta x = 0.01$ and velocity $\widetilde{c}$.

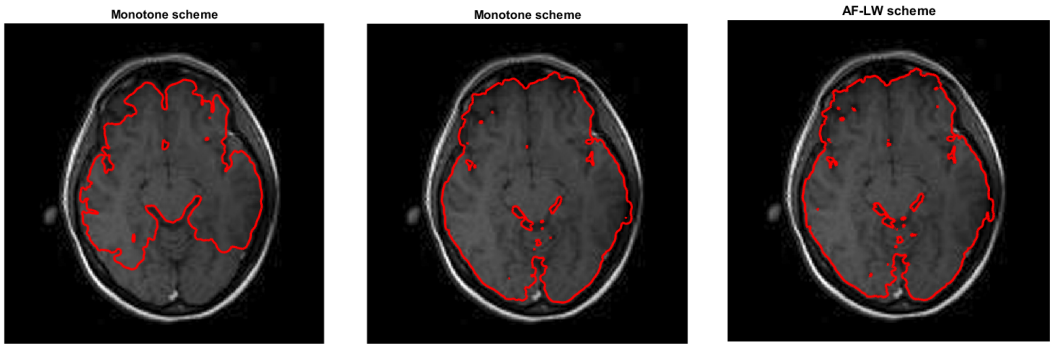

Figure 4.18.: (Test 7a) Datum 2. Plots of the final front using the monotone scheme with $tol = 0.00005$, $N_i = 264$ (left), and with $tol = 0.00001$, $N_i = 495$ (middle), and using the AF-LW scheme with $tol = 0.00005$, $N_i = 431$ (right), all using L^1 norm in the stopping criterion, with $p = 4$ and $N = 5$, $\Delta x = 0.01$ and velocity $\widetilde{c}$.

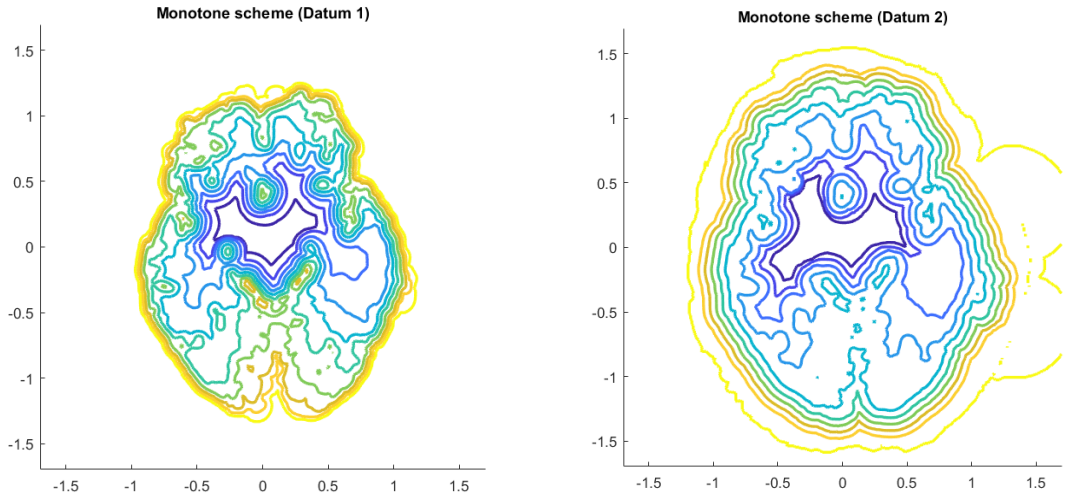

Figure 4.19.: (Test 7a) Contour plots of the the final representations using the monotone scheme and velocity $\tilde{c}$, with Datum 1 (left) and Datum 2, $N_i = 495$, (right).

Table 4.19.: (Test 7b) Errors and number of iterations using L^1 norm.

$\widetilde{c}$	$\Delta x = 0.02$			**Monotone**			**AF-LW**		
	tol	p	N	N_i	$P\text{-}Err_{rel}$	$P\text{-}Err_1$	N_i	$P\text{-}Err_{rel}$	$P\text{-}Err_1$
	0.00005	5	3	228	0.0118	0.2476	262	0.0078	0.1628

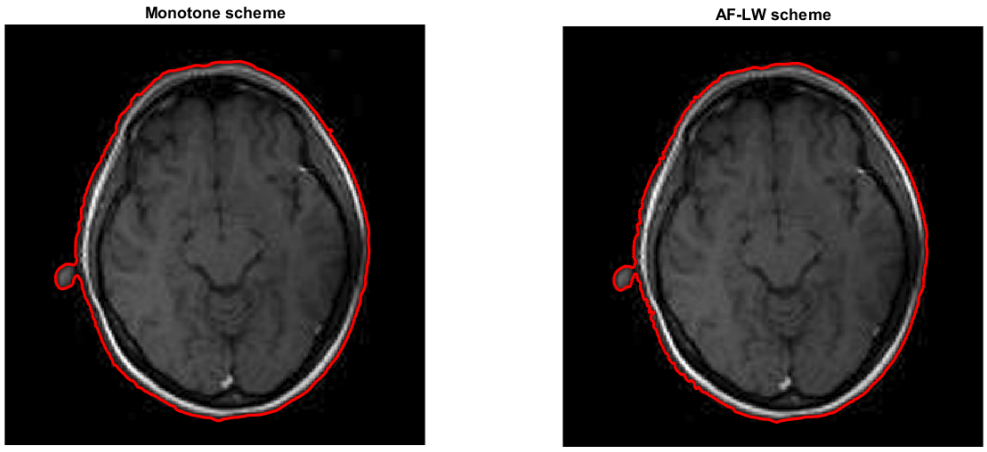

Figure 4.20.: (Test 7b) Plots of the final front using the monotone scheme (left) and the AF-LW scheme (right) with velocity $\widetilde{c}$.

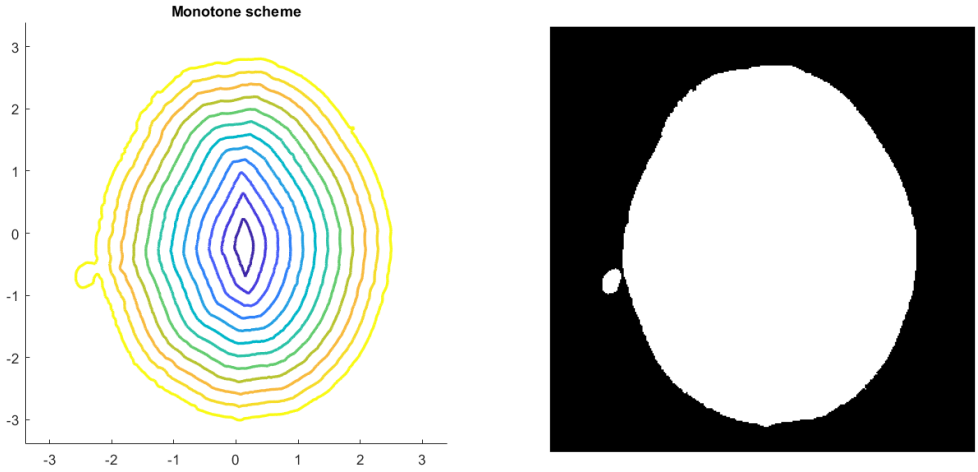

Figure 4.21.: (Test 7b) On the left, contour plot of the the final representations using the monotone scheme, and on the right the "artificial" mask used to compute the error..

Test 8-9 Wrist fractures

We continue with the biomedical examples by testing a series of fractures of human bones. For evident reasons we approach the problem only in the case of the front shrinking from the frame of the image, trying to capture the relevant (in some sense) boundaries of the considered profiles. The aim is the same, that is to show the reliability of the modified model and the implementation proposed, trying also to exploit some advantage using the adaptive filtered scheme over the monotone scheme.

In the first two tests we work on images of wrist fractures, for which the "relevant" boundary to be segmented it is not very clear. In any case, since it seems a rather interesting situation, we perform the simulation and report the results in Figures 4.22 and 4.23. In both tests, the AF-LW scheme inspects the boundary more deeply, being able to segment also some bones of the wrist in the first case, and with higher accuracy, as can be noted focusing on the area of the fingers in both cases. In the figures we also reported the contour plots of the final representations, from we can observe that the gradient of the initial data is preserved, except in proximity of the protrusions between the fingers, where the level set collapse into each other. This behavior is most probably due to the non-convexity of the level sets.

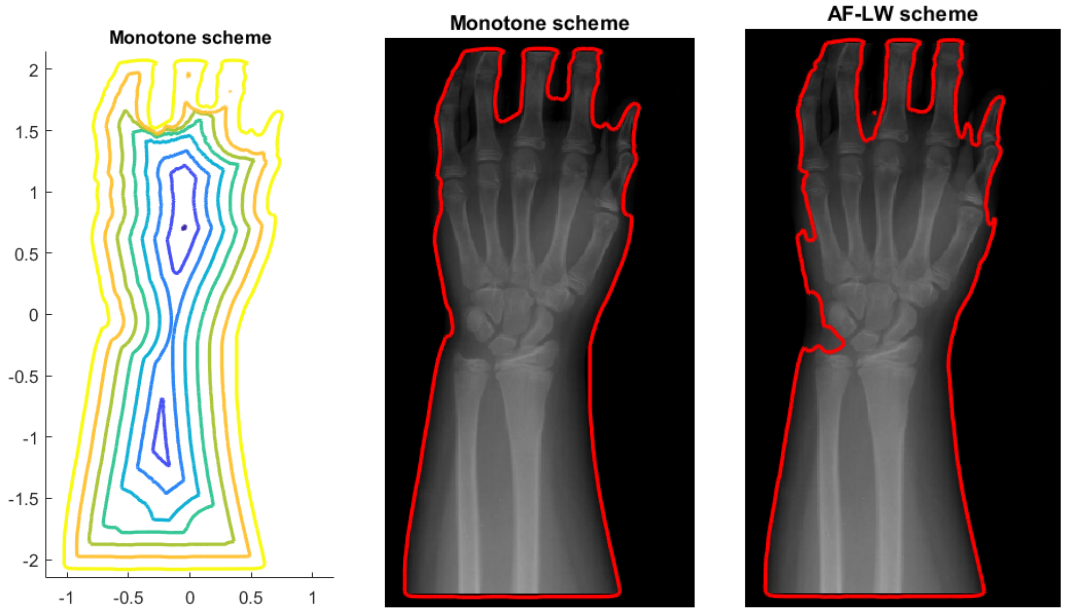

Figure 4.22.: (Test 8b) From left to right: final representation and plot of the final front using the monotone scheme ($N_i = 160$), and final front using the AF-LW scheme ($N_i = 246$), with L^1 norm and $tol = 0.00008$, $p = 4$, $N = 5$, $\Delta x = 0.01$ and velocity $\widetilde{c}$.

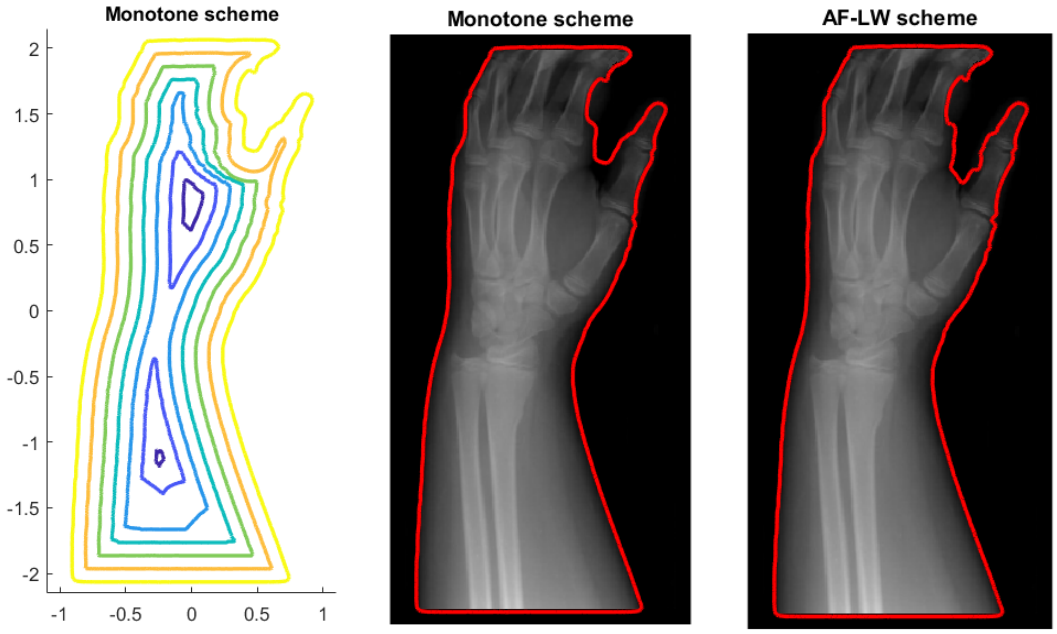

Figure 4.23.: (Test 9b) From left to right: final representation and plot of the final front using the monotone scheme ($N_i = 191$), and final front using the AF-LW scheme ($N_i = 223$), with L^1 norm and $tol = 0.00002$, $p = 5$, $N = 3$, $\Delta x = 0.01$ and velocity $\widetilde{c}$.

Test 10. Foot fracture

The following test concerns the image of a foot, where the boundary to be recovered is much more clear. It is noteworthy that also in this case we have another boundary clearly visible (the shape of the entire foot), although it is less marked w.r.t. the previous tests on wrist fractures. Consequently, choosing the parameters in an appropriate manner (especially N), we are able to recover almost the full shape of the bones, at least using the AF-LW scheme, whereas the monotone scheme stops the evolution too soon, being more sensitive to the slowing effects of the shaded (not relevant) boundary (see Figure 4.24).

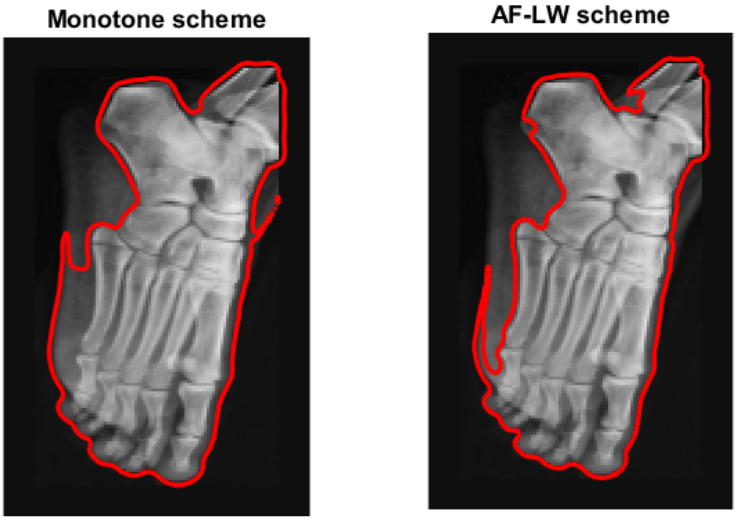

Figure 4.24.: (Test 10b) Plots of the final front using the monotone scheme ($N_i = 156$), and the AF-LW scheme ($N_i = 243$), with L^1 norm and $tol = 0.00005$, $p = 4$, $N = 5$, $\Delta x = 0.02$ and velocity $\widetilde{c}$.

Test 11. Hip bone fracture

The last test is concerning a hip bone fracture. Here we simply approximate the shape of the bones from the outside, then we compute the error using the same trick of Test 7, as shown in Figure 4.26. In the same figure we have also the contour plot of the final representation. It is interesting to notice that the gradient of the representation is preserved in most part of the surface, except for the deep protrusion, where the level sets collapse into each other.

From Table 4.20 we can detect a rather clear behavior. For the tested choices of the parameters, the monotone scheme is clearly more stable, achieving convergence for all tolerance cases. Nevertheless, if we compare the (best) results which are, in particular, comparable in terms of N_i, we see that the adaptive filtered schemes has a better accuracy in terms of our computed error. On the other hand, for the second choice of the tolerance the front keeps shrinking loosing the entire shape. It is worth to

point out that if $tol = 0.00001$ neither the front governed by the monotone scheme is able to stop correctly.

Table 4.20.: (Test 11b) Errors and number of iterations using L^1 norm.

$\widetilde{c}$	$\Delta x = 0.01$			**Monotone**			**AF-LW**		
	tol	p	N	N_i	$P\text{-}Err_{rel}$	$P\text{-}Err_1$	N_i	$P\text{-}Err_{rel}$	$P\text{-}Err_1$
	0.00004	3	4	176	0.0271	0.3571	305	3.87e-04	0.0051
	0.00002	3	4	306	0.0034	0.0444	–	X	X

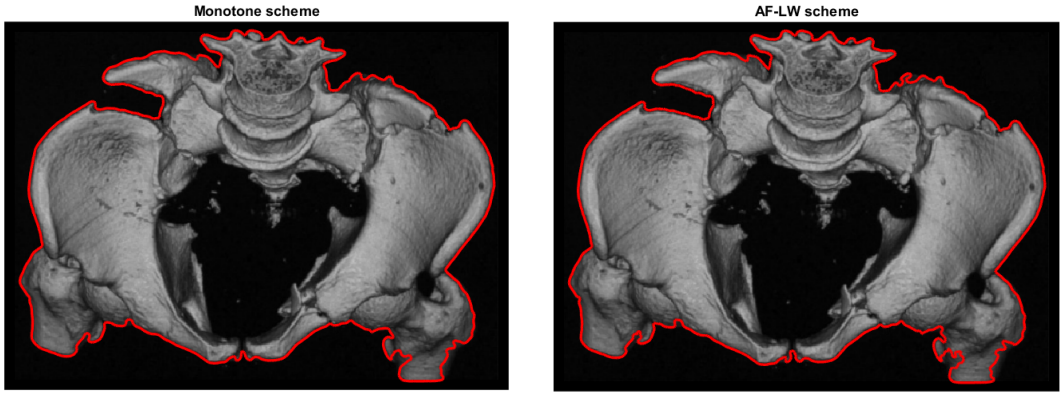

Figure 4.25.: (Test 11b) Plots of the final front using the monotone scheme, $N_i = 306$ (left) and the AF-LW scheme, $N_i = 306$ (right).

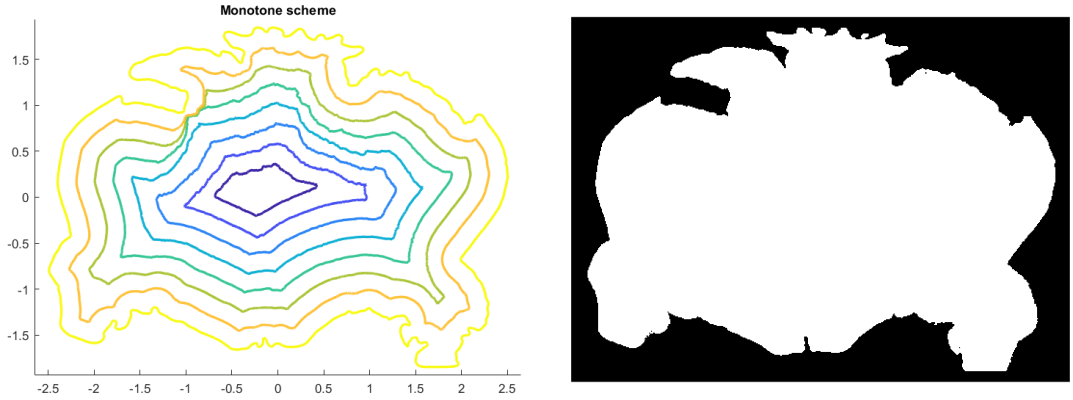

Figure 4.26.: (Test 11b) On the left, contour plot of the the final representations using the monotone scheme. On the right the "artificial" mask used to compute the error.

Test 12. Pneumonia

Finally, we conclude our numerical tests by proposing a more complex disposition of the initial condition, which naturally leads to collapsing and merging fronts. The initial datum here is composed by four equal paraboloids (each defined as Datum 1) or cones (defined as Datum 2), placed as shown in Figure 4.27, with 0-level sets composed by circles of radius $r = 0.125$.

From Figures 4.28 and 4.29 it is evident that the AF-LW scheme produces better results using both initial conditions. Moreover, for both schemes the results vary changing the initial condition, with better results coming from the distance function. The good behavior of the scheme and of the modified model in the case of a merging fronts is, again, testified by the contour plots of the final representations, shown in Figure 4.30.

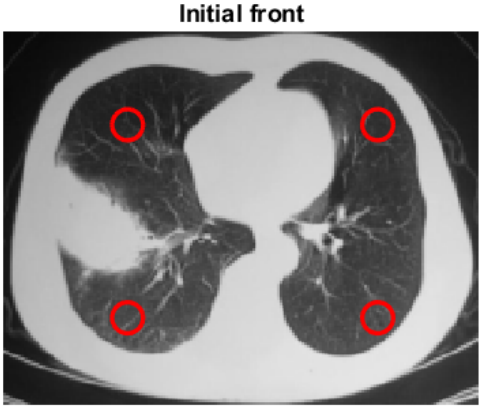

Figure 4.27.: (Test 12a) Initial front composed of four separate circles of radius $r = 0.125$.

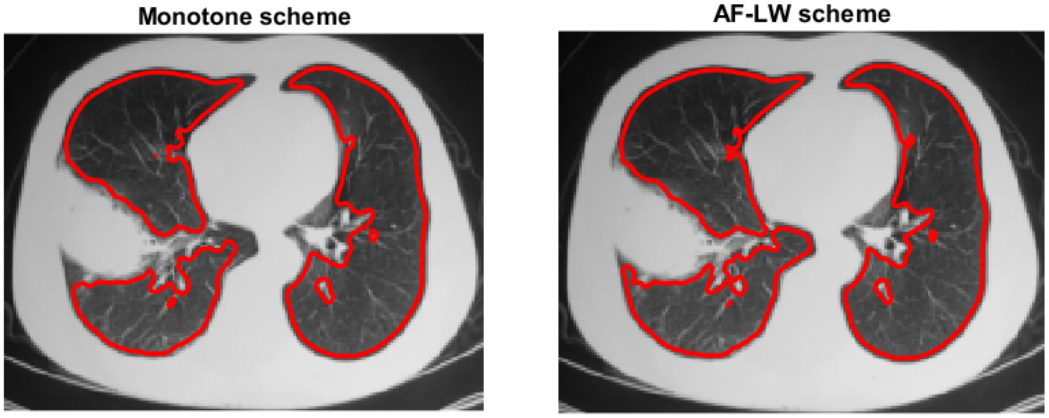

Figure 4.28.: (Test 12a) Datum 1. Plots of the final front using the monotone scheme ($N_i = 185$), and the AF-LW scheme ($N_i = 194$), with L^1 norm and $tol = 0.00001$, $p = 4$, $N = 5$, $\Delta x = 0.02$ and velocity $\widetilde{c}$.

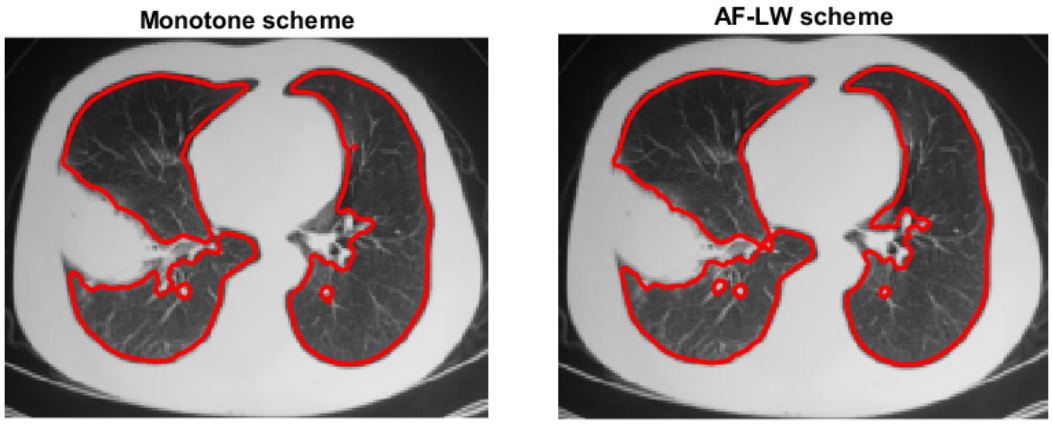

Figure 4.29.: (Test 12a) Datum 2. Plots of the final front using the monotone scheme ($N_i = 413$), and the AF-LW scheme ($N_i = 589$), with L^1 norm and $tol = 0.00001$, $p = 4$, $N = 5$, $\Delta x = 0.02$ and velocity $\widetilde{c}$.

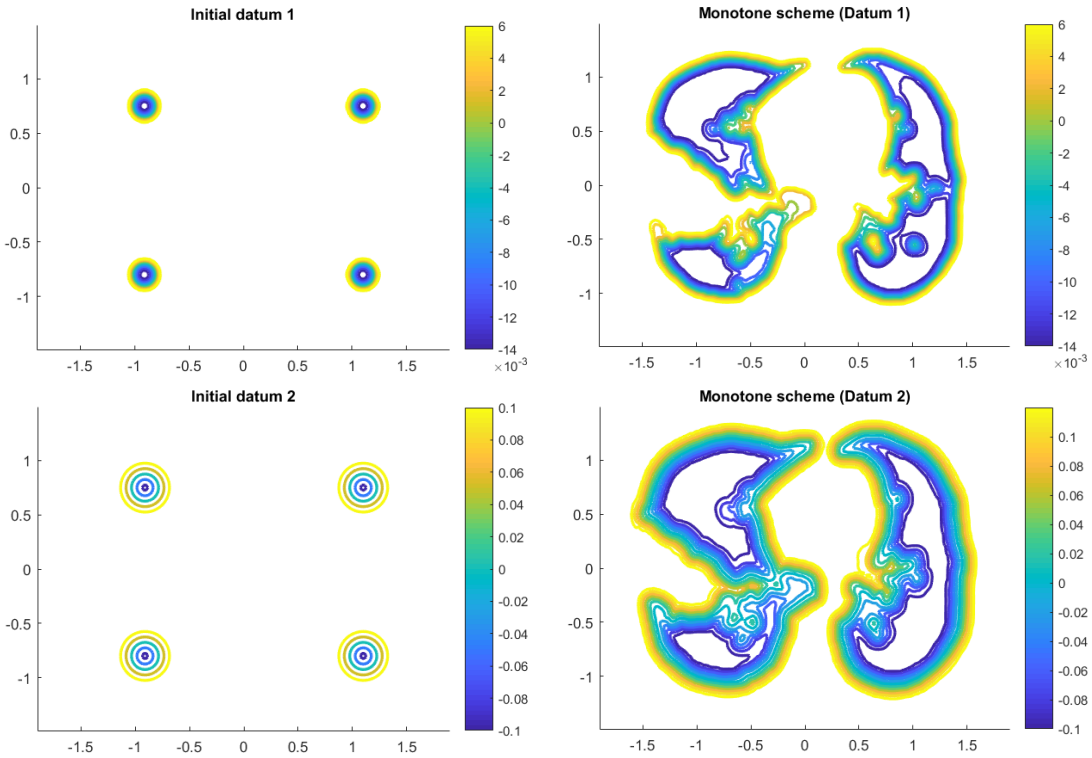

Figure 4.30.: (Test 12a) Contour plots of the initial datum (left) and the final representations (right), using the monotone scheme. Top: Datum 1, bottom: Datum 2.

4.5. Conclusions

In this chapter we have presented an efficient procedure for the segmentation of synthetic and real images, based on the level set method. Thanks to the new definition of the velocity $\tilde{c}$ we were able to achieve interesting results in many complex situations, mainly regarding biomedical applications. The modification of the velocity is rather straightforward and has been justified by the method of characteristic. Moreover, our preliminary implementation, although clearly needs some improvement, has given promising responses. The modified model let us choose the numerical scheme more freely, since it allows continuous solutions, and then show the possible advantages offered by the high-order Adaptive Filtered Scheme with respect to the basic monotone scheme.

Future developments will involve a deeper inspection of the evolution, trying to relax the assumption on the non-intersection of the characteristics. Probably, one way to overcome this shortcoming is to define the function $d(u)$ according to (4.19). Moreover, a more stable procedure to reconstruct the point used to define $\tilde{c}$ is to be inspected, mainly w.r.t. the computation of ∇u (which is not well defined if the surface develops a singularity). Finally, it would be ideal to devise a formula for the computation of the tolerance of the stopping criterion, eventually dependent only on the initial condition.

Bibliography

[AA00] S. Augoula, R. Abgrall, *High order numerical discretization for Hamilton-Jacobi equations on triangular meshes*, J. Sci. Comput., **15** (2), (2000), 197-229.

[A09] R. Abgrall, *Construction of simple, stable and convergent high order scheme for steady first order Hamilton-Jacobi equation*, SIAM J. Sci. Comput., **31**, 2419-2446, 2009.

[ABM10] F. Aràndiga, A. M. Belda, P. Mulet, *Point-value WENO multiresolution applications to stable image compression*, J. Sci. Comput., **43** (2010), pp. 158–182.

[BCD97] M. Bardi, I. Capuzzo-Dolcetta, Optimal Control and Viscosity Solution of Hamilton-Jacobi-Bellmann Equations, Birkhauser, Boston Heidelberg 1997

[BF90] M. Bardi, M. Falcone, *An approximation scheme for the minimum time function*, SIAM J. Control Optim., **28** (1990), 950-965.

[B98] G. Barles, Solutions de viscositè des equations d'Hamilton-Jacobi, Springer-Verlag, 1998.

[BS91] G. Barles, P. E. Souganidis, *Convergence of approximation schemes for fully nonlinear second order equations*, Asympt. Anal., **4** (1991), 271–283.

[BT01] G. Barles, A. Tourin, *Commutation properties of semigroups for first-order Hamilton-Jacobi equations and application to multi-time equations*, Indiana U. Math. J. **50** (2001), 1523–1544.

[BCS13] O. Bokanowski, Y. Cheng, and C.-W. Shu, *A discontinuous galerkin scheme for front propagation with obstacle*, Numer. Math., **126**(2) (2013),1-31, 2013.

[BFFKZ14] O. Bokanowsi, M.Falcone, R.Ferretti, L.Grüne, D.Kalise, H.Zidani, *Value iteration convergence of ε-monotone schemes for stationary Hamilton-Jacobi equations*, Discrete Contin. Dyn, Syst. Se. A, **35** (2015), 4041-4070.

[BFS16] O. Bokanowski, M. Falcone, S. Sahu, *An efficient filtered scheme for some first order Hamilton-Jacobi-Bellman equations*, SIAM Journal on Scientific Computing. **38**(1) (2016), A171-A195.

[BCCD08] R. Borges, M. Carmona, B. Costa, W.S. Don, *An improved weighted essentially non-oscillatory scheme for hyperbolic conservation laws*, J. Comput. Phys., 227 (6) (2008) 3101–3211.

[CCD11] M. Carmona, B. Costa, W.S. Don, *High order weighted essentially non-oscillatory WENO-Z schemes for hyperbolic conservation laws*, J. Comput. Phys., **230** (2011), 1766-1792.

[CL83] M. G. Crandall, P. L. Lions, *Viscosity solutions of Hamilton-Jacobi equations*, Trans. Amer. Math. Soc., **277** (1983), 1-42.

[CL84] M. G. Crandall, P. L. Lions, *Two Approximations of solutions of Hamilton-Jacobi equations*, Math. Comput. **43**, 1 (1984).

[CFF10] E. Carlini, M. Falcone, R. Ferretti, *Convergence of a large time-step scheme for mean curvature motion*, Interf. Free Bound., **12** (2010), 409-441.

[CFN95] L. Corrias, M. Falcone, R.Natalini, *Numerical schemes for conservation laws via Hamilton-Jacobi equations*, Math. Comp. **64**, (1995), 555-580.

[CFR05] E. Carlini, R. Ferretti, G. Russo, *A weighted essentially nonoscillatory, large-time step scheme for Hamilton-Jacobi equations*, SIAM J. Sci. Comput., **27**, (2005), 1071-1091.

[CIL92] M. G. Crandall, I. Ishii, P. L. Lions, *User's guide to viscosity solutions of second order partial differential equations*, Bull. Amer. Math. Soc., **27** (1992), 1-67.

[CS89] B. Cockburn, C.-W. Shu, *TVB Runge-Kutta local projection discontinuous Galerkin*

nite element method for conservation laws iii: One-dimensional systems, J. Comput. Phys., **84** (1989), 90-113.

[DFS08] J.D. Durou, M. Falcone, M. Sagona, *Numerical methods for shape from shading: a new survey with benchmarks*, Comput. Vis. Image Underst. **109**(1) (2008), 22–43.

[E10] L. C. Evans, Partial Differential Equations, Second Edition. Amer. Math. Soc. Graduate Studies in Mathematics, **19**, Providence 2010.

[FGL94] M. Falcone, T. Giorgi, P. Loreti, *Level set of viscosity solutions: some applications to front and rendez-vous problem*, SIAM J. Appl. Math., **54** (1994), 1335-1354.

[FF94] M. Falcone, R. Ferretti, *Discrete-time high-order schemes for viscosity solutions of Hamilton-Jacobi equations*, Numer. Math. **67**, 315 (1994).

[FF14] M. Falcone, R. Ferretti, Semi-Lagrangian Approximation schemes for linear and Hamilton-Jacobi equations, SIAM, 2014

[FPT18] M. Falcone, G. Paolucci, S. Tozza, *Adaptive Filtered schemes for first order Hamilton-Jacobi equations.* In: Proc. ENUMATH 2017, to appear

[FPTa] M. Falcone, G.Paolucci, S. Tozza, *Convergence of Adaptive Filtered schemes for Hamilton-Jacobi equations*, in preparation.

[FPTb] M. Falcone, G. Paolucci, S. Tozza, *Smoothness indicators for multidimensional Adaptive Filtered Schemes*, in preparation.

[FPTc] M. Falcone, G. Paolucci, S. Tozza *A High-Order Scheme for Image Segmentation via a modified Level-Set method*, in preparation.

[FT16] S. Tozza M. Falcone, *Analysis and Approximation of Some Shape-from-Shading Models for Non-Lambertian Surfaces*, J. Math. Imaging Vis. **55** (2016)

[F02] R. Ferretti, *Convergence of semi-Lagrangian approximations to convex Hamilton-Jacobi equations under (very) large Courant numbers*, SIAM J. Numer. Anal., **40** (2002), 2240-2253.

[FMT12] U. S. Fjordholm, S. Mishra, E. Tadmor, *Arbitrarily high order accurate entropy stable essentially non-oscillatory schemes for systems of conservation laws*, SIAM J. Numer. Anal., **50** (2012), 423-444.

[FO13] B. D. Froese, A. M. Oberman, *Convergent filtered schemes for the Monge-Ampère partial differential equation*, SIAM J. Numer. Anal., **51** (2013), 423-444.

[HEOC87] A. Harten, B. Engquist, S. Osher, and S.R. Chakravarty, *Uniformly high order essentially non-oscillatory schemes*, J. Comput. phys., **4** (1987) 231-303.

[HAP05] A. K. Henrick, T. D. Aslam, J. M. Powers, *Mapped weighted essentially non-oscillatory schemes: Achieving optimal order near critical points*, J. Comput. Phys., **207** (2005), pp. 542–567.

[HB89] B.K.P. Horn, M.J. Brooks (eds.), Shape from Shading, The MIT Press, Cambridge. Massachussetts, 1989.

[JP00] G. Jiang, D.-P. Peng, *Weighted ENO schemes for Hamilton-Jacobi equations*, SIAM Journal on Scientific Computing, **21** (2000), 2126-2143.

[JS96] G. Jiang, C.-W. Shu, *Efficient implementation of weighted ENO schemes*, Journal of Computational Physics, **126** (1996), 202-228.

[JX98] S. Jin, Z. Xin, *Numerical Passage from System of Conservation Laws to Hamilton-jacobi Equations, and a Relaxation Scheme*, SIAM J. on Num. Anal., **35** n.6, 2385-2404 (1998).

[KWT88] M. Kass, A. Witkin, D. Terzopoulos, *Snakes: Active Contour Models*, International Journal of Computer Vision, pp. 321—331, 1988.

[KNP01] A. Kurganov, S. Noelle, G. Petrova, *Semi-discrete central-upwind scheme for hyperbolic conservation laws and Hamilton-Jacobi equations*, SIAM J. Sci. Comput., **23** (2001), 707–740.

[L82] P. L. Lions, Generalized solution of Hamilton-Jacobi equations, Pitman, London, 1982.

[LR86] P.-L. Lions, J. Rochet, *Hopf formula and multi-time hamilton-jacobi equations*, Proc Amer Math Soc., **96** (1986), 79-84.

[LS95] P. L. Lions, P. Souganidis, *Convergence of MUSCL and filtered schemes for scalar conservation laws and Hamilton–Jacobi equations*, Num. Math. **69**, 441–470 (1995).

[LT01] C. T. Lin, E. Tadmor, L^1 *stability and error estimates for Hamilton–Jacobi solutions*, Num. Math., **87** (2001), 701–735.

[MSV93] R. Malladi, J. A. Sethian, B. C. Vemuri, *Shape modeling with front propagation: A level set approach*, Center for Pure and Applied Mathematics, Report PAM-589, Univ. of California, Berkeley, August 1993.

[ObSa15] A.M. Oberman, T. Salvador, *Filtered schemes for Hamilton-Jacobi equations: a simple construction of convergent accurate difference schemes.* Journal of Computational Physics., 284:367-388, 2015.

[OSe88] S. Osher, J. A. Sethian, *Fronts propagating with curvature dependent speed: Algorithms based on Hamilton-Jacobi formulation*, Journal of Computational Physics, **79**, pp. 12-49, 1988.

[OSh91] S. Osher, C.-W. Shu, *High-order essentially non-oscillatory schemes for Hamilton–Jacobi equations.* SIAM J. Numer. Anal. **28**, 907 (1991).

[Se85] J. A. Sethian, *Curvature and the evolution of fronts*, Commun. in Mathematical Physics, **101**, pp. 487-499, 1985.

[Se90] J. A. Sethian, *Numerical algorithms for propagating interfaces: Hamilton-Jacobi equations and conservation laws*, Journal of Differential Geometry, **31**, pp. 131-161, 1990.

[Se96] J.A. Sethian, Level Set Method. Evolving interfaces in geometry, fluid mechanics,computer vision, and materials science, Cambridge Monographs on Applied and Computational Mathematics, **3**, Cambridge University Press, Cambridge, 1996.

[Sh98] C.-W. Shu, *Essentially non-oscillatory and weighted essentially non-oscillatory schemes for hyperbolic conservation laws*, in "Advanced Numerical Approximation of Nonlinear Hyperbolic Equations (Cetraro, 1997)", Lecture Notes in Math., 1697, Springer-Verlag, Berlin, 1998.

[So85a] P. E. Souganidis, *Approximation schemes for viscosity solutions of Hamilton-Jacobi equations*, J. Diff. Eqs., **59** (1985), 1-43.

[So85b] P. E. Souganidis, *Product formulas and max-min representations (with applications to differential games) for viscosity solutions of Hamilton-Jacobi equations*, J. Nonlinear Anal. Theory Appl., **9** (1985) 217-257.

[T06] A. Tourin, *Splitting methods for Hamilton-Jacobi equations*, Numer. Methods Partial Differential Equations, **22** (2006), 381–396.

A. Technical results

For completeness we also give the proofs of the properties of the undivided differences and the binomial coefficients we used in the proof of Proposition 2.1.

Lemma A.1. *Let us assume $i \geq 1$ and write $f_h[\cdot]$ for the undivided difference of a function f. Then, it holds*

$$f_h[0,\dots,i] = \sum_{j=0}^{i-l} \binom{i-l}{j} (-1)^{i-l-j} f_h[j,\dots,j+l], \qquad \text{for } l = 0,\dots,i. \tag{A.1}$$

Moreover, we have that

$$\sum_{j=0}^{n} \binom{i}{j} (-1)^{i-j} = \begin{cases} \binom{i-1}{n}(-1)^{i-n} & \text{for } n < i \\ 0 & \text{for } n = i. \end{cases} \tag{A.2}$$

Proof. Let us start from the proof of (A.1) and let us proceed by induction on i.

Firstly, let us notice that for $l = i$ the identity is trivially satisfied, whence the case $i = 0$ follows directly. Then, for any $l = 0,\dots,i-1$, suppose that the statement holds for $i-1$ and for $i > 0$ let us compute,

$$\begin{aligned}
f_h[0,\dots,i] &= f_h[1,\dots,i] - f_h[0,\dots,i-1] && \text{by definition of } f_h[\cdot] \\
&= \sum_{j=0}^{i-l-1} \binom{i-l-1}{j} (-1)^{i-l-1-j} f_h[j+1,\dots,j+1+l] \\
&\quad - \sum_{j=0}^{i-l-1} \binom{i-l-1}{j} (-1)^{i-l-1-j} f_h[j,\dots,j+l] && \text{by inductive hyp.} \\
&= f_h[i-l,\dots,i] + (-1)^{i-l} f_h[0,\dots,l] \\
&\quad + \sum_{j=1}^{i-l-1} \binom{i-l-1}{j-1} (-1)^{i-l-j} f_h[j,\dots,j+l] \\
&\quad + \sum_{j=1}^{i-l-1} \binom{i-l-1}{j} (-1)^{i-l-j} f_h[j,\dots,j+l] \\
&= f_h[i-l,\dots,i] + (-1)^{i-l} f_h[0,\dots,l] \\
&\quad + \sum_{j=1}^{i-l-1} \binom{i-l}{j} (-1)^{i-l-j} f_h[j,\dots,j+l] && \binom{n-1}{k-1} + \binom{n-1}{k} = \binom{n}{k} \\
&= \sum_{j=0}^{i-l} \binom{i-l}{j} (-1)^{i-l-j} f_h[j,\dots,j+l],
\end{aligned}$$

as we wanted.

***Remark* A.1.** To simplify the notation we have stated the result for $f_h[0,\dots,i]$ but the proof clearly holds for $f_h[k,\dots,k+i]$, $\forall k$. In the second identity of the previous chain we have assumed this fact applying the inductive hypothesis on both terms.

Let us focus now on the second relation of the lemma (A.2) and proceed again by induction, but this time on $n : 0 \leq n < i$. For $n = 0$ we have $(-1)^i = (-1)^i$, than the identity holds. Suppose that (A.2) holds for $n-1 < i-1$ and compute

$$\begin{aligned}
\sum_{j=0}^{n}\binom{i}{j}(-1)^{i-j} &= \sum_{j=0}^{n-1}\binom{i}{j}(-1)^{i-j} + \binom{i}{n}(-1)^{1-n} \\
&= \binom{i-1}{n-1}(-1)^{i+1-n} + \binom{i}{n}(-1)^{i-n} && \text{by inductive hyp.} \\
&= \binom{i-1}{n-1}(-1)^{i+1-n} - \left[\binom{i-1}{n} + \binom{i-1}{n-1}\right](-1)^{i+1-n} \\
&= \binom{i-1}{n}(-1)^{i-n}.
\end{aligned}$$

For $n = i$ instead, from what we have just seen we can easily compute

$$\begin{aligned}
\sum_{j=0}^{i}\binom{i}{j}(-1)^{i-j} &= \sum_{j=0}^{i-1}\binom{i}{j}(-1)^{i-j} + (-1)^{i-i} \\
&= \binom{i-1}{i-1}(-1)^{i-i+1} + 1 \\
&= -1 + 1 = 0.
\end{aligned}$$

□

For the last result is better to prove first the following technical lemma.

Lemma A.2. *Let $x \in \mathbb{R}$ and $n \in \mathbb{N}$, $n \geq 1$. Then, for $t \in \mathbb{N}$, $t \geq 1$*

$$(x+n)^t = \sum_{j=0}^{t-1}(x+j)(x+n)^{t-j-1}(n)_j + (n)_t, \tag{A.3}$$

where $(x)_k = x(x-1)\cdots(x-k+1)$ is the falling factorial ($(x)_0 = 1$). Whence it follows directly that, for $i, t \in \mathbb{N} : i \geq 1$,

$$\sum_{n=0}^{i}\binom{i}{n}(-1)^{i-n}(x+n)^t = \begin{cases} 0 & \text{for } 0 \leq t < i \\ i! & \text{for } t = i. \end{cases} \tag{A.4}$$

Moreover, for $t > i$,

$$\sum_{n=0}^{i}\binom{i}{n}(-1)^{i-n}(x+n)^t = \sum_{j=0}^{\min\{t-i,i\}}(x+j)(i)_j \sum_{n=0}^{i-j}\binom{i-j}{n}(-1)^{i-n-j}((x+j)+n)^{t-1-j}.$$

Proof. Let us proceed again by induction, this time on t. For $t = 1$ we have $x+n = x+n$ so the identity is verified. Then suppose the (A.3) true for $t > 1$ and compute

$$\begin{aligned}
(x+n)^{t+1} &= (x+n)^t(x+n) \\
&= \left[\sum_{j=0}^{t-1}(x+j)(x+n)^{t-1-j}(n)_j + (n)_t\right](x+n) \qquad \text{by inductive hyp.} \\
&= \sum_{j=0}^{t-1}(x+j)(x+n)^{t-j}(n)_j + (n)_t(x+t+n-t) \\
&= \sum_{j=0}^{t}(x+j)(x+n)^{t-j}(n)_j + (n)_{t+1},
\end{aligned}$$

as we wanted.

Let us pass to the relation (A.4) and proceed by induction on $t < i$. For $t = 0$ the identity holds by Lemma A.1. Suppose that (A.4) holds for $t-1 < i-1$ and compute

$$\begin{aligned}
\sum_{n=0}^{i}\binom{i}{n}(-1)^{i-n}(x+n)^t &= \sum_{n=0}^{i}\binom{i}{n}(-1)^{i-n}\left[\sum_{j=0}^{t-1}(x+j)(x+n)^{t-1-j}(n)_j + (n)_t\right] \\
&= \sum_{j=0}^{t-1}(x+j)(i)_j\sum_{n=j}^{i}\binom{i-j}{n-j}(-1)^{i-n}(x+n)^{t-1-j} \\
&\quad + (i)_t\sum_{n=t}^{i}\binom{i-t}{n-t}(-1)^{i-n} \\
&= \sum_{j=0}^{t-1}(x+j)(i)_j\underbrace{\sum_{n=0}^{i-j}\binom{i-j}{n}(-1)^{i-n-j}((x+j)+n)^{t-1-j}}_{=0 \text{ by inductive hypothesis}} \\
&\quad + (i)_t\underbrace{\sum_{n=0}^{i-t}\binom{i-t}{n}(-1)^{i-n-t}}_{=0 \text{ by Lemma A.1}} = 0.
\end{aligned}$$

Finally, for $t = i$ and repeating the same computations we notice that the last term it is not null but

$$(i)_i\sum_{n=i}^{i}\binom{i-i}{n-i} = (i)_i = i(i-1)\cdots(i-i+1) = i!\,,$$

as we wanted.

For the last relation it suffices to repeat again the same expansions we have just done and notice that if $j > i$, then $(n)_j = 0$ for $n = 0, \ldots, i$, and moreover if $t - i < i$,

$$\sum_{n=0}^{i-j}\binom{i-j}{n}(-1)^{i-n-j}((x+j)+n)^{t-1-j} = 0 \qquad \text{per } t-i < j \le i,$$

for what we have seen in the previous case. □

Corollary A.3. *Let $f \in C^{r+1}([x_{j+k-r}, x_{j+k}])$ for $k \in \{0, \dots, r-1\}$, $r > 1$. Then, for $i = 1, \dots, r$ and $s = 0, \dots, i$,*

$$f_h[(k-r+s), \dots, (k-r+s+i)] = h^i f^{(i)}(x_j) + o(h^i). \tag{A.5}$$

Proof. Defining $x = k - r + s$, if we write the (A.1) for $l = 0$ we have

$$\begin{aligned} f_h[x, \dots, x+i] &= \sum_{n=0}^{i} \binom{i}{n} (-1)^{i-n} f_h[x+n] \\ &= \sum_{n=0}^{i} \binom{i}{n} (-1)^{i-n} f(x_j + (x+n)h). \end{aligned}$$

Then, developing with Taylor up to order r we can write

$$f(x_j + (x+n)h) = \sum_{t=0}^{r} f^{(t)}(x_j) \frac{[h(x+n)]^t}{t!} + o(h^r),$$

whence

$$\begin{aligned} f_h[x, \dots, x+i] &= \sum_{n=0}^{i} \binom{i}{n} (-1)^{i-n} \sum_{t=0}^{r} f^{(t)}(x_j) \frac{[h(x+n)]^t}{t!} + o(h^r) \\ &= \sum_{t=0}^{r} f^{(t)}(x_j) \frac{h^t}{t!} \sum_{n=0}^{i} \binom{i}{n} (-1)^{i-n} (x+n)^t + o(h^r) \\ &= h^i f^{(i)}(x_j) + \sum_{t=i+1}^{r} f^{(t)}(x_j) \frac{h^t}{t!} \sum_{n=0}^{i} \binom{i}{n} (-1)^{i-n} (x+n)^t + o(h^r) \\ &= h^i f^{(i)}(x_j) + o(h^i), \end{aligned}$$

having exploited (A.4) in the third identity of the chain. □